

# 数学 IB 演習 ( 第 6 回 ) の略解

## 目次

1. 問 1 の解答	1
2. 問 1 について	2
3. 問 1 の解答についてと	2
4. 問 2 の解答	4
5. 多変数関数の Taylor 展開について	5
6. Taylor 多項式の特徴付け	9
7. 全微分可能とは (数学的な定義について)*	12
8. 臨界点とは	14
9. 極値の判定法について	17
10. 問 3 の解答	21
11. Jacobi 行列とは	22
12. 逆関数定理とは	24
13. 写像の微分とは	26
14. 合成写像の微分則について	29

### 1. 問 1 の解答

第 5 回の問 2 のところでも見たように、一般に、級数  $\sum_{n=1}^{\infty} a_n$  に対して、

$$M = \lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}},$$

あるいは、すべての自然数  $n \in \mathbb{N}$  に対して、 $a_n \neq 0$  であるときには、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

により定まる級数  $\sum_{n=1}^{\infty} |a_n|$  の「仮想的な公比」 $M$  を考えるとき、

$$\begin{cases} M < 1 \implies \sum_{n=1}^{\infty} a_n \text{ は絶対収束する.} \\ M > 1 \implies \sum_{n=1}^{\infty} a_n \text{ は発散する.} \end{cases}$$

(1)

となることが分かります。そこで、この「級数の収束判定法」をもとに調べてみます。

$$(1) \quad a_n = \frac{n^a}{n!} \text{ とすると,}$$

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \frac{\frac{(n+1)^a}{(n+1)!}}{\frac{n^a}{n!}} \\ &= \frac{1}{n+1} \cdot \left(1 + \frac{1}{n}\right)^a \end{aligned}$$

となるので、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 0 < 1$$

となることが分かります。よって、(1) 式より、与えられた級数は収束することが分かります。

$$(2) \quad a_n = \left(\frac{n}{n+1}\right)^{n^2} \text{ とすると,}$$

$$\begin{aligned} |a_n|^{\frac{1}{n}} &= \left(\frac{n}{n+1}\right)^n \\ &= \frac{1}{\left(1 + \frac{1}{n}\right)^n} \end{aligned}$$

となることが分かります。いま、

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

となることに注意すると、\*1)

$$M = \lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} = \frac{1}{e} < 1$$

\*1) このことは、例えば、 $\log$  をとって、

$$\log \left(1 + \frac{1}{n}\right)^n = \frac{\log \left(1 + \frac{1}{n}\right) - \log 1}{\frac{1}{n}}$$

と書き換えてから、 $n \rightarrow \infty$  の極限を考えると、これは丁度「 $\log x$  という関数の  $x = 1$  での微分になっている」ことから分かります。

となることが分かります。<sup>\*2)</sup> よって、(1) 式より、与えられた級数は収束することが分かります。

$$(3) \quad a_n = \frac{(a+1)(2a+1)\cdots(na+1)}{(b+1)(2b+1)\cdots(nb+1)} \text{ とすると,}$$

$$a_{n+1} = \frac{(n+1)a+1}{(n+1)b+1} \cdot a_n$$

と書けるので、

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{a + \frac{1}{n+1}}{b + \frac{1}{n+1}}$$

となることが分かります。したがって、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{a}{b}$$

となることが分かります。よって、(1) 式より、 $\frac{a}{b} < 1$  ならば、与えられた級数は収束し、 $\frac{a}{b} > 1$  ならば、与えられた級数は発散することが分かります。また、 $a = b$  のときには、 $M = 1$  となるので、(1) 式からは判定できませんが、この場合には、勝手な自然数  $n \in \mathbb{N}$  に対して、 $a_n = 1$  となるので、与えられた級数は発散することが分かります。

以上より、

$$\begin{cases} a < b \implies \text{与えられた級数は収束する。} \\ a \geq b \implies \text{与えられた級数は発散する。} \end{cases}$$

となることが分かります。

$$(4) \quad a_n = \frac{1}{n+1} \cdot \log\left(1 + \frac{1}{n}\right) \text{ とします。第 2 回の問 3 で見たように、} \log(1+x) \text{ は,}$$

$$\begin{aligned} \log(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \\ &= x \cdot \left(1 - \frac{x}{2} + \frac{x^2}{3} - \cdots\right) \end{aligned}$$

というように Taylor 展開されることに注意すると、

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \left| \frac{\frac{1}{n+2} \cdot \log\left(1 + \frac{1}{n+1}\right)}{\frac{1}{n+1} \cdot \log\left(1 + \frac{1}{n}\right)} \right| \\ &= \left| \frac{\frac{1}{n+2} \cdot \frac{1}{n+1} \cdot \left(1 - \frac{1}{2(n+1)} + \cdots\right)}{\frac{1}{n+1} \cdot \frac{1}{n} \cdot \left(1 - \frac{1}{2n} + \cdots\right)} \right| \end{aligned}$$

と表わせることが分かります。したがって、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$$

となることが分かりますから、この場合も (1) 式で

<sup>\*2)</sup>  $e = 2.718\cdots$  でした。

は判定することができません。

そこで、直接  $a_n$  の大きさを見積もることを考えてみます。いま、 $x \geq 0$  のとき、

$$\log(1+x) \leq x$$

となることに注意すると、<sup>\*3)</sup>  $a_n$  は、

$$0 < a_n = \frac{1}{n+1} \cdot \log\left(1 + \frac{1}{n}\right) \leq \frac{1}{n(n+1)}$$

というように見積もれることが分かります。したがって、部分  $S_N$  を考えると、

$$\begin{aligned} S_N &= \sum_{n=1}^N \frac{1}{n+1} \cdot \log\left(1 + \frac{1}{n}\right) \\ &\leq \sum_{n=1}^N \frac{1}{n(n+1)} \\ &= \sum_{n=1}^N \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{N+1} \\ &< 1 \end{aligned}$$

となることが分かりますから、 $\{S_N\}_{N=1,2,\dots}$  は「頭打ち」になる単調増加数列になることが分かります。よって、与えられた級数は収束することが分かります。

## 2. 問 1 について

問 1 は、皆さんに、(1) 式という「級数の収束判定法」に慣れてもらおうと思って出題しました。(3) で  $a = b$  の場合や、(4) などの場合には、「仮想的な公比」 $M$  は  $M = 1$  となってしまうので、この判定法では判定できないこととなります。このような場合には、積分の値や大きさの分かる級数と「大きさ比べ」をするなどして、個別に対処する必要があることに注意して下さい。

## 3. 問 1 の解答について

第 1 回の解答の最初のところでも注意しましたが、それぞれの問題に対する解答は一通りではありません。そこで、問 1 の解答について、少し補足しておくことにします。

まず、(2) についてですが、上で挙げた解答では、

<sup>\*3)</sup> これは、例えば、 $f(x) = \log(1+x) - x$  として、 $f(x)$  の増減を調べてみることで分かります。このように、与えられた関数  $f(x)$  に対して、 $f(x)$  の Taylor 展開の最初の何項かを持つてくることで、 $f(x)$  の大きさを見積もれることがよくあります。

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$$

となることから、 $M = \frac{1}{e} < 1$  となることを結論しました。ここで、仮に、 $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$  となることが分からなかったとしても、 $\left(1 + \frac{1}{n}\right)^n$  を二項展開して、最初の二項に注目すれば、

$$\left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \cdots + \left(\frac{1}{n}\right)^n \geq 2$$

となることが分かりますから、

$$|a_n|^{\frac{1}{n}} \leq \frac{1}{2}$$

というように評価できることが分かります。そこで、この評価式を用いて、部分積  $\{S_N\}_{N=1,2,\dots}$  が「頭打ち」となる単調増加数列になることを示しても構いません。

また、この (2) の例では、

$$M = \lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}}$$

という表示を用いる方が簡単ですが、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

という表示を用いても、例えば、次のように計算することができます。いま、

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \frac{\left(\frac{n+1}{n+2}\right)^{(n+1)^2}}{\left(\frac{n}{n+1}\right)^{n^2}} \\ &= \frac{\left(1 + \frac{1}{n}\right)^{n^2}}{\left(1 + \frac{1}{n+1}\right)^{(n+1)^2}} \end{aligned}$$

と書き直してから、両辺の  $\log$  を取ると、

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

となることが分かります。ここで、

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

というように Taylor 展開されることに注意すると、 $n$  が大きいときに、

$$\begin{aligned} \log \left| \frac{a_{n+1}}{a_n} \right| &= n^2 \left\{ \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \cdots \right\} \\ &\quad - (n+1)^2 \left\{ \frac{1}{n+1} - \frac{1}{2(n+1)^2} + \frac{1}{3(n+1)^3} - \cdots \right\} \end{aligned}$$

数学 IB 演習

$$\begin{aligned} &= \left\{ n - \frac{1}{2} + \frac{1}{3n} - \cdots \right\} \\ &\quad - \left\{ (n+1) - \frac{1}{2} + \frac{1}{3(n+1)} - \cdots \right\} \\ &= -1 + \frac{1}{3n} + \cdots - \frac{1}{3(n+1)} + \cdots \end{aligned}$$

となることが分かるので、

$$\lim_{n \rightarrow \infty} \log \left| \frac{a_{n+1}}{a_n} \right| = -1$$

となることが分かります。これより、

$$M = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{e}$$

となることが分かります。

次に、(4) についてですが、上の解答で挙げたように、

$$a_n \leq \frac{1}{n(n+1)}$$

と評価した後で、さらに、 $\frac{1}{n(n+1)}$  を、

$$\frac{1}{n(n+1)} \leq \frac{1}{n^2}$$

と評価して、

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < +\infty$$

となることから、部分積  $\{S_N\}_{N=1,2,\dots}$  が「頭打ち」になる単調増加数列になることを結論しても構いません。

また、以前、この数学 IB 演習をやっていたときに、山村正樹さんが、(4) について、面白い解答をしていたので紹介します。いま、与えられた級数を、

$$\begin{aligned} &\sum_{n=1}^{\infty} \frac{1}{n+1} \log \left(1 + \frac{1}{n}\right) \\ &= \sum_{n=1}^{\infty} \frac{1}{n+1} \{\log(n+1) - \log n\} \end{aligned}$$

と書き換えてみます。このとき、右辺の式をじっと睨んで、 $x_n = \log n$  とおいてみると、

$$\frac{1}{n+1} \{\log(n+1) - \log n\} = e^{-x_{n+1}} (x_{n+1} - x_n)$$

と書き換えられることが分かります。そこで、

$$f(x) = e^{-x}$$

のグラフを描いて「面積比べ」をしてみると、勝手な自然数  $n \in \mathbb{N}$  に対して、

$$e^{-x_{n+1}} (x_{n+1} - x_n) \leq \int_{x_n}^{x_{n+1}} e^{-x} dx$$

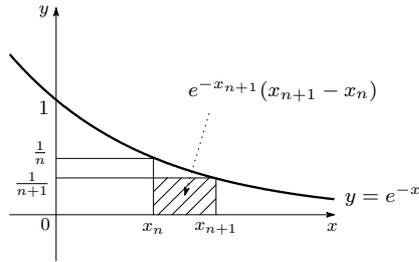


図1 それぞれの自然数  $n$  に対して、 $\int_{x_n}^{x_{n+1}} e^{-x} dx$  は  $e^{-x_{n+1}}(x_{n+1} - x_n)$  より大きい。

となることが分かります (図1を参照)。したがって、与えられた級数の部分和は、

$$\begin{aligned} S_N &= \sum_{n=1}^N \frac{1}{n+1} \log\left(1 + \frac{1}{n}\right) \\ &= \sum_{n=1}^N \frac{1}{n+1} \{\log(n+1) - \log n\} \\ &= \sum_{n=1}^N e^{-x_{n+1}}(x_{n+1} - x_n) \\ &\leq \sum_{n=1}^N \int_{x_n}^{x_{n+1}} e^{-x} dx \\ &\leq \int_0^{x_{N+1}} e^{-x} dx \\ &\leq \int_0^{\infty} e^{-x} dx = 1 \end{aligned}$$

と評価できることが分かります。これより、部分和  $\{S_N\}_{N=1,2,\dots}$  は「頭打ち」になる単調増加数列であることが分かりますから、与えられた級数は収束することが分かります。

#### 4. 問2の解答

(1)  $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$  を計算してみると、

$$\begin{cases} \frac{\partial f}{\partial x} = 2x - y - 1 \\ \frac{\partial f}{\partial y} = -x + 4y - 2 \end{cases}$$

となることが分かります。そこで、 $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$  を解いてみると、 $f$  の臨界点は、

$$(x, y) = \left(\frac{6}{7}, \frac{5}{7}\right)$$

であることが分かります。

また、 $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial y^2}$  を計算してみると、

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = 2 \\ \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -1 \\ \frac{\partial^2 f}{\partial y^2} = 4 \end{cases}$$

となるので、ヘッシアン  $H_f(x, y)$  は、

$$H_f(x, y) = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}$$

となることが分かります。よって、臨界点  $(x, y) = \left(\frac{6}{7}, \frac{5}{7}\right)$  でのヘッシアンは、

$$H_f\left(\frac{6}{7}, \frac{5}{7}\right) = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}$$

となることが分かります。

(2) 同様に、 $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$  を計算してみると、

$$\begin{cases} \frac{\partial f}{\partial x} = 3x^2 - 3y \\ \frac{\partial f}{\partial y} = 3y^2 - 3x \end{cases}$$

となることが分かります。そこで、 $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$  を解いてみると、 $f$  の臨界点は、

$$(x, y) = (0, 0), (1, 1)$$

であることが分かります。

また、

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = 6x \\ \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -3 \\ \frac{\partial^2 f}{\partial y^2} = 6y \end{cases}$$

となるので、ヘッシアン  $H_f(x, y)$  は、

$$H_f(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$$

となることが分かります。よって、臨界点  $(x, y) = (0, 0), (1, 1)$  でのヘッシアンは、それぞれ、

$$H_f(0, 0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix},$$

$$H_f(1, 1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

となることが分かります。

(3) 同様に、 $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$  を計算してみると、

$$\begin{cases} \frac{\partial f}{\partial x} = 4x^3 - 4y \\ \frac{\partial f}{\partial y} = -4x + 4y \end{cases}$$

となることが分かります。そこで、 $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$  を解いてみると、 $f$  の臨界点は、

$$(x, y) = (0, 0), (1, 1), (-1, -1)$$

であることが分かります。

また、

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = 12x^2 \\ \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -4 \\ \frac{\partial^2 f}{\partial y^2} = 4 \end{cases}$$

となるので、ヘッシアン  $H_f(x, y)$  は、

$$H_f(x, y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 4 \end{pmatrix}$$

となることが分かります。よって、臨界点  $(x, y) = (0, 0), (1, 1), (-1, -1)$  でのヘッシアンは、それぞれ、

$$H_f(0, 0) = \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix},$$

$$H_f(1, 1) = H_f(-1, -1) = \begin{pmatrix} 12 & -4 \\ -4 & 4 \end{pmatrix}$$

となることが分かります。

- (4) 同様に、 $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$  を計算してみると、

$$\begin{cases} \frac{\partial f}{\partial x} = (1 - xy - y^2)e^{-xy} \\ \frac{\partial f}{\partial y} = (1 - x^2 - xy)e^{-xy} \end{cases}$$

となることが分かります。そこで、 $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$  を解いてみると、 $f$  の臨界点は、

$$(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$$

であることが分かります。

また、

$$\begin{cases} \frac{\partial^2 f}{\partial x^2} = (xy^2 + y^3 - 2y)e^{-xy} \\ \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = (x^2y + xy^2 - 2x - 2y)e^{-xy} \\ \frac{\partial^2 f}{\partial y^2} = (x^2y + x^3 - 2x)e^{-xy} \end{cases}$$

となるので、ヘッシアン  $H_f(x, y)$  は、

$$H_f(x, y) = e^{-xy} \begin{pmatrix} xy^2 + y^3 - 2y & x^2y + xy^2 - 2x - 2y \\ x^2y + xy^2 - 2x - 2y & x^2y + x^3 - 2x \end{pmatrix}$$

となることが分かります。よって、臨界点  $(x, y) =$

$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$  でのヘッシアンは、それぞれ、

$$H_f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}e} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix},$$

$$H_f\left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}e} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

となることが分かります。

## 5. 多変数関数の Taylor 展開について

これまでも何度か述べてきましたが、「理解の易しい多項式の力を借りて、一般の関数の様子を理解することを試みる」ということが微積分学における最も基本的な考え方のひとつです。より具体的には、一般の関数を「多項式の姿」に「化かし」、化かされた「多項式の姿」を通して関数の様子を理解することを試みるというわけです。実際、一変数関数に対しては、第2回のところで、一般の関数を「おつりの項」付きで「多項式の姿」に「化かす」ことができるということを見ました。また、第3回の間4のところでは、「接線を描いて関数の増減を調べる」ということや「極値の判定」ということが、「化か」された「一次式の姿」や「二次式の姿」を眺めることで、一般の関数の大まかな様子を調べているということであると解釈できることを見ました。

こうした考え方は、一変数関数に対してだけでなく、多変数関数に対しても自然に拡張することができ、一変数関数のときと同様の「戦略」を通して、多変数関数の大まかな様子を理解することができるようになります。そこで、以下では、こうした事柄を順番に見ていこうと思いますが、手始めに、この節では、一般の多変数関数を「多項式の姿」に「化かす」という問題を考えてみることにします。考え方の本質は、二変数関数のときにすべて現われていますから、第5回るときと同様に、以下では、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対して説明してみることになります。<sup>\*4)</sup>

そこで、まず、一変数関数のときと同様に、二変数の滑らかな関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  が、勝手にひとつ与えられているとして、関数  $f(x, y)$  が、

$$f(x, y) = \sum_{k, m=0}^{\infty} c_{k, m} x^k y^m$$

<sup>\*4)</sup> この二変数関数の場合を良く理解すれば、二変数以上の多変数関数の場合にもどうすれば良いのかが分かるのではないかと思います。

$$= c_{0,0} + c_{1,0}x + c_{0,1}y + c_{2,0}x^2 + c_{1,1}xy + c_{0,2}y^2 + \dots \quad (2)$$

というように「多項式の姿」に「化ける」としたら、どのような姿に「化ける」のが一番もっともらしいのかということを考えてみることにします。<sup>\*5)</sup> そこで、「滑らかな関数に対しては、高階の偏導関数は偏微分を行なう変数の順番によらない」ということに注意して、一変数関数のときと同様に、(2) 式の両辺を  $x$  に関して  $k$  回、 $y$  に関して  $m$  回偏微分してから、 $(x, y) = (0, 0)$  としてみると、

$$\frac{\partial^{k+m} f}{\partial x^k \partial y^m}(0, 0) = k!m! \cdot c_{k,m} \quad (3)$$

となることが分かります。したがって、(3) 式から、

$$c_{k,m} = \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(0, 0)$$

となることが分かります。以上から、二変数の滑らかな関数  $f(x, y)$  が「多項式の姿」に「化ける」としたら、

$$\begin{aligned} f(x, y) &= \sum_{k,m=0}^{\infty} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(0, 0) x^k y^m \\ &= f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + \frac{\partial^2 f}{\partial x \partial y}(0, 0)xy \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0)y^2 + \dots \end{aligned} \quad (4)$$

