2008 年度全学自由研究ゼミナール

多変数関数の微分：第 10 回

6 月 24 日 清野和彦

7 合成関数とその微分公式

7.1 合成関数の微分だけ特別扱いするのはなぜか

高校で 1 変数関数の微分を勉強したとき、まず微分や導関数を定義し、次に多項式や三角関数などのよく出会う関数の導関数を具体的に求め、最後に積、商、合成関数、逆関数の微分法を調べました。これらの「公式」が手に入ったことによって、具体的に微分できる関数がいくつかあるとき、それらを足し引きしてできる関数はもちろん、掛けたり割ったり合成したり逆関数を作ったりしても微分できることになり、「式一本」で書ける関数は何でも微分できるようになったわけです。このように、積や合成関数の微分公式があると、微分できる関数の世界が一気に広がります。

事情は多変数でも同じなはずです。前節までで多変数関数の微分を定義したので、次にすべきことは、基本的な多変数関数の偏導関数を求めることと、積や合成関数の微分公式を得ることでしょう。

ところが、偏微分は注目した変数以外は定数と見なすことによる 1 変数関数としての微分ですし、基本的な多変数関数（つまり、式で書ける多変数関数）は式で書ける 1 変数関数をいくつかそれぞれ別の変数を持つ関数として用意し、それらを足したり引いたり掛けたり割ったり合成したりしてできるものなので、基本的な多変数関数が具体的に与えられたら、それを偏微分することは 1 変数関数の微分の知識の範囲内で完全にできてしまうわけです。

例えば、$f(x,y)=xy\sin x\cos y$ とすると、x で偏微分するときは $y\cos y$ は定数扱い、y で偏微分するときは $x\sin x$ は定数扱いなので、

$$\frac{\partial f}{\partial x}(x,y)=y\cos y\frac{d}{dx}(x\sin x)=y\cos y(\sin x+x\cos x)$$
$$\frac{\partial f}{\partial y}(x,y)=x\sin x\frac{d}{dy}(y\cos y)=x\sin x(\cos y-y\sin y)$$

というふうに 1 変数関数の積の微分法で計算できてしまいます。商の微分も全く同様です。また、例えば $\sin z$ に $z=x^2y^3$ を合成した関数 $g(x,y)=\sin x^2y^3$ を考えても、やはり x で偏微分するときは y^3 は定数扱い（つまり x^2 の単なる係数)、y で偏微分するときは x^2 は定数扱い（つまり y^3 の単なる係数）なので、

$$\frac{\partial g}{\partial x}(x,y)=2xy^3\cos x^2y^3,\qquad \frac{\partial g}{\partial y}(x,y)=3x^2y^2\cos x^2y^3$$

と計算できます。なお、多変数関数には逆関数は存在しないので[1]、「多変数関数そのものについての逆関数の微分法」というものは存在しません。（多変数で逆関数に当たる概念は三つくらい後の節で出てくる予定です。）

多変数関数の微分には偏微分以外にも方向微分や 1 次近似（全微分可能ということ）があるではないかと思うかも知れませんが、基本的な関数はすべて C^1-級、つまり偏導関数が連続関数なので、これらの微分の概念はすべて同じものの別な姿だと思ってよいわけですから（前節を参照してください）、やはり 1 変数関数の微分の範囲内に収まってしまいます。

しかし、上の計算は 1 変数関数のときとはちょっとだけ違っている面があります。というのは、1 変数関数のときには、基本的な関数の導関数を求めるのとは**別に**積、商、合成関数の微分法を「**公式**」として用意したのでした。だから、例えば大学に来て初めてであった逆三角関数のようなものでも、高校のときに用意しておいた公式だけで微分を計算することができた（or る）わけです。ところが、上でお見せした偏微分の計算は、偏微分しようとする関数の中に現れている 1 変数関数が導関数を既に知っている関数だったからできたように見えます。つまり、多変数関数の偏微分そのものに対する「公式」にはなっていないわけです。具体的な関数の姿を使わない理論として偏微分の間の関係を考えるには、いわば抽象的な「公式」を求めておかなければなりません。

とは言え、積や商の微分公式は実はほとんど 1 変数のときのままです。実際、例えば $f(x,y)$ と $g(x,y)$ の y に定数 b を代入してできる x の 1 変数関数を $\varphi(x) = f(x,b)$, $\psi(x) = g(x,b)$ と書くことにすれば、$f(x,y)$ と $g(x,y)$ の積の (a,b) における x による偏微分は、偏微分の定義から

$$\begin{aligned}\frac{\partial(fg)}{\partial x}(a,b) &= \frac{d(\varphi\psi)}{dx}(a)\\ &= \frac{d\varphi}{dx}(x)\psi(a) + \varphi(a)\frac{d\psi}{dx}(a)\\ &= \frac{\partial f}{\partial x}(a,b)g(a,b) + f(a,b)\frac{\partial g}{\partial x}(a,b)\end{aligned}$$

となります。（二つ目の等号で 1 変数関数の積の微分法を使っています。）x での偏微分を下付添え字で f_x, g_x などと表し、「どこでの偏微分か」を表す (a,b) を省略して書けば、1 変数関数の積の微分法の式にもっと似た形の

$$(fg)_x = f_x g + f g_x$$

で表すことができます。商の微分も全く同様で、$1/f(x,y)$ の (a,b) における x に

[1] $z = f(x,y)$ の逆関数とは、$(x,y) = (g_1(z), g_2(z))$ であって $f(g_1(z), g_2(z)) = z$ と $g_1(f(x,y)) = x, g_2(f(x,y)) = y$ を満たすものということになりますが、$(x,y) = (g_1(z), g_2(z))$ というのは z をパラメタとするパラメタ**曲線**ですので、$f(x,y)$ の定義域（平面内で広がりを持つ部分）全体を値域とすることはできません。よって、多変数関数には逆関数は存在しません。

よる偏微分は、上の計算と同じ記号を流用すると、

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{1}{f}\right)(a,b) &= \frac{d}{dx}\left(\frac{1}{\varphi}\right)(a) \\ &= -\frac{1}{(\varphi(a))^2}\frac{d\varphi}{dx}(a) \\ &= -\frac{1}{(f(a,b))^2}\frac{\partial f}{\partial x}(a,b)\end{aligned}$$

となります。（二つ目の等号で 1 変数関数の商の微分法を使っています。）つまり、

$$\left(\frac{1}{f}\right)_x = -\frac{f_x}{f^2}$$

と、1 変数関数のときと全く同じ式になっているわけです。

というわけで、合成関数の微分公式を求めることだけが残るわけです。もちろん、これも積や商のように簡単に求まるならそれだけのための節を立てたりはしません。合成関数の微分を特別扱いする理由は

- 1 変数関数の場合より合成関数の概念そのものが捉えにくい。
- 証明が難しく、何をしているのかわかりにくい。
- 合成関数の微分公式をよく見ることで、多変数**写像**の微分のあるべき姿が見えてくる。

と言った、どれも重要なものばかりです。

7.2 合成関数の微分公式

まず、合成関数の微分公式を証明抜きで紹介します。多変数の合成関数には多くの変数と多くの関数が出てくるので、記号の使い方もここで合わせて決めてしまいましょう。読んでいて記号で混乱したらここに戻ってみて下さい。

1 変数関数が二つ、例えば $f(x), g(x)$ があったとき、もしも $g(x)$ の値域が $f(x)$ の定義域に含まれているなら、$g(x)$ を $f(x)$ の x のところに入れることができます。ただ、このように書くと x が二重の意味を持ってしまう（g の独立変数としての意味と、g の従属変数としての意味）ので、紛れがないように文字を決めましょう。普通「$y = g(x)$ を $f(y)$ に代入する」というふうに記号を決めると思いますが、このあと 2 変数関数を考えるときに文字が足りなくなってしまうので、次のように 1 変数のときからギリシャ文字を使わせてもらうことにします。

あとで 2 変数関数を考えるときに変数は 2 つ一組で必要になります。よく使う組は (x,y), (u,v), (s,t) などでしょう。このゼミでは (u,v) の組はベクトルの成分として使ってきたので、(s,t) を使うことにしましょう。変数が一つしかないとき

は普通 s より t を使うと思いますが、2 変数になったときの混乱を避けるために、1 変数のときは s を使うことにします。x に対応するギリシャ文字は ξ、y に対応するギリシャ文字は η なので、x に代入される関数を $\xi(s)$ や $\xi(s,t)$、y に代入される関数を $\eta(s)$ や $\eta(s,t)$ としましょう。ξ や η という文字になじみの薄い人も多いと思いますが、是非ここで慣れてください。「x に対応 x に対応」とか「y に対応 y に対応」と心の中で唱えながらつき合ってもらえればわりとすぐ慣れると思います。

7.2.1　1 変数関数に 1 変数関数を入れた場合

1 変数関数 $f(x)$ に 1 変数関数 $\xi(s)$ を合成するということは、イメージ的には図 1 のようになるでしょう。 これは「s に $f(\xi(s))$ を対応させる」という歴とした

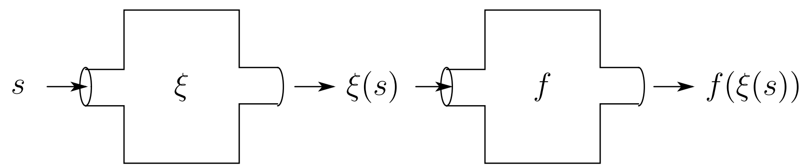

図 1: 「1 変数関数を二つ合成する」ということのイメージ

関数ですので、関数らしい記号を決めておくのがよいでしょう。「$f(\xi(s))$ でよいのでは?」と思うかも知れませんが、$f(\xi(s))$ という記号には「関数そのものを表す部分」がないのが問題です。関数 $f(x)$ は (x) を省いて f と書くことができる、つまり f が関数そのものを表しているように、$f(x)$ に $x=\xi(s)$ を代入してできる関数にも、このようにして定義された関数だということが一目で分かるような「関数そのものを表す記号」を決めたいわけです。$f(\xi(s))$ がダメな理由は大きな括弧があるからなので、それを省いてやったらどうか、と思うかも知れませんが、そうすると $f\xi(s)$ となって積と紛らわしくなってしまいます。そこで、f と ξ の間に「積ではなくて合成だよ」ということを表す印を入れてやればよいのではないかと思えてくるでしょう。積は「$\cdot$」で表すことが多いので、もう少し偉そうに「$\circ$」にするのが一般的です。つまり、

関数 $f\circ\xi$ を

$$f\circ\xi(s)=f(\xi(s))$$

で定義する

というわけです (図 2)。 普通関数の名前は f とか φ とか一文字で表すことが多いので違和感があるかも知れませんが、例えば sin の三文字で一つの関数の名前で

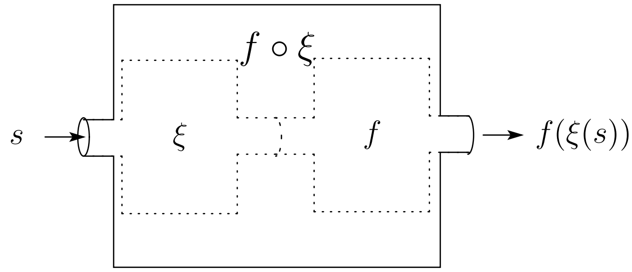

図 2: 出力口と入力口を貼り付けて全体を大きな箱に入れる

あるように $f\circ\xi$ の「三文字」で一つの関数の名前なのだ、もっといえば $f\circ\xi$ で「一文字」なのだと思って下さい。

f の独立変数と ξ の独立変数を別な文字で表したように、物理などの具体的な状況を考えるときには f の独立変数と ξ の独立変数は別な意味を持つ量を考えることが多いと思います。このような場合には、定義域（を含む $\mathbb{R}$）や値域（を含む $\mathbb{R}$）をはっきり書き表す「写像の記法」を使うと状況がよりはっきりするでしょう。

$$f\circ\xi\colon \mathbb{R}\xrightarrow{\quad\xi\quad}\mathbb{R}\xrightarrow{\quad f\quad}\mathbb{R}$$

となります。この記法やイメージ図 1 と $f\circ\xi$ や $f(\xi(s))$ という記号では f と ξ の順序が逆になっていることに注意してください。

この場合の合成関数の微分公式は高校のときに学んだとおり、

$$\frac{d(f\circ\xi)}{ds}(s)=\frac{df}{dx}(\xi(s))\frac{d\xi}{ds}(s) \qquad \text{すなわち} \qquad (f\circ\xi)'(s)=f'(\xi(s))\xi'(s)$$

です。変数 s を省いて「関数そのもの」の等式として書くと

$$\frac{d(f\circ\xi)}{ds}=\left(\frac{df}{dx}\circ\xi\right)\cdot\frac{d\xi}{ds} \qquad \text{すなわち} \qquad (f\circ\xi)'=(f'\circ\xi)\cdot\xi'$$

となります[2]。

注意. 上の書き表し方はかなり丁寧というかくどい書き方です。多くの場合 $x=\xi(s)$ を代入して

$$\frac{d(f\circ\xi)}{ds}(s)=\frac{df}{dx}(x)\frac{d\xi}{ds}(s) \qquad \text{あるいは} \qquad (f\circ\xi)'(s)=f'(x)\xi'(s)$$

と書くでしょう。さらに、s と x が $x=\xi(s)$ で結びついていることは分かり切っているとして省いてしまい、

$$\frac{d(f\circ\xi)}{ds}=\frac{df}{dx}\frac{d\xi}{ds} \qquad \text{あるいは} \qquad (f\circ\xi)'=f'\cdot\xi'$$

[2] $\cdot$ は掛け算です。(s) がある場合のように $\cdot$ なしでももちろんよいのですが、わかりにくいようなので (s) なしの方では付けることにしました。

と書くことの方が多いかも知れません。右側の書き方では、f は x の関数なのだから f' は x での微分、ξ は s の関数なのだから ξ' は s による微分しかあり得ないので誤解の余地はないわけです。そのことを逆手に取ると、左側の書き方ではどの変数による微分かをはっきり書いているのだから、$f \circ \xi$ のことを f と書いてしまっても誤解は起こりえないとも言えます。つまり f そのものは x でしか微分できないのだから、df/ds と書くことによって f と ξ の合成関数を微分していることまで表せていると考えるわけです。さらに 関数 ξ は x に代入されるわけですから、ξ と書かずに x と書いてしまってもこれまた誤解の余地はないはずです。文字の節約のためにも x に代入される関数を $x(s)$ と書く方が普通でしょう。というわけで、

$$\frac{df}{ds} = \frac{df}{dx}\frac{dx}{ds}$$

という、物理の本などでよく見かける書き方が得られるわけです。

この書き方は、df や dx をただの数だと思って普通に分数の掛け算をすると両辺が一致するので印象に残りやすいのですが、これから見るように、**多変数関数の合成関数の微分公式ではそうはなっていない**ので、かえって間違いの元になる可能性もあります。便利な記憶法にはあまり頼らず、次節以降で説明する証明を理解するようにしましょう。★

問題 28. $\varphi(s) = \sin(e^s)$ を微分せよ。

7.2.2　1 変数関数に 2 変数関数を入れた場合

次に 2 変数関数を合成することを考えたいのですが、「入れ物」の方の関数 f と中に入る方の関数 ξ（と η）の両方を一遍に 2 変数関数にすると混乱しやすいと思うので、徐々に変数の数を増やして行きましょう。

まず $f(x)$ に $x = \xi(s,t)$ を合成することを考えます。この場合は 1 変数関数に 1 変数関数を合成する場合と大差ありません。実際、記号は

$$f \circ \xi(s,t) = f(\xi(s,t))$$

というように全く同じで済むし、イメージ図も図 3 を真ん中でくっつけて図 4 になるだけで、1 変数関数に 1 変数関数を合成するときとの違いは、一番左側のインプット口が二つあるということだけです。

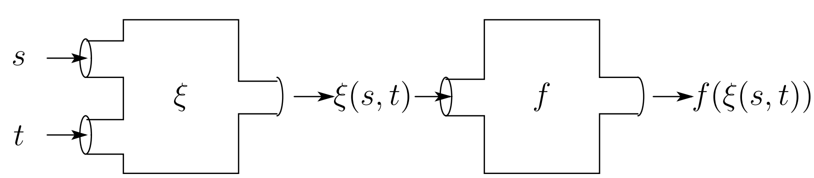

図 3: 2 変数関数の出力を 1 変数関数に入力する

写像の記法では、二つの変数 s,t を平面上の点の座標のように考えて

$$f \circ \xi \colon \mathbb{R}^2 \xrightarrow{\ \xi\ } \mathbb{R} \xrightarrow{\ f\ } \mathbb{R}$$

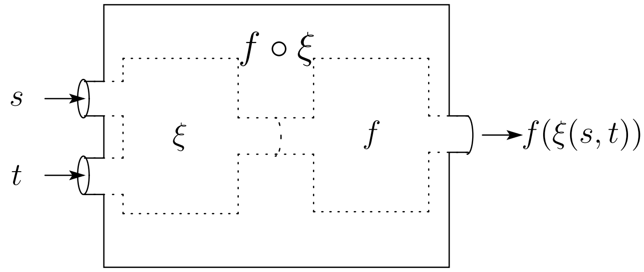

図 4: 出力口と入力口を貼り付けて全体を大きな箱に入れる

と書きます。

この場合の合成関数の微分法は次のようになります。

$$\frac{\partial(f\circ\xi)}{\partial s}(s,t)=\frac{df}{dx}(\xi(s,t))\frac{\partial\xi}{\partial s}(s,t),\qquad \frac{\partial(f\circ\xi)}{\partial t}(s,t)=\frac{df}{dx}(\xi(s,t))\frac{\partial\xi}{\partial t}(s,t)$$

すなわち

$$(f\circ\xi)_s(s,t)=f'(\xi(s,t))\xi_s(s,t),\qquad (f\circ\xi)_t(s,t)=f'(\xi(s,t))\xi_t(s,t)$$

です。関数そのものの等式として書くなら、

$$\frac{\partial(f\circ\xi)}{\partial s}=\frac{df}{dx}\circ\xi\cdot\frac{\partial\xi}{\partial s},\qquad \frac{\partial(f\circ\xi)}{\partial t}=\frac{df}{dx}\circ\xi\cdot\frac{\partial\xi}{\partial t}$$

すなわち

$$(f\circ\xi)_s=(f'\circ\xi)\cdot\xi_s,\qquad (f\circ\xi)_t=(f'\circ\xi)\cdot\xi_t$$

となります。f は 1 変数関数なので、その微分には df/dx や f' など 1 変数関数の微分の記号を使っています。

合成関数が 2 変数なので s による偏微分と t による偏微分の二つあって 1 変数関数同士のときより複雑に見えるかも知れませんが、一つ一つを見れば 1 変数関数同士のときの公式とほとんど同じであることがわかるでしょう。その理由は偏微分の定義を思い出してみればすぐわかるので、1 変数関数の合成関数の微分公式を認めることにしてここで証明をしてしまってもよいくらいなのですが、話が前後するのを避けるために、証明はすべて次節以降にまわすことにします。

注意. 1 変数関数同士のときの注意の最後に書いた記法を使ってみましょう。つまり、$f\circ\xi$ のことも f と書いてしまい、ξ のことは x と書いてしまうわけです。すると、例えば s での偏微分は

$$\frac{\partial f}{\partial s}=\frac{df}{dx}\frac{\partial x}{\partial s}$$

となります。これは d と ∂ の違いを無視すれば普通の分数の間の等式になっています。★

問題 29. $\varphi(s,t) = \sin(s^2 t)$ とする。
(1) この節で紹介した公式を使わずに、1 変数関数の合成関数の微分公式の範囲内で二つの偏導関数を計算せよ。（つまり、このゼミでこれまでやってきたように偏微分せよ。）
(2) $\varphi(s,t)$ を $f(x) = \sin x$ に $\xi(s,t) = s^2 t$ を合成した合成関数と見ることによりこの節で紹介した公式を使って二つの偏導関数を計算し、(1) で計算したものと一致していることを確認せよ。

7.2.3　2 変数関数に 1 変数関数を 2 つ入れた場合

f が 2 変数関数 $f(x,y)$ で、これに二つの 1 変数関数 $\xi(s), \eta(s)$ を合成する場合を考えましょう。この場合は今までとは違って少し複雑になります。イメージ図 5 を見て下さい。

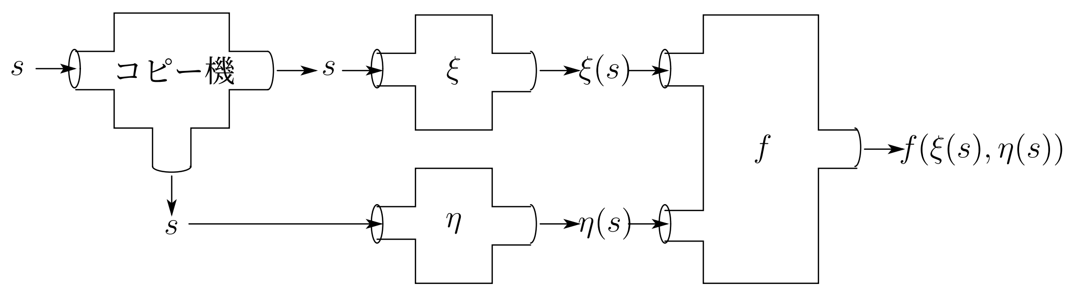

図 5: 2 変数関数に 1 変数関数を二つ合成するには「コピー機」が必要

注意. ここで、中に入れる関数を $\xi(s), \eta(t)$ と別な独立変数を持つ 1 変数関数にしてはいけません。「関数を合成する」ということの意味は、

元の独立変数（たち）を別な独立変数（たち）の従属変数と見なす

という、いわば「一般化された変数変換」なので、$f(\xi(s), \eta(t))$ という関数は

元の独立変数 x, y を新しい独立変数 s, t の従属変数と見なす

ということになってしまいます。つまり、この場合 $\xi(s)$ という関数は s の 1 変数関数ではなく、s と t を変数に持つ 2 変数関数だがたまたま t には依存していない、と考えなくてはならないということになるわけです。だから、$f(\xi(s), \eta(t))$ という合成関数はこのあとで紹介する「2 変数関数に 2 変数関数を合成する」という場合に入ることになります。★

さて、これまでと同じようにイメージ図 5 の出力口と入力口を貼り付けて一つの関数を作ってみましょう。すると、図 6 のようになります。 これまでは大きな箱（つまり新しい関数）の中には古い関数たちが入っているだけでしたが、この場合は**「コピー機」まで中に入ってしまっている**ところが違います。これまでの 2

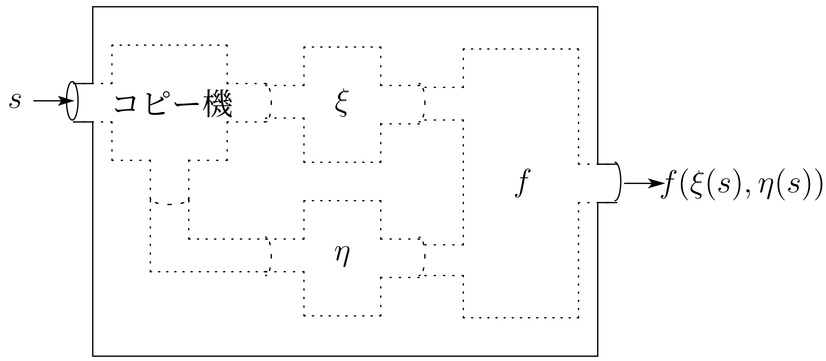

図 6: 大きな箱の中にコピー機まで入ってしまっているところが今までと違う

つの場合には、新しい独立変数 s（や t）の従属変数は x ただ一つだったのに、この場合には s の従属変数が x と y の二つあるところが大きな違いを生むのです。そのことが原因で $\circ$ を使った合成関数の「名付け」もうまくいきません。上手く名前を付けるには、関数（すなわち行き先が $\mathbb{R}$）をはみ出して、ξ と η を合わせて一つの写像（今の場合は行き先が $\mathbb{R}^2$）と見なさなければならないのです。

しかし、その見方は前節で「合成関数の微分を特別扱いする意味」を三つ挙げたうちの最後のものに深く関連しており、合成関数というものを見る視点を変えなければ理解できないので、あとで改めて説明することにし、この節では合成関数の微分公式を紹介するだけにします。

さて、$f\circ\xi$ のような記号がないので、

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

と記号を変えることにします。（φ は f に当たるギリシャ文字です[3]。）すると、

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

すなわち、

$$\varphi'(s)=f_x(\xi(s),\eta(s))\xi'(s)+f_y(\xi(s),\eta(s))\eta'(s)$$

となります。ξ と η は 1 変数関数なので $d\xi/ds$ や ξ' という 1 変数関数の微分の記号を使うことは前の場合と同じです。

既に指摘したように、関数の範疇に留まるならこの場合は $\circ$ を使った合成関数の書き方ができないので、ここから (s) を取り除いて関数そのものの間の等式を作ることはできませんが、$x=\xi(s), y=\eta(s)$ はわかっているものとしてしてしまえば、

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

[3]正確には physics や photograph の ‘ph’ に当たるのだと思います。

あるいは

$$\varphi' = f_x \cdot \xi' + f_y \cdot \eta'$$

と書くことができます。

合成関数の中にコピー機が入っていること、つまり x も y も s に依存していることから、s **で微分するだけなのに** x **による偏微分の項と** y **に偏微分の項の両方が出てきてしまう**のです。

> f が x だけの関数だった場合の 1 変数関数の合成関数の微分公式と f が y だけの関数だった場合の 1 変数関数の合成関数の微分公式の両方が足されてしまう

と見れば自然に見えてくるのではないでしょうか。

注意. このタイプの合成は力学でよく出会います。平面の各点 (x, y) にある「量」$f(x, y)$ が与えられている一方、その平面上を時刻 t に $(\xi(t), \eta(t))$ という場所にいるように運動している粒子があったとき、時刻 t に粒子の居場所に対応する「量」を対応させることで関数 $\varphi(t)$ ができるわけです。例によって物理では φ という新しい文字を避けて f と書き、ξ と η のことも x, y と書いてしまうので、

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

となります。これでも誤解の余地はないというわけです。さらに、時刻による微分は上付きの点で表すことが多いので、

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

と書いたりします。★

問題 30. $\varphi(s) = e^s \sin s$ とする。
(1) s の 1 変数関数として普通に微分せよ。
(2) $\varphi(s)$ を $f(x, y) = xy$ に $\xi(s) = e^s$ と $\eta(s) = \sin s$ を合成した合成関数と見ることによりこの節で紹介した公式を使って導関数を計算し、(1) で計算したものと一致していることを確認せよ。

7.2.4 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 1 変数関数を入れた場合」と「1 変数関数に 2 変数関数を入れた場合」で実質的な違いがなかったのと同じ理由で、「2 変数関数に 1 変数関数を 2 つ入れた場合」と今から述べる「2 変数関数に 2 変数関数を 2 つ入れ場合」にも実質的な違いはありません。新しい変数が一つ増えたのでもう一つコピー機が必要になり、合成関数の中にはコピー機が 2 台入っていることになりますが、例えば s という変数で偏微分する場合には、s の通るコピー機はそのうちの 1 台だけなので

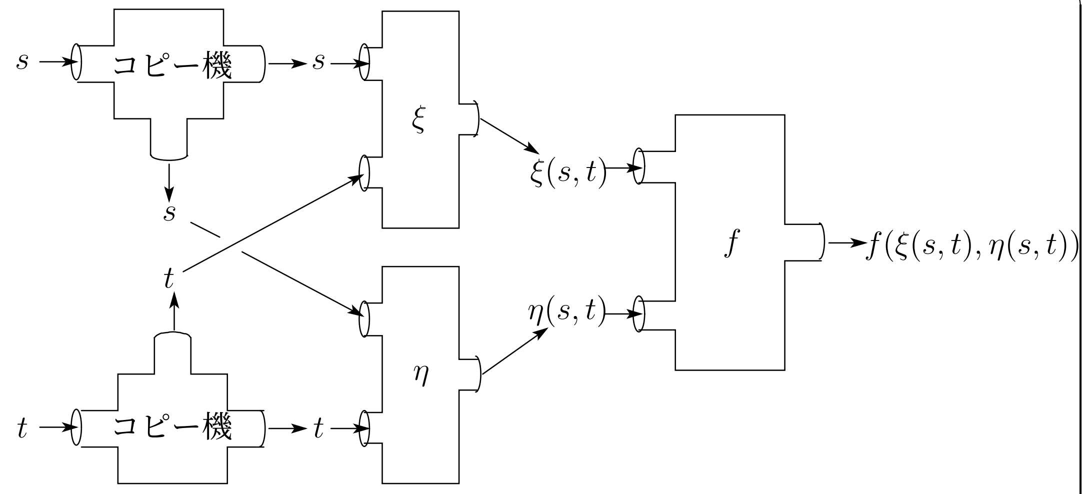

図 7: コピー機が 2 台必要

「2 変数関数に 1 変数関数を 2 つ入れた場合」と同じになるわけです。イメージ図は図 7 と図 8 です。

「2 変数関数に 1 変数関数を 2 つ入れた場合」と同様、写像を使わず関数の範疇だけで考えるなら f と ξ と η と $\circ$ という記号で合成関数を表すことはできないので、φ を使って

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

と表すことにします。すると、

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

すなわち、

$$\varphi_s(s,t) = f_x(\xi(s,t), \eta(s,t))\xi_s(s,t) + f_y(\xi(s,t), \eta(s,t))\eta_s(s,t)$$
$$\varphi_t(s,t) = f_x(\xi(s,t), \eta(s,t))\xi_t(s,t) + f_y(\xi(s,t), \eta(s,t))\eta_t(s,t)$$

となります。$x = \xi(s), y = \eta(s)$ はわかっているものとしてして省略した書き方をすると、

$$\frac{\partial \varphi}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial \xi}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial \eta}{\partial s}$$
$$\frac{\partial \varphi}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial \xi}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial \eta}{\partial t}$$

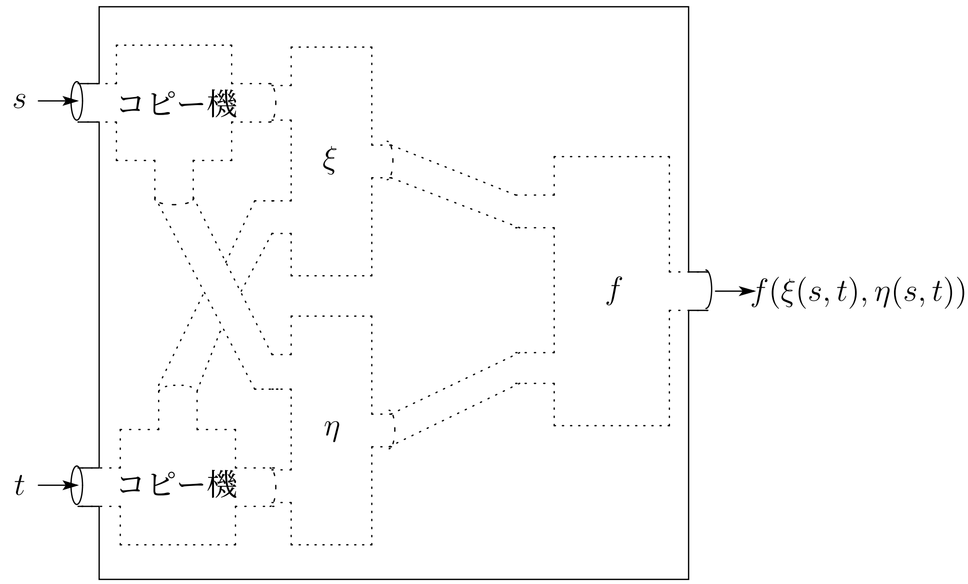

図 8: コピー機が 2 台入っているが、それぞれの変数の通るコピー機は 1 台だけ

あるいは

$$\varphi_s = f_x \cdot \xi_s + f_y \cdot \eta_s$$
$$\varphi_t = f_x \cdot \xi_t + f_y \cdot \eta_t$$

となります。

ゴチャゴチャしてわかりにくく見えるかも知れません。しかし、考えてみれば偏微分というのは微分しない変数のことは定数と考えるのですから、s で偏微分する場合 t は定数扱いなので公式の中に t による偏微分は出てくるはずがありません。一方、「2 変数関数に 1 変数関数を 2 つ入れる場合」と同様、s という変数は x と y の両方を通じて f にかかわっているのですから、s での偏微分には x による偏微分と y による偏微分の両方が入っていなければおかしいということになります。そう思ってから公式を見直すと、特にゴチャゴチャしているようには見えなくなるのではないかと思います。

注意. このような f と s,t との関わり方