2008 年度全学自由研究ゼミナール

多変数関数の微分：第 11 回

7 月 1 日 清野和彦

7.5 合成関数の微分公式の証明

偏微分は 1 変数関数の微分の定義から全く安直に定義されるものであるにも拘わらず、多変数関数の微分として十分な情報を持っていること、正確には、問題にしている多変数関数が全微分可能、つまり 1 次近似を持つならば、その 1 次近似式は二つの偏微分の値によって完全に決定されることを第 5 節で証明しました。ということは、合成関数の微分公式も、高校のときに学んだ 1 変数関数の合成関数の微分公式の証明と同じように証明することも可能なはずだし、また、1 次近似の視点から証明することも可能なはずです。ここでは、後者の方法、つまり、関数が全微分可能な場合、偏微分の値は「1 次近似（すなわち接平面の式）の係数である」という視点から証明することにします。

証明はたくさんの文字を含む複雑な式の極限値を計算する、という外見上は面倒で退屈なものです。式変形に翻弄されるうちに何がなんだか分からなくなってしまうかも知れません。しかし、やろうとしていることは

合成関数の 1 次近似は、それぞれの関数の 1 次近似の合成だ

ということを証明することに尽きます。これが分かれば、合成関数の 1 次近似を定義に従って直接求めなくても、それぞれの関数の 1 次近似を合成してやるという**簡単なこと**（だって、1 次式に 1 次式を入れるだけですから!）で求められてしまうことになり、合成関数の微分公式は、この「1 次式に 1 次式を入れる」という手続きを係数だけ取り出して書き下したものになっているということが結論されるからです。だから、一つ一つの式変形にあまりこだわらずに、全体として「1 次近似の合成が合成関数の 1 次近似になっていることの証明」なのだということをいつも忘れないようにすることが大切です。

7.5.1 1 変数関数に 1 変数関数を入れた場合

すぐに多変数関数の場合を扱ってもよいのですが、はじめての考え方ですので、焦らずに 1 変数関数の場合から順番に示して行きましょう。

1 変数関数の場合、第 5.1 節（第 5 回 4 ページの式 (4) のあたり[1]）で説明したよ

[1]ただし、記号の使い方がちょっと違ってしまっています。第 5 節と同じ記号の使い方にしようとすると、合成する関数（中に入れる方の関数）についての記号がどうしても足りなくなってしまうのです。申し訳ありませんが、第 5 節を参照する場合には気を付けて読んで下さい。

うに、a という数が $f'(x_0)$ と一致することは

$$p(x) = f(x_0) + a(x - x_0)$$

という 1 次式が

$$\lim_{x \to x_0} \frac{f(x) - p(x)}{x - x_0} = 0 \tag{1}$$

を満たすことと**同値**でした。つまり、

> 関数 $f(x)$ と $x = x_0$ に対し、条件(1) を満たす 1 次近似は存在したとしても**一つだけ**であり、そのような $p(x)$ が存在することは f が x_0 で微分可能なことと同値で、そのとき
>
> $$p(x) = f(x_0) + f'(x_0)(x - x_0)$$
>
> である。

ということが成り立ちます。そして、この $p(x)$ のことを $f(x)$ の $x = x_0$ における **1 次近似**と呼びました。ここでやりたいことは、上で説明したことのうち、特に

> 微分可能な関数と微分したい点を決めると 1 次近似は**ただ一つに定まる**

ということに注目して合成関数の微分公式を証明することです。

さて、$f(x)$ に $x = \xi(s)$ を入れた関数 $f \circ \xi(s)$ の $s = s_0$ における 1 次近似を考えたいわけですが、もちろん、今 $f(x)$ は $x_0 = \xi(s_0)$ で、$\xi(s)$ は $s = s_0$ で微分可能と仮定するので、$f(x)$ の x_0 における 1 次近似

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

と $\xi(s)$ の s_0 における 1 次近似

$$\varpi(s) = \xi'(s_0)(s - s_0) + x_0$$

がただ一つだけ存在しています[2]。一方、われわれが欲しいのは、$f \circ \xi$ の s_0 における 1 次近似です。もちろん、f が $x_0 = \xi(s_0)$ で、ξ が s_0 で微分可能であれば $f \circ \xi$ も s_0 で微分可能であることも証明しなければならないことであり、もちろんあとで証明しますが、そのことは認めることにすると、$f \circ \xi$ の s_0 における 1 次近似 $P(s)$ は

$$P(s) = (f \circ \xi)(s_0) + (f \circ \xi)'(s_0)(s - s_0)$$

[2] $p(x)$ の p は多項式を意味する ‘polynomial’ の頭文字を取りました。また ϖ は p に当たるギリシャ文字 π の異体字です。(だから ϖ もパイと読みます。) π だと円周率みたいになってしまうので異体字にしました。見たこともなくてイヤだと思いますが、結構慣れるものですのでちょっと我慢してみて下さい。

となります。

ここで、$f\circ\xi(s)=f(\xi(s))$ という合成の定義と、1 次式に 1 次式を代入しても 1 次式のままであることに目を付けて、$p(x)$ に $x=\varpi(s)$ を入れてみることを試みてみましょう。つまり、

$p\circ\varpi(s)$ が $f\circ\xi(s)$ の s_0 における 1 次近似なのではないか

と考えてみようというわけです。実際に $p\circ\varpi(s)$ を具体的に計算してみると、

$$\begin{aligned}p\circ\varpi(s)&=f'(x_0)(\varpi(s)-x_0)+f(x_0)\\&=f(x_0)+f'(x_0)(x_0+\xi'(s_0)(s-s_0)-x_0)\\&=f(x_0)+f'(x_0)\xi'(s_0)(s-s_0)\end{aligned}$$

となりますので、もしも予想通り $f\circ\xi$ の 1 次近似 $P(s)$ が f の 1 次近似と ξ の 1 次近似の合成 $p\circ\xi(s)$ と一致するなら、

$$(f\circ\xi)'(s_0)=f'(x_0)\xi'(s_0)$$

である、という欲しかった合成関数の微分公式が得られるわけです。

このことだけで、もはや「合成の 1 次近似はそれぞれの 1 次近似の合成」ということを証明なしで認めてもよいような気分になってしまいそうですが、数学ですから、やはり面倒でも証明しないわけにはいきません。

何をすれば証明したことになるのかというと、

$$\lim_{s\to s_0}\frac{f\circ\xi(s)-p\circ\varpi(s)}{s-s_0}=0 \tag{2}$$

を示すことです。なぜなら、これが成り立つことは $p\circ\varpi(s)$ が $f\circ\xi(s)$ の $s=s_0$ における 1 次近似であることを意味し、1 次近似がただ一つしか存在し得ないことから $P(s)=p\circ\varpi(s)$ が結論されるからです。

それでは、式(2) を示しましょう。といっても、式(2) をじっと見ているだけではどうやって証明すればよいかなかなか思いつけませんから、グラフを通じて図形的に考えてみましょう。まず、中に入る方の関数 $x=\xi(s)$ のグラフと $s=s_0$ における接線である $x=\varpi(s)$ のグラフを思い浮かべます。示したいことは、$x=\xi(s)$ のグラフを $y=f(x)$ で動かしてできる曲線である $y=f(\xi(s))=f\circ\xi(s)$ のグラフの $s=s_0$ における接線が、$x=\varpi(s)$ という直線を $y=p(x)$ という 1 次式で動かしてできる直線である $y=p(\varpi(s))=p\circ\varpi(s)$ のグラフと一致することです。そこで、この二つのグラフの「仲介者」として

$x=\xi(s)$ のグラフを 1 次式 $y=p(x)$ で動かしてできる曲線である $y=p(\xi(s))=p\circ\xi(s)$ のグラフ

を考えてみることにしましょう[3]。この「仲介者」を使って、

> $y = f \circ \xi(s)$ の $s = s_0$ における接線は $y = p \circ \varpi(s)$ である

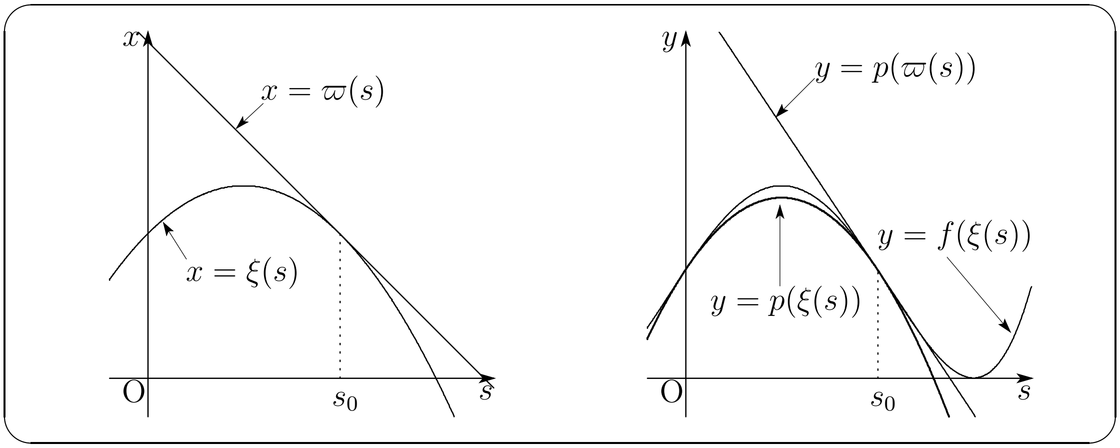
ということを、

> $y = f \circ \xi(s)$ の $s = s_0$ における接線と $y = p \circ \xi(s)$ の $s = s_0$ における接線は一致する

ということと、

> $y = p \circ \xi(s)$ の $s = s_0$ における接線は $y = p \circ \varpi(s)$ である

ということの二つに分けて証明しようという方針です。



それでは、証明しましょう。まず、

$$\begin{aligned}\lim_{s \to s_0} \frac{f \circ \xi(s) - p \circ \varpi(s)}{s - s_0} &= \lim_{s \to s_0} \frac{f \circ \xi(s) - p \circ \xi(s) + p \circ \xi(s) - p \circ \varpi(s)}{s - s_0} \\ &= \lim_{s \to s_0} \left(\frac{f(\xi(s)) - p(\xi(s))}{s - s_0} + \frac{p(\xi(s)) - p(\varpi(s))}{s - s_0} \right)\end{aligned}$$

というように「仲介者」を使って「分子の水増し」をします。ここで、最後の極限を第 1 項と第 2 項に分けて考えてみましょう。

まず、第 1 項について。極限を取る前の式を

$$\frac{f(\xi(s)) - p(\xi(s))}{s - s_0} = \frac{f(\xi(s)) - p(\xi(s))}{\xi(s) - \xi(s_0)} \frac{\xi(s) - \xi(s_0)}{s - s_0}$$

---

[3] $x = \xi(s)$ のグラフを $y = p(x)$ で動かす代わりに、$x = \varpi(s)$ のグラフを $y = f(x)$ で動かしたものを「仲介者」とすることも可能です。ただし、そうすると ϖ が ξ の 1 次近似であることを直接利用することができなくなってしまうという難点があり、証明は難しくなってしまいます。興味のある方は考えてみて下さい。

と変形しましょう。ξ は s_0 で微分可能と仮定しているので、

$$\lim_{s \to s_0} \frac{\xi(s) - \xi(s_0)}{s - s_0} = \frac{d\xi}{ds}(s_0)$$

です。また、微分可能とうことは連続、すなわち $\lim_{s \to s_0} \xi(s) = \xi(s_0) = x_0$ が成り立っているので、$x = \xi(s)$ と置き換えることにより、

$$\lim_{s \to s_0} \frac{f(\xi(s)) - p(\xi(s))}{\xi(s) - \xi(s_0)} = \lim_{x \to x_0} \frac{f(x) - p(x)}{x - x_0} = 0$$

となります[4]。なぜ 0 になるかというと、$p(x)$ は $f(x)$ の x_0 における 1 次近似だからです。以上より、第 1 項については

$$\begin{aligned}
\lim_{s \to s_0} \frac{f(\xi(s)) - p(\xi(s))}{s - s_0} &= \lim_{s \to s_0} \frac{f(\xi(s)) - p(\xi(s))}{\xi(s) - \xi(s_0)} \frac{\xi(s) - \xi(s_0)}{s - s_0} \\
&= \left(\lim_{s \to s_0} \frac{f(\xi(s)) - p(\xi(s))}{\xi(s) - \xi(s_0)} \right) \left(\lim_{s \to s_0} \frac{\xi(s) - \xi(s_0)}{s - s_0} \right) \\
&= 0 \cdot \frac{d\xi}{ds}(s_0) = 0
\end{aligned}$$

となることがわかりました。(これは「$y = f \circ \xi(s)$ の $s = s_0$ における接線と $y = p \circ \xi(s)$ の $s = s_0$ における接線は一致する」ということを証明したことに当たります。)

次に、第 2 項について。$p(x)$ は 1 次式なのですから、具体的に

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

と書いて計算してみましょう。すると、

$$\begin{aligned}
\frac{p(\xi(s)) - p(\varpi(s))}{s - s_0} &= \frac{f(x_0) + f'(x_0)(\xi(s) - x_0) - f(x_0) - f'(x_0)(\varpi(s) - x_0)}{s - s_0} \\
&= f'(x_0) \frac{\xi(s) - \varpi(s)}{s - s_0}
\end{aligned}$$

となります。そして、$\varpi(s)$ が $\xi(s)$ の s_0 における 1 次近似であることから、

$$\lim_{s \to s_0} \frac{p(\xi(s)) - p(\varpi(s))}{s - s_0} = f'(x_0) \lim_{s \to s_0} \frac{\xi(s) - \varpi(s)}{s - s_0} = f'(x_0) \cdot 0 = 0$$

---

[4] これでほとんど良いのですが、一つだけ問題があります。それは s_0 のどんなに近くにも $\xi(s) = \xi(s_0)$ となる s が存在する場合、$\xi(s) - \xi(s_0)$ で割ることができないという問題です。$\xi(s)$ が s_0 の近くで定数関数であるか、無限に振動して $\xi(s_0)$ と同じ値を無限に取ってしまう場合です。この場合には、分子の $f(\xi(s)) - p(\xi(s))$ がほとんど 0 なので、$\xi(x) - \xi(s_0)$ を分子分母に掛けるという工夫をしなくても元の極限値が 0 であることが結論できます。このようなことをいちいち気にしていると何をやっているのか大筋が掴みにくくなってしまうので、今後はこのような特別な場合は気にせずに進むことにします。

となります。（これは「$y = p \circ \xi(s)$ の $s = s_0$ における接線は $y = p \circ \varpi(s)$ である」ということを証明したことに当たります。）

以上より、第 1 項も第 2 項も 0 に収束したので全体として 0 に収束します。これで式(2) が成り立つことが分かり、1 変数関数の合成関数の微分公式 $(f \circ \xi)'(s_0) = f'(x_0) \cdot \xi'(s_0)$ が証明できました。

注意. 1 次近似という考え方を使わずに、微分係数の定義から直接証明しようとすると次のようになります。

$$\begin{aligned}\frac{df \circ \xi}{ds}(s_0) &= \lim_{s \to s_0} \frac{f \circ \xi(s) - f \circ \xi(x_0)}{s - s_0} = \lim_{s \to s_0} \frac{f(\xi(s)) - f(\xi(s_0))}{\xi(s) - \xi(s_0)} \frac{\xi(x) - \xi(s_0)}{s - s_0} \\ &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \lim_{s \to s_0} \frac{\xi(s) - \xi(s_0)}{s - s_0} = \frac{df}{dx}(x_0) \frac{d\xi}{ds}(s_0)\end{aligned}$$

はっきり言って、1 次近似を使った証明より簡単ですね。どこが違うかというと、微分係数の定義式を使ったこの計算は

$$\lim_{s \to s_0} \frac{f(\xi(s)) - f(x_0))}{s - s_0} = f'(x_0)\xi'(s_0) \tag{3}$$

であることを証明したのに対し、1 次近似を使った照明では右辺を左辺に移項して分子に載せた

$$\lim_{s \to s_0} \frac{f(\xi(s)) - f(x_0) - f'(x_0)\xi'(s_0)(s - s_0)}{s - s_0} = 0 \tag{4}$$

を証明したのです。このように、1 変数の範囲では 1 次近似の視点から考えるメリットは（証明の手間という点では）ありません。しかし、多変数になると 1 次近似の方法が途端に威力を発揮するのです。なぜなら、上の微分係数の定義を使った計算は多変数関数にそのまま拡張することができないので一つ一つの偏微分を 1 変数関数の微分と見て適用するしかないのですが、1 次近似の視点はそのまま多変数関数の場合に適用できて、多変数関数の微分にもっともふさわしい全微分可能性という概念になっているからです。★

7.5.2　1 変数関数に 2 変数関数を入れた場合

この場合が 1 変数関数に 1 変数関数を入れる場合と本質的に違わないことは公式を紹介したときに強調したところですが、

- 2 変数関数でも 1 次近似は一つしかないことを確認する。
- 二つの偏微分が同時に求まってしまう場面を見る。

ということを簡単な場合で済ませておくのがよいと思うので、一足飛びに進むのはやめて、ゆっくりと証明して行くことにしましょう。

まず、2 変数関数における「**1 次近似の一意性**」について確認します。それは第 5.2.1 節の定理 1（第 5 回プリント 6 ページ）

> $f(x,y)$ が (x_0,y_0) で全微分可能なとき、$f(x,y)$ は (x_0,y_0) で x でも y でも偏微分可能であり、1 次近似 $p(x,y)$ は
>
> $$p(x,y)=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$
>
> である。

というものです。（ここでの記号に合わせて (a,b) を (x_0,y_0) に変えました。また、元の定理 1 には方向微分に関する部分もあるのですが、今は必要ないので省きました。）つまり、$f(x,y)$ が (x_0,y_0) で全微分可能なときは、2 変数の 1 次式 $p(x,y)$ が

$$\lim_{(x,y)\to(x_0,y_0)}\frac{f(x,y)-p(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}}=0 \tag{5}$$

を満たすことと

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

であることは**同値である**というわけで、特に、式(5) を満たす 1 次式 $p(x,y)$ は**ただ一つしか存在しない**ということになります。このことを使って「1 変数関数に 2 変数関数を入れた場合」の合成関数の微分公式を証明しようというわけです。

$f(x)$ に $x=\xi(s,t)$ を合成した関数 $f\circ\xi(s,t)$ の (s_0,t_0) における偏微分の値を**両方同時に**計算します。$f(x)$ は $x_0=\xi(s_0,t_0)$ で微分可能、$\xi(s,t)$ は (s_0,t_0) で**全微分可能**と仮定します。そして、$f(x)$ の x_0 における 1 次近似を $p(x)$、$\xi(s,t)$ の (s_0,t_0) における 1 次近似を $\varpi(s,t)$ としましょう。具体的には、

$$p(x)=f(x_0)+\frac{df}{dx}(x_0)(x-x_0) \tag{6}$$

および

$$\varpi(s,t)=x_0+\frac{\partial\xi}{\partial s}(s_0,t_0)(s-s_0)+\frac{\partial\xi}{\partial t}(s_0,t_0)(t-t_0) \tag{7}$$

です。（$\xi(s_0,t_0)=x_0$ であることに注意してください。）証明したいことは、

> $p(x)$ に $x=\varpi(s,t)$ を入れて得られる 1 次式 $p\circ\varpi(s,t)$ が $f\circ\xi(s,t)$ の (s_0,t_0) における 1 次近似になっている

ということ、つまり、

$$\lim_{(s,t)\to(s_0,t_0)}\frac{f\circ\xi(s,t)-p\circ\varpi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}=0 \tag{8}$$

が成り立つことです。

これが証明できたとします。すると、

$$p \circ \varpi(s,t) = f(x_0) + \frac{\partial(f \circ \xi)}{\partial s}(s_0,t_0)(s-s_0) + \frac{\partial(f \circ \xi)}{\partial t}(s_0,t_0)(t-t_0)$$

が成り立つことになります。なぜなら、1 次近似はこの式の右辺の関数**しかない**からです。一方、$p \circ \varpi(s,t)$ を式(6) に式(7) を代入することで具体的に計算してみると、

$$\begin{aligned}
p \circ \varpi(s,t) &= f(x_0) + \frac{df}{dx}(x_0)(\varpi(s,t) - x_0) \\
&= f(x_0) + \frac{df}{dx}(x_0)\left(x_0 + \frac{\partial\xi}{\partial s}(s_0,t_0)(s-s_0) + \frac{\partial\xi}{\partial t}(s_0,t_0)(t-t_0) - x_0\right) \\
&= f(x_0) + \frac{df}{dx}(x_0)\frac{\partial\xi}{\partial s}(s_0,t_0)(s-s_0) + \frac{df}{dx}(x_0)\frac{\partial\xi}{\partial t}(s_0,t_0)(t-t_0)
\end{aligned}$$

となります。この二つを比較して、

$$\begin{aligned}
\frac{\partial(f \circ \xi)}{\partial s}(s_0,t_0) &= \frac{df}{dx}(x_0)\frac{\partial\xi}{\partial s}(s_0,t_0) \\
\frac{\partial(f \circ \xi)}{\partial t}(s_0,t_0) &= \frac{df}{dx}(x_0)\frac{\partial\xi}{\partial t}(s_0,t_0)
\end{aligned}$$

という欲しかった合成関数の微分公式が二つ一遍に得られる、というわけです。

それでは式(8) を証明しましょう。

まず、この式の左辺の分子に前節と同様の「水増し」をすると、

$$\begin{aligned}
&\lim_{(s,t)\to(s_0,t_0)} \frac{f\circ\xi(s,t) - p\circ\varpi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} \\
&= \lim_{(s,t)\to(s_0,t_0)} \frac{f(\xi(s,t)) - p(\xi(s,t)) + p(\xi(s,t)) - p(\varpi(s,t))}{\sqrt{(s-s_0)^2+(t-t_0)^2}} \\
&= \lim_{(s,t)\to(s_0,t_0)} \left(\frac{f(\xi(s,t)) - p(\xi(s,t))}{\xi(s,t)-\xi(s_0,t_0)} \frac{\xi(s,t)-\xi(s_0,t_0)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} \right. \\
&\qquad \left. + \frac{p(\xi(s,t)) - p(\varpi(s,t))}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\right)
\end{aligned}$$

となります。

最後の式の第 1 項の第 1 因子[5]は、$\xi(s,t)$ が (s_0,t_0) で**全微分可能だから連続**でもあることと、$p(x)$ が $f(x)$ の x_0 における 1 次近似であることから、

$$\lim_{(s,t)\to(s_0,t_0)} \frac{f(\xi(s,t)) - p(\xi(s,t))}{\xi(s,t)-\xi(s_0,t_0)} = \lim_{x\to x_0}\frac{f(x)-p(x)}{x-x_0} = 0$$

[5]因子 (factor) という言葉は、例えば $(x+1)(x-1)(x+2)$ というふうにいくつかの式が掛けられているときの $x+1$、$x-1$、$x+2$ 一つ一つのことを言います。

となります。一方、第 2 因子は 1 変数関数同士の場合と違って極限は存在するとは限りません。どのようになっているか詳しく見てみましょう。分子を

$$\xi(s,t)-\xi(s_0,t_0)=\xi(s,t)-\varpi(s,t)+\frac{\partial\xi}{\partial s}(s_0,t_0)(s-s_0)+\frac{\partial\xi}{\partial t}(s_0,t_0)(t-t_0)$$

と変形します。$\varpi(s,t)$ が $\xi(s,t)$ の 1 次近似であることから

$$\lim_{(s,t)\to(s_0,t_0)}\frac{\xi(s,t)-\varpi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}=0$$

となることはわかっているので、問題は残りの部分の極限です。記号が長くてわかりにくいので、

$$\frac{\partial\xi}{\partial s}(s_0,t_0)=a,\quad \frac{\partial\xi}{\partial t}(s_0,t_0)=b,\quad s-s_0=h,\quad t-t_0=k$$

とおくことにしましょう。調べたいのは、$(h,k)\to(0,0)$ のときの

$$\frac{ah+bk}{\sqrt{h^2+k^2}}$$

の振る舞いです。極限が存在しないことは、例えば $k=0$ として $h\to+0$ とすると a になるのに、$h\to-0$ とすると $-a$ になり、また $h=0$ として $k\to+0$ とすると b に、$k\to-0$ とすると $-b$ になる、つまり (h,k) の $(0,0)$ への値の近づけ方で極限値が変わってしまうことからわかります。($a=b=0$ のときだけ 0 に収束しますが。)

しかし、$|h|\le\sqrt{h^2+k^2}$, $|k|\le\sqrt{h^2+k^2}$ であることから、$(0,0)$ でない任意の (h,k) に対して、

$$\left|\frac{ah+bk}{\sqrt{h^2+k^2}}\right|\le|a|\frac{|h|}{\sqrt{h^2+k^2}}+|b|\frac{|k|}{\sqrt{h^2+k^2}}\le|a|+|b|$$

が成り立っています。よって、問題の極限の第 1 項は、全体としては

$$\begin{aligned}&\lim_{(s,t)\to(s_0,t_0)}\left|\frac{f(\xi(s,t))-p(\xi(s,t))}{\xi(s,t)-\xi(s_0,t_0)}\frac{\xi(s,t)-\xi(s_0,t_0)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\right|\\&\le\lim_{(s,t)\to(s_0,t_0)}\left|\frac{f(\xi(s,t))-p(\xi(s,t))}{\xi(s,t)-\xi(s_0,t_0)}(|a|+|b|)\right|=0\cdot(|a|+|b|)=0\end{aligned}$$

となって、はさみうちの原理により 0 に収束します。

一方、第 2 項の方は、分子が

$$\begin{aligned}&p(\xi(s,t))-p(\varpi(s,t))\\&=f(x_0)+f'(x_0)(\xi(s,t)-x_0)-f(x_0)-f'(x_0)(\varpi(s,t)-x_0)\\&=f'(x_0)(\xi(s,t)-\varpi(s,t))\end{aligned}$$

と具体的に計算でき、$\varpi(s,t)$ が $\xi(s,t)$ の (s_0,t_0) における 1 次近似であることから、

$$\begin{aligned}\lim_{(s,t)\to(s_0,t_0)} \frac{p(\xi(s,t)) - p(\varpi(s,t))}{\sqrt{(s-s_0)^2+(t-t_0)^2}} &= f'(x_0) \lim_{(s,t)\to(s_0,t_0)} \frac{\xi(s,t)-\varpi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} \\ &= f'(x_0)\cdot 0 = 0\end{aligned}$$

となります。

以上により、式(8) の成り立つこと、つまり、$p\circ\varpi(s,t)$ が $f\circ\xi(s,t)$ の (s_0,t_0) における 1 次近似であることが証明できました。

7.5.3 2 変数関数に 1 変数関数を 2 つ入れた場合

次に、2 変数関数に 1 変数関数を 2 つ入れる場合を 1 次近似の方法で考えてみましょう。公式を紹介したときに述べたように、この場合が多変数関数になったことによって合成関数が難しくなる本質的な場合です。（合成関数の中にコピー機が入ってしまうからでした。）だから、この場合の証明をしっかり理解することが重要です。

なお、この場合は合成関数が 1 変数なので、「2 つの公式が一遍に得られる」という恩恵はありません。だから、偏微分の定義に従って計算するより証明が見通しよく簡単になるかどうかがポイントです。そこで、この場合の偏微分の定義に従った計算による証明を付録として最後に付けておきますので、興味のある人は目を通してみてください。

$f(x,y)$ に $x=\xi(s)$ と $y=\eta(s)$ を入れた合成関数を、$\circ$ を使った記法では書けないので $\varphi(s)$ と書くことにしました。

$$\varphi(s) = f(\xi(s),\eta(s))$$

です。f は全微分可能、ξ と η も微分可能、$x_0=\xi(s_0), y_0=\eta(s_0)$ として、$f(x,y)$ の (x_0,y_0) での 1 次近似を $p(x,y)$、$\xi(s)$ の s_0 での 1 次近似を $\varpi(s)$、$\eta(s)$ の s_0 での 1 次近似を $\rho(s)$ とします[6]。具体的には、

$$p(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \tag{9}$$

$$\varpi(s) = x_0 + \frac{d\xi}{ds}(s_0)(s-s_0) \tag{10}$$

$$\rho(s) = y_0 + \frac{d\eta}{ds}(s_0)(s-s_0) \tag{11}$$

[6] ρ はギリシャ語のアルファベットで ϖ の次の文字で、「ロー」と読みます。ローマ字の ‘r’ に当たる文字です。ギリシャ文字には q に当たる文字がないので、p の次の次の r に当たる文字を使うことにしました。

です。$p(x,y)$ に $x=\varpi(s), y=\rho(s)$ を入れた合成関数（s の 1 次式です）も $\circ$ を使った書き方では書けないので、

$$\psi(s)=p(\varpi(s),\rho(s))$$

とおくことにします。示したいことは、例によって $\psi(s)$ が $\varphi(s)$ の s_0 における 1 次近似であること、つまり、

$$\lim_{s\to s_0}\frac{\varphi(s)-\psi(s)}{|s-s_0|}=0 \tag{12}$$

が成り立つことです。（あとの式変形を楽にするために分母に絶対値を付けておきました。極限値が 0 の場合には絶対値が付いていてもいなくても同じなので付けても大丈夫です。）

これが証明できたとします。すると、

$$\psi(s)=f(x_0,y_0)+\frac{d\varphi}{ds}(s_0)(s-s_0)$$

が成り立つことになります。なぜなら $\varphi(s)$ の 1 次近似はこの式の右辺の関数**しかない**からです。一方、式(9) に式(10) と(11) を入れて $\psi(s)$ を具体的に計算してみると、

$$\begin{aligned}\psi(s)&=p(\varpi(s),\rho(s))\\&=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)\left(x_0+\frac{d\xi}{ds}(s_0)(s-s_0)-x_0\right)\\&\qquad+\frac{\partial f}{\partial y}(x_0,y_0)\left(y_0+\frac{d\eta}{ds}(s_0)(s-s_0)-y_0\right)\\&=f(x_0,y_0)+\left(\frac{\partial f}{\partial x}(x_0,y_0)\frac{d\xi}{ds}(s_0)+\frac{\partial f}{\partial y}(x_0,y_0)\frac{d\eta}{ds}(s_0)\right)(s-s_0)\end{aligned}$$

となります。この二つを比較して、

$$\frac{d\psi}{ds}(s_0)=\frac{\partial f}{\partial x}(x_0,y_0)\frac{d\xi}{ds}(s_0)+\frac{\partial f}{\partial y}(x_0,y_0)\frac{d\eta}{ds}(s_0)$$

という欲しかった合成関数の微分公式が得られる、というわけです。

それでは、式(12) を証明しましょう。

極限を取る前の式をいつものように「分子の水増し」をしてから変形して行くと、

$$\begin{aligned}\frac{\varphi(s)-\psi(s)}{|s-s_0|}&=\frac{f(\xi(s),\eta(s))-p(\varpi(s),\rho(s))}{|s-s_0|}\\&=\frac{f(\xi(s),\eta(s))-p(\xi(s),\eta(s))+p(\xi(s),\eta(s))-p(\varpi(s),\rho(s))}{|s-s_0|}\\&=\frac{f(\xi(s),\eta(s))-p(\xi(s),\eta(s))}{\sqrt{(\xi(s)-x_0)^2+(\eta(s)-y_0)^2}}\frac{\sqrt{(\xi(s)-x_0)^2+(\eta(s)-y_0)^2}}{|s-s_0|}\\&\quad+\frac{p(\xi(s),\eta(s))-p(\varpi(s),\rho(s))}{|s-s_0|}\end{aligned}$$

となります。

最後の式の第 1 項の第 2 因子で $s \to s_0$ とすると、$\xi(s)$ と $\eta(s)$ が微分可能であることから、

$$\begin{aligned}\lim_{s\to s_0}\frac{\sqrt{(\xi(s)-x_0)^2+(\eta(s)-y_0)^2}}{|s-s_0|} &= \lim_{s\to s_0}\sqrt{\frac{(\xi(s)-x_0)^2+(\eta(s)-y_0)^2}{(s-s_0)^2}}\\ &= \sqrt{\left(\lim_{s\to s_0}\frac{\xi(s)-x_0}{s-s_0}\right)^2+\left(\lim_{s\to s_0}\frac{\eta(s)-y_0}{s-s_0}\right)^2}\\ &= \sqrt{\left(\frac{d\xi}{ds}(s_0)\right)^2+\left(\frac{d\eta}{ds}(s_0)\right)^2}\end{aligned}$$

となります。

第 1 項の第 1 因子について考えてみましょう。$\xi(s)$ と $\eta(s)$ が微分可能だから連続でもあるので、$s \to s_0$ のとき $(\xi(s), \eta(s))$ は (x_0, y_0) に収束します。今、$p(x, y)$ は $f(x, y)$ の (x_0, y_0) における 1 次近似なので、(x, y) がどのように (x_0, y_0) に近づこうとも

$$\frac{f(x,y)-p(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} \to 0$$

が成り立っています。よって、第 1 項の第 1 因子は

$$\lim_{s\to s_0}\frac{f(\xi(s),\eta(s))-p(\xi(s),\eta(s))}{\sqrt{(\xi(s)-x_0)^2+(\eta(s)-y_0)^2}} = 0$$

となります。

以上より、$s \to s_0$ のとき最後の式の第 1 項は 0 に収束することが分かりました。

次に、第 2 項について考えてみましょう。$p(x, y)$ は 1 次式(9) なのですから、分子を具体的に展開してしまいましょう。目がチカチカするのを防ぐために、

$$\frac{\partial f}{\partial x}(x_0,y_0) = a, \qquad \frac{\partial f}{\partial y}(x_0,y_0) = b$$

とおくことにします。すると、

$$\begin{aligned}p(\xi(s),\eta(s)) - p(\varpi(s),\rho(s)) &= f(x_0,y_0) + a(\xi(s)-x_0) + b(\eta(s)-y_0)\\ &\qquad - f(x_0,y_0) - a(\varpi(s)-x_0) - b(\rho(s)-y_0)\\ &= a(\xi(s)-\varpi(s)) + b(\eta(s)-\rho(s))\end{aligned}$$

となります。$\varpi(s)$ は $\xi(s)$ の、$\rho(s)$ は $\eta(s)$ の s_0 における 1 次近似なので、結局、第 2 項で $s \to s_0$ とすると、

$$\begin{aligned}\lim_{s\to s_0}\frac{p(\xi(s),\eta(s))-p(\varpi(s),\rho(s))}{|s-s_0|} &= a\lim_{s\to s_0}\frac{\xi(s)-\varpi(s)}{|s-s_0|} + b\lim_{s\to s_0}\frac{\eta(s)-\rho(s)}{|s-s_0|}\\ &= a\cdot 0 + b\cdot 0 = 0\end{aligned}$$

となります。

以上で第1項も第2項も0に収束することが分かったので、証明したかった式(12)、つまり $\psi(s)$ が $\varphi(s)$ の s_0 における1次近似であることが証明できました。

問題 32. $f(x,y)=\sqrt{4-x^2-y^2}$, $\xi(s)=e^s$, $\eta(s)=e^{-s}$ とする。
(1) 合成関数 $\varphi(s)=f(\xi(s),\eta(s))$ の $s=0$ における接線の傾きを幾何学的な考察から求めよ。
(2) パラメタ曲線 $(\xi(s),\eta(s))$ の $s=0$ における接線（すなわち $\xi(s)$ の $s=0$ における1次近似を x 成分、$\eta(s)$ の $s=0$ における1次近似を y 成分とする直線）を $f(x,y)$ の $(1,1)$ における接平面の式（つまり1次近似）に代入してできる直線の傾きを求めよ。
(2) 合成関数の微分公式を使って $\varphi(s)$ の $s=0$ における微分の値を求めよ。

7.5.4 付録：前節の場合の偏微分の定義に従った証明

比較のために、前節と同じことを1次近似ではなく偏微分の定義で証明してみます。この節は余談ですので、先を急ぐひとは 17 ページの第 7.5.5 節にとんで下さい。

$\varphi'(s_0)$ を微分の定義式で書くと、

$$\lim_{s\to s_0}\frac{\varphi(s)-\varphi(s_0)}{s-s_0}=\lim_{s\to s_0}\frac{f(\xi(s),\eta(s))-f(\xi(s_0),\eta(s_0))}{s-s_0} \tag{13}$$

となります。このあとどう式変形をすればよいでしょうか。

しかし、この式(13)をぼんやりと眺めていると、**1変数関数の積の微分を導いたときの式**（第2回10〜11ページ）を思い出しませんか? 実際、$f(x,y)=xy$ なら $\varphi(s)=\xi(s)\eta(s)$ となって $\varphi(s)$ の微分は単なる積の微分になります。そこで、積の微分法の証明からヒントを得るために、積の微分法の証明のポイントを思い出してみましょう。

$$\begin{aligned}\frac{\xi(s)\eta(s)-\xi(s_0)\eta(s_0)}{s-s_0}&=\frac{\xi(s)\eta(s)-\xi(s_0)\eta(s)+\xi(s_0)\eta(s)-\xi(s_0)\eta(s_0)}{s-s_0}\\&=\frac{\xi(s)-\xi(s_0)}{s-s_0}\eta(s)+\xi(s_0)\frac{\eta(s)-\eta(s_0)}{s-s_0}\end{aligned}$$

というふうに、分子にわざと $-\xi(s_0)\eta(s)+\xi(s_0)\eta(s)$ という項を水増しするところがミソでした。そこで、式(13)でも同じように、

$$-f(\xi(s_0),\eta(s))+f(\xi(s_0),\eta(s))$$

を間に入れてみましょう。すると、

$$\begin{aligned}
&\frac{\varphi(s)-\varphi(s_0)}{s-s_0} \\
&= \frac{f(\xi(s),\eta(s))-f(\xi(s_0),\eta(s_0))}{s-s_0} \\
&= \frac{f(\xi(s),\eta(s))-f(\xi(s_0),\eta(s))+f(\xi(s_0),h(s))-f(\xi(s_0),\eta(s_0))}{s-s_0} \\
&= \frac{f(\xi(s),\eta(s))-f(\xi(s_0),\eta(s))}{\xi(s)-\xi(s_0)}\frac{\xi(s)-\xi(s_0)}{s-s_0} \qquad (14)\\
&\qquad + \frac{f(\xi(s_0),\eta(s))-f(\xi(s_0),\eta(s_0))}{\eta(s)-\eta(s_0)}\frac{\eta(s)-\eta(s_0)}{s-s_0} \qquad (15)
\end{aligned}$$

となります。

最後の式(14)+(15) で $s\to s_0$ としましょう。ξ と η がどちらも s_0 で微分可能なら、第 1 項(14) の第 2 因子と第 2 項(15) の第 2 因子は、それぞれ

$$\lim_{s\to s_0}\frac{\xi(s)-\xi(s_0)}{s-s_0}=\frac{d\xi}{ds}(s_0), \qquad \lim_{s\to s_0}\frac{\eta(s)-\eta(s_0)}{s-s_0}=\frac{d\eta}{ds}(s_0)$$

と収束します。また、η が s_0 で微分可能なら η は s_0 で連続、つまり、$\lim\limits_{s\to s_0}\eta(s)=\eta(s_0)$ が成り立っていますから、f が (x_0,y_0) で y によって偏微分可能なら、式(15) の第 1 因子で $y=\eta(s)$ とし $s\to s_0$ とすると

$$\lim_{s\to s_0}\frac{f(\xi(s_0),\eta(s))-f(\xi(s_0),\eta(s_0))}{\eta(s)-\eta(s_0)}=\lim_{y\to y_0}\frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}=\frac{\partial f}{\partial y}(x_0,y_0)$$

となります。残るは(14) の第 1 因子だけですが、これが問題です。

もしも(14) の第 1 因子の分子が $f(\xi(s),\eta(s_0))-f(\xi(s_0),\eta(s_0))$ であれば、つまり η の中身が s でなく s_0 なら、式(15) の第 1 因子と全く同様の計算で $s\to s_0$ のとき $f_x(x_0,y_0)$ という目標の値になってくれるのですが、η の中身が s_0 ではなく s なので安直に $s\to s_0$ の極限を取ることができません。「そんなことないんじゃない? $s\to s_0$ のとき $\eta(s)\to y_0$ なんだから、$\eta(s)$ を最初から $\eta(s_0)=y_0$ に取り替えておいても問題ないのでは?」と思われるかも知れません。つまり

$$\frac{f(\xi(s),\eta(s))-f(\xi(s_0),\eta(s))}{\xi(s)-\xi(s_0)}\neq\frac{f(\xi(s),\eta(s_0))-f(\xi(s_0),\eta(s_0))}{\xi(s)-\xi(s_0)}$$

だけれど、$s\to s_0$ の極限を取ってしまえばそんな「小さな違い」はなくなって、めでたく $f_x(\xi(s_0),\eta(s_0))$ に収束するのではないかと。

ところが、これは**一般には成り立ちません**。例を見ておきましょう。この関数は第 4 回の問題 13 で方向微分は可能だが「幸せの関係式」が成り立たない（つまり全微分可能でない）例として既に調べたことのある関数です。

例 4. $f(x,y)$ を

$$f(x,y) = \begin{cases} \dfrac{y(3x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ \\ 0 & (x,y) = (0,0) \end{cases}$$

とし、$\xi(s) = \eta(s) = s$, $s_0 = 0$ で考えてみましょう。$(\xi(s_0), \eta(s_0)) = (0,0)$ ですので、まず $(0,0)$ での x による偏微分の値を計算すると、

$$\frac{\partial f}{\partial x}(0,0) = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{x\to 0} \frac{0-0}{x} = 0$$

となって $f_x(0,0) = 0$ です。η の中身が s でなく s_0 だった場合の第 1 項第 1 因子は

$$\lim_{s\to 0} \frac{f(\xi(s), \eta(0)) - f(\xi(0), \eta(0))}{\xi(s)-\xi(0)} = \lim_{s\to 0} \frac{f(s,0)-f(0,0)}{s} = \lim_{s\to 0} 0 = 0$$

となってちゃんと $f_x(0,0)$ と一致します。一方、$\eta(s)$ を使った本物の第 1 項第 1 因子を計算すると、

$$\begin{aligned} \lim_{s\to 0} \frac{f(\xi(s), \eta(s)) - f(\xi(0), \eta(s))}{\xi(s)-\xi(0)} &= \lim_{s\to 0} \frac{f(s,s)-f(0,s)}{s} \\ &= \lim_{s\to 0} \frac{1}{s}\left(\frac{3s^3-s^3}{s^2+s^2} - \frac{-s^3}{s^2}\right) = \lim_{s\to 0} 2 = 2 \end{aligned}$$

となって $f_x(0,0) = 0$ に一致しません。 ■

これで、$f(x,y)$ が x で偏微分可能というだけではこの場合の合成関数の微分公式が成り立たないことになってしまいました。とは言え、勘のいい人はお気づきだと思いますが、f が C^1 級なら大丈夫です。

注意. 前節で既に証明してあるように f が全微分可能なら成り立つのですが、ここでの証明方法で全微分可能性だけしか仮定しないと本筋でないところで難しくなってしまうので、ここでは C^1 級として話を進めます。★

f が C^1 級なら、(14) の第 1 因子は期待通り $f_x(x_0, y_0)$ に収束することを証明しましょう。ポイントは、式(14) の第 1 因子の分子の中の $f(\xi(s), \eta(s))$ を $f(\xi(s), \eta(s_0))$ と、$f(\xi(s_0), \eta(s))$ を $f(\xi(s_0), \eta(s_0)$ と結びつけるために $f(x,y)$ を y だけを変数とする 1 変数関数と見なして平均値の定理を使うところです。どういうことかというと、一般に、微分可能な 1 変数関数 $g(y)$ と $y \neq y_0$ について、平均値の定理

$$\frac{g(y)-g(y_0)}{y-y_0} = g'(c) \quad \text{の成り立つ } c \text{ が } y \text{ と } y_0 \text{ の間に存在する}$$

を

$$g(y) = g(y_0) + g'(c)(y-y_0) \quad \text{の成り立つ } c \text{ が } y \text{ と } y_0 \text{ の間に存在する}$$

というふうに

$$g(y) \text{ を } g(y_0) \text{ で近似する関係式}$$

と見なすのです。

それでは証明しましょう。

証明. まず、$f(x,y)$ を y だけの関数だと思って 1 変数関数の平均値の定理を使います。y の 1 変数関数と思っていることをはっきりさせるために

$$\psi_x(y) := f(a,y)$$

というふうに x を小さく書いて y の 1 変数関数 ψ_x を定義します。(添え字の x は偏微分の記号ではなく、「x を決めるごとに決まる」ということを忘れないために書いたただの記号です。) すると、f が C^1 級であることから、特に $\psi_x(y)$ は微分可能なので平均値の定理を適用できます。つまり、$y \neq b$ である b と y に対し、

$$\frac{\psi_x(y) - \psi_x(b)}{y-b} = \frac{d\psi_x}{dy}(c)$$

を満たす c が y と b の間に存在します。これを ψ を使わず f で書くと

$$\frac{f(x,y) - f(x,b)}{y-b} = \frac{\partial f}{\partial y}(x,c)$$

となる c が b と y の間に存在する、となります。ただし、この c は b や y だけでなく x にも依存していることに注意してください。

ここで x として $\xi(s)$ と $\xi(s_0) = x_0$ を、y として $\eta(s)$ を、b として $\eta(s_0) = y_0$ と取ると、

$$\begin{aligned}
\frac{f(\xi(s),\eta(s)) - f(\xi(s),\eta(s_0))}{\eta(s)-\eta(s_0)} &= \frac{\partial f}{\partial y}(\xi(s),c) \\
\frac{f(\xi(s_0),\eta(s)) - f(\xi(s_0),\eta(s_0))}{\eta(s)-\eta(s_0)} &= \frac{\partial f}{\partial y}(\xi(s_0),c')
\end{aligned}$$

を満たす c と c' が $\eta(s)$ と $\eta(s_0)$ の間に存在することが分かります。それぞれ分母を払って整理すると、

$$\begin{aligned}
f(\xi(s),\eta(s)) &= f(\xi(s),\eta(s_0)) + \frac{\partial f}{\partial y}(\xi(s),c)(\eta(s)-\eta(s_0)) \\
f(\xi(s_0),\eta(s)) &= f(\xi(s_0),\eta(s_0)) + \frac{\partial f}{\partial y}(\xi(s_0),c')(\eta(s)-\eta(s_0))
\end{aligned}$$

となります。この 2 式をの右辺を式(14) の第 1 因子の分子に入れると、

$$\begin{aligned}
&\frac{f(\xi(s),\eta(s)) - f(\xi(s_0),\eta(s))}{\xi(s)-\xi(s_0)} \\
&= \frac{\{f(\xi(s),\eta(s_0)) + f_y(\xi(s),c)(\eta(s)-\eta(s_0))\} - \{f(\xi(s_0),\eta(s_0)) + f_y(\xi(s_0),c')(\eta(s)-\eta(s_0))\}}{\xi(s)-\xi(s_0)} \\
&= \frac{f(\xi(s),\eta(s_0)) - f(\xi(s_0),\eta(s_0))}{\xi(s)-\xi(s_0)} + \frac{\partial f}{\partial y}(\xi(s),c)\frac{\eta(s)-\eta(s_0)}{\xi(s)-\xi(s_0)} - \frac{\partial f}{\partial y}(\xi(s_0),c')\frac{\eta(s)-\eta(s_0)}{\xi(s)-\xi(s_0)} \qquad (16)
\end{aligned}$$

と変形できます。ここで $s \to s_0$ の極限をとりましょう。ξ も η も $s = s_0$ で微分可能であるとすると、

$$
\begin{aligned}
\lim_{s \to s_0} \frac{\eta(s) - \eta(s_0)}{\xi(s) - \xi(s_0)} &= \lim_{s \to s_0} \frac{\eta(s) - \eta(s_0)}{s - s_0} \frac{s - s_0}{\xi(s) - \xi(s_0)} \\
&= \left(\lim_{s \to s_0} \frac{\eta(s) - \eta(s_0)}{s - s_0} \right) \left(\lim_{s \to s_0} \frac{s - s_0}{\xi(s) - \xi(s_0)} \right) = \frac{\eta'(s_0)}{\xi'(s_0)}
\end{aligned}
$$

となります。一方、微分可能なら連続でもあるので $\lim_{s \to s_0} \xi(s) = \xi(s_0) = x_0$ および $\lim_{s \to s_0} \eta(s) = \eta(s_0) = y_0$ であり、c や c' は $\eta(s)$ と $\eta(s_0) = y_0$ の間の数なので、$s \to s_0$ のとき c も c' も $\eta(s_0) = y_0$ に収束します。今 f は C^1 級としているので、特に f_y は (x_0, y_0) で連続ですから

$$
\lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s), c) = \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s_0), c')
$$

となります。以上の二つのことから、式(16) の後の二つの項は

$$
\begin{aligned}
&\lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s), c) \frac{\eta(s) - \eta(s_0)}{\xi(s) - \xi(s_0)} \\
&= \left(\lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s), c) \right) \left(\lim_{s \to s_0} \frac{\eta(s) - \eta(s_0)}{\xi(s) - \xi(s_0)} \right) = \frac{\partial f}{\partial y}(x_0, y_0) \frac{\eta'(s_0)}{\xi'(s_0)} \\
&\lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s_0), c') \frac{\eta(s) - \eta(s_0)}{\xi(s) - \xi(s_0)} \\
&= \left(\lim_{s \to s_0} \frac{\partial f}{\partial y}(\xi(s_0), c') \right) \left(\lim_{s \to s_0} \frac{\eta(s) - \eta(s_0)}{\xi(s) - \xi(s_0)} \right) = \frac{\partial f}{\partial y}(x_0, y_0) \frac{\eta'(s_0)}{\xi'(s_0)}
\end{aligned}
$$

と同じ値に収束します。よって、$s \to s_0$ の極限をとると式(16) の後の二つの項は打ち消し合うので、

$$
\begin{aligned}
\lim_{s \to s_0} \frac{f(\xi(s), \eta(s)) - f(\xi(s_0), \eta(s))}{\xi(s) - \xi(s_0)} &= \lim_{s \to s_0} \frac{f(\xi(s), \eta(s_0)) - f(\xi(s_0), \eta(s_0))}{\xi(s) - \xi(s_0)} \\
&= \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = \frac{\partial f}{\partial x}(x_0, y_0)
\end{aligned}
$$

となります。これで目標が示せました。 □

以上で証明完了です。1 次近似の視点からの証明と比較してみて下さい。

7.5.5 2 変数関数に 2 変数関数を 2 つ入れた場合

「1 変数関数に 2 変数関数を入れた場合」の議論と「2 変数関数に 1 変数関数を 2 つ入れた場合」の議論をくっつけると「2 変数関数に 2 変数関数を 2 つ入れた場合」の証明になるのですが、全く同じことを繰り返すのも能がないので、ここで

はもとの関数と 1 次近似との差である**剰余**にスポットを当てて証明してみましょう。もちろんやることは今までと同じなのですが、剰余という考え方に慣れることは大切です。

なお、これから紹介する証明は**そのまま**一般の場合に適用することができます。興味のある人は確認してみて下さい。

$f(x,y),\ \xi(s,t),\ \eta(s,t)$ をすべて全微分可能な関数とし、$x=\xi(s,t),\ y=\eta(s,t)$ を $f(x,y)$ に合成して得られる合成関数を $\varphi(s,t)$ としましょう。

$$\varphi(s,t)=f(\xi(s,t),\eta(s,t))$$

です。

$x_0=\xi(s_0,t_0), y_0=\eta(s_0,t_0)$ とし、$f(x,y)$ の (x_0,y_0) における 1 次近似を $p(x,y)$、$\xi(s,t)$ と $\eta(s,t)$ の (s_0,t_0) における 1 次近似をそれぞれ $\varpi(s,t),\rho(s,t)$ とします。また、$p(x,y),\varpi(s,t),\rho(s,t)$ を具体的な 1 次式として書いたときの見た目をスッキリさせるために、

$$\begin{aligned}
\frac{\partial f}{\partial x}(x_0,y_0)&=a, &\qquad \frac{\partial f}{\partial y}(x_0,y_0)&=b,\\
\frac{\partial \xi}{\partial s}(s_0,t_0)&=\alpha_1, &\qquad \frac{\partial \xi}{\partial t}(s_0,t_0)&=\alpha_2,\\
\frac{\partial \eta}{\partial s}(s_0,t_0)&=\beta_1, &\qquad \frac{\partial \eta}{\partial t}(s_0,t_0)&=\beta_2
\end{aligned}$$

と置くことにします。これを使うと、

$$p(x,y)=f(x_0,y_0)+a(x-x_0)+b(y-y_0) \tag{17}$$

$$\varpi(s,t)=x_0+\alpha_1(s-s_0)+\alpha_2(t-t_0) \tag{18}$$

$$\rho(s,t)=y_0+\beta_1(s-s_0)+\beta_2(t-t_0) \tag{19}$$

となります。

$p(x,y)$ に $x=\varpi(s,t)$ と $y=\rho(s,t)$ を合成してできる s,t の 1 次式も合成関数の記号で書くことができないので $\psi(s,t)$ と書くことにします。つまり、

$$\psi(s,t)=p(\varpi(s,t),\rho(s,t))$$

です。示したいことは、$\psi(s,t)$ が $\varphi(s,t)$ の (s_0,t_0) における 1 次近似であること、すなわち

$$\lim_{(s,t)\to(s_0,t_0)}\frac{\varphi(s,t)-\psi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}=0 \tag{20}$$

が成り立つことです。

これが証明できたとします。すると、

$$\psi(s,t)=f(x_0,y_0)+\frac{\partial \varphi}{\partial s}(s_0,t_0)(s-s_0)+\frac{\partial \varphi}{\partial t}(s_0,t_0)(t-t_0)$$

が成り立つことになります。なぜなら、$\varphi(s,t)$ の 1 次近似はこの式の右辺**しかない**からです。一方、式(17) に式(18) と(19) を代入して $\psi(s,t)$ を具体的に計算すると、

$$\begin{aligned}\psi(s,t) &= p(\varpi(s,t),\rho(s,t))\\ &= f(x_0,y_0) + a(x_0+\alpha_1(s-s_0)+\alpha_2(t-t_0)-x_0)\\ &\qquad\qquad + b(y_0+\beta_1(s-s_0)+\beta_2(t-t_0)-y_0)\\ &= f(x_0,y_0) + (a\alpha_1+b\beta_1)(s-s_0) + (a\alpha_2+b\beta_2)(t-t_0)\end{aligned}$$

となります。二つの式を比較して、$a,b,\alpha_1,\alpha_2,\beta_1,\beta_2$ を元に戻すと、

$$\begin{aligned}\frac{\partial\varphi}{\partial s}(s_0,t_0) &= \frac{\partial f}{\partial x}(x_0,y_0)\frac{\partial\xi}{\partial s}(s_0,t_0) + \frac{\partial f}{\partial y}(x_0,y_0)\frac{\partial\eta}{\partial s}(s_0,t_0)\\ \frac{\partial\varphi}{\partial t}(s_0,t_0) &= \frac{\partial f}{\partial x}(x_0,y_0)\frac{\partial\xi}{\partial t}(s_0,t_0) + \frac{\partial f}{\partial y}(x_0,y_0)\frac{\partial\eta}{\partial t}(s_0,t_0)\end{aligned}$$

という欲しかった合成関数の微分公式が二つ一遍に得られる、というわけです。

それでは、最後に式(20) を証明しましょう.

まず、$f(x,y)$, $\xi(s,t)$, $\eta(s,t)$ のそれぞれの**剰余**を $r(x,y),\sigma(s,t),\tau(s,t)$ とします。つまり、

$$\begin{aligned}r(x,y) &:= f(x,y) - p(x,y)\\ \sigma(s,t) &:= \xi(s,t) - \varpi(s,t)\\ \tau(s,t) &:= \eta(s,t) - \rho(s,t)\end{aligned}$$

と定義するということです。$p(x,y)$ が $f(x,y)$ の (x_0,y_0) における 1 次近似であり、$\varpi(s,t)$ が $\xi(s,t)$ の、$\rho(s,t)$ が $\eta(s,t)$ の (s_0,t_0) における 1 次近似であるということは、剰余を使って

$$\begin{aligned}\lim_{(x,y)\to(x_0,y_0)} \frac{r(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} &= 0\\ \lim_{(s,t)\to(s_0,t_0)} \frac{\sigma(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} &= 0\\ \lim_{(s,t)\to(s_0,t_0)} \frac{\tau(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} &= 0\end{aligned}$$

が成り立つことと書き表すことができます。

これらの記号を使うと、

$$
\begin{aligned}
&\varphi(s,t)\\
&= f(\xi(s,t),\eta(s,t))\\
&= p(\xi(s,t),\eta(s,t)) + r(\xi(s,t),\eta(s,t))\\
&= f(x_0,y_0) + a(\xi(s,t)-x_0) + b(\eta(s,t)-y_0) + r(\xi(s,t),\eta(s,t))\\
&= f(x_0,y_0) + a(\varpi(s,t)+\sigma(s,t)-x_0) + b(\rho(s,t)+\tau(s,t)-y_0)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + r(\xi(s,t),\eta(s,t))\\
&= f(x_0,y_0) + a(\alpha_1(s-s_0)+\alpha_2(t-t_0)+\sigma(s,t))\\
&\qquad\qquad\quad + b(\beta_1(s-s_0)+\beta_2(t-t_0)+\tau(s,t)) + r(\xi(s,t),\eta(s,t))\\
&= f(x_0,y_0) + (a\alpha_1+b\beta_1)(s-s_0) + (a\alpha_2+b\beta_2)(t-t_0)\\
&\qquad\qquad\qquad\quad + a\sigma(s,t) + b\tau(s,t) + r(\xi(s,t),\eta(s,t))
\end{aligned}
$$

となります。よって、1 次近似の合成

$$
\begin{aligned}
\psi(s,t) &:= p(\varpi(s,t),\rho(s,t))\\
&= f(x_0,y_0) + (a\alpha_1+b\beta_1)(s-s_0) + (a\alpha_2+b\beta_2)(t-t_0)
\end{aligned}
$$

を $\varphi(s,t)$ から引いて $\sqrt{(s-s_0)^2+(t-t_0)^2}$ で割ると

$$
\begin{aligned}
&\frac{\varphi(s,t)-\psi(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&= \frac{a\sigma(s,t)+b\tau(s,t)+r(\xi(s,t),\eta(s,t))}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&= a\frac{\sigma(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} + b\frac{\tau(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&\qquad + \frac{r(\xi(s,t),\eta(s,t))}{\sqrt{(\xi(s,t)-x_0)^2+(\eta(s,t)-y_0)^2}}\frac{\sqrt{(\xi(s,t)-x_0)^2+(\eta(s,t)-y_0)^2}}{\sqrt{(s-s_0)^2+(t-t_0)^2}}
\end{aligned}
$$

となります。ここで、$\sigma(s,t)$ と $\tau(s,t)$ は 1 次近似の剰余なので、既に書いたとおり

$$
\lim_{(s,t)\to(s_0,t_0)} \frac{\sigma(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} = \lim_{(s,t)\to(s_0,t_0)} \frac{\tau(s,t)}{\sqrt{(s-s_0)^2+(t-t_0)^2}} = 0
$$

が成り立っています。また、ξ も η も全微分可能であることから連続でもあるので、

$$
\lim_{(s,t)\to(s_0,t_0)} \xi(s,t) = x_0, \qquad \lim_{(s,t)\to(s_0,t_0)} \eta(s,t) = y_0
$$

です。よって、

$$
\begin{aligned}
&\lim_{(s,t)\to(s_0,t_0)} \frac{r(\xi(s,t),\eta(s,t))}{\sqrt{(\xi(s,t)-x_0)^2+(\eta(s,t)-y_0)^2}}\\
&\qquad = \lim_{(x,y)\to(x_0,y_0)} \frac{r(x,y)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = 0
\end{aligned}
$$

が成り立ちます。問題は

$$\frac{\sqrt{(\xi(s,t)-x_0)^2+(\eta(s,t)-y_0)^2}}{\sqrt{(s-s_0)^2+(t-t_0)^2}}$$

で $(s,t)\to(s_0,t_0)$ としたときどうなるかです。これは「1 変数関数に 2 変数関数を入れた場合」のときと同様に「収束はしないが有界」になります。キチンと証明しましょう。

証明. まず、任意の 2 実数 u,v に対して

$$\max\{|u|,|v|\}\le\sqrt{u^2+v^2}\le|u|+|v|$$

が成り立つことに注意しましょう。これと三角不等式を使うと、

$$\begin{aligned}
&\frac{\sqrt{(\xi(s,t)-x_0)^2+(\eta(s,t)-y_0)^2}}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&\le\frac{|\xi(s,t)-x_0|+|\eta(s,t)-y_0|}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&\le\frac{|a_1||s-s_0|+|a_2||t-t_0|+|\sigma(s,t)|+|b_1||s-s_0|+|b_2||t-t_0|+|\tau(s,t)|}{\sqrt{(s-s_0)^2+(t-t_0)^2}}\\
&\le(|a_1|+|b_1|)\frac{|s-s_0|}{|s-s_0|}+(|a_2|+|b_2|)\frac{|t-t_0|}{|t-t_0|}\\
&\qquad+\frac{|\sigma(s,t)|}{\sqrt{(s-s_0)^2+(t-t_0)^2}}+\frac{|\tau(s,t)|}{\sqrt{(s-s_0)^2+(t-t_0)^2}}
\end{aligned}$$

となります。最後の二つの項は σ と τ が 1 次近似の剰余項であることから $(s,t)\to(s_0,t_0)$ のとき 0 に収束します。よって、全体としては $(s,t)\to(s_0,t_0)$ のとき収束はしないとしても 絶対値が $|a_1|+|a_2|+|b_1|+|b_2|$ を超えないので有界です。 □

以上より、証明したかった式(20)、すなわち $\psi(s,t)$ が $\varphi(s,t)$ の (s_0,t_0) における 1 次近似であることが証明できました。

7.6 1 次近似と行列による記法

7.6.1 写像と変位

第 7.3 節で、合成関数の微分公式は行列の積を使うときれいにまとめられるということを紹介しました。そこにはそのことの説明はしないと書いてしまいましたが、やはりちょっとだけ「何故行列を使うときれいにまとまるのか」ということの理由を説明してみます。

2 変数関数 $f(x,y)$ に二つの 2 変数関数 $x=\xi(s,t)$, $y=\eta(s,t)$ を入れた合成関数 $\varphi(s,t)=f(\xi(s,t),\eta(s,t))$ の場合の微分公式を行列でまとめたものは

$$\begin{bmatrix}\frac{\partial\varphi}{\partial s} & \frac{\partial\varphi}{\partial t}\end{bmatrix}=\begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{bmatrix}\begin{bmatrix}\frac{\partial\xi}{\partial s} & \frac{\partial\xi}{\partial t}\\ \frac{\partial\eta}{\partial s} & \frac{\partial\eta}{\partial t}\end{bmatrix} \tag{21}$$

でした。

さて、ここで $d\varphi=[\varphi_s\ \varphi_t]$ や $df=[f_x\ f_y]$ の意味を思い出してみましょう。どこまで思い出すのかというと、$d\varphi$ や df を 1 行 2 列の行列だと思うのは、その行列は「変位ベクトル」という 2 次元縦ベクトルに左から掛けることによって初めて意味を持つのだ、ということを思い出すのです。すると、上の式(21) の左辺の $d\varphi$ は (s,t) という独立変数の空間での変位ベクトルに左から掛けることになります。ということは、右辺も全体として (s,t) 空間での変位ベクトルに左から掛かることになりますが、そうすると、(s,t) 空間での変位ベクトルに掛かっているのは 2 行 2 列の行列になってしまいます。しかも、その掛け算をした結果できあがるベクトルには左から df が掛かってきています。df を左から掛けることに意味のあるベクトルは (x,y) という変数の空間における変位ベクトルでした。ということは、2 行 2 列の行列

$$\begin{bmatrix}\frac{\partial\xi}{\partial s} & \frac{\partial\xi}{\partial t}\\ \frac{\partial\eta}{\partial s} & \frac{\partial\eta}{\partial t}\end{bmatrix}$$

は

(s,t) 空間の変位に左から掛けると (x,y) 空間の変位になる

という意味を持つものになっていていなければ、上の式変形は単なる形式的なものに過ぎなかったことになってしまいます。

ここで、**写像**という概念を登場させましょう。2 変数関数に 1 変数関数や 2 変数関数を入れる場合、記号 $\circ$ を使った合成関数の表し方がうまく通用しませんでした。その理由は「コピー機」が邪魔だったからです。この「コピー機」が消えてしまうような設定を作り出してくれるのが写像という概念です。なんだか恐ろしげな言い回しになってしまっていますが、何のことはない、二つの変数 x と y を (x,y) とまとめて一つのもののように扱ったのと同様にして、$\xi(s,t)$ と $\eta(s,t)$ もまとめてしまおうというだけのことです。つまり、

s と t を決めると $\xi(s,t)$ と $\eta(s,t)$ の値がそれぞれ決まる

と考える代わりに、

(s,t) を決めると $(\xi(s,t),\eta(s,t))$ が決まる

と考えるだけのことです。ただし、微分を考える場合、$(\xi(s,t),\eta(s,t))$ という値そのものではなく、変位ベクトルが重要であり、**変位ベクトルは縦ベクトル**であるような記法をここでは採用しているので、はじめから

$$X(s,t) := \begin{bmatrix} \xi(s,t) \\ \eta(x,t) \end{bmatrix}$$

というふうに縦に並べることにします。左辺に書いた $X(s,t)$ は (s,t) に 2 次元縦ベクトルを対応させる写像

$$X\colon \mathbb{R}^2 \to \mathbb{R}^2$$

を一文字で表す[7]ために決めた記号です。値がベクトルなので上に矢印を付けてもよいですが、面倒なのでやめました。

このような写像 X を導入すると、まず、

$$\varphi(s,t) = f \circ X(s,t)$$

という合成関数の記号が復活します。また、そのような形式的なことばかりではなく、X が

(s,t) 空間の点に (x,y) 空間の点を対応させる

写像であることから、X の「微分」と呼べるものがあるとすれば、それは

(s,t) 空間の**変位ベクトル**に (x,y) 空間の**変位ベクトル**を対応させる

というまさに欲しかったものになるはずです。そう思ってもう一度問題の行列を眺めてみると、第 1 行は

x 成分の変位を与える 1 行 2 列の行列 $d\xi$

になっているし、第 2 行は

y 成分の変位を与える 1 行 2 列の行列 $d\eta$

になっていますから、行列全体としては

(x,y) の変位を縦ベクトルとして与える行列

になっているわけです。

以上により、問題の行列は X の微分と呼ぶにふさわしい行列であると言えます。そこで、ちゃんと記号を決めるべきなのですが、d という記号は関数にしか使わ

[7] ξ の大文字の Ξ を使おうかとも思ったのですが、あまりにもなじみがないでしょうから、対応するローマ字の方の大文字にしておきました。

ないという習慣が（もちろん意味があって）決まっているので dX と書くわけにはいかないのです。そこで、1 変数関数のときの記号を流用して $X'(s,t)$ と書く、という選択肢が一つあります。それから、この行列を初めて問題にした（と言われている）ヤコビ（Jacobi）という人の頭文字をとって $J_X(s,t)$ と書く記法もよく使われます。実際、この行列そのものの名前は「X の微分」ではなく「X の**ヤコビ行列**」という方が普通です。以上、記号は

$$X'(s,t) = J_X(s,t) = \begin{bmatrix} \dfrac{\partial \xi}{\partial s} & \dfrac{\partial \xi}{\partial t} \\ \dfrac{\partial \eta}{\partial s} & \dfrac{\partial \eta}{\partial t} \end{bmatrix}$$

ということになります。これを使うと合成関数の微分公式は

$$d(\varphi \circ X) = df \cdot J_X$$

とまとめることができます。

ここまでは「2 変数関数に 2 変数関数を 2 つ入れた場合」に限って説明しましたが、全く同じ行列を使った整理が「n 変数関数に m 変数関数を n 個入れた場合」に成り立つことも実際に計算をしてもらえれば分かります。つまり、

$$f(x_1, x_2, \ldots, x_m)$$

に

$$x_i = \xi_i(s_1, s_2, \ldots, s_n) \qquad i = 1, 2, \ldots, m$$

を入れることを、

$$X(s_1, s_2, \ldots, s_n) := \begin{bmatrix} \xi_1(s_1, s_2, \ldots, s_n) \\ \xi_2(s_1, s_2, \ldots, s_n) \\ \vdots \\ \xi_m(s_1, s_2, \ldots, s_n) \end{bmatrix}$$

という写像 $X\colon \mathbb{R}^n \to \mathbb{R}^m$ との合成だと考えると、X の微分、すなわちヤコビ行列を

$$X'(s_1, \ldots, s_n) = J_X(s_1, \ldots, s_n) = \begin{bmatrix} \dfrac{\partial \xi_1}{\partial s_1} & \dfrac{\partial \xi_1}{\partial s_2} & \cdots & \dfrac{\partial \xi_1}{\partial s_n} \\ \dfrac{\partial \xi_2}{\partial s_1} & \dfrac{\partial \xi_2}{\partial s_2} & \cdots & \dfrac{\partial \xi_2}{\partial s_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \xi_m}{\partial s_1} & \dfrac{\partial \xi_m}{\partial s_2} & \cdots & \dfrac{\partial \xi_m}{\partial s_n} \end{bmatrix}$$

という m 行 n 列の行列で定義すれば（と言うか、こうであるべきであって）、合成関数の微分公式は

$$d(f \circ X) = df \cdot J_X$$

と表されることになります。

このように一般の m, n で考えてみると、m や n が 1 であって悪い理由は何もないことに気づくでしょう。つまり、単なる関数も「$\mathbb{R}$ への**写像**」じゃないか、というわけです。そう考えると、

> 2 変数関数 $f(x, y)$ の微分、つまりヤコビ行列は 1 行 2 列の行列 df そのもの

であり、

> 1 変数関数 $f(x)$ の微分、つまりヤコビ行列は $f'(x)$ を唯一の成分とする **1 行 1 列の行列**

だったんだ、というわけです。

ここまで一般的に考えるなら、記号も d とか J とかはやめてシンプルにダッシュで統一してしまうことにすれば、

$$\mathbb{R}^n \xrightarrow{\quad g \quad} \mathbb{R}^m \xrightarrow{\quad f \quad} \mathbb{R}^l$$

に対して（もちろん n, m, l は 1 でも O.K. です）、

$$(f \circ g)' = (f' \circ g) \cdot g'$$

が成り立つことになります。1 変数関数のときの公式が完全に再現されるわけです。

7.6.2 写像の 1 次近似

以上で行列にまとめることに意味があることは分かっていただけたと思います。ここで、行列が 1 次近似という視点からもっと積極的な意味を持っていることを、証明抜きで結果だけ説明します。

2 変数の 1 次式を二つ用意し、それを並べて $\mathbb{R}^2$ から $\mathbb{R}^2$ への写像を作ります。例えば、

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 s + b_1 t + c_1 \\ a_2 s + b_2 t + c_2 \end{bmatrix}$$

としましょう。これの右辺は

$$\begin{bmatrix} a_1 s + b_1 t + c_1 \\ a_2 s + b_2 t + c_2 \end{bmatrix} = \begin{bmatrix} a_1 s + b_1 t \\ a_2 s + b_2 t \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

というように行列の積とベクトルの足し算を使って書き表せます。ここで、

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \vec{s} = \begin{bmatrix} s \\ t \end{bmatrix}, \qquad \vec{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \qquad A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

と記号を決めると、

$$\vec{x} = A\vec{s} + \vec{c} \tag{22}$$

というように 1 変数の 1 次関数と全く同じ見た目で表すことができます。このように、一般の写像においても 1 次関数に当たる写像は行列とベクトルを使って式(22)のように書けるのです。

さて、いきなり結論ですが、写像 $X\colon \mathbb{R}^n \to \mathbb{R}^m$ の $\vec{s}_0 \in \mathbb{R}^n$ における 1 次近似（つまり X に「一番似ている」1 次関数）は、ヤコビ行列 $J_X(\vec{s}_0)$ を使って

$$J_X(\vec{s}_0)(\vec{s} - \vec{s}_0) + X(\vec{s}_0)$$

と書けるのです。例えば、$\xi(s_0, t_0) = x_0$, $\eta(s_0, t_0) = y_0$ とすると、写像

$$X(s,t) = \begin{bmatrix} \xi(s,t) \\ \eta(s,t) \end{bmatrix}$$

の (s_0, t_0) における一次近似は

$$\begin{bmatrix} \dfrac{\partial \xi}{\partial s}(s_0,t_0) & \dfrac{\partial \xi}{\partial t}(s_0,t_0) \\ \dfrac{\partial \eta}{\partial t}(s_0,t_0) & \dfrac{\partial \eta}{\partial t}(s_0,t_0) \end{bmatrix} \begin{bmatrix} s - s_0 \\ t - t_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

なのです。

7.7 逆写像の微分と変数変換

1 変数関数のときには存在する場合があったのに、多変数関数では存在し得ない重要な概念に**逆関数**というものがありました。しかし、写像まで対象を広げておけば**逆写像**として復活します。逆写像を持つ写像が自然にでてくるのは直交座標と極座標の兼ね合いを考える場合などの**変数変換**においてです。物理における変数変換の重要性は言うまでもないでしょう。この節では、逆写像の微分公式を説明した後、それを使って変数変換に対する微分の振る舞いについて考えます。

7.7.1 逆写像の微分公式

n 変数関数が m 個あったとき、それの決める写像が逆写像を持つ、つまり、

$$y_1 = f_1(x_1, \ldots, x_n), \ldots, y_m = f_m(x_1, \ldots, x_n) \tag{23}$$

が逆に解けて

$$x_1 = g_1(y_1, \ldots, y_m), \ldots, x_n = g_n(y_1, \ldots, y_m) \tag{24}$$

となったとしましょう。つまり、すべての $i = 1, 2, \ldots, m$ について

$$f_i(g_1(y_1, \ldots, y_m), \ldots, g_n(y_1, \ldots, y_m)) = y_i$$

が成り立ち、かつ、すべての $j = 1, 2, \ldots, n$ について

$$g_j(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) = x_j$$

が成り立っているとします。

m 個の関数(23) たちの決める $\mathbb{R}^n$ から $\mathbb{R}^m$ への写像を

$$F\colon \mathbb{R}^n \ni \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \to \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix} \in \mathbb{R}^m$$

とし、n 個の関数(24) たちの決める $\mathbb{R}^m$ から $\mathbb{R}^n$ への写像を

$$G\colon \mathbb{R}^m \ni \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \to \begin{bmatrix} g_1(y_1, \ldots, y_m) \\ \vdots \\ g_n(y_1, \ldots, y_m) \end{bmatrix} \in \mathbb{R}^n$$

とすると、「逆に解ける」とは

$$F \circ G = \text{「}\mathbb{R}^m \text{ の恒等写像」}, \qquad G \circ F = \text{「}\mathbb{R}^n \text{ の恒等写像」}$$

の両方が成り立つことと同じです。この 2 式の両辺を微分する、つまりヤコビ行列にすると、合成関数の微分公式によって任意の点で

$$J_F \cdot J_G = E_m, \qquad J_G \cdot J_F = E_n$$

が成り立ちます。ただし E_m および E_n はそれぞれ m 次および n 次の単位行列です。($\mathbb{R}^n$ の恒等写像とは

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \longmapsto \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

と対応させること、つまり第 i 成分が x_i という関数でできている写像のことなので、ヤコビ行列の (i, j) 成分は

$$\frac{\partial x_i}{\partial x_j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

すなわち単位行列 E_n になります。）つまり、任意の点で J_f と J_g は互いに逆行列だというわけです。よって、特に J_f は正方行列であり、$n = m$ となります[8]。

写像の記法では、このような G のことを「F の逆写像」と呼び F^{-1} と書きますから、それに従って「逆写像の微分公式」は

$$J_{F^{-1}} = {J_F}^{-1} \tag{25}$$

となります。もちろん右辺は「J_F という行列の逆行列」という意味です。**1 変数関数の逆関数の微分法と全く同じ形**になるわけです。

2 変数のときを成分で書いてみましょう。添え字で書くのはやめて、

$$F = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}, \qquad F^{-1} = \begin{bmatrix} \varphi(s,t) \\ \psi(s,t) \end{bmatrix}$$

としましょう。すると、

$$J_F = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}, \qquad J_{F^{-1}} = \begin{bmatrix} \frac{\partial \varphi}{\partial s} & \frac{\partial \varphi}{\partial t} \\ \frac{\partial \psi}{\partial s} & \frac{\partial \psi}{\partial t} \end{bmatrix}$$

なので、逆写像の微分法の公式は

$$\begin{bmatrix} \frac{\partial \varphi}{\partial s} & \frac{\partial \varphi}{\partial t} \\ \frac{\partial \psi}{\partial s} & \frac{\partial \psi}{\partial t} \end{bmatrix} = \frac{1}{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}} \begin{bmatrix} \frac{\partial g}{\partial y} & -\frac{\partial f}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial f}{\partial x} \end{bmatrix}$$

となります。成分で書けば、例えば

$$\frac{\partial \varphi}{\partial s} = \frac{\frac{\partial g}{\partial y}}{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}}$$

などとなって、わけわかりませんね。行列で表示することの便利さが実感できると思います。

7.7.2 変数変換

逆写像の微分の公式が手に入ったので、「微分の変数変換」が簡単に計算できます。

[8]「$AB = E_n$ かつ $BA = E_m$ となるならば $n = m$、A は正則行列で $B = A^{-1}$」という事実については数学 II（線型代数）の階数のところで学びます。ここでは認めさせてください。

まず、1 変数関数の場合で考えてみましょう。$f(x)$ を $x=\xi(s)$ により s の関数に変数変換してみましょう。どちらの変数で考えているかハッキリさせるために、$\xi(s)$ を入れた合成関数を $g(s)$ と書くことにします。つまり、$g(s)=f(\xi(s))$ です。合成関数の微分法により

$$\frac{dg}{ds}(s)=\frac{df}{dx}(\xi(s))\frac{d\xi}{ds}(s)$$

となりますが、この式の登場人物のうち dg/ds と $d\xi/ds$ は s の関数で、df/dx は（今は x に $\xi(s)$ が入っているから s の関数ですが、本来は）x の関数です。だから、この公式は s の関数を x の関数と s の関数で表しているわけで、少々据わりが悪い感じがします。

ここで、もし $\xi(s)$ が逆関数 $s=\sigma(x)$ を持つなら、逆関数の微分法から

$$\frac{d\xi}{ds}(s)=\frac{1}{\dfrac{d\sigma}{dx}(\xi(s))}$$

ですので、

$$\frac{dg}{ds}(s)=\frac{\dfrac{df}{dx}(x)}{\dfrac{d\sigma}{dx}(x)}$$

と、s での微分をを（本来）x を変数とする関数として表すことができます。（等号の意味は、「右辺の x に $\xi(s)$ を代入するか、左辺の s に $\sigma(s)$ を代入すれば両辺が同じ関数になる」という意味です。）もちろん、同様に、

$$\frac{df}{dx}(x)=\frac{\dfrac{dg}{ds}(s)}{\dfrac{d\xi}{ds}(s)}$$

と、x での微分を（本来）s を変数とする関数として表すこともできます。

ヤコビ行列を使えば、多変数関数の変数変換に対しても全く同じ計算が可能になります。f を $x_1,\ldots,x_n$ を変数とする関数とし、$\xi_1(s_1,\ldots,s_n),\ldots,\xi_n(s_1,\ldots,s_n)$ という「逆に解ける」関数の組によって f の変数を $x_1,\ldots,x_n$ から $s_1,\ldots,s_n$ に変換してみましょう。上と同様、どちらの変数で考えているのかわからなくならないために、$s_1,\ldots,s_n$ に変数を変換したあとの関数を $g(s_1,\ldots,s_n)$ と書くことにします。変数変換の写像を一文字で $\Phi\colon\mathbb{R}^n\to\mathbb{R}^n$ と書くことにする、詳しくは、

$$\Phi(s_1,\ldots,s_n)=\begin{bmatrix}\xi_1(s_1,\ldots,s_n)\\ \vdots\\ \xi_n(s_1,\ldots,s_n)\end{bmatrix}$$

とすると、

$$g = f \circ \Phi$$

なので、合成写像の微分法から

$$J_g(s_1, \ldots, s_n) = J_f(x_1, \ldots, x_n) J_\Phi(s_1, \ldots, s_n)$$

が得られる一方、Φ の逆写像 Φ^{-1} があって、

$$f = g \circ \Phi^{-1}$$

が成り立っているので、

$$J_f(x_1, \ldots, x_n) = J_g(s_1, \ldots, s_n) J_{\Phi^{-1}}(x_1, \ldots, x_n)$$

という行列の式が得られます。ただし、J_f と J_g は 1 行 n 列の行列、つまり df と dg を表す行列です。1 変数のときと同様、J_f と $J_{\Phi^{-1}}$ の本来の変数は $x_1, \ldots, x_n$ で、J_g と J_Φ の本来の変数は $s_1, \ldots, s_n$ です。しかし、逆写像の微分公式により、J_Φ と $J_{\Phi^{-1}}$ は互いに相手の逆行列ですので、上の二つの式は

$$J_g(s_1, \ldots, s_n) = J_f(x_1, \ldots, x_n) \big(J_{\Phi^{-1}}(x_1, \ldots, x_n) \big)^{-1}$$

及び

$$J_f(x_1, \ldots, x_n) = J_g(s_1, \ldots, s_n) \big(J_\Phi(s_1, \ldots, s_n) \big)^{-1} \tag{26}$$

となります。

問題 33. 極座標 $x = r\cos\theta, y = r\sin\theta$ を考える。これは全単射ではないので無制限では変数変換になっていないが、例えば $r > 0, 0 \leq \theta < 2\pi$ と制限すれば全単射なので変数変換である。以下、この範囲で考えることにする。また、2 変数関数 $f(x, y)$ に対して $g(r, \theta)$ を

$$g(r, \theta) = f(r\cos\theta, r\sin\theta)$$

で定義する。

(1) f が全微分可能なとき、$f_x(r\cos\theta, r\sin\theta)$ を g_r, g_θ を使って表せ。

(2) f が C^2 級のとき、$f_{xx}(r\cos\theta, r\sin\theta) + f_{yy}(r\cos\theta, r\sin\theta)$ を g の偏微分や 2 階偏微分を使って表せ。

問題 32 の解答

(1) 関数 $z = f(x, y)$ のグラフは半径 2 の球面の上側ですので、点 (x, y) が原点に近ければ近いほど $f(x, y)$ の値は大きくなります。実際、原点と点 (x, y) の距離は $\sqrt{x^2 + y^2}$ であり、$f(x, y)$ は

$$f(x, y) = \sqrt{4 - \left(\sqrt{x^2 + y^2}\right)^2}$$

と書き換えられるので、$\sqrt{x^2 + y^2}$ が小さいほど f の値は大きくなります。一方、xy 平面上の曲線 $(\xi(s), \eta(s)) = (e^s, e^{-s})$ は双曲線 $xy = 1$ の第 1 象限の部分ですので、$s = 0$ のときの点 $(1, 1)$ はこの曲線上で原点に最も近い点です。よって、合成関数 $\varphi(s)$ は $s = 0$ のときに最大値をとります。

微分可能な 1 変数関数は最大値をとる点で微係数は 0 です。（つまり接線は独立変数 s の軸に平行です。）つまり $\varphi'(0) = 0$ です。

(1) 別解 $\xi(s)$ と $\eta(s)$ は $\eta(-s) = \xi(s)$ という関係にあります。（このことからもちろん $xi(-s) = \eta(s)$ でもあります。一方、$f(x, y)$ は

$$f(y, x) = \sqrt{4 - y^2 - x^2} = \sqrt{4 - x^2 - y^2} = f(x, y)$$

となって x と y を取り替えても変わりません。よって、

$$\varphi(-s) = f(\xi(-s), \eta(-s)) = f(\eta(s), \xi(s)) = f(\xi(s), \eta(s)) = \varphi(s)$$

となり $\varphi(s)$ は偶関数であることが分かります。つまり $z = \varphi(s)$ のグラフを z 軸を折り線として折りたたむとピッタリ重なるということです。このことは $s = 0$ における接線に対しても成り立つので、その接線は s 軸に平行です。よって $\varphi'(s) = 0$ となります。

注意. もちろん $\varphi(s)$ を直接微分することで $\varphi'(0) = 0$ を得ることができますが、「幾何学的なから求めよ」という条件を付けたのは、上の二つのような考察をしてもらうためでした。「きちっとした計算」と「幾何学的なイメージ」の両方をいつも心がけるようにすると、より理解が深まるのではないかと思います。★

(2) $\xi(s) = e^s$ ですので $\xi'(s) = e^s$ です。よって、$s = 0$ における $\xi(s)$ の 1 次近似 $\varpi(s)$ は

$$\varpi(s) = \xi(0) + \xi'(0)(s - 0) = 1 + s$$

となります。同様に、$\eta(s) = e^{-s}$ の 1 次近似式 $\rho(s)$ は、$\eta'(s) = -e^{-s}$ であることから

$$\rho(s) = \eta(0) + \eta'(0)(s - 0) = 1 - s$$

となります。一方、$f(x,y)$ の $(\xi(0),\eta(0))=(1,1)$ における接平面は、$z=f(x,y)$ のグラフが原点を中心とする球面であることから、$(1,1,f(1,1))=(1,1,\sqrt{2})$ を通り $\begin{bmatrix}1\\1\\\sqrt{2}\end{bmatrix}$ を法線ベクトルとする平面です。よって、点 (x,y,z) が接平面の点であるための必要十分条件は

$$0=\begin{bmatrix}1\\1\\\sqrt{2}\end{bmatrix}\bullet\begin{bmatrix}x-1\\y-1\\z-\sqrt{2}\end{bmatrix}=x-1+y-1+\sqrt{2}z-2$$

を満たすことです。よって、$f(x,y)$ の $(1,1)$ における 1 次近似 $p(x,y)$ は

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

となります。（もちろん、$f(x,y)$ の $(1,1)$ における二つの偏微分の値を計算して $p(x,y)$ を求めても結構です。

$$\begin{aligned}p(x,y)&=f(1,1)+f_x(1,1)(x-1)+f_y(1,1)(y-1)\\&=\sqrt{2}-\frac{1}{\sqrt{2}}(x-1)-\frac{1}{\sqrt{2}}(y-1)\end{aligned}$$

と、当然同じ結果になります。）$p(x,y)$ に $x=\varpi(s)$ と $y=\rho(s)$ を代入すると、

$$p(\varpi(s),\rho(s))=\sqrt{2}-\frac{1}{\sqrt{2}}(1+s-1)-\frac{1}{\sqrt{2}}(1-s-1)=\sqrt{2}$$

という値が $\sqrt{2}$ の定数関数になります。よって傾きは 0 です。

(3) $\varphi(s)$ は 2 変数関数に 1 変数関数を二つ入れた合成関数ですので、合成関数の微分公式は

$$\frac{d\varphi}{ds}(s)=\frac{\partial f}{\partial x}(\xi(s),\eta(s))\frac{d\xi}{ds}(s)+\frac{\partial f}{\partial y}(\xi(s),\eta(s))\frac{d\eta}{ds}(s)$$

となります。右辺の登場人物を具体的に計算すると、

$$\frac{\partial f}{\partial x}(x,y)=-\frac{x}{\sqrt{4-x^2-y^2}},\quad \frac{\partial f}{\partial y}(x,y)=-\frac{y}{\sqrt{4-x^2-y^2}}$$

および

$$\frac{d\xi}{ds}(s)=e^s,\qquad \frac{d\eta}{ds}(x)=-e^{-s}$$

となります。これらに $s=0,\ x=\xi(0)=1,\ y=\eta(0)=1$ を代入して合成関数の微分公式に戻してやると、

$$\frac{d\varphi}{ds}(0)=-\frac{1}{\sqrt{2}}e^0-\frac{1}{\sqrt{2}}(-e^0)=0$$

となります。

以上のように、幾何学的に「こうあるべき」という考察、1 次近似の合成、合成関数の微分公式すべての値が一致しました。 □

問題 33 の解答

(1) 今の場合に式(26) を具体的に書くとヤコビ行列で

$$\left[\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y}\right]=\left[\frac{\partial g}{\partial r},\ \frac{\partial g}{\partial \theta}\right]\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}^{-1}$$

となります。成分では

$$\begin{aligned}\frac{\partial f}{\partial x}&=\frac{\partial g}{\partial r}\cos\theta-\frac{1}{r}\frac{\partial g}{\partial \theta}\sin\theta\\ \frac{\partial f}{\partial y}&=\frac{\partial g}{\partial r}\sin\theta+\frac{1}{r}\frac{\partial g}{\partial \theta}\cos\theta\end{aligned}$$

です。 □

念のためにこの式の意味を言いますと、左辺の f_x や f_y は (x,y) の関数ですが、そこに $x=r\cos\theta, y=r\sin\theta$ を入れてやると右辺の関数になる、というわけです。

(2) $f_{xx}=(f_x)_x$ ですので、f_x に (1) で得た式を二回使うことにより、

$$\begin{aligned}f_{xx}&=\frac{\partial f_x}{\partial x}\\ &=\frac{\partial}{\partial r}\left(g_r\cos\theta-\frac{1}{r}g_\theta\sin\theta\right)\cos\theta-\frac{1}{r}\frac{\partial}{\partial\theta}\left(g_r\cos\theta-\frac{1}{r}g_\theta\sin\theta\right)\sin\theta\\ &=\left(g_{rr}\cos\theta+\frac{1}{r^2}g_\theta\sin\theta-\frac{1}{r}g_{r\theta}\sin\theta\right)\cos\theta\\ &\qquad-\frac{1}{r}\left(g_{\theta r}\cos\theta-g_r\sin\theta-\frac{1}{r}g_{\theta\theta}\sin\theta-\frac{1}{r}g_\theta\cos\theta\right)\sin\theta\\ &=g_{rr}\cos^2\theta-\frac{1}{r}g_{r\theta}2\sin\theta\cos\theta+\frac{1}{r^2}g_{\theta\theta}\sin^2\theta+\frac{1}{r}g_r\sin^2\theta+\frac{1}{r^2}g_\theta 2\sin\theta\cos\theta\end{aligned}$$

となります。同様に、

$$
\begin{aligned}
f_{yy} &= \frac{\partial}{\partial y} f_y \\
&= \frac{\partial}{\partial r}\left(g_r \sin\theta + \frac{1}{r} g_\theta \cos\theta\right)\sin\theta + \frac{1}{r}\frac{\partial}{\partial\theta}\left(g_r \sin\theta + \frac{1}{r} g_\theta \cos\theta\right)\cos\theta \\
&= \left(g_{rr}\sin\theta - \frac{1}{r^2} g_{\theta\theta}\cos\theta + \frac{1}{r} g_{r\theta}\cos\theta\right)\sin\theta \\
&\qquad + \frac{1}{r}\left(g_{\theta r}\sin\theta + g_r\cos\theta + \frac{1}{r} g_{\theta\theta}\cos\theta - \frac{1}{r} g_\theta \sin\theta\right)\cos\theta \\
&= g_{rr}\sin^2\theta + \frac{1}{r} g_{r\theta} 2\sin\theta\cos\theta + \frac{1}{r^2} g_{\theta\theta}\cos^2\theta + \frac{1}{r} g_r \cos^2\theta - \frac{1}{r^2} g_\theta 2\sin\theta\cos\theta
\end{aligned}
$$

です。この二つを足しあわせて、

$$
\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial\theta^2} + \frac{1}{r}\frac{\partial g}{\partial r}
$$

となります。 □

$f(x,y)$ は（微分さえできれば）何でも良かったので、

$$
\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}
$$

と書けます。これは、**ラプラス作用素 $\boldsymbol{\partial^2/\partial x^2 + \partial^2/\partial y^2}$ を極座標で表す**公式で、力学などでよく出会うものです。