という姿に「化ける」のが一番もっともらしいということが分かりました。

こうして二変数関数  $f(x, y)$  が「化ける」べき「多項式の姿」に「当たり」がつかしました。ただし、(4) 式のように、いきなり「次数が無限大の多項式の姿」に「化かす」ことを考えると、一変数関数のときと同様に、一般には、(4) 式の等号が成り立つとは限らないというような問題も出てきますから、状況をより良く理解するためには、剰余項付きで「次数が有限の多項式の姿」に「化かす」ことが大切です。

そこで、一変数関数のときと同様の考察を行ないたいわけですが、第 2 回の問 4 のところでも触たように、本質的なアイデアは一変数関数に対する Taylor 展開の考察でほぼ尽きています。すなわち、勝手な滑らかな関数  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  と、勝手な自然数  $n \in \mathbb{N}$  に対して、

$$\begin{aligned} \varphi(t) &= \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + \dots \\ &\quad + \frac{\varphi^{(n)}(0)}{n!}t^n + R_n(t) \end{aligned}$$

というように、剰余項付きで「次数が有限の多項式の姿」に「化かす」ことができるということではほぼ尽きています。ここで、剰余項  $R_n(t)$  は、第 2 回の問 4 のところで見たとように、微積分学の基本定理から始めて部分積分を繰り返すという方針を取ると、

$$R_n(t) = \frac{1}{n!} \int_0^t (t-s)^n \varphi^{(n+1)}(s) ds \quad (5)$$

という表示を持つことが分かるのでした。さらに、(5) 式の右辺の積分に対して、「積分に関する平均値の定理」を適用してみると、

$$R_n(t) = \frac{\varphi^{(n+1)}(\theta)}{(n+1)!} t^{n+1} \quad (6)$$

となるような実数  $\theta \in \mathbb{R}$  が  $0$  と  $t$  の間に存在することが分かるのでした。<sup>\*6)</sup> 特に、 $t = 1$  としてみると、勝手な滑らかな関数  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  と、勝手な自然数  $n \in \mathbb{N}$  に対して、

$$\begin{aligned} \varphi(1) &= \varphi(0) + \varphi'(0) + \frac{\varphi''(0)}{2!} + \dots \\ &\quad + \frac{\varphi^{(n)}(0)}{n!} + \frac{\varphi^{(n+1)}(\theta)}{(n+1)!} \end{aligned} \quad (7)$$

となるような実数  $\theta \in \mathbb{R}$  が  $0$  と  $1$  の間に存在することが分かります。そこで、ここでは、この (7) 式をもとにして、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対する「Taylor の定理」を導くことにします。

そこで、いま、 $(a, b) \in \mathbb{R}^2$  という点を、勝手にひとつ取ってきて、「 $f(a, b)$  という値がどのような表示を持つのか」ということを考察してみることにします。そのために、以下では、 $(a, b) \in \mathbb{R}^2$  はひとつ値の定まった「定数」であると考えて議論を進めることにします。これでは抽象的で分かりにくいと思われる方は、 $(a, b)$  のところに、 $(a, b) = (1, 2)$  などの具体的な数を代入して考えてみて下さい。

そこで、問題を一変数関数の場合に帰着するために、 $(0, 0), (a, b) \in \mathbb{R}^2$  という二点を結ぶ直線

$$c(t) = (ta, tb)$$

を考えると、この直線上での関数  $f(x, y)$  の値を  $\varphi(t)$  と書くことにします (図 2 を参照)。すなわち、

\*5) ここで、添字に「 $l$ 」という文字を使ってしまうと、後で「 $l$ 」を「 $1$ 」と読み間違えてしまうという余計な混乱を起こすといけないので、二変数の多項式の次数を表わす添字として、 $(k, l)$  ではなく、 $(k, m)$  を用いることにしました。

\*6) 第 3 回の問 4 のところで見たとように、上手い関数に対してロルの定理を適用するという方針を取ることで、直接、(6) 式の表示を得ることもできます。

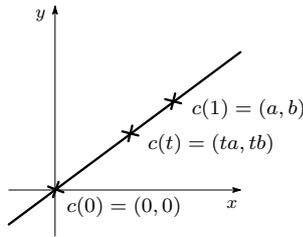


図 2  $\mathbb{R}^2$  上の二点  $(0, 0)$ ,  $(a, b)$  を結ぶ直線  $c(t)$  上で関数  $f(x, y)$  の値を考える.

$$\begin{aligned}\varphi(t) &= f(c(t)) \\ &= f(ta, tb)\end{aligned}\quad (8)$$

という一変数関数を補助的に考えてみることにします.\*7) このとき, (8) 式で与えられる  $\varphi(t)$  という一変数関数に対して, 上の (7) 式を当てはめるとどうなるのかということを考えてみます. そのためには, (8) 式により与えられる関数  $\varphi(t)$  に対して,  $\varphi'(t)$ ,  $\varphi''(t)$ , ... などを求める必要があります. そこで, 少し様子を探ってみるために, まずは,  $\varphi'(t)$  について考えてみることにします.

いま,  $t_0 \in \mathbb{R}$  を, 勝手にひとつ取ってきたときに, 関数  $\varphi(t)$  の  $t = t_0$  における微分係数  $\varphi'(t_0)$  とは, 定義により,

$$\varphi'(t_0) = \lim_{h \rightarrow 0} \frac{\varphi(t_0 + h) - \varphi(t_0)}{h}$$

ということでした. 我々の場合,  $\varphi(t)$  は (8) 式で与えられていますから,

$$\varphi'(t_0) = \lim_{h \rightarrow 0} \frac{f((t_0 + h)a, (t_0 + h)b) - f(t_0a, t_0b)}{h}$$

というように表わせることが分かります. ここで,

$$(x_0, y_0) = (t_0a, t_0b)$$

と書くことにすると,

$$\varphi'(t_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, x_0 + hb) - f(x_0, y_0)}{h} \quad (9)$$

となりますが, (9) 式の右辺は, 第 5 回の問 1 のところで,  $C^1$  級の関数の接平面を求めるために「 $(x, y) = (x_0, y_0)$  における関数  $f(x, y)$  の  $\mathbf{v} = (a, b)$  方向の

\*7) 変数  $t$  を時間であると考えて, 時刻  $t = 0$  で原点  $(0, 0)$  にいた点粒子が, 等速直線運動をして時刻  $t = 1$  で  $(a, b)$  という点に至ると考えるとイメージしやすいかもしれません. このとき,  $\varphi(t)$  は, 時刻  $t$  で点粒子がいる場所での関数  $f(x, y)$  の値ということになります.

接線の傾き」を考察したときに現われた式と同じ式であることに注意します. したがって, そのときの結果から,

$$\begin{aligned}\varphi'(t_0) &= \frac{\partial f}{\partial x}(x_0, y_0) \cdot a + \frac{\partial f}{\partial y}(x_0, y_0) \cdot b \\ &= \frac{\partial f}{\partial x}(t_0a, t_0b) \cdot a + \frac{\partial f}{\partial y}(t_0a, t_0b) \cdot b\end{aligned}$$

となることが分かります. ここで,  $t_0 \in \mathbb{R}$  は勝手な値で良かったので, 「変数」らしく, 再び,  $t_0$  を  $t$  と表わすことにすると,

$$\varphi'(t) = \frac{\partial f}{\partial x}(ta, tb) \cdot a + \frac{\partial f}{\partial y}(ta, tb) \cdot b \quad (10)$$

となることが分かりました. これを関数  $f(x, y)$  だけを用いて表わせば,

$$\frac{d}{dt} f(ta, tb) = \frac{\partial f}{\partial x}(ta, tb) \cdot a + \frac{\partial f}{\partial y}(ta, tb) \cdot b \quad (11)$$

ということになります.\*8)

そこで, 次に,  $\varphi'(t)$  に対する (10) 式という表示を用いて,  $\varphi''(t)$  を求めることを考えてみます. そのためには,

$$\frac{d}{dt} \left\{ \frac{\partial f}{\partial x}(ta, tb) \right\}$$

などを求める必要がありますが, 関数  $\frac{\partial f}{\partial x}$  を  $f$  だと思って,  $f \rightsquigarrow \frac{\partial f}{\partial x}$  と置き換えて, (11) 式を適用すると,

$$\frac{d}{dt} \left\{ \frac{\partial f}{\partial x}(ta, tb) \right\} = \frac{\partial^2 f}{\partial x^2}(ta, tb) \cdot a + \frac{\partial^2 f}{\partial y \partial x}(ta, tb) \cdot b$$

となることが分かります. 同様にして,

$$\frac{d}{dt} \left\{ \frac{\partial f}{\partial y}(ta, tb) \right\}$$

を求めると,

$$\frac{d}{dt} \left\{ \frac{\partial f}{\partial y}(ta, tb) \right\} = \frac{\partial^2 f}{\partial x \partial y}(ta, tb) \cdot a + \frac{\partial^2 f}{\partial y^2}(ta, tb) \cdot b$$

となることが分かりますから, 結局,

$$\begin{aligned}\varphi''(t) &= \frac{\partial^2 f}{\partial x^2}(ta, tb) \cdot a^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(ta, tb) \cdot ab \\ &\quad + \frac{\partial^2 f}{\partial y^2}(ta, tb) \cdot b^2\end{aligned}\quad (12)$$

\*8) ここで, chain rule と呼ばれる多変数関数の場合の「合成関数の微分則」をご存じの方は, chain rule を用いて (11) 式を得ただと考えてもらっても構いません. 実際, 上の議論は chain rule に対する議論の特別な場合になっているわけです. なお, chain rule については, 14 節で説明されていますので, 興味のある方は, そちらも参照してみてください.

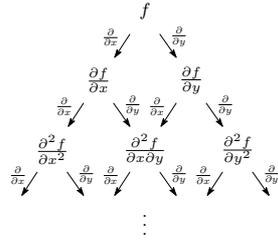


図3  $t$  で微分する毎に現われる関数のパターンは、二項係数に対するパスカルの三角形と同じパターンになっている。

となることが分かります。<sup>\*9)</sup>

以下、同様に、 $\varphi'''(t)$  くらいまで計算してみると、一般に、勝手な自然数  $n \in \mathbb{N}$  に対して、

$$\varphi^{(n)}(t) = \sum_{k=0}^n \frac{n!}{k!(n-k)!} \frac{\partial^n f}{\partial x^k \partial y^{n-k}}(ta, tb) \cdot a^k b^{n-k} \quad (13)$$

となることが予想できます。例えば、 $t$  で微分する毎にどのような関数が現われるのかということ、図3のように表わして考えると、ちょうど二項係数に対するパスカルの三角形と同じパターンになっていることが分かります。この点に注意すると、(13) 式の右辺の係数に二項係数が表われることが納得できるかもしれません。このように(13) 式が予想できてしまえば、後は、数学的帰納法を用いて、実際に(13) 式が成り立っていることを証明することができます。<sup>\*10)</sup> そこで、これらの結果を(7) 式に当てはめてみると、 $f(a, b)$  という値に対して、

$$\begin{aligned} f(a, b) &= \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(0, 0) a^k b^m + R_n \\ &= f(0, 0) + \frac{\partial f}{\partial x}(0, 0) a + \frac{\partial f}{\partial y}(0, 0) b \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0) a^2 + \frac{\partial^2 f}{\partial x \partial y}(0, 0) ab \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0) b^2 + \dots + R_n \end{aligned}$$

\*9) ここで、 $f(x, y)$  は滑らかな関数なので、第5回の問1のところで見たように、

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

となることを用いました。

\*10) 皆さん、確かめてみてください。ちなみに、(13) 式は、

$$\varphi^{(n)}(t) = \left\{ \left( a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y} \right)^n f \right\} (ta, tb)$$

というように表わすこともできます。皆さんの中にも、この形の方が「見やすい」と思われる方がいるかもしれません。

$$R_n = \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \frac{\partial^{n+1} f}{\partial x^k \partial y^m}(0a, 0b) a^k b^m$$

という表示が得られることが分かりました。

上では、 $(a, b) \in \mathbb{R}^2$  はひとつ値の定まった「定数」であると考えて議論しましたが、 $(a, b) \in \mathbb{R}^2$  という値は何でも良かったので、それぞれの点  $(a, b) \in \mathbb{R}^2$  に対して、上の議論をそれぞれ行なったのだと考えてみます。すると、 $(a, b)$  の値に応じて、一般には、 $\theta$  の値も変わり得ますから、それぞれの点  $(a, b)$  に対する  $\theta$  の値を  $\theta_{(a,b)}$  と書き替えることで、上の展開式が成り立っていることとなります。このことに注意して、 $(a, b)$  という文字を、改めて  $(x, y)$  と書き直すことにすると、結局、勝手な点  $(x, y) \in \mathbb{R}^2$  に対して、

$$\begin{aligned} f(x, y) &= \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(0, 0) x^k y^m \\ &\quad + R_n(x, y) \\ &= f(0, 0) + \frac{\partial f}{\partial x}(0, 0) x + \frac{\partial f}{\partial y}(0, 0) y \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(0, 0) x^2 + \frac{\partial^2 f}{\partial x \partial y}(0, 0) xy \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(0, 0) y^2 + \dots + R_n(x, y) \end{aligned}$$

$$\begin{aligned} R_n(x, y) &= \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \frac{\partial^{n+1} f}{\partial x^k \partial y^m}(\theta_{(x,y)}x, \theta_{(x,y)}y) x^k y^m \end{aligned}$$

となるような実数  $\theta_{(x,y)} \in \mathbb{R}$  が0と1の間に存在することが分かりました。こうして(7) 式をもとにして、二変数関数に対する Taylor の定理を導くことができました。

ここで、 $(\theta_{(x,y)}x, \theta_{(x,y)}y)$  は  $(0, 0)$  と  $(x, y)$  を結ぶ線分上の点であることに注意して、関数の値を考えた点を  $p = (x, y) \in \mathbb{R}^2$ 、Taylor 展開の中心点である原点を  $o = (0, 0) \in \mathbb{R}^2$  と書くことにすると、勝手な点  $p = (x, y) \in \mathbb{R}^2$  に対して、

$$\begin{aligned} f(p) &= \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(o) x^k y^m \\ &\quad + R_n(p) \end{aligned}$$

$$R_n(p) = \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \frac{\partial^{n+1} f}{\partial x^k \partial y^m}(q) x^k y^m$$

となるような点  $q \in \mathbb{R}^2$  が原点  $o$  と点  $p$  を結ぶ線分上に存在するということ、Taylor の定理の主張を少

しだけコンパクトな形で表現することができます。このような「幾何学的な表現」を用いると、 $(x, y)$  というような点の座標を書かずに済ませることができるので、一般の多変数関数に対する Taylor の定理も比較的スッキリした形で表現することができます。

また、より一般に、 $p_0 = (x_0, y_0) \in \mathbb{R}^2$  のまわりでの Taylor 展開を考えれば、 $(x_0, y_0), (x, y) \in \mathbb{R}^2$  という二点を結ぶ直線を考えて、その直線上での関数  $f(x, y)$  の値を与える

$$\varphi(t) = f(x_0 + t(x - x_0), y_0 + t(y - y_0)) \quad (14)$$

という一変数関数を補助的に考えれば良いということになります。すなわち、(8) 式の代わりに、(14) 式に対して、同じ議論を繰り返せば良いわけです。すると、この場合には、勝手な点  $p = (x, y) \in \mathbb{R}^2$  に対して、

$$f(p) = \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(p_0) (x-x_0)^k (y-y_0)^m + R_n(p) \quad (15)$$

$$R_n(p) = \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(q) (x-x_0)^k (y-y_0)^m$$

となるような点  $q \in \mathbb{R}^2$  が点  $p_0$  と点  $p$  を結ぶ線分上に存在するという主張が得られることが分かります。<sup>\*11)</sup> 一変数関数のときと同様に、(15) 式の右辺に現われる多項式

$$P_n(x, y) = \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(p_0) (x-x_0)^k (y-y_0)^m$$

を、点  $p_0 = (x_0, y_0)$  における関数  $f(x, y)$  の Taylor 多項式と呼びます。

## 6. Taylor 多項式の特徴付け

さて、第 3 回の問 3 と問 4 のところでは、一変数関数の場合に、滑らかな関数  $f: \mathbb{R} \rightarrow \mathbb{R}$  に対して、 $n \in \mathbb{N}, a \in \mathbb{R}$  として、 $x = a$  における関数  $f(x)$  の Taylor 多項式

$$P_n(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!} (x-a)^n$$

は、 $k = 0, 1, 2, \dots, n$  に対して、

\*11) 皆さん、確かめてみて下さい。

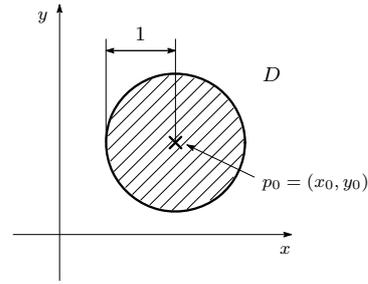


図 4 点  $p_0$  を中心とする半径 1 の円板  $D$  を考える。

$$\lim_{x \rightarrow a} \frac{f(x) - P_n(x)}{(x-a)^k} = 0 \quad (16)$$

が成り立つような  $n$  次の多項式として一意的に特徴付けられることを見ました。また、この特徴付けを用いると、 $x = a$  における Taylor 多項式  $P_n(x)$  は、 $n$  次の多項式の中で、「 $|x-a| \ll 1$  のとき、関数  $f(x)$  を最も良く近似する多項式」であることが分かるのでした。すなわち、 $x = a$  における Taylor 多項式  $P_n(x)$  とは、「 $x = a$  の近く」だけを考えてときに、関数  $f(x)$  が  $n$  次の多項式の中で自分自身に最も良く似た「姿」に「化け」ようとした「姿」であると解釈することができます。また、このことが、Taylor 多項式  $P_n(x)$  の様子を調べることにより、「 $x = a$  の近く」での関数  $f(x)$  の大まかな様子が理解できるということの理論的な根拠を与えているのでした。

そこで、ここでは、多変数関数の Taylor 多項式に対しても、全く同様の考察を行なうことができるということを確認してみることにします。5 節と同様に、話を具体的にするために、以下では、二変数関数の場合に説明することにします。<sup>\*12)</sup>

いま、 $n \in \mathbb{N}, p_0 = (x_0, y_0) \in \mathbb{R}^2$  として、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式を、

$$P_n(x, y) = \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(p_0) (x-x_0)^k (y-y_0)^m$$

と表わすことにします。このとき、点  $p_0$  を中心とする半径 1 の円板

$$D = \{p = (x, y) \in \mathbb{R}^2 \mid \|p - p_0\| \leq 1\}$$

を考えて (図 4 を参照)、 $k, m \in \mathbb{N}$  として、円板  $D$  上での関数  $\left| \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(x, y) \right|$  の最大値を、

\*12) 三変数以上の多変数関数の場合にも全く同様の議論ができます。

$$M_{k,m} = \max_{p \in D} \left| \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(p) \right|$$

と表わすことにします。<sup>\*13)</sup> ただし、ユークリッド平面  $\mathbb{R}^2$  における点  $p = (x, y)$  と点  $p_0 = (x_0, y_0)$  の間の距離を、

$$\|p - p_0\| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

と表わしました。いま、Taylor の定理から、

$$f(x, y) = P_n(x, y) + R_n(x, y) \quad (17)$$

$$R_n(x, y) = \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(q) (x-x_0)^k (y-y_0)^m \quad (18)$$

となるような点  $q \in \mathbb{R}^2$  が点  $p_0$  と点  $p$  を結ぶ線分上に存在することが分かりますが、 $p \in D$  とすると、線分  $\overline{p_0 p}$  も円板  $D$  に含まれることが分かりますから、

$$q \in D$$

となることが分かります。したがって、

$$|x - x_0|, |y - y_0| \leq \|p - p_0\|$$

となることにも注意すると、

$$\begin{aligned} & \left| \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(q) (x-x_0)^k (y-y_0)^m \right| \\ &= \left| \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(q) \right| \cdot |x-x_0|^k \cdot |y-y_0|^m \\ &\leq M_{k,m} \cdot \|p - p_0\|^{k+m} \end{aligned} \quad (19)$$

となることが分かります。よって、

$$C = \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \cdot M_{k,m}$$

とすると、(17) 式、(18) 式、(19) 式から、 $p \in D$  に対して、

$$\begin{aligned} |f(x, y) - P_n(x, y)| &= |R_n(x, y)| \\ &\leq \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \cdot \left| \frac{\partial^{k+m} f}{\partial x^k \partial y^m}(q) (x-x_0)^k (y-y_0)^m \right| \end{aligned}$$

<sup>\*13)</sup> 以下では、 $p \rightarrow p_0$  という極限を考察するので、円板  $D$  の半径は、1 でなくとも、 $\frac{1}{2}$  でも、 $\frac{1}{100}$  でも、何でも構いません。

$$\begin{aligned} &\leq \sum_{\substack{0 \leq k, m \\ k+m=n+1}} \frac{1}{k!m!} \cdot M_{k,m} \cdot \|p - p_0\|^{k+m} \\ &= C \cdot \|p - p_0\|^{n+1} \end{aligned}$$

となることが分かります。以上から、 $p = (x, y) \in D$  に対して、

$$|f(x, y) - P_n(x, y)| \leq C \cdot \|p - p_0\|^{n+1} \quad (20)$$

という評価式が成り立つことが分かりました。

この (20) 式という評価式を用いると、一変数関数のときと同様に、点  $p_0 = (x_0, y_0)$  における関数  $f(x, y)$  の Taylor 多項式  $P_n(x, y)$  を、次のように特徴付けることができることが分かります。いま、 $k \in \mathbb{N}$  として、(20) 式から、 $p = (x, y) \in D$  に対して、

$$\frac{|f(x, y) - P_n(x, y)|}{\|p - p_0\|^k} \leq C \cdot \|p - p_0\|^{n-k+1} \quad (21)$$

となることが分かります。よって、(21) 式の両辺で、 $p \rightarrow p_0$  なる極限を考えると、 $k = 0, 1, 2, \dots, n$  に対して、

$$\lim_{p \rightarrow p_0} \frac{|f(x, y) - P_n(x, y)|}{\|p - p_0\|^k} = 0 \quad (22)$$

となることが分かります。この (22) 式が、二変数関数の場合の (16) 式の対応物であるということになります。<sup>\*14)</sup>

そこで、次に、 $k = 0, 1, 2, \dots, n$  に対して、(22) 式を成り立たせるような  $n$  次の多項式は Taylor 多項式  $P_n(x, y)$  しか存在しないということを確かめてみることにします。いま、 $Q_n(x, y)$  を  $n$  次の多項式として、 $Q_n(x, y)$  も、 $k = 0, 1, 2, \dots, n$  に対して、

$$\lim_{p \rightarrow p_0} \frac{|f(x, y) - Q_n(x, y)|}{\|p - p_0\|^k} = 0 \quad (23)$$

という式を満たしていると仮定してみます。このとき、 $Q_n(x, y) - P_n(x, y)$  という多項式を考えてみます。<sup>\*15)</sup> いま、

$$Q_n(p) - P_n(p) = (f(p) - P_n(p)) - (f(p) - Q_n(p))$$

と書き直してみると、

<sup>\*14)</sup> 「直線  $\mathbb{R}$  上の  $\frac{f(x) - P_n(x)}{(x-a)^k}$  という点と 0 という点の間の距離が 0 に近づく」と解釈すると、(16) 式は、

$$\left( \lim_{x \rightarrow a} \frac{|f(x) - P_n(x)|}{|x - a|^k} \right) = \lim_{x \rightarrow a} \left| \frac{f(x) - P_n(x)}{(x-a)^k} \right| = 0$$

このように書き直せることに注意して下さい。

<sup>\*15)</sup> スペースを節約するために、以下では、 $f(x, y)$  などを  $f(p)$  などとも表わすことにします。

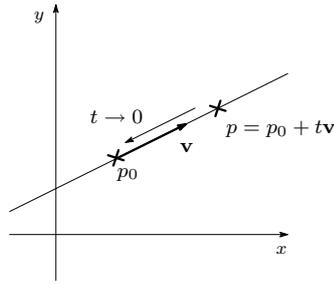


図 5 点  $p$  を  $\mathbf{v}$  という方向から点  $p_0$  に近づけてみる.

$$|Q_n(p) - P_n(p)| \leq |f(p) - P_n(p)| + |f(p) - Q_n(p)|$$

となることが分かりますから、

$$\frac{|Q_n(p) - P_n(p)|}{\|p - p_0\|^k} \leq \frac{|f(p) - P_n(p)|}{\|p - p_0\|^k} + \frac{|f(p) - Q_n(p)|}{\|p - p_0\|^k} \quad (24)$$

となることが分かります。よって、(22) 式、(23) 式、(24) 式から、 $k = 0, 1, 2, \dots, n$  に対して、

$$\lim_{p \rightarrow p_0} \frac{|Q_n(p) - P_n(p)|}{\|p - p_0\|^k} = 0 \quad (25)$$

となることが分かります。

そこで、いま、

$$Q_n(p) - P_n(p) = \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} c_{k,m} (x - x_0)^k (y - y_0)^m$$

と表わしてみます。<sup>\*16)</sup> このとき、 $\mathbf{0} \neq \mathbf{v} = (a, b) \in \mathbb{R}^2$  を、勝手にひとつ取ってきて、 $\mathbf{v}$  という方向から点  $p$  を点  $p_0$  に近づけるとどうなるかということを考えてみることにします。すなわち、

$$p = p_0 + t\mathbf{v}$$

として、 $t \rightarrow 0$  としたときに、(25) 式からどのようなことが分かるのかということを考えてみることにします (図 5 を参照)。このとき、

$$Q_n(p) - P_n(p) = \sum_{\substack{0 \leq k, m \\ 0 \leq k+m \leq n}} c_{k,m} a^k b^m \cdot t^{k+m}$$

と表わせることが分かりますが、

$$\varphi_l = \sum_{\substack{0 \leq k, m \\ k+m=l}} c_{k,m} a^k b^m, \quad (l = 0, 1, \dots, n) \quad (26)$$

<sup>\*16)</sup> 以下、示したいことは、すべての係数  $c_{k,m}$  が  $c_{k,m} = 0$  となるということです。

と書くことにして、

$$Q_n(p) - P_n(p) = \varphi_0 + \varphi_1 t + \dots + \varphi_n t^n \quad (27)$$

というように、 $t$  のべきで整理して表わすことにします。

そこで、まず、 $k = 0$  として、(25) 式の条件を調べてみると、

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} |Q_n(p) - P_n(p)| \\ &= \lim_{t \rightarrow 0} |\varphi_0 + \varphi_1 t + \dots + \varphi_n t^n| \\ &= |\varphi_0| \end{aligned}$$

となることが分かりますから、 $\varphi_0 = 0$  となることが分かります。次に、

$$\|p - p_0\| = \|t\mathbf{v}\| = |t| \cdot \|\mathbf{v}\|$$

となることに注意して、 $k = 1$  として、(25) 式の条件を調べてみると、

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} \frac{|Q_n(p) - P_n(p)|}{\|p - p_0\|} \\ &= \frac{1}{\|\mathbf{v}\|} \cdot \lim_{t \rightarrow 0} \left| \frac{\varphi_1 t + \varphi_2 t^2 + \dots + \varphi_n t^n}{t} \right| \\ &= \frac{1}{\|\mathbf{v}\|} \cdot \lim_{t \rightarrow 0} |\varphi_1 + \varphi_2 t + \dots + \varphi_n t^{n-1}| \\ &= \frac{1}{\|\mathbf{v}\|} \cdot |\varphi_1| \end{aligned}$$

となることが分かりますから、 $\varphi_1 = 0$  となることが分かります。<sup>\*17)</sup> 以下、同様にして、(25) 式の条件を順番に調べてみると、結局、

$$\varphi_0 = \varphi_1 = \dots = \varphi_n = 0 \quad (28)$$

となることが分かります。<sup>\*18)</sup> 上では、 $\mathbf{v} = (a, b)$  を、勝手にひとつ固定して議論を進めましたが、 $\mathbf{v} = (a, b) \in \mathbb{R}^2$  はどんな値でも構わなかったので、(28) 式から、(26) 式で与えられる  $\varphi_l$  を  $(a, b)$  の関数であると見たときに、関数  $\varphi_l(a, b)$  は零関数になることが分かります。よって、すべての係数  $c_{k,m}$  は  $c_{k,m} = 0$  となることが分かりますから、<sup>\*19)</sup>

<sup>\*17)</sup>  $k = 0$  での考察から  $\varphi_0 = 0$  であることが分かっていますから、ここでも  $\varphi_0 = 0$  として議論しました。

<sup>\*18)</sup> 皆さん、確かめてみてください。

<sup>\*19)</sup> 例えば、

$$0 = \varphi_l(a, b) = \sum_{\substack{0 \leq k, m \\ k+m=l}} c_{k,m} a^k b^m$$

という式の両辺を、 $a$  に関して  $k$  回、 $b$  に関して  $m$  回偏微分してから、 $(a, b) = (0, 0)$  としてみると、 $c_{k,m} = 0$  となることが分かります。

$$Q_n(p) - P_n(p) = 0$$

となることが分かります。したがって、

$$Q_n(x, y) = P_n(x, y)$$

となることが分かります。

以上から、一変数関数のときと同様に、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_n(x, y)$  は「 $k = 0, 1, 2, \dots, n$  に対して、(22) 式が成り立つような  $n$  次の多項式」として一意的に特徴付けることができるということが分かりました。

この特徴付けの式を用いると、 $P_n(x, y)$  以外の  $n$  次の多項式  $Q_n(x, y)$  に対して、 $m \in \mathbb{N}$ ,  $C' \in \mathbb{R}$  として、 $p = (x, y) \in D$  のときに、

$$|f(x, y) - Q_n(x, y)| \leq C' \cdot \|p - p_0\|^m$$

という評価式が成り立ったとしても、評価式における  $\|p - p_0\|$  のべきである  $m$  は、常に、 $m \leq n$  となることが分かります。<sup>\*20)</sup> 第 3 回の問 4 のところで見たように、一般に、 $P(x, y)$  を多項式として、

$$|f(x, y) - P(x, y)| \leq C \cdot \|p - p_0\|^m$$

という形の評価式を考えたときに、 $\|p - p_0\| \ll 1$  という状況だけを問題にしている場合には、定数  $C$  の大きさがどんなものであれ、「近似の良さ」は  $\|p - p_0\|$  のべきである  $m$  の大きさによって定まると考えられます。したがって、二変数関数の場合にも、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_n(x, y)$  は、 $n$  次の多項式の中で、「 $\|p - p_0\| \ll 1$  のとき、関数  $f(x, y)$  を最も良く近似する多項式」であることが分かりました。

また、一変数関数の場合に、第 3 回の問 4 のところで行なった「幅つきグラフ」の議論を繰り返すと、今の場合、勝手な点  $p = (x, y) \in \mathbb{R}^2$  に対して、

$$\varepsilon_n(x, y) = C \cdot \|p - p_0\|^{n+1}$$

という「幅」を付けて、関数  $f(x, y)$  や多項式  $P(x, y)$  のグラフを描くということになります。そこで、関数  $f(x, y)$  の点  $p_0$  のまわりでの Taylor 多項式  $P_n(x, y)$  を考えると、(20) 式から、 $p = (x, y) \in D$  に対して、

$$|f(x, y) - P_n(x, y)| \leq \varepsilon_n(x, y)$$

となることが分かりますから、 $f(x, y)$  と  $P_n(x, y)$  の  $\varepsilon_n(x, y)$  という幅のついた「幅つきグラフ」は、円板  $D$  上で、どのような点  $p = (x, y) \in D$  に対しても、重なり合う部分を持つことが分かります。

一方、 $Q_n(x, y)$  を  $Q_n(x, y) \neq P_n(x, y)$  となる  $n$  次の多項式とすると、一変数関数のときと同様に、上の Taylor 多項式の特徴付けの式を用いて議論すると、点  $p_0 = (x_0, y_0) \in \mathbb{R}^2$  のいくらでも近くに、

$$|f(x', y') - Q_n(x', y')| > 3\varepsilon_n(x', y') \quad (29)$$

となるような点  $p' = (x', y') \in \mathbb{R}^2$  が存在することが分かります。<sup>\*21)</sup> よって、 $f(x, y)$  と  $Q_n(x, y)$  の「幅つきグラフ」は、点  $p_0$  のいくらでも近くに重ならない部分を持つということが分かります。

いま、「幅つきグラフ」の幅  $\varepsilon_n(x, y)$  は、

$$\lim_{(x, y) \rightarrow (x_0, y_0)} \varepsilon_n(x, y) = 0$$

となりますから、 $f(x, y)$  と  $P_n(x, y)$  の「幅つきグラフ」が、円板  $D$  上の勝手な点  $p = (x, y) \in D$  に対して、重なり合う部分を持つということは、 $(x, y) \rightarrow (x_0, y_0)$  のときに、 $z = f(x, y)$  のグラフと  $z = P_n(x, y)$  のグラフは、事実上、見分けがつかなくなるということの意味していると解釈することができます。一方、 $Q_n(x, y)$  を Taylor 多項式  $P_n(x, y)$  以外の  $n$  次の多項式とすると、グラフの幅を  $\varepsilon_n(x, y)$  として「幅つきグラフ」を考えている限りは、 $f(x, y)$  と  $Q_n(x, y)$  の「幅つきグラフ」は、点  $p_0$  のいくらでも近くに重なり合わない部分を持ってしまふのでした。その意味で、 $z = Q_n(x, y)$  のグラフは  $z = f(x, y)$  のグラフに「あまり近くない」と言えますから、関数  $f(x, y)$  の Taylor 多項式  $P_n(x, y)$  は、 $n$  次の多項式の中で、「 $\|p - p_0\| \ll 1$  のときに、関数  $f(x, y)$  に最も良く「姿」が似ている多項式」であると考えられます。

#### 7. 全微分可能とは (数学的な定義について)\*

さて、6 節の結果を用いると、第 5 回の解説の中では少し曖昧な形で済ましていた「全微分可能である」ということの数学的に正確な定義を与えることができます。そこで、ここでは、この問題について、少し考えてみることにします。

\*20) このことは、第 3 回の問 4 のところで、一変数関数に対して行なったのと全く同じ議論で示すことができます。

\*21) 興味のある方は、第 3 回の問 4 のところで説明した一変数関数の場合の議論を参考にして、確かめて下さい。

第5回の問1のところでは、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対して、点  $p_0 = (x_0, y_0) \in \mathbb{R}^2$  において、関数  $f(x, y)$  のグラフに「接平面が描ける」ときに、関数  $f(x, y)$  は点  $p_0$  において全微分可能であると定義しました。ここで、「ある平面が関数  $f(x, y)$  のグラフに接している」というのは直感的な表現であり、上の定義は「誰にとっても誤解の生じる可能性がない曖昧さのない定義」ではないことに注意して下さい。そこで、「全微分可能である」ということの数学的に正確な定義を与えるためには、「平面が関数  $f(x, y)$  のグラフに接している」ということを「数学的に曖昧さのない形」で言い換える必要がありますが、6節での議論の中にそのためのヒントが隠されています。

いま、 $n = 1$  として、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  を考えてみると、

$$\begin{aligned} P_1(x, y) - f(p_0) &= \frac{\partial f}{\partial x}(p_0)(x - x_0) + \frac{\partial f}{\partial y}(p_0)(y - y_0) \quad (30) \end{aligned}$$

となることが分かります。一方、第5回の問1のところで見たとように、 $C^1$  級の関数  $f(x, y)$  の点  $p_0$  における接平面の式は、

$$\begin{aligned} z - f(p_0) &= \frac{\partial f}{\partial x}(p_0)(x - x_0) + \frac{\partial f}{\partial y}(p_0)(y - y_0) \quad (31) \end{aligned}$$

で与えられるのでした。そこで、(30) 式と (31) 式を見比べてみると、 $z = P_1(x, y)$  のグラフがちょうど点  $p_0$  における接平面になっていることが分かります。

一般に、 $A, B, C \in \mathbb{R}$  として、

$$Q(x, y) = Ax + By + C \quad (32)$$

という二変数の一次の多項式を考えると、 $z = Q(x, y)$  のグラフは  $\mathbb{R}^3$  内の平面を定めることが、例えば、次のようにして分かります。いま、 $(x_0, y_0) \in \mathbb{R}^2$  を、勝手にひとつ取ってきて、 $z_0 = Q(x_0, y_0)$  と定めて、 $z = Q(x, y)$  という式を、

$$z - z_0 = A(x - x_0) + B(y - y_0) \quad (33)$$

というように書き直してみます。<sup>\*22)</sup>すると、第5回の問1のところと同様に、点  $(x_0, y_0, z_0) \in \mathbb{R}^3$  からの「ずれ」を表わすベクトルを

<sup>\*22)</sup> すなわち、 $z = Ax + By + C$  という式の両辺から、 $z_0 = Ax_0 + By_0 + C$  という式の両辺を引き算してみました。

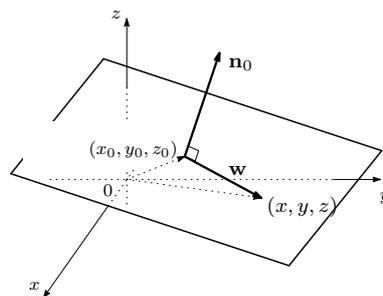


図6 (33) 式は、点  $(x_0, y_0, z_0) \in \mathbb{R}^3$  を通り、ベクトル  $\mathbf{n}_0$  に直交するような平面を表わしている。

$$\mathbf{w} = (x - x_0, y - y_0, z - z_0)$$

として、

$$\mathbf{n}_0 = (-A, -B, 1)$$

と書くことにすると、(33) 式は、 $\mathbb{R}^3$  上の内積  $\langle \cdot, \cdot \rangle$  を用いて、

$$\langle \mathbf{n}_0, \mathbf{w} \rangle = 0$$

という形に表わすことができます。よって、(33) 式は、点  $(x_0, y_0, z_0) \in \mathbb{R}^3$  を通り、ベクトル  $\mathbf{n}_0$  と直交する平面を表わしていることが分かります (図6参照)。

一方、 $\mathbb{R}^3$  内の平面は、平面上の点  $(x_0, y_0, z_0)$  と平面に直交する  $\mathbf{0}$  でないベクトル  $\mathbf{n}_0 = (\alpha, \beta, \gamma)$  を、勝手にひとつずつ取ってくと、

$$\alpha(x - x_0) + \beta(y - y_0) + \gamma(z - z_0) = 0 \quad (34)$$

という式によって表わせることが分かります。このとき、 $\gamma \neq 0$  であると仮定すると、

$$A = -\frac{\alpha}{\gamma}, B = -\frac{\beta}{\gamma}, C = z_0 + \frac{\alpha x_0 + \beta y_0}{\gamma}$$

として、(34) 式は、

$$z = Ax + By + C$$

という形に書き直せることが分かりますから、このような平面は (32) 式のような一次の多項式  $Q(x, y)$  のグラフとして実現できることが分かります。以上から、 $\mathbb{R}^3$  内の平面と (32) 式のような一次の多項式  $Q(x, y)$  が「ほぼ」対応することが分かりました。<sup>\*23)</sup>そこで、

<sup>\*23)</sup> ここで、「ほぼ」と言ったのは、(34) 式において、 $\gamma = 0$  となる場合が抜け落ちているからですが、 $\gamma = 0$  の場合には、対応する  $\mathbb{R}^3$  内の平面は  $z$  軸と平行な直線を集めたような平面になってしまいますから、以下の議論では、最初から、このような平面を「関数  $f(x, y)$  の接平面の候補」の中から除外して考えることにします。

以下では、 $\mathbb{R}^3$  内の平面と (32) 式のような一次の多項式  $Q(x, y)$  を同一視して考えることにします。

さて、6 節で見たように、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  は、数ある一次の多項式の中で、「 $\|p - p_0\| \ll 1$  のとき、関数  $f(x, y)$  を最も良く近似する多項式」であることが分かります。このことは、数ある  $\mathbb{R}^3$  内の平面の中で、「点  $p_0$  の近くで、関数  $f(x, y)$  のグラフに最も良く「姿」が似ている平面」が「Taylor 多項式  $P_1(x, y)$  のグラフとして得られる平面」であるということを意味しています。また、上で注意したように、この平面は、前回、いくらか直感的な議論にもとづいて求めた「接平面」と一致していることも分かります。

そこで、こうした事実を念頭において、「点  $p_0$  において関数  $f(x, y)$  のグラフに接する」ということの言い換えとして、「点  $p_0$  の近くで、関数  $f(x, y)$  を最も良く近似する」ということが採用できるのではないかと考えられました。ただし、このままでは、まだ少し議論しづらいので、実際には、この「近似の良さ」を保証する「Taylor 多項式の特徴付けの式」が「全微分可能である」ということの数学的な定義として採用されています。

6 節で見たように、関数  $f(x, y)$  が滑らかな関数の場合には、 $k = 0, 1$  に対して、

$$\lim_{p \rightarrow p_0} \frac{|f(x, y) - Q(x, y)|}{\|p - p_0\|^k} = 0 \quad (35)$$

となるような一次の多項式  $Q(x, y)$  は、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  しか存在しないことが分かります。そこで、一般に、(滑らかとは限らない) 二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対して、 $k = 0, 1$  に対して、(35) 式が成り立つような一次の多項式  $Q(x, y)$  が存在するときに、関数  $f(x, y)$  は点  $p_0$  において全微分可能であると定義されます。また、このとき、「関数  $Q(x, y)$  のグラフとして得られる平面」を点  $p_0$  における関数  $f(x, y)$  のグラフの接平面と呼びます。6 節で行なった議論を繰り返すと、このような一次の多項式  $Q(x, y)$  は、存在すれば一意的に定まることが分かります。<sup>\*24)</sup>

ここで、 $p_0 = (x_0, y_0)$  として、一次の多項式  $Q(x, y)$  を、

$$Q(x, y) = A(x - x_0) + B(y - y_0) + C$$

\*24) 皆さん、確かめてみて下さい。

という形で表わすことにすると、

$$\lim_{p \rightarrow p_0} |f(x, y) - Q(x, y)| = |f(p_0) - C|$$

となることが分かりますから、 $k = 0$  に対する (35) 式の条件は、

$$C = f(p_0)$$

という式と数学的な内容は同じであることが分かります。そこで、最初から、 $C = f(p_0)$  と選んで、 $k = 1$  に対する (35) 式の条件だけを考えるのが普通です。すなわち、普通、微積分学の教科書では、

$$\lim_{p \rightarrow p_0} \frac{|f(p) - f(p_0) - A(x - x_0) - B(y - y_0)|}{\|p - p_0\|} = 0$$

となる実数  $A, B \in \mathbb{R}$  が存在するとき、「関数  $f(x, y)$  は点  $p_0$  において全微分可能である」というように定義されています。ただし、普段、皆さんが接するような関数は「式一発」で書けるような滑らかな関数のことが多いでしょうし、滑らかな関数  $f(x, y)$  に対しては、すでに見たように、「関数  $f(x, y)$  は  $\mathbb{R}^2$  上の勝手な点  $p_0 \in \mathbb{R}^2$  において全微分可能であり、 $Q(x, y)$  は点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  で与えられる」ことが分かりますから、皆さんが、上の定義にもとづいて、実際に全微分可能かどうかを考察しなければいけない場面に出会うことは、(数学の定期試験においての他は、) 余りないのではないかと思います。

## 8. 臨界点とは

さて、問題文でも触れたように、 $\mathbb{R}^2$  上の滑らかな関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対して、

$$\frac{\partial f}{\partial x}(p_0) = \frac{\partial f}{\partial y}(p_0) = 0 \quad (36)$$

となるような点  $p_0 \in \mathbb{R}^2$  を、関数  $f(x, y)$  の臨界点 (critical point) と呼びます。<sup>\*25)</sup> 5 節で見たように、一般に、 $\mathbb{R}^2$  上の点  $p_0 = (x_0, y_0) \in \mathbb{R}^2$  に対して、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  は、

$$P_1(x, y) = f(p_0) + \frac{\partial f}{\partial x}(p_0)(x - x_0) + \frac{\partial f}{\partial y}(p_0)(y - y_0)$$

という式によって与えられますから、臨界点とは、

$$P_1(x, y) = f(p_0)$$

というように、Taylor 多項式  $P_1(x, y)$  が定数関数に

\*25) 全く同様に、三変数以上の多変数関数  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  に対しても、すべての偏微分係数が 0 になる点として「臨界点」が定義されます。

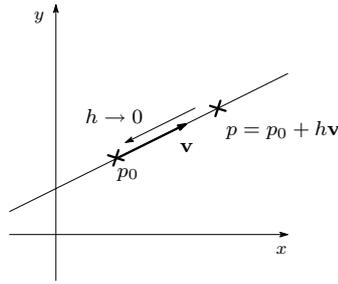


図7 点  $p$  を  $\mathbf{v}$  という方向から点  $p_0$  に近づけてみる.

なるような点であると言い換えることができます. また, 7節でも注意したように, 一次の多項式  $P_1(x, y)$  のグラフが点  $p_0$  における関数  $f(x, y)$  のグラフの接平面を与え, この平面の法線ベクトルは,

$$\mathbf{n}_0 = \left( -\frac{\partial f}{\partial x}(p_0), -\frac{\partial f}{\partial y}(p_0), 1 \right)$$

という式によって与えられることがわかりますから, 臨界点とは,

$$\mathbf{n}_0 = (0, 0, 1)$$

というように, 接平面の法線ベクトル  $\mathbf{n}_0$  が  $z$  軸の方向を向くような点, すなわち, 接平面が  $xy$  平面と平行になり, その傾きが 0 となるような点であるとも言えることができます.

そこで, (36) 式の条件を, もう少し良く理解するために, ここで, 方向微分について少し思い出してみることになります. いま,  $\mathbb{R}^2$  上の関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対して, 点  $p_0 \in \mathbb{R}^2$  において,  $\mathbf{v} = (a, b) \in \mathbb{R}^2$  の定める方向の接線の傾き

$$\lim_{h \rightarrow 0} \frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \quad (37)$$

を考えてみます (図7を参照). このような状況では,  $\mathbf{v} \in \mathbb{R}^2$  は点  $p_0$  を始点とするベクトルであると考えの方が自然ですから, この点を強調するために, 点  $p_0$  を始点とするベクトル全体の集合を考えて,

$$T_{p_0}\mathbb{R}^2 = \{\mathbf{v} \mid \mathbf{v} \text{ は点 } p_0 \text{ を始点とするベクトル.}\}$$

と表わすことにします. この集合  $T_{p_0}\mathbb{R}^2$  を  $\mathbb{R}^2$  の点  $p_0$  における接空間 (tangent space) と呼びます (図8を参照). この記号を用いると,  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$  と考えることができます.

そこで, いま,  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$  として, (37) 式で与えられる極限が存在するとします. このとき, この極限値を,

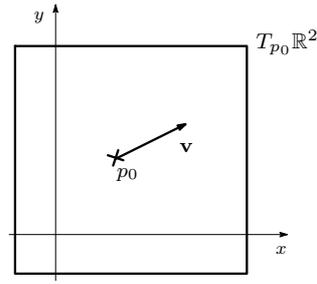


図8  $\mathbb{R}^2$  の点  $p_0$  における接空間  $T_{p_0}\mathbb{R}^2$ .

$$(df)_{p_0}(\mathbf{v}) = \lim_{h \rightarrow 0} \frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h}$$

と表わし, 点  $p_0$  における  $\mathbf{v}$  方向の方向微分係数と呼びます. また, 勝手なベクトル  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$  に対して,  $\mathbf{v}$  方向の方向微分係数  $(df)_{p_0}(\mathbf{v})$  が存在するときに, 関数  $f(x, y)$  は点  $p_0$  において方向微分可能であると言えます. このとき,  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$  に対して,  $(df)_{p_0}(\mathbf{v}) \in \mathbb{R}$  を対応させることにより,

$$(df)_{p_0}: T_{p_0}\mathbb{R}^2 \rightarrow \mathbb{R} \quad (38)$$

という写像が定義できることとなります.

第5回の問1のところで見たとおり, 一般には, (38) 式の写像  $(df)_{p_0}$  は複雑な形になり得ますが, 関数  $f(x, y)$  が  $C^1$  級の関数の場合には,  $(df)_{p_0}$  はとても簡単な写像になります. 実際, この場合,  $\mathbf{v} = (a, b) \in T_{p_0}\mathbb{R}^2$  に対して,

$$(df)_{p_0}(\mathbf{v}) = \frac{\partial f}{\partial x}(p_0)a + \frac{\partial f}{\partial y}(p_0)b \quad (39)$$

というように定数項のない一次式の形で与えられるのでした. いま,  $\mathbf{v}$  を,

$$\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in T_{p_0}\mathbb{R}^2$$

というように縦ベクトルで表わすことにすると, (39) 式は行列の掛け算を用いて,

$$(df)_{p_0}(\mathbf{v}) = \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

というように書き直せることがわかります. すなわち, この場合,  $(df)_{p_0}: T_{p_0}\mathbb{R}^2 \rightarrow \mathbb{R}$  は,

$$\begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix} \quad (40)$$

という行列を掛け算することによって定まる写像であ

ることが分かります。<sup>\*26)</sup> この (線型) 写像  $(df)_{p_0}$  を関数  $f(x, y)$  の点  $p_0$  における全微分と呼びます。<sup>\*27)</sup>

そこで、以下では、(38) 式の写像  $(df)_{p_0}$  と (40) 式で与えられる行列を同一視して、

$$(df)_{p_0} = \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix} \quad (41)$$

など書くことにします。この記号を用いると、点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  は、

$$P_1(x, y) = f(p_0) + (df)_{p_0} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

というように表わすことができます。これを、一変数関数の場合の対応する式、

$$P_1(x) = f(x_0) + f'(x_0)(x - x_0)$$

と比べてみると、

$$\begin{aligned} f'(x_0) &\longleftrightarrow (df)_{p_0} \\ (x - x_0) &\longleftrightarrow \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \end{aligned}$$

というように対応していることが分かります。<sup>\*28)</sup> したがって、二変数関数の場合への「微分」の概念の「正しい拡張」は、「偏微分」ではなく、「全微分」 $(df)_{p_0}$  であることが分かります。すなわち、一変数関数の場合の「微分」 $f'(x_0)$  を 1 行 1 列の行列  $(f'(x_0))$  と同一視して考えることにすると、

$$\begin{array}{c} f: \mathbb{R} \rightarrow \mathbb{R} \\ \downarrow \text{多変数化} \\ f: \mathbb{R}^2 \rightarrow \mathbb{R} \end{array}$$

というように、関数の変数を多変数化したときに、それに伴って「微分」の概念が、

$$\begin{array}{c} \left( \frac{df}{dx}(x_0) \right) \\ \downarrow \text{列が増える} \\ (df)_{p_0} = \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix} \end{array}$$

\*26) 線型代数学では、このような写像を線型写像と呼びます。

\*27) ただし、「全微分」という用語は「偏微分」という用語と対比させるために用いられた歴史的な用語であり、以下で見るように、現在では、 $(df)_{p_0}$  が多変数関数の場合の「微分」の概念の自然な拡張であると認識されているので、「全微分」のことを、単に「微分」と呼ぶのが普通です。

\*28) ここで、 $(x - x_0)$  と  $\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$  は、いずれも、「展開の中心点」 $x_0 \in \mathbb{R}$ 、あるいは、 $p_0 = (x_0, y_0) \in \mathbb{R}^2$  から、「値を考えたい点」点  $x \in \mathbb{R}$ 、あるいは、 $p = (x, y) \in \mathbb{R}^2$  へのズレを表わしていると解釈できることに注意して下さい。

というように拡張されることが分かります。そこで、以下では、全微分  $(df)_{p_0}$  のことを、単に「微分」と呼ぶことにします。このとき、

$$\frac{\partial f}{\partial x}(p_0) = \frac{\partial f}{\partial y}(p_0) = 0 \iff (df)_{p_0} = 0$$

と解釈できますから、<sup>\*29)</sup> 臨界点とは「微分が消える点」であると解釈できることが分かります。

さて、皆さん良くご存じのように、一変数関数の場合には、

$$x_0 \in \mathbb{R} \text{ が極値点.} \implies f'(x_0) = 0 \quad (42)$$

という事実に注目して、 $f'(x_0) = 0$  という式を解くことで極値点の候補を求めることができます。二変数関数の場合にも、(42) 式と同様の事実が成り立つことが、例えば、次のようにして分かります。

そこで、いま、 $p_0 \in \mathbb{R}^2$  が関数  $f(x, y)$  の極小点であるとします。<sup>\*30)</sup> すなわち、点  $p_0$  を中心とした十分小さな半径  $0 < r_0 \in \mathbb{R}$  の円板

$$D = \{p \in \mathbb{R}^2 \mid \|p - p_0\| \leq r_0\}$$

を考えると、勝手な点  $p \in D$  に対して、

$$f(p) \geq f(p_0) \quad (43)$$

が成り立っているとします。このとき、 $v \in T_{p_0}\mathbb{R}^2$  を、勝手にひとつ取ってきて、点  $p_0$  における  $v$  方向の方向微分係数  $(df)_{p_0}(v)$  がどうなるかということを考えてみます。すると、(43) 式から、勝手な実数  $h \in \mathbb{R}$  に対して、

$$f(p_0 + hv) - f(p_0) \geq 0 \quad (44)$$

となることが分かります。<sup>\*31)</sup> よって、(44) 式から、 $h > 0$  であるとすると、

$$\frac{f(p_0 + hv) - f(p_0)}{h} \geq 0 \quad (45)$$

となることが分かります。したがって、 $h > 0$  の方向から  $h$  を 0 に近づけたのだと考えると、(45) 式から、

\*29) ここで、 $(df)_{p_0} = 0$  という式の右辺に現われる 0 は「零行列」を表わしていることに注意して下さい。

\*30) 極大点である場合も全く同様に議論できます。

\*31) 正確には、「 $p_0 + hv \in D$  となっている場合には」という付加的な条件が必要ですが、以下では、 $h \rightarrow 0$  という極限を考えるので、この点についてはいちいち断らないことにします。気になる方は、「以下では、 $|h| \leq \frac{r_0}{\|v\|}$  となる実数  $h \in \mathbb{R}$  しか考えないことにします」と最初に宣言してから議論を進めているのだと考えてみて下さい。

$$(df)_{p_0}(\mathbf{v}) = \lim_{h \rightarrow 0} \frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \geq 0 \quad (46)$$

となることが分かります。全く同様に、 $h < 0$  である  
とすると、(44) 式から、

$$\frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \leq 0 \quad (47)$$

となることが分かります。したがって、 $h < 0$  の方向  
から  $h$  を 0 に近づけたのだと考えると、(47) 式から、

$$(df)_{p_0}(\mathbf{v}) = \lim_{h \rightarrow 0} \frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \leq 0 \quad (48)$$

となることが分かります。よって、(46) 式と (48) 式  
から、

$$(df)_{p_0}(\mathbf{v}) = 0 \quad (49)$$

となることが分かります。ここで、 $\mathbf{v}$  は、どんなベク  
トルでも良かったので、勝手なベクトル  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$   
に対して、(49) 式が成り立つことが分かります。した  
がって、

$$(df)_{p_0} = 0$$

となることが分かりますから、<sup>\*32)</sup>

$$p_0 \in \mathbb{R}^2 \text{ が極値点} \implies (df)_{p_0} = 0 \quad (50)$$

となることが分かりました。以上から、一変数関数の  
ときと同様に、二変数関数に対しても、「微分が消える  
点」である「臨界点」を求めることにより、極値点の  
候補を求めることができることが分かります。

## 9. 極値の判定法について

さて、これまで見てきたように、二変数関数  $f(x, y)$   
に対して、「接平面の様子を調べることにより、点  $p_0$  の  
近くでの関数  $f(x, y)$  の大まかな様子を理解する」と  
いうことは、「点  $p_0$  における関数  $f(x, y)$  の Taylor  
多項式  $P_1(x, y)$  を用いて、

$$f(x, y) \doteq P_1(x, y)$$

というように近似して、一次の多項式  $P_1(x, y)$  の様  
子を調べることにより、点  $p_0$  の近くでの関数  $f(x, y)$

\*32) 例えば、 $\mathbf{v}$  として、

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

と選んで、(49) 式の条件を考えると、 $(df)_{p_0}$  という行列の  
1 行 1 列成分、1 行 2 列成分が共に 0 となることが分かり  
ます。

の大まかな様子を理解する」ということであると解釈  
できることが分かります。

このとき、 $p_0 = (x_0, y_0) \in \mathbb{R}^2$  が関数  $f(x, y)$  の臨  
界点であるとする、 $P_1(x, y) = f(x_0, y_0)$  となってい  
まい、

$$f(x, y) \doteq f(x_0, y_0)$$

という情報しか得られないことが分かります。したがっ  
て、この場合、もう少し詳しい情報を得るためには、

$$f(x, y) \doteq P_2(x, y)$$

というように、さらに近似を上げて考察する必要があ  
ります。第 3 回の問 4 のところで見たように、一変数  
関数の場合には、 $f'(x_0) = 0$  となる点  $x_0$  の近くで、

$$f(x) \doteq P_2(x)$$

というように、さらに近似を上げて考えるということは、

$$f(x) - f(x_0) \doteq \frac{1}{2} f''(x_0)(x - x_0)^2 \quad (51)$$

というように近似して考えるということでした。また、  
 $f''(x_0) \neq 0$  のときには、(51) 式の右辺に現われる  
二次関数の様子を調べることにより、点  $x_0$  が極大点  
なのか、あるいは、極小点なのかという「極値の判定」  
ができるのでした。そこで、ここでは、二変数関数に対  
して、同様の考察を行なってみることにします。<sup>\*33)</sup>

そこで、いま、 $p_0 = (x_0, y_0) \in \mathbb{R}^2$  が関数  $f(x, y)$  の  
臨界点であるとし、すると、5 節で見たように、臨界  
点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_2(x, y)$   
は、

$$P_2(x, y) = f(p_0) + \frac{1}{2} \left\{ \frac{\partial^2 f}{\partial x^2}(p_0) X^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(p_0) XY + \frac{\partial^2 f}{\partial y^2}(p_0) Y^2 \right\}$$

となることが分かりますから、

$$f(x, y) \doteq P_2(x, y)$$

と近似して考えるということは、

$$\begin{aligned} f(x, y) - f(x_0, y_0) & \\ \doteq \frac{1}{2} \left\{ \frac{\partial^2 f}{\partial x^2}(p_0) X^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(p_0) XY + \frac{\partial^2 f}{\partial y^2}(p_0) Y^2 \right\} & \end{aligned} \quad (52)$$

と近似して考えることであるということが分かります。

\*33) 三変数以上の多変数関数に対しても、全く同様の考察がで  
きます。

ただし、式を見やすくするために、

$$X = x - x_0, Y = y - y_0$$

と書くことにしました。<sup>\*34)</sup> ここで、

$$H_f(p_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p_0) & \frac{\partial^2 f}{\partial x \partial y}(p_0) \\ \frac{\partial^2 f}{\partial y \partial x}(p_0) & \frac{\partial^2 f}{\partial y^2}(p_0) \end{pmatrix} \quad (53)$$

という行列を考えると、(52) 式は、行列の積を用いて、

$$\begin{aligned} f(x, y) - f(x_0, y_0) \\ \doteq \frac{1}{2} \begin{pmatrix} X & Y \end{pmatrix} H_f(p_0) \begin{pmatrix} X \\ Y \end{pmatrix} \end{aligned} \quad (54)$$

と表わせることが分かります。<sup>\*35)</sup> ここで現れた (53) 式で与えられる行列  $H_f(p_0)$  を関数  $f(x, y)$  の臨界点  $p_0$  におけるヘッシアン (Hessian) と呼びます。いま、(51) 式の右边を、

$$f(x) - f(x_0) \doteq \frac{1}{2} (x - x_0) f''(x_0) (x - x_0) \quad (55)$$

というように書き直して、(54) 式と (55) 式を見比べてみると、

$$f''(x_0) \longleftrightarrow H_f(p_0)$$

と対応していることが分かります。よって、「極値の判定」で活躍した  $f''(x_0)$  という値の二変数関数における対応物がヘッシアン  $H_f(p_0)$  であることが分かります。<sup>\*36)</sup>

ここでは、少し抽象的だと思われる方のために、ここで、問 2 の (4) の例をもとに、何をやっているのかということを見つめることにします。この場合、関数  $f(x, y)$  は、

$$f(x, y) = (x + y) e^{-xy}$$

という式で与えられているのでした。このとき、臨界点  $(x_0, y_0) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$  のまわりで、関数  $f(x, y)$  がどのように見えるのかということを考えてみます。そこで、前と同様に、この臨界点からのズレを  $(X, Y) \in \mathbb{R}^2$  と表わすことにして、 $x = \frac{1}{\sqrt{2}} + X, y = \frac{1}{\sqrt{2}} + Y$  を

\*34) ここで、 $\mathbb{R}^2$  の座標を  $(x, y) \rightsquigarrow (X, Y)$  と取り替えて考えるということは、 $\mathbb{R}^2$  の原点を  $(0, 0) \rightsquigarrow (x_0, y_0)$  と取り替えて考えるということです。

\*35) 皆さん、右辺の行列の積を計算して確かめて下さい。

\*36) 8 節と同様に、 $f''(x_0)$  という値を 1 行 1 列の行列 ( $f''(x_0)$ ) と同一視して考えると、「二変数化」にともない「二階の偏微分係数」 $f''(x_0)$  という概念が 1 行 1 列の行列から 2 行 2 列の行列に拡張されたと考えることができます。

$f(x, y)$  に代入してみると、

$$\begin{aligned} f\left(\frac{1}{\sqrt{2}} + X, \frac{1}{\sqrt{2}} + Y\right) \\ = \left(\frac{1}{\sqrt{2}} + X + \frac{1}{\sqrt{2}} + Y\right) e^{-\left(\frac{1}{\sqrt{2}} + X\right)\left(\frac{1}{\sqrt{2}} + Y\right)} \\ = (\sqrt{2} + X + Y) e^{-\frac{1}{2}} e^{-\frac{1}{\sqrt{2}}(X+Y) - XY} \end{aligned}$$

となることが分かります。ここで、指数関数  $e^x$  の Taylor 展開

$$e^x = 1 + x + \frac{1}{2!}x^2 + \dots$$

を用いて、 $e^{-\frac{1}{\sqrt{2}}(X+Y) - XY}$  の部分を展開して、 $X, Y$  について二次式の部分までを考えてみると、

$$\begin{aligned} e^{1/2} f\left(\frac{1}{\sqrt{2}} + X, \frac{1}{\sqrt{2}} + Y\right) \\ = (\sqrt{2} + X + Y) \left\{ 1 - \frac{1}{\sqrt{2}}(X + Y) - XY \right. \\ \left. + \frac{1}{2} \left( \frac{1}{\sqrt{2}}(X + Y) + XY \right)^2 + \dots \right\} \\ = \sqrt{2} \left\{ 1 - \frac{1}{\sqrt{2}}(X + Y) - XY + \frac{1}{4}(X + Y)^2 + \dots \right\} \\ + (X + Y) \left\{ 1 - \frac{1}{\sqrt{2}}(X + Y) + \dots \right\} \\ = \sqrt{2} + \frac{1}{2\sqrt{2}} \{ (X + Y)^2 - 4XY - 2(X + Y)^2 \} + \dots \\ = \sqrt{2} - \frac{1}{2\sqrt{2}} \{ X^2 + 6XY + Y^2 \} + \dots \end{aligned}$$

となることが分かります。したがって、 $X, Y$  について二次式の項までを考えると、

$$\begin{aligned} f\left(\frac{1}{\sqrt{2}} + X, \frac{1}{\sqrt{2}} + Y\right) - \sqrt{\frac{2}{e}} \\ \doteq -\frac{1}{2} \cdot \frac{1}{\sqrt{2e}} \begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \end{aligned}$$

と表わせることが分かりますが、これは、問 2 の (4) で求めたヘッシアンの式に他なりません。

以上より、臨界点のまわりで関数  $f(x, y)$  の様子を調べることは、勝手にひとつ与えられた対称行列<sup>\*37)</sup>

$$H = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

に対して、

$$\begin{aligned} F(X, Y) &= \begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \quad (56) \\ &= AX^2 + 2BXY + CY^2 \quad (57) \end{aligned}$$

という式により定まる二次関数  $F(X, Y)$  の様子を調

\*37) すなわち、行列の成分が対角線に関して対称になっている行列のことです。

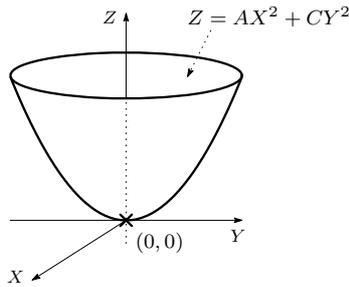


図9  $A > 0, C > 0$  のとき,  $(X, Y) = (0, 0)$  は, 二次関数  $F(X, Y) = AX^2 + CY^2$  の極小点になる.

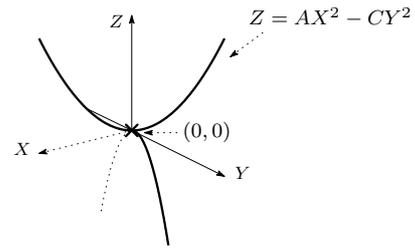


図10  $A > 0, C < 0$  のとき, 二次関数  $F(X, Y) = AX^2 - CY^2$  のグラフは馬の鞍のような形になる.

べることに帰着することが分かりました.\*38) こうした関数は二次形式と呼ばれていて, 線型代数学の知識を用いることで, その性質をより良く理解することができるのが分かっています. それに関して詳しいことは, いずれ, 線型代数学の方で学んでもらうことにして, ここでは, どのようなことを問題にして調べるのかということを少し説明してみることにします.

いま, 勝手な対称行列  $H$  に対して, (57) 式により定まる二次関数  $F(X, Y)$  の様子をより良く理解したいわけですが, そのために, まず, 様子が簡単に分かるような場合を考察してみることにします. すなわち,  $H$  が対角行列になっている場合に, 二次関数  $F(X, Y)$  の様子を調べてみることにします. このとき,  $B = 0$  となりますから, 二次関数  $F(X, Y)$  は,

$$F(X, Y) = AX^2 + CY^2$$

となります. すると, この表示から, 例えば,  $A, C > 0$  である場合には, どの方向に進んでも値は増えることが分かりますから,  $(X, Y) = (0, 0)$  が二次関数  $F(X, Y)$  の極小点になっていることが分かります (図9も参照). 同様に,  $A, C < 0$  のときには,  $(X, Y) = (0, 0)$  が二次関数  $F(X, Y)$  の極大点になっていることが分かります. また,  $A > 0, C < 0$  であるとする, 関数  $F(X, Y)$  のグラフは,  $X$  軸上では  $X^2$  のような形,  $Y$  軸上では  $-Y^2$  のような形をしていますから, 全体としては, 馬の鞍のような形をしていることが分かります (図10も参照).\*39) このように,  $H$  が対角行列の場合には, 二次関数  $F(X, Y)$  の様子は良く分かり

\*38) 第5回の問1のところで見たとように,  $C^2$  級の関数に対しては  $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$  となるので, そのヘッシアンは常に対称行列になることが分かります.

\*39) この意味で, このような臨界点を「鞍点」と呼ぶことがあります.

ます.

一方,  $B \neq 0$  のときは,  $F(X, Y)$  の式には  $2BXY$  という項が現われますから, 二次関数  $F(X, Y)$  のグラフの様子は, 式の形からはすぐには分かりません. そこで, この場合にも, 上のような「簡単な場合に考察を帰着できないか」と考えることは自然なことに思われます. このとき, アイデアは,  $F(X, Y)$  の形が簡単でない\*40)のは, 「座標の取り方が悪い」からであり, 「座標  $(X, Y)$  を上手く取り替える」ことで, 二次関数  $F(X, Y)$  を上のような「見やすい形」に変換できるのではないかと考えてみることで.

そこで, いま,  $a, b, c, d \in \mathbb{R}$  として,

$$\begin{cases} X = aX' + bY' \\ Y = cX' + dY' \end{cases} \quad (58)$$

という座標変換を考えてみます. ここで,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

と書くことにすると, (58) 式は,

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X' \\ Y' \end{pmatrix} = M \begin{pmatrix} X' \\ Y' \end{pmatrix} \quad (59)$$

と表わすことができます.\*41) このとき, 二次関数  $F(X, Y)$  を, 新しい座標  $(X', Y')$  で表わしてみると, (56) 式と (59) 式から,

\*40) すなわち,  $B \neq 0$  ということです.

\*41) ここで, 座標変換と言うからには,  $(X', Y')$  も  $(X, Y)$  について解けないといけません, このことは,  $M$  が正則行列であるということ, すなわち,  $M$  には逆行列が存在するという意味を意味します.

$$F(X, Y) = \begin{pmatrix} X' & Y' \end{pmatrix} {}^t M H M \begin{pmatrix} X' \\ Y' \end{pmatrix} \quad (60)$$

となることが分かります。ただし、行列  $M$  の転置行列 (transpose matrix) を、

$${}^t M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

と表わしました。よって、(60) 式から、(59) 式という座標変換のもとで、対称行列  $H$  は、

$$H \mapsto {}^t M H M$$

と「姿を変える」ことが分かります。

この変換公式から、座標  $(X, Y)$  を取り替えることにより、二次関数  $F(X, Y)$  を「見やすい形」にするという問題は、与えられた対称行列  $H$  に対して、 ${}^t M H M$  が対角行列という「見やすい形」になるような正則行列  $M$  を見つけるという問題に帰着することが分かります。実は、線型代数学の知識を用いて考察すると、勝手な対称行列  $H$  に対して、常にこのような正則行列  $M$  を見つけることができるということが分かります。すなわち、勝手な二次形式  $F(X, Y)$  は、適当な座標変換を行なうことで、常に「見やすい形」に変換できることが分かります。しかも、このような「上手い座標系」として、もともとの  $X$  軸、 $Y$  軸のように「座標軸が互いに直交するようなもの」が取れるということも分かります。<sup>\*42)</sup> こうした事実については、線型代数学で、「固有値」、「固有ベクトル」、「対称行列の直交行列による対角化」などの事柄を学んだ後に立ち返ってみて下さい。

もし、「座標軸が直交する」ということまでは要求しなくても、順番に「平方完成」してゆくことで、このような座標系を簡単に見つけることができます。例えば、 $A \neq 0$  なら、

$$\begin{aligned} AX^2 + 2BXY + CY^2 \\ = A \left( X + \frac{B}{A} Y \right)^2 + \frac{AC - B^2}{A} Y^2 \end{aligned} \quad (61)$$

と変形できることが分かりますから、

$$X' = X + \frac{B}{A} Y, \quad Y' = Y$$

と座標変換すれば良いことが分かります。また、 $C \neq 0$  なら、

<sup>\*42)</sup> これらの事柄は、二次元に限らず、一般の次元で成り立ちます。

$$\begin{aligned} AX^2 + 2BXY + CY^2 \\ = \frac{AC - B^2}{C} X^2 + C \left( Y + \frac{B}{C} X \right)^2 \end{aligned}$$

と変形できることが分かりますし、 $A = C = 0$  のときは、

$$(X + Y)^2 - (X - Y)^2 = 4XY$$

であることに注意すると、

$$2BXY = \frac{1}{2} B(X + Y)^2 - \frac{1}{2} B(X - Y)^2 \quad (62)$$

と変形できることが分かります。

二次関数  $F(X, Y)$  の様子を調べるためには、このような形でも十分なことが多いので、微積分学の教科書では、上のように「順番に平方完成した形」を用いて「極値の判定」を議論していることも多いです。最初に  $B = 0$  の場合として見たように「平方完成した後に現われる係数が両方とも同符号なら、 $(X, Y) = (0, 0)$  は極値点になっている」ことが分かりますから、例えば、(61) 式から、「 $A > 0, AC - B^2 > 0$  なら  $(X, Y) = (0, 0)$  は極小点である」ということが分かります。また、「平方完成した後に現われる係数が互いに異符号なら、 $(X, Y) = (0, 0)$  は鞍点になっている」ことが分かりますから、例えば、同じく (61) 式から、「 $A > 0, AC - B^2 < 0$  なら  $(X, Y) = (0, 0)$  は鞍点である」ということが分かります。あるいは、(62) 式に注目すると、「 $A = C = 0, B \neq 0$  なら  $(X, Y) = (0, 0)$  は鞍点である」ということが分かります。

例えば、上で考えた問 2 の (4) の例では、

$$\begin{aligned} f\left(\frac{1}{\sqrt{2}} + X, \frac{1}{\sqrt{2}} + Y\right) - \sqrt{\frac{2}{e}} \\ \doteq -\frac{1}{2\sqrt{2}e} (X^2 + 6XY + Y^2) \quad (63) \\ = -\frac{1}{2\sqrt{2}e} \{(X + 3Y)^2 - 8Y^2\} \end{aligned}$$

というように平方完成できますから、臨界点  $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$  は「鞍点」であることが分かります。<sup>\*43)</sup> 興味のある方は、問 2 のそれぞれの例で、臨界点の近くで、関数  $f(x, y)$  がどのような形をしているのかということを考察してみてください。

さて、関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  の大まか様子を調べるとい

<sup>\*43)</sup> 一変数関数のときと同様に、Taylor の定理を用いて、(63) 式の「 $\doteq$ 」を「 $=$ 」に置き換えて議論すると、きちんと証明することができます。興味のある方は、第 3 回の問 4 のところで説明した一変数関数の場合の議論を参考にして、「きちんとした議論」を考えてみてください。

うももとの観点から考えると、臨界点  $p_0$  でのヘッシアン  $H_f(p_0)$  を「対角化」した後で、対角成分に  $0$  が現われる場合には、「 $0$  に対応する方向では、関数  $f(x, y)$  の値は二次式の近似で見たときには定数である」ということが分かるだけなので、さらに詳しい情報を得るためには、

$$f(x, y) \doteq P_3(x, y)$$

と近似して、 $P_3(x, y)$  の様子を調べる必要があります。<sup>\*44)</sup> ただし、三次式以上の多項式関数に対しては、ここで考えたような「上手い一般論」は知られていません。<sup>\*45)</sup> そこで、この場合には個別に対処しないといけないこととなります。その意味で、「二次式まで考えれば決着がつく」ような関数は「性質の良い関数」であると考えられます。これは、「勝手の臨界点  $p_0$  に対して、臨界点  $p_0$  におけるヘッシアン  $H_f(p_0)$  を「対角化」した後で、対角成分に  $0$  が現われない」ということですが、実は、このことは「勝手の臨界点  $p_0$  に対して、臨界点  $p_0$  におけるヘッシアンの行列式  $\det H_f(x_0, y_0)$  が  $0$  でない」ということと同じこととなります。<sup>\*46)</sup> 数学では、このような臨界点のことを非退化な臨界点 (non-degenerate critical point) と呼び、「臨界点すべてが非退化になっている」ような「性質の良い関数」を Morse 関数と呼びます。

実は、 $\mathbb{R}^n$  のような「まっすぐな空間」上だけでなく、球面のような「曲がった空間」上でも微積分を行なうことができるのですが、「曲がった空間」上で関数の「臨界点の様子」を調べてみることで「空間の形」を調べることができるということが分かっています。すなわち、「臨界点がいくつ存在するか」というような「関数のあり方」と「どのような「穴」が空いているのか」というような「空間の形」とが密接に関係しているということが、二十世紀の半ば以降、数学者の間で広く認識されるようになり、現在では、例えば、Morse 理論と呼ばれる形でまとめられています。Morse 理論とは Morse 関数という「性質の良い関数の力」を借りて、三次元、四次元、… といった「高次元の見えない空間の様子」を「見ようとする」ものなのですが、現在の

\*44) それでも駄目なら、さらに、 $P_4(x, y)$  の様子、 $P_5(x, y)$  の様子、… というように、より高い次数の多項式関数の様子を調べる必要があります。

\*45) そもそも、そうした「上手い一般論」があり得るのかということさえも知られていません。

\*46) これらの事柄についても、いずれ線型代数で学んで下さい。

幾何学において、にわかには視覚化して捉えることができないような「高次元の空間」の「形を見る」ことに大いに役立っています。こうした事柄に興味を持たれた方は、現在学んでいる多変数関数の微積分学に対する理解を深めるとともに、「空間の形」を捉える手段である位相幾何学の言葉を学んでから、将来、Morse 理論に挑戦してみると良いのではないかと思います。

## 10. 問3の解答

$X = \frac{x}{x+y}$ ,  $Y = \frac{x+y}{x+y+z}$ ,  $Z = x + y + z$  と書いて、それぞれの関数の偏微分を計算してみると、 $p = (x, y, z) \in \mathbb{R}^3$  として、Jacobi 行列  $J_f(p)$  は、

$$\begin{aligned} J_f(p) &= \begin{pmatrix} \frac{\partial X}{\partial x}(p) & \frac{\partial X}{\partial y}(p) & \frac{\partial X}{\partial z}(p) \\ \frac{\partial Y}{\partial x}(p) & \frac{\partial Y}{\partial y}(p) & \frac{\partial Y}{\partial z}(p) \\ \frac{\partial Z}{\partial x}(p) & \frac{\partial Z}{\partial y}(p) & \frac{\partial Z}{\partial z}(p) \end{pmatrix} \\ &= \begin{pmatrix} \frac{y}{(x+y)^2} & \frac{-x}{(x+y)^2} & 0 \\ \frac{z}{(x+y+z)^2} & \frac{z}{(x+y+z)^2} & \frac{-(x+y)}{(x+y+z)^2} \\ 1 & 1 & 1 \end{pmatrix} \end{aligned}$$

となることが分かります。<sup>\*47)</sup> また、Jacobi 行列  $J_f(p)$  の行列式は、<sup>\*48)</sup> 例えば、

$$\begin{aligned} \det J_f(p) &= \frac{1}{(x+y)^2(x+y+z)^2} \cdot \begin{vmatrix} y & -x & 0 \\ z & z & -(x+y) \\ 1 & 1 & 1 \end{vmatrix} \\ &= \frac{1}{(x+y)^2(x+y+z)^2} \cdot \begin{vmatrix} x+y & -x & 0 \\ 0 & z & -(x+y) \\ 0 & 1 & 1 \end{vmatrix} \\ &= \frac{(x+y)}{(x+y)^2(x+y+z)^2} \cdot \begin{vmatrix} z & -(x+y) \\ 1 & 1 \end{vmatrix} \\ &= \frac{1}{(x+y)(x+y+z)^2} \cdot \begin{vmatrix} x+y+z & -(x+y) \\ 0 & 1 \end{vmatrix} \\ &= \frac{(x+y+z)}{(x+y)(x+y+z)^2} \\ &= \frac{1}{(x+y)(x+y+z)} \end{aligned}$$

というように計算することができます。

\*47) ここで、 $\frac{\partial X}{\partial x}$  や  $\frac{\partial Y}{\partial x}$ ,  $\frac{\partial Y}{\partial y}$  などを計算するときに、

$$X = \frac{x}{x+y} = 1 - \frac{y}{x+y}, \quad Y = \frac{x+y}{x+y+z} = 1 - \frac{z}{x+y+z}$$

など書き直してから計算すると、偏微分の計算の間違いが少なくなるかもしれません。

\*48) この行列式を「Jacobian」と呼んだりします。

次に、 $D$  が、写像  $f$  により、どのような集合に写されるのかということを考えてみます。そのために、まず、 $x > 0, y > 0, z > 0$  であるとしたときに、

$$\begin{cases} X = \frac{x}{x+y} \\ Y = \frac{x+y}{x+y+z} \\ Z = x+y+z \end{cases}$$

がどのような条件を満たすことになるのかということを考えてみます。すると、定義から、

$$0 < X, 0 < Y, 0 < Z \quad (64)$$

となることがすぐに分かります。また、

$$x < x+y, x+y < x+y+z$$

となることに注意すると、

$$X < 1, Y < 1 \quad (65)$$

となることも分かります。そこで、さらに、 $x+y+z < 1$  であるとする、

$$Z < 1 \quad (66)$$

となることが分かりますから、結局、(64) 式、(65) 式、(66) 式から、

$$\begin{cases} x > 0 \\ y > 0 \\ z > 0 \\ x+y+z < 1 \end{cases} \implies \begin{cases} 0 < X < 1 \\ 0 < Y < 1 \\ 0 < Z < 1 \end{cases} \quad (67)$$

となることが分かります。

一方、 $x, y, z$  は、 $X, Y, Z$  を用いて、

$$\begin{cases} x = XYZ \\ y = (1-X)YZ \\ z = (1-Y)Z \end{cases}$$

と表わせることに注意すると、逆に、 $0 < X < 1, 0 < Y < 1, 0 < Z < 1$  のとき、

$$x > 0, y > 0, z > 0 \quad (68)$$

となることが分かります。さらに、 $Z < 1$  とすれば、 $x+y+z = Z$  なので、

$$x+y+z < 1 \quad (69)$$

となることが分かります。よって、(68) 式、(69) 式

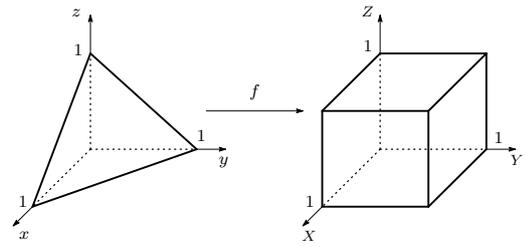


図 11 写像  $f$  によって、領域  $D$  は、立方体の内部に移される。

から、

$$\begin{cases} 0 < X < 1 \\ 0 < Y < 1 \\ 0 < Z < 1 \end{cases} \implies \begin{cases} x > 0 \\ y > 0 \\ z > 0 \\ x+y+z < 1 \end{cases} \quad (70)$$

となることも分かりました。したがって、(67) 式、(70) 式より、

$$D = \left\{ (x, y, z) \in \mathbb{R}^3 \mid \begin{array}{l} x > 0, y > 0, z > 0, \\ x+y+z < 1 \end{array} \right\}$$

は、写像  $f$  により、

$$\left\{ (X, Y, Z) \in \mathbb{R}^3 \mid \begin{array}{l} 0 < X < 1, 0 < Y < 1, \\ 0 < Z < 1 \end{array} \right\}$$

という立方体の内部に写されることが分かりました (図 11 を参照)。

興味のある方は、 $D$  が三角錐の内部であることと、写像  $f$  により、この三角錐の内部が立方体の内部にどのように幾何学的に写されているのかということも考えてみて下さい。

## 11. Jacobi 行列とは

さて、問 2 のところでは、「接平面の様子を調べることにより、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  の様子を調べる」ということを考えました。これは、 $p_0 = (x_0, y_0) \in \mathbb{R}^2$  として、「点  $p_0$  における関数  $f(x, y)$  の Taylor 多項式  $P_1(x, y)$  を用いて、

$$f(x, y) \doteq P_1(x, y)$$

というように近似して、一次の多項式  $P_1(x, y)$  の様子を調べることにより、点  $p_0$  の近くでの関数  $f(x, y)$  の大まかな様子を理解する」ということでした。そこで、ここでは、同じアイデアを  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  という写像に対して適用するとどうということになるのかと

いうことを考えてみることにします。具体的なイメージを持ちやすいように、ここでは、問3の例に合せて、 $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  という場合に説明することにします。<sup>\*49)</sup>

そこで、いま、 $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  という写像が与えられているとします。これでは抽象的だと思われる方は、問3の例を考えてもらっても構いません。<sup>\*50)</sup> また、問3の解答で使った記号と合わせるために、 $(x, y, z) \in \mathbb{R}^3$  に対して、 $f(x, y, z) \in \mathbb{R}^3$  の成分を、

$$f(x, y, z) = (X(x, y, z), Y(x, y, z), Z(x, y, z))$$

と表わすことにします。普通は、

$$f(x, y, z) = (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z))$$

などと表わすことも多く、とても便利な記法なのですが、慣れないうちは文字が似ていると「いらぬ混乱」をしてしまう恐れもあるので、上のように表わすことにします。すると、「写像  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  を与える」ということは「 $X(x, y, z), Y(x, y, z), Z(x, y, z)$  という三つの関数を与える」ということと同じことになりませんが、関数  $X(x, y, z), Y(x, y, z), Z(x, y, z)$  がすべて滑らかな関数であるときに、写像  $f$  は滑らかな写像と呼ばれます。

さて、写像  $f$  の成分である  $X, Y, Z$  は普通の三変数関数ですから、二変数関数のときと同様に、それぞれの関数に対して、「接空間を描いて」関数の大まかな様子を調べてみるということが出来ます。この節の最初でも触れたように、これは、与えられた関数の近似として、Taylor 展開の一次式の項までを取り、一次の近似多項式の様子を調べることにより、関数の大まかな様子を理解するというものでした。

そこで、いま、点  $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$  を、勝手にひとつ取ってきて、点  $p_0$  の近くで、写像  $f(x, y, z)$  がどのように見えるのかということを考察してみることになります。すると、二変数関数のときと同様に、Taylor 展開の一次式の項までを取って近似してみると、点  $p_0$  の近くで、関数  $X, Y, Z$  は、それぞれ、

$$X(x, y, z) \doteq X(p_0) + \frac{\partial X}{\partial x}(p_0)(x-x_0) + \frac{\partial X}{\partial y}(p_0)(y-y_0) + \frac{\partial X}{\partial z}(p_0)(z-z_0)$$

<sup>\*49)</sup> 一般の写像  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  の場合でも、全く同様に考えることが出来ます。

<sup>\*50)</sup> 問3の例の場合には、本当は、 $f: D \rightarrow \mathbb{R}^3$  というように、写像  $f$  の定義域が  $\mathbb{R}^3$  全体ではなく、その部分集合  $D$  となっているわけですが、以下の議論では、この差は全く関係ありません。

$$\begin{aligned} Y(x, y, z) &\doteq Y(p_0) + \frac{\partial Y}{\partial x}(p_0)(x-x_0) \\ &\quad + \frac{\partial Y}{\partial y}(p_0)(y-y_0) + \frac{\partial Y}{\partial z}(p_0)(z-z_0) \\ Z(x, y, z) &\doteq Z(p_0) + \frac{\partial Z}{\partial x}(p_0)(x-x_0) \\ &\quad + \frac{\partial Z}{\partial y}(p_0)(y-y_0) + \frac{\partial Z}{\partial z}(p_0)(z-z_0) \end{aligned}$$

というように見えることが分かります。<sup>\*51)</sup> 行列の記号を用いると、これらの式は、

$$\begin{aligned} &\begin{pmatrix} X(x, y, z) \\ Y(x, y, z) \\ Z(x, y, z) \end{pmatrix} - \begin{pmatrix} X(p_0) \\ Y(p_0) \\ Z(p_0) \end{pmatrix} \\ &\doteq \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) & \frac{\partial X}{\partial z}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) & \frac{\partial Y}{\partial z}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) & \frac{\partial Z}{\partial z}(p_0) \end{pmatrix} \begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \quad (71) \end{aligned}$$

というようにまとめて書き表わすことが出来ます。

このことは、 $f$  の定義域である  $\mathbb{R}^3$  の原点と  $f$  の行き先である  $\mathbb{R}^3$  の原点を、それぞれ、 $p_0 = (x_0, y_0, z_0)$  と  $f(p_0) = (X(p_0), Y(p_0), Z(p_0))$  に取り直して、

$$\begin{aligned} (x', y', z') &= (x-x_0, y-y_0, z-z_0) \\ (X', Y', Z') &= (X-X(p_0), Y-Y(p_0), Z-Z(p_0)) \end{aligned}$$

という新しい座標で表わすと、点  $p_0 \in \mathbb{R}^3$  の近くでは、写像  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  は、

$$J_f(p_0) = \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) & \frac{\partial X}{\partial z}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) & \frac{\partial Y}{\partial z}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) & \frac{\partial Z}{\partial z}(p_0) \end{pmatrix}$$

という行列を掛け算することによって定まる(線型)写像のように見えるということを意味しています。ここで現われた  $J_f(p_0)$  という行列を点  $p_0 \in \mathbb{R}^3$  における写像  $f$  の Jacobi 行列と呼びます。すなわち、Jacobi 行列とは、滑らかな写像  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  に対して、Taylor 展開の一次式の項までを考えて近似したときに現われる行列であると理解することが出来ます。

これでは抽象的だと思われる方のために、問3の例を用いて、何を考えているのかということをも具体的に見てみることにします。すると、問3の例では、写像  $f(x, y, z)$  は、

$$f(x, y, z) = \left( \frac{x}{x+y}, \frac{x+y}{x+y+z}, x+y+z \right)$$

という式により与えられていました。そこで、例とし

<sup>\*51)</sup> 二変数関数のときと同様に、スペースを節約するために、 $(x_0, y_0, z_0)$  を  $p_0$  と表わすことにしました。

て、点  $p_0 = (1, 0, 0) \in \mathbb{R}^3$  の近くで、写像  $f$  がどのように見えるのかということを考えてみることにします。すなわち、写像  $f$  を「 $(x-1), y, z$  の多項式の姿」に「化かす」ことを考えてみます。<sup>\*52)</sup>

そこで、

$$\begin{aligned} X &= \frac{x}{x+y} \\ &= 1 - \frac{y}{x+y} \\ &= 1 - \frac{y}{1 + \{(x-1) + y\}} \end{aligned}$$

と表わして、

$$\frac{1}{1+T} = 1 - T + T^2 - \dots$$

という Taylor 展開を用いると、 $T = (x-1) + y$  として、

$$\begin{aligned} X &= 1 - \frac{y}{1 + \{(x-1) + y\}} \\ &= 1 - y \cdot [1 - \{(x-1) + y\} + \dots] \\ &= 1 - y + (x-1)y + y^2 + \dots \end{aligned} \quad (72)$$

となることが分かります。同様に、

$$\begin{aligned} Y &= \frac{x+y}{x+y+z} \\ &= 1 - \frac{z}{x+y+z} \\ &= 1 - \frac{z}{1 + \{(x-1) + y + z\}} \\ &= 1 - z \cdot [1 - \{(x-1) + y + z\} + \dots] \\ &= 1 - z + (x-1)z + yz + z^2 + \dots \end{aligned} \quad (73)$$

$$\begin{aligned} Z &= x + y + z \\ &= 1 + (x-1) + y + z \end{aligned} \quad (74)$$

となることが分かります。したがって、 $(x-1), y, z$  について一次式の部分までを考えると、(72) 式、(73) 式、(74) 式から、点  $p_0 = (1, 0, 0) \in \mathbb{R}^3$  の近くで、写像  $f$  は、

$$\begin{pmatrix} X-1 \\ Y-1 \\ Z-1 \end{pmatrix} \doteq \begin{pmatrix} -y \\ -z \\ (x-1) + y + z \end{pmatrix}$$

\*52) 本当は、 $p_0 = (1, 0, 0) \notin D$  なので、写像  $f$  の定義域を、

$$\{(x, y, z) \in \mathbb{R}^3 \mid x+y \neq 0, x+y+z \neq 0\}$$

に拡張して考えているのだと思って下さい。

$$= \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y \\ z \end{pmatrix} \quad (75)$$

というように見えることが分かりました。

もちろん、(75) 式の右辺に現れた行列は、問 3 で求めた Jacobi 行列  $J_f(p)$  に  $p = (1, 0, 0)$  を代入したものに他なりません。興味のある方は、一般に、 $p_0 = (x_0, y_0, z_0) \in D$  に対して、例えば、上のようなやり方で、 $X, Y, Z$  を「 $(x-x_0), (y-y_0), (z-z_0)$  の多項式の姿」に「化かす」ことを試みることで、問 3 で求めた Jacobi 行列  $J_f(p)$  に  $p = (x_0, y_0, z_0)$  を代入した行列が得られることを確かめてみて下さい。

## 12. 逆関数定理とは

さて、一変数関数の場合に、Taylor 展開から得られる一次の近似多項式の様子を調べることで、もともとの関数の大まかな様子が理解できるということは、皆さん、良くご存じのことではないかと思えます。実際、第 3 回の問 4 のところでは、一変数関数  $f: \mathbb{R} \rightarrow \mathbb{R}$  に対して、例えば、「ある点  $x_0 \in \mathbb{R}$  で  $f'(x_0) > 0$  となるならば、 $x_0$  の近くで、関数  $f(x)$  は単調増加である」という事実を、次のように考えました。

いま、点  $x_0$  における関数  $f(x)$  の Taylor 多項式

$$P_1(x) = f(x_0) + f'(x_0)(x - x_0)$$

を考えると、 $P_1(x)$  は一次関数ですから、「 $f'(x_0) > 0$  なら  $P_1(x)$  は単調増加関数になる」ということが立ち所に分かります。すると、「点  $x_0$  の近くでは、関数  $f(x)$  は  $P_1(x)$  という単調増加な一次関数のように見える」のですから、「関数  $f(x)$  の方も  $x_0$  の近くで単調増加になっているのではないか」と期待されます。実際、Taylor の定理を用いると、この期待が正しいことを確かめることができるのでした。こうした事実の存在は、例えば、単調増加性といった「関数の定性的な性質」は Taylor 展開から得られる一次の近似多項式に十分良く反映されていることを示唆しています。

ここで、もともとの関数  $f(x)$  に対しては「 $x_0$  の近くでは」という「但し書き」が付いていることに注意して下さい。そもそも、Taylor 展開における二次式以上の項を落として考えたのは、「 $x$  が  $x_0$  に十分に近ければ、 $(x - x_0)$  に関する高次の項はとても小さいから無視しても良からう」と考えたからでした。したがっ

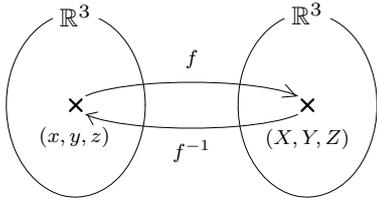


図 12 写像  $f(x, y, z)$  によって  $(x, y, z)$  と  $(X, Y, Z)$  が一対一に対応しているときには、写像  $f$  の逆写像  $f^{-1}(X, Y, Z)$  を考えることができる。

て、 $x_0$  に余り近くないところでは、この近似が妥当ではなくなりますから、 $x_0$  から遠い場所での関数  $f(x)$  の振る舞いが、点  $x_0$  における Taylor 多項式  $P_1(x)$  にうまく反映されるという保証は、一般には、全くないということになります。そのような理由で、「 $x_0$  の近くでは」という「但し書き」が付いています。

さて、「関数の定性的な性質」は Taylor 展開から得られる一次の近似多項式に十分良く反映されているだろうという直感は、何も一変数関数に限ったものではありません。例えば、滑らかな写像  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  に対しては、写像  $f$  の「定性的な性質」は Jacobian 行列を掛け算することにより定まる（線型）写像  $J_f(p_0): \mathbb{R}^3 \rightarrow \mathbb{R}^3$  に十分良く反映されると期待することは自然なことに思われます。実際、こうした期待が「正しい」ということを示すような定理がいくつか知られているのですが、その中の代表的なものとして「逆関数定理」というものがあります。

いま、 $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  という写像に対して、勝手にひとつ点  $(X, Y, Z) \in \mathbb{R}^3$  を取ってきたときに、 $f(x, y, z) = (X, Y, Z)$  となるような点  $(x, y, z) \in \mathbb{R}^3$  がただひとつだけ見つかると仮定してみます。このような状況では、写像  $f$  によって、定義域である  $\mathbb{R}^3$  の点と値域である  $\mathbb{R}^3$  の点がぴったり一対一に対応しているので、逆に、 $(X, Y, Z) \in \mathbb{R}^3$  に対して、 $f(x, y, z) = (X, Y, Z)$  となるような点  $(x, y, z)$  を対応させるという写像を考えることができます（図 12 を参照）。この写像を、写像  $f$  の逆写像と呼び、 $f^{-1}$  と表わしたり、変数を明記したい場合には、 $f^{-1}(X, Y, Z)$  などと表わしたりします。直感的には、写像

$$f(x, y, z) = (X(x, y, z), Y(x, y, z), Z(x, y, z))$$

が「逆写像を持つ」とは「 $(x, y, z)$  が  $(X, Y, Z)$  の関数として、逆に解ける」ということです。このような逆写像が存在するかどうかを考えると有効なのが

「逆関数定理」です。

いま、(71) 式より、 $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$  という点のまわりでは、写像  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  は、

$$\begin{pmatrix} X(x, y, z) \\ Y(x, y, z) \\ Z(x, y, z) \end{pmatrix} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} \doteq J_f(p_0) \begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \quad (76)$$

というように見えるのでした。<sup>\*53)</sup> このとき、(76) 式の右辺である Jacobian 行列  $J_f(p_0)$  を掛け算することによって定まる（線型）写像  $J_f(p_0): \mathbb{R}^3 \rightarrow \mathbb{R}^3$  に対しては、「逆写像が存在するかどうかという問題」を「Jacobian 行列  $J_f(p_0)$  が正則行列<sup>\*54)</sup>かどうかという問題」に読み替えることができます。さらに、線型代数学の知識を用いれば、「Jacobian 行列の行列式  $\det J_f(p_0)$  が 0 かどうか」を調べることで、この問題の正否を簡単に判定することもできます。

そこで、いま、Jacobian 行列  $J_f(p_0)$  が正則行列であると仮定してみます。すると、(76) 式より、

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \doteq J_f(p_0)^{-1} \begin{pmatrix} X-X_0 \\ Y-Y_0 \\ Z-Z_0 \end{pmatrix} \quad (77)$$

というように、近似的には、 $(x, y, z)$  が  $(X, Y, Z)$  の関数として、逆に解けてしまうことが分かります。逆関数定理とは、このように一次式による近似で、(76) 式が (77) 式のように解けてしまうと、「点  $p_0 \in \mathbb{R}^3$  の近くでは」、近似なしに  $(x, y, z)$  が  $(X, Y, Z)$  の関数として、逆に解けてしまうということを主張する定理です。すなわち、一次式の近似に対して逆写像が存在するときには、「点  $p_0 \in \mathbb{R}^3$  の近くでは、もとの写像  $f$  に対しても逆写像が存在する」ということを主張する定理です。

皆さんも、いずれこうした定理を学ぶことになると思いますが、そのときに「確かに、こういうことは期待できそうだ」と思えるように、今のうちに考え方の基礎を、皆さんなりに納得しておかれたら良いのではないかと思います。興味のある方は、問 3 で考えた写像  $f$  の式を逆に解いた

$$g(X, Y, Z) = (XYZ, (1-X)YZ, (1-Y)Z)$$

という写像  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  について、それぞれの点

<sup>\*53)</sup> ここで、 $f(p_0) = (X(p_0), Y(p_0), Z(p_0)) = (X_0, Y_0, Z_0)$  と書きました。

<sup>\*54)</sup> すなわち、逆行列のある行列のことです。

$(x, y, z) \in \mathbb{R}^3$  に対して,

$$\{(X, Y, Z) \in \mathbb{R}^3 \mid g(X, Y, Z) = (x, y, z)\}$$

という集合がどうなるのかということを調べてみて下さい。また、それぞれの点  $(X, Y, Z) \in \mathbb{R}^3$  における写像  $g$  の Jacobi 行列  $J_g(X, Y, Z)$  の行列式を求めてみて下さい。そして、この行列式の値が 0 でないような点のまわりでは、 $g$  の逆写像が存在するのに、行列式の値が 0 となるような点のまわりでは、 $g$  の逆写像が存在しないことを確かめてみて下さい。

### 13. 写像の微分とは

さて、8 節では、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対する「微分」の概念の自然な拡張は「全微分」と呼ばれる

$$(df)_{p_0} = \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix}$$

という 1 行 2 列の行列であることを注意しました。また、11 節では、 $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  という写像に対して、Jacobi 行列と呼ばれる

$$J_f(p_0) = \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) & \frac{\partial X}{\partial z}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) & \frac{\partial Y}{\partial z}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) & \frac{\partial Z}{\partial z}(p_0) \end{pmatrix}$$

という 3 行 3 列の行列を考えました。これらの概念の間の関係をより良く理解するために、ここでは、一変数関数から始めて、関数の概念が拡張されるに伴って、「微分」の概念がどのように拡張されるのかということを少し反省してみることにします。

上でも注意したように、8 節では、

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

↓ 多変数化

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

というように、関数の変数を多変数化したときに、それに伴って「微分」の概念が、

$$\begin{pmatrix} \frac{df}{dx}(x_0) \end{pmatrix} \xrightarrow{\text{列が増える}} (df)_{p_0} = \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix}$$

というように拡張されることを見ました。そこで、次に、

$$f: \mathbb{R} \rightarrow \mathbb{R} \xrightarrow{\text{ベクトル値}} f: \mathbb{R} \rightarrow \mathbb{R}^3$$

というように、関数の値域の方を「ベクトル値」化したときに、「微分」の概念がどのように拡張されるのか

ということを考えてみることにします。このとき、注目すべき点は、

$$\frac{df}{dx}(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (78)$$

という一変数関数  $f: \mathbb{R} \rightarrow \mathbb{R}$  に対する微分係数の定義式は、一変数のベクトル値関数  $f: \mathbb{R} \rightarrow \mathbb{R}^3$  に対しても、そのままの形で意味を持つということです。すなわち、いま、ベクトル値関数  $f: \mathbb{R} \rightarrow \mathbb{R}^3$  の成分を、

$$f(x) = \begin{pmatrix} X(x) \\ Y(x) \\ Z(x) \end{pmatrix}$$

と表わすことにすると、

$$f(x_0 + h) - f(x_0) = \begin{pmatrix} X(x_0 + h) - X(x_0) \\ Y(x_0 + h) - Y(x_0) \\ Z(x_0 + h) - Z(x_0) \end{pmatrix}$$

となることが分かりますから、

$$\frac{f(x_0 + h) - f(x_0)}{h} = \begin{pmatrix} \frac{X(x_0 + h) - X(x_0)}{h} \\ \frac{Y(x_0 + h) - Y(x_0)}{h} \\ \frac{Z(x_0 + h) - Z(x_0)}{h} \end{pmatrix} \quad (79)$$

となることが分かります。したがって、(78) 式と (79) 式から、ベクトル値関数  $f: \mathbb{R} \rightarrow \mathbb{R}^3$  の点  $x_0 \in \mathbb{R}$  における微分係数は、

$$\frac{df}{dx}(x_0) = \begin{pmatrix} \frac{dX}{dx}(x_0) \\ \frac{dY}{dx}(x_0) \\ \frac{dZ}{dx}(x_0) \end{pmatrix}$$

という式によって与えられることが分かります。すなわち、

$$f: \mathbb{R} \rightarrow \mathbb{R} \xrightarrow{\text{ベクトル値}} f: \mathbb{R} \rightarrow \mathbb{R}^3$$

というように、関数の値を「ベクトル値」化したときに、それに伴って「微分」の概念が、

$$\left( \frac{df}{dx}(x_0) \right) \xrightarrow{\text{行が増える}} \frac{df}{dx}(x_0) = \begin{pmatrix} \frac{dX}{dx}(x_0) \\ \frac{dY}{dx}(x_0) \\ \frac{dZ}{dx}(x_0) \end{pmatrix}$$

というように拡張されることが分かりました。

そこで、次に、ベクトル値関数  $f: \mathbb{R} \rightarrow \mathbb{R}^3$  の多変数化、あるいは、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  の「ベクトル値」化として、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  という写像の点  $p_0 \in \mathbb{R}^2$  における「全微分」 $(df)_{p_0}$  を考えてみます。すると、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  の成分を、

$$f(x) = \begin{pmatrix} X(x, y) \\ Y(x, y) \\ Z(x, y) \end{pmatrix}$$

と表わすことにして、

$$\begin{aligned} (df)_{p_0} &= \begin{pmatrix} (dX)_{p_0} \\ (dY)_{p_0} \\ (dZ)_{p_0} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) \end{pmatrix} \end{aligned}$$

となることが分かります。これは、点  $p_0 \in \mathbb{R}^2$  における写像  $f$  の Jacobi 行列に他なりません。

以上から、一変数関数から始めて、

$$\begin{array}{ccc} f: \mathbb{R} \rightarrow \mathbb{R} & \xrightarrow{\text{ベクトル値}} & f: \mathbb{R} \rightarrow \mathbb{R}^3 \\ \downarrow \text{多変数化} & & \downarrow \text{多変数化} \\ f: \mathbb{R}^2 \rightarrow \mathbb{R} & \xrightarrow{\text{ベクトル値}} & f: \mathbb{R}^2 \rightarrow \mathbb{R}^3 \end{array}$$

というように関数の概念が拡張されるに伴って、

$$\begin{array}{ccc} \left( \frac{df}{dx} \right) & \xrightarrow{\text{行が増える}} & \begin{pmatrix} \frac{dX}{dx} \\ \frac{dY}{dx} \\ \frac{dZ}{dx} \end{pmatrix} \\ \downarrow \text{列が増える} & & \downarrow \text{列が増える} \\ \left( \frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right) & \xrightarrow{\text{行が増える}} & \begin{pmatrix} \frac{\partial X}{\partial x} & \frac{\partial X}{\partial y} \\ \frac{\partial Y}{\partial x} & \frac{\partial Y}{\partial y} \\ \frac{\partial Z}{\partial x} & \frac{\partial Z}{\partial y} \end{pmatrix} \end{array}$$

というように「微分」の概念が拡張されることが分かります。<sup>\*55)</sup> これらの行列は、伝統的には、

$$\begin{array}{ccc} \text{微分} & \xrightarrow{\text{行が増える}} & \text{(ベクトル値関数の) 微分} \\ \downarrow \text{列が増える} & & \downarrow \text{列が増える} \\ \text{全微分} & \xrightarrow{\text{行が増える}} & \text{Jacobi 行列} \end{array}$$

というように、それぞれ異なる名前で呼ばれていましたが、現在では、写像  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  の微分として統一的理解されています。

そこで、写像  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  の微分とは何かということをも、もう一度、見直してみます。以下では、話を具体的にするために、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  という写像に対して、写像  $f$  の微分とは何かということの説明

<sup>\*55)</sup> ここで、スペースの節約のために、微分を考える点である  $x_0 \in \mathbb{R}$  や  $p_0 \in \mathbb{R}^2$  を省略して表わすことにしました。

してみることにします。<sup>\*56)</sup>

さて、上では、

$$f: \mathbb{R} \rightarrow \mathbb{R} \xrightarrow{\text{ベクトル値}} f: \mathbb{R} \rightarrow \mathbb{R}^3$$

という関数の値域の「ベクトル値」化にあたり、点  $x_0 \in \mathbb{R}$  における関数  $f$  の微分係数という概念が、(78) 式という見かけの上では全く同じ定義式を用いて理解できることを注意しました。そこで、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  の微分とは何かということを考えるに当たっても、

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} \xrightarrow{\text{ベクトル値}} f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

という図式に注目して、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  に対する微分概念である「全微分」の定義式をもとにして考えてみることにします。

すると、8 節で見たように、 $\mathbb{R}^2$  の点  $p_0$  における接空間を、

$$T_{p_0}\mathbb{R}^2 = \{ \mathbf{v} \mid \mathbf{v} \text{ は点 } p_0 \text{ を始点とするベクトル.} \}$$

として、二変数関数  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  の点  $p_0 \in \mathbb{R}^2$  における全微分  $(df)_{p_0}$  は、それぞれのベクトル  $\mathbf{v} \in T_{p_0}\mathbb{R}^2$  に対して、

$$(df)_{p_0}(\mathbf{v}) = \lim_{h \rightarrow 0} \frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \quad (80)$$

という点  $p_0$  における  $\mathbf{v}$  方向の方向微分係数を対応させる写像

$$(df)_{p_0}: T_{p_0}\mathbb{R}^2 \rightarrow \mathbb{R}$$

として理解することができるのでした。すなわち、 $T_{p_0}\mathbb{R}^2 = \mathbb{R}^2$  と同一視して、

$$\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in T_{p_0}\mathbb{R}^2 = \mathbb{R}^2$$

と表わして、(80) 式の右辺の極限值を計算してみると、第 5 回の問 1 のところで見たように、

$$\begin{aligned} (df)_{p_0}(\mathbf{v}) &= \frac{\partial f}{\partial x}(p_0) \cdot a + \frac{\partial f}{\partial y}(p_0) \cdot b \\ &= \left( \frac{\partial f}{\partial x}(p_0) \quad \frac{\partial f}{\partial y}(p_0) \right) \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

となることが分かるので、

$$(df)_{p_0} = \left( \frac{\partial f}{\partial x}(p_0) \quad \frac{\partial f}{\partial y}(p_0) \right)$$

<sup>\*56)</sup> 一般の  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  という写像に対しても全く同様の議論ができます。

というように同一視ができるというわけです。<sup>\*57)</sup>

そこで、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  の点  $p_0$  における Jacobi 行列に対して、(80) 式をもとにして、同様の解釈ができるのかどうかということを考えてみることにします。<sup>\*58)</sup> いま、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  の成分を、

$$f(p) = \begin{pmatrix} X(p) \\ Y(p) \\ Z(p) \end{pmatrix}$$

と表わすことにすると、前と同様に、

$$\frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} = \begin{pmatrix} \frac{X(p_0 + h\mathbf{v}) - X(p_0)}{h} \\ \frac{Y(p_0 + h\mathbf{v}) - Y(p_0)}{h} \\ \frac{Z(p_0 + h\mathbf{v}) - Z(p_0)}{h} \end{pmatrix}$$

となるのが分かりますから、

$$\begin{aligned} (df)_{p_0}(\mathbf{v}) &= \begin{pmatrix} (dX)_{p_0}(\mathbf{v}) \\ (dY)_{p_0}(\mathbf{v}) \\ (dZ)_{p_0}(\mathbf{v}) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) \cdot a + \frac{\partial X}{\partial y}(p_0) \cdot b \\ \frac{\partial Y}{\partial x}(p_0) \cdot a + \frac{\partial Y}{\partial y}(p_0) \cdot b \\ \frac{\partial Z}{\partial x}(p_0) \cdot a + \frac{\partial Z}{\partial y}(p_0) \cdot b \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

となるのが分かります。したがって、二変数関数の場合と同様に、(80) 式で定義される写像を、

$$(df)_{p_0}: T_{p_0}\mathbb{R}^2 \rightarrow \mathbb{R}^3 \quad (81)$$

として、

$$(df)_{p_0} = \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) \end{pmatrix}$$

と同一視できることが分かります。

こうして、写像  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  の Jacobi 行列も、(80) 式をもとにして、同様の解釈ができることが分かります。

<sup>\*57)</sup> 線形代数学の言葉使いを用いると、 $T_{p_0}\mathbb{R}^2 = \mathbb{R}^2$  という「座標付け」のもとでの「線型写像  $(df)_{p_0}$  の表現行列」が、

$$\begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) \end{pmatrix}$$

であるということになります。

<sup>\*58)</sup> 前と同様、(80) 式は、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  という写像に対しても、そのままの形で意味がある式であることに注意して下さい。

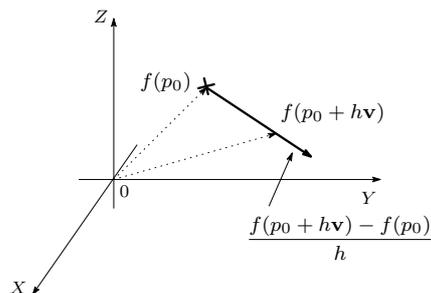


図 13 勝手な実数  $h \in \mathbb{R}$  に対して、 $\frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h}$  は  $f(p_0)$  を始点とするベクトルであると考えられる。

した。ただし、(80) 式の定義式をもう少し注意深く眺めると、(81) 式の写像  $(df)_{p_0}$  の値域である  $\mathbb{R}^3$  も、単に  $\mathbb{R}^3$  であるというよりも、むしろ、 $\mathbb{R}^3$  の点  $f(p_0)$  における接空間  $T_{f(p_0)}\mathbb{R}^3$  であると考えの方がより自然であることが、次のようにして分かります。

いま、 $\mathbb{R}^3$  の点  $f(p_0 + h\mathbf{v})$  や  $f(p_0)$  を  $\mathbb{R}^3$  のベクトルと同一視して考えると、 $f(p_0 + h\mathbf{v}) - f(p_0)$  は点  $f(p_0)$  を始点とするベクトルとなるのが分かります。したがって、そのベクトルを  $\frac{1}{h}$  倍することによって得られる  $\frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h}$  も点  $f(p_0)$  を始点とするベクトルであると考えることができます。すなわち、勝手な実数  $h \in \mathbb{R}$  に対して、

$$\frac{f(p_0 + h\mathbf{v}) - f(p_0)}{h} \in T_{f(p_0)}\mathbb{R}^3 \quad (82)$$

であると考えことができます (図 13 を参照)。よって、(82) 式で与えられるベクトルの極限值である  $(df)_{p_0}(\mathbf{v})$  も、

$$(df)_{p_0}(\mathbf{v}) \in T_{f(p_0)}\mathbb{R}^3$$

というように、 $T_{f(p_0)}\mathbb{R}^3$  の元であると考えの方がより自然であることが分かります。

以上から、(80) 式で定義される写像  $(df)_{p_0}$  は、

$$(df)_{p_0}: T_{p_0}\mathbb{R}^2 \rightarrow T_{f(p_0)}\mathbb{R}^3 \quad (83)$$

という写像であると考えの方がより自然であることが分かりました。この (83) 式の (線型) 写像  $(df)_{p_0}$  を点  $p_0$  における写像  $f$  の微分と呼びます。<sup>\*59)</sup> この記号を用いると、写像  $f$  に対して、Taylor 展開の一次式

<sup>\*59)</sup> 前と同様に、線形代数学の言葉使いを用いると、 $T_{p_0}\mathbb{R}^2 = \mathbb{R}^2$ ,  $T_{f(p_0)}\mathbb{R}^3 = \mathbb{R}^3$  という「座標付け」のもとでの「線型写像  $(df)_{p_0}$  の表現行列」が Jacobi 行列  $J_f(p_0)$  であるということになります。

の項までを取って考えた近似式も,

$$f(p_0 + \mathbf{v}) \doteq f(p_0) + (df)_{p_0}(\mathbf{v}) \quad (84)$$

というように簡明な形で表わせることが分かります。この (84) 式は「点  $p_0$  から  $\mathbf{v}$  だけずれた点  $p_0 + \mathbf{v}$  が写像  $f$  によって写される点  $f(p_0 + \mathbf{v})$  は、第一近似で、点  $f(p_0)$  から  $(df)_{p_0}(\mathbf{v})$  だけずれた点になる」ということを表わしています。

#### 14. 合成写像の微分則について

一変数関数の場合に、具体的な関数に対して微分の計算を行なおうと思ったときに、合成関数の微分則がとても役に立つということは、皆さん良くご存じのことではないかと思います。多変数関数や、より一般に、写像の場合にも事情は全く同じで、具体的な微分の計算や理論的な考察を行なうに当たり、「合成関数の微分則」がとても役に立ちます。そこで、ここでは、多変数関数や写像に対する「合成関数の微分則」について少し考えることにします。すなわち、 $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $h: \mathbb{R}^l \rightarrow \mathbb{R}^m$  という二つの滑らかな写像に対して、 $g$  と  $h$  の合成写像を  $f = g \circ h: \mathbb{R}^l \rightarrow \mathbb{R}^n$  とするとき、合成写像  $f = g \circ h$  の微分を写像  $g$  の微分と写像  $h$  の微分を用いて表わすことを考えてみようと思います。<sup>\*60)</sup>

そこで、まず、 $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $h: \mathbb{R} \rightarrow \mathbb{R}^2$  という場合について考えてみることにします。<sup>\*61)</sup> すなわち、

$$f = g \circ h: \mathbb{R} \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$$

という状況について考えてみることにします。このとき、関数  $g$  の変数を  $(x, y) \in \mathbb{R}^2$ 、関数  $h$  の変数を  $t \in \mathbb{R}$  として、 $h(t)$  の成分を、

$$h(t) = (\varphi(t), \psi(t))$$

と表わすことにします。すると、

$$\begin{aligned} f(t) &= g(h(t)) \\ &= g(\varphi(t), \psi(t)) \end{aligned}$$

となりますが、問題は「 $\frac{df}{dt}$  を  $\frac{\partial g}{\partial x}$ ,  $\frac{\partial g}{\partial y}$ ,  $\frac{d\varphi}{dt}$ ,  $\frac{d\psi}{dt}$  を用いて表わせ」ということになります。

そこで、 $\mathbb{R}$  上の点  $t_0 \in \mathbb{R}$  を、勝手にひとつ取ってき

<sup>\*60)</sup> 以下では、例えば、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  というような写像も「ベクトル値の二変数関数」などを見なすことにして、「関数」という言葉と「写像」という言葉を余り区別せずに用いることにします。

<sup>\*61)</sup>  $g: \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $h: \mathbb{R} \rightarrow \mathbb{R}^m$  という場合も、全く同様に議論することができます。

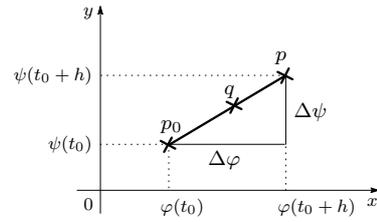


図 14 (88) 式を成り立たせるような点  $q$  が線分  $\overline{p_0 p}$  上に存在する。

て、点  $t_0$  における関数  $f(t)$  の微分係数  $\frac{df}{dt}(t_0)$  がどのように表わされるのかということを考えてみることにします。すなわち、

$$\frac{df}{dt}(t_0) = \lim_{h \rightarrow 0} \frac{f(t_0 + h) - f(t_0)}{h}$$

という極限值がどのように表わされるのかということを考えてみることにします。

いま、 $f(t)$  の定義から、

$$f(t_0 + h) - f(t_0) = g(h(t_0 + h)) - g(h(t_0)) \quad (85)$$

と表わせることに注意して、

$$\begin{cases} p = h(t_0 + h) = (\varphi(t_0 + h), \psi(t_0 + h)) \\ p_0 = h(t_0) = (\varphi(t_0), \psi(t_0)) \end{cases} \quad (86)$$

と表わすことにします。また、点  $p$  と点  $p_0$  の間の座標の差を、それぞれ、

$$\begin{cases} \Delta\varphi = \varphi(t_0 + h) - \varphi(t_0) \\ \Delta\psi = \psi(t_0 + h) - \psi(t_0) \end{cases} \quad (87)$$

と表わすことにします。すると、二変数関数  $g(x, y)$  に対する Taylor の定理から、

$$g(p) = g(p_0) + \frac{\partial g}{\partial x}(q) \Delta\varphi + \frac{\partial g}{\partial y}(q) \Delta\psi \quad (88)$$

となるような点  $q$  が線分  $\overline{p_0 p}$  上に存在することが分かります (図 14 を参照)。よって、(85) 式、(86) 式、(88) 式から、

$$\begin{aligned} \frac{f(t_0 + h) - f(t_0)}{h} &= \frac{g(p) - g(p_0)}{h} \\ &= \frac{\partial g}{\partial x}(q) \cdot \frac{\Delta\varphi}{h} + \frac{\partial g}{\partial y}(q) \cdot \frac{\Delta\psi}{h} \quad (89) \end{aligned}$$

というように書き直せることが分かります。そこで、 $h \rightarrow 0$  なる極限を考えてみると、点  $q$  は線分  $\overline{p_0 p}$  上の点であり、 $h \rightarrow 0$  のとき、 $p \rightarrow p_0 \in \mathbb{R}^2$  となることが分かりますから、

$$q \rightarrow p_0 \in \mathbb{R}^2 \quad (90)$$

となることが分かります。また、(87) 式から、

$$\begin{cases} \frac{\Delta\varphi}{h} = \frac{\varphi(t_0+h) - \varphi(t_0)}{h} \\ \frac{\Delta\psi}{h} = \frac{\psi(t_0+h) - \psi(t_0)}{h} \end{cases}$$

と表わせることが分かりますから、

$$\begin{cases} \lim_{h \rightarrow 0} \frac{\Delta\varphi}{h} = \frac{d\varphi}{dt}(t_0) \\ \lim_{h \rightarrow 0} \frac{\Delta\psi}{h} = \frac{d\psi}{dt}(t_0) \end{cases} \quad (91)$$

となることが分かります。よって、(89) 式、(90) 式、(91) 式から、

$$\begin{aligned} \frac{df}{dt}(t_0) &= \frac{\partial g}{\partial x}(p_0) \cdot \frac{d\varphi}{dt}(t_0) + \frac{\partial g}{\partial y}(p_0) \cdot \frac{d\psi}{dt}(t_0) \\ &= \frac{\partial g}{\partial x}(h(t_0)) \cdot \frac{d\varphi}{dt}(t_0) + \frac{\partial g}{\partial y}(h(t_0)) \cdot \frac{d\psi}{dt}(t_0) \end{aligned} \quad (92)$$

となることが分かります。ここで、 $t_0 \in \mathbb{R}$  は、勝手な点で良かったので、再び、変数らしく、 $t_0 \rightsquigarrow t$  と書き直すことにすると、結局、(92) 式から、

$$\frac{df}{dt}(t) = \frac{\partial g}{\partial x}(h(t)) \cdot \frac{d\varphi}{dt}(t) + \frac{\partial g}{\partial y}(h(t)) \cdot \frac{d\psi}{dt}(t) \quad (93)$$

となることが分かりました。いま、関数  $f(t)$  を合成関数らしく  $g(h(t))$  というように表わし、 $h(t)$  も  $(\varphi(t), \psi(t))$  というように成分を用いて表わすことにすると、(93) 式は、

$$\begin{aligned} \frac{d}{dt}g(\varphi(t), \psi(t)) &= \frac{\partial g}{\partial x}(\varphi(t), \psi(t)) \cdot \frac{d\varphi}{dt}(t) \\ &\quad + \frac{\partial g}{\partial y}(\varphi(t), \psi(t)) \cdot \frac{d\psi}{dt}(t) \end{aligned} \quad (94)$$

と表わすことができます。この (94) 式が、今の場合の合成関数の微分則ということになります。

ここで、さらに、 $h(t)$  の成分も  $(\varphi(t), \psi(t)) \rightsquigarrow (x(t), y(t))$  と書き直すことにすると、(94) 式は、

$$\begin{aligned} \frac{d}{dt}g(x(t), y(t)) &= \frac{\partial g}{\partial x}(x(t), y(t)) \cdot \frac{dx}{dt}(t) \\ &\quad + \frac{\partial g}{\partial y}(x(t), y(t)) \cdot \frac{dy}{dt}(t) \end{aligned} \quad (95)$$

というように表わされることになり、合成関数の微分を計算するやり方が、次のように、よりハッキリするかもしれません。皆さん良くご存じのように、一変数関数の場合には、

$$\frac{d}{dt}g(x(t)) = \frac{dg}{dx}(x(t)) \cdot \frac{dx}{dt}(t) \quad (96)$$

というように合成関数の微分を計算することができます。ここで、(96) 式の右辺は「まず、関数  $g$  を変数  $x$

•  $g$  が一変数関数のとき

$$\begin{aligned} \frac{d}{dt}g(x(t)) &= \frac{dg}{dx} \cdot \frac{dx}{dt} \\ &= \begin{matrix} \textcircled{g} \\ \textcircled{x} \textcircled{t} \end{matrix} \end{aligned}$$

•  $g$  が二変数関数のとき

$$\begin{aligned} \frac{d}{dt}g(x(t), y(t)) &= \frac{\partial g}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial g}{\partial y} \cdot \frac{dy}{dt} \\ &= \begin{matrix} \textcircled{g} \\ \textcircled{x} \textcircled{t} \end{matrix} + \begin{matrix} \textcircled{g} \\ \textcircled{y} \textcircled{t} \end{matrix} \end{aligned}$$

図 15 関数  $g$  が一変数関数の場合の合成関数の微分則をひとつの「鎖」のように考えると、関数  $g$  が多変数関数の場合の微分則はこれらの「鎖」を繋げたようなものに見える。

で微分して、次に、関数  $x$  を変数  $t$  で微分する」ということを意味していますが、これを、図 15 のようにひとつの「鎖」のように考えてみることにします。すると、関数  $g$  が二変数関数の場合には、 $x$  と  $y$  という関数  $g$  の各変数ごとに用意された「鎖」を繋げたようなものとして、(95) 式の右辺が解釈できることが分かります (図 15 を参照)\*62)。その意味で、(95) 式のような合成関数の微分則のことを連鎖律 (chain rule) と呼んだりします。

そこで、次に、 $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $h: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  という場合について考えてみることにします。\*63) すなわち、

$$f = g \circ h: \mathbb{R}^2 \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$$

という状況について考えてみることにします。このとき、前と同様に、関数  $g$  の変数を  $(x, y) \in \mathbb{R}^2$ 、関数  $h$  の変数を  $(s, t) \in \mathbb{R}^2$  として、 $h(s, t)$  の成分を、

$$h(s, t) = (\varphi(s, t), \psi(s, t))$$

と表わすことにします。すると、合成関数  $f(s, t)$  は、

$$\begin{aligned} f(s, t) &= g(h(s, t)) \\ &= g(\varphi(s, t), \psi(s, t)) \end{aligned} \quad (97)$$

と表わせることになります。

そこで、二変数関数  $f(s, t)$  の偏導関数を求めたいわけですが、ここで、偏微分を計算するには「微分する変数以外の変数は定数だと思って、一変数関数として微

\*62) 全く同様に、関数  $g$  が  $n$  変数関数の場合には、各変数ごとに用意された  $n$  個の「鎖」が繋がることになります。

\*63)  $g: \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $h: \mathbb{R}^l \rightarrow \mathbb{R}^m$  という場合も、全く同様に議論することができます。

分すればよい」ということに注意します。すなわち、例えば、 $\frac{\partial f}{\partial s}(s, t)$  という偏導関数を計算するには、 $t$  は定数であると考えて、二変数関数  $f(s, t)$  を、あたかも、

$$f(s) = g(\varphi(s), \psi(s))$$

というような一変数の合成関数と見なして微分を計算すれば良いわけです。上で見たように、このような形の合成関数の微分は、(93) 式、あるいは、(94) 式のように計算できますから、本当は  $t$  も変数であるということに注意して、今の場合に (93) 式を適用してみると、

$$\begin{aligned} \frac{\partial f}{\partial s}(s, t) &= \frac{\partial g}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial s}(s, t) \\ &+ \frac{\partial g}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial s}(s, t) \end{aligned} \quad (98)$$

となることが分かります。全く同様に考えると、 $t$  という変数に関する偏導関数も、

$$\begin{aligned} \frac{\partial f}{\partial t}(s, t) &= \frac{\partial g}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial t}(s, t) \\ &+ \frac{\partial g}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial t}(s, t) \end{aligned} \quad (99)$$

というように計算できることが分かります。したがって、(98) 式、(99) 式から、合成関数  $f(s, t)$  の全微分が求まることになります。

そこで、最後に、 $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $h: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  という場合について考えてみることにします。<sup>\*64)</sup> すなわち、

$$f = g \circ h: \mathbb{R}^2 \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

という状況について考えてみることにします。このとき、前と同様に、関数  $g$  の変数を  $(x, y) \in \mathbb{R}^2$  として、関数  $g(x, y)$  の成分を、

$$g(x, y) = (X(x, y), Y(x, y))$$

と表わし、関数  $h$  の変数を  $(s, t) \in \mathbb{R}^2$  として、 $h(s, t)$  の成分を、

$$h(s, t) = (\varphi(s, t), \psi(s, t))$$

と表わすことにします。すると、

$$\begin{aligned} f(s, t) &= g(h(s, t)) \\ &= (X(\varphi(s, t), \psi(s, t)), Y(\varphi(s, t), \psi(s, t))) \end{aligned}$$

となることが分かりますから、合成関数  $f(s, t)$  の成分を、

$$f(s, t) = (f_1(s, t), f_2(s, t))$$

<sup>\*64)</sup>  $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $h: \mathbb{R}^l \rightarrow \mathbb{R}^m$  という場合も、全く同様に議論することができます。

と表わすことにすると、

$$\begin{cases} f_1(s, t) = X(\varphi(s, t), \psi(s, t)) \\ f_2(s, t) = Y(\varphi(s, t), \psi(s, t)) \end{cases} \quad (100)$$

となることが分かります。

そこで、合成関数  $f(s, t)$  の微分を求めるためには、 $f(s, t)$  の成分である  $f_1(s, t)$  や  $f_2(s, t)$  という関数の偏導関数を計算する必要がありますが、(100) 式から、これらの関数は、(97) 式のような形の合成関数であることに注意します。よって、これらの関数に対して、(98) 式、(99) 式を適用することで、

$$\begin{aligned} \frac{\partial f_1}{\partial s}(s, t) &= \frac{\partial X}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial s}(s, t) \\ &+ \frac{\partial X}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial s}(s, t) \end{aligned} \quad (101)$$

$$\begin{aligned} \frac{\partial f_1}{\partial t}(s, t) &= \frac{\partial X}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial t}(s, t) \\ &+ \frac{\partial X}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial t}(s, t) \end{aligned} \quad (102)$$

$$\begin{aligned} \frac{\partial f_2}{\partial s}(s, t) &= \frac{\partial Y}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial s}(s, t) \\ &+ \frac{\partial Y}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial s}(s, t) \end{aligned} \quad (103)$$

$$\begin{aligned} \frac{\partial f_2}{\partial t}(s, t) &= \frac{\partial Y}{\partial x}(h(s, t)) \cdot \frac{\partial \varphi}{\partial t}(s, t) \\ &+ \frac{\partial Y}{\partial y}(h(s, t)) \cdot \frac{\partial \psi}{\partial t}(s, t) \end{aligned} \quad (104)$$

となることが分かります。したがって、(101) 式、(102) 式、(103) 式、(104) 式から、合成関数  $f(s, t)$  の Jacobi 行列が求まることになります。

以上から、連鎖律を用いて、合成写像の微分が計算できることが分かりましたが、こうして得られた合成写像の微分則の意味をより良く理解するために、これらの結果を 13 節で見た「写像の微分」を用いて表わしてみることにします。

すると、

$$f = g \circ h: \mathbb{R} \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

という場合には、

$$\begin{aligned} (df)_t &= \left( \frac{df}{dt}(t) \right) \\ (dg)_p &= \left( \frac{\partial g}{\partial x}(p) \quad \frac{\partial g}{\partial y}(p) \right) \\ (dh)_t &= \left( \frac{d\varphi}{dt}(t) \right) \\ &\quad \left( \frac{d\psi}{dt}(t) \right) \end{aligned}$$

として、(93) 式は、

$$(df)_t = (dg)_{h(t)} \cdot (dh)_t$$

というように表わせることが分かります。<sup>\*65)</sup>

<sup>\*65)</sup> 皆さん、確かめてみて下さい。

また,

$$f = g \circ h : \mathbb{R}^2 \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$$

という場合には,

$$\begin{aligned} (df)_{(s,t)} &= \begin{pmatrix} \frac{\partial f}{\partial s}(s,t) & \frac{\partial f}{\partial t}(s,t) \end{pmatrix} \\ (dg)_p &= \begin{pmatrix} \frac{\partial g}{\partial x}(p) & \frac{\partial g}{\partial y}(p) \end{pmatrix} \\ (dh)_{(s,t)} &= \begin{pmatrix} \frac{\partial \varphi}{\partial s}(s,t) & \frac{\partial \varphi}{\partial t}(s,t) \\ \frac{\partial \psi}{\partial s}(s,t) & \frac{\partial \psi}{\partial t}(s,t) \end{pmatrix} \end{aligned}$$

として, (98) 式, (99) 式は,

$$(df)_{(s,t)} = (dg)_{h(s,t)} \cdot (dh)_{(s,t)}$$

というように表わせることが分かります.\*66)

さらに,

$$f = g \circ h : \mathbb{R}^2 \xrightarrow{h} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

という場合には,

$$\begin{aligned} (df)_{(s,t)} &= \begin{pmatrix} \frac{\partial f_1}{\partial s}(s,t) & \frac{\partial f_1}{\partial t}(s,t) \\ \frac{\partial f_2}{\partial s}(s,t) & \frac{\partial f_2}{\partial t}(s,t) \end{pmatrix} \\ (dg)_p &= \begin{pmatrix} \frac{\partial X}{\partial x}(p) & \frac{\partial X}{\partial y}(p) \\ \frac{\partial Y}{\partial x}(p) & \frac{\partial Y}{\partial y}(p) \end{pmatrix} \\ (dh)_{(s,t)} &= \begin{pmatrix} \frac{\partial \varphi}{\partial s}(s,t) & \frac{\partial \varphi}{\partial t}(s,t) \\ \frac{\partial \psi}{\partial s}(s,t) & \frac{\partial \psi}{\partial t}(s,t) \end{pmatrix} \end{aligned}$$

として, (101) 式, (102) 式, (103) 式, (104) 式は,

$$(df)_{(s,t)} = (dg)_{h(s,t)} \cdot (dh)_{(s,t)}$$

というように表わせることが分かります.\*67)

そこで, 一般に,

$$f = g \circ h : \mathbb{R}^l \xrightarrow{h} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^n \quad (105)$$

という場合に, 全く同様の考察を行なってみると,  $p \in \mathbb{R}^l$

として,

$$(df)_p = (dg)_{h(p)} \cdot (dh)_p \quad (106)$$

となることが分かります. すなわち, 「合成写像の微分則」とは「(105) 式のような合成写像の点  $p \in \mathbb{R}^l$  における微分写像

$$(df)_p : T_p \mathbb{R}^l \rightarrow T_{f(p)} \mathbb{R}^n = T_{g(h(p))} \mathbb{R}^n$$

が,

\*66) 皆さん, 確かめてみて下さい.

\*67) 皆さん, 確かめてみて下さい.

$$(df)_p : T_p \mathbb{R}^l \xrightarrow{(dh)_p} T_{h(p)} \mathbb{R}^m \xrightarrow{(dg)_{h(p)}} T_{g(h(p))} \mathbb{R}^n$$

というように, それぞれの関数  $g, h$  の微分写像の合成写像になっている」という事実を表わしていることが分かります.

さて, 13 節の最後では, 写像の微分の記号を用いると,

$$f : \mathbb{R}^l \rightarrow \mathbb{R}^n$$

という写像に対して, 点  $p \in \mathbb{R}^l$  における写像  $f$  の一次の Taylor 多項式による近似が,  $\mathbf{v} \in T_p \mathbb{R}^l$  として,

$$f(p + \mathbf{v}) \doteq f(p) + (df)_p(\mathbf{v}) \quad (107)$$

というように簡明な形で表わせることを注意しました. 興味のある方は, (107) 式と同様にして, 点  $h(p) \in \mathbb{R}^m$  における写像  $g$  の一次の Taylor 多項式による近似の式と, 点  $p \in \mathbb{R}^l$  における写像  $h$  の一次の Taylor 多項式による近似の式を書き下してみして下さい. そして, これらの式を用いて, 直接, (106) 式を確かめることができるかどうかということを考えてみて下さい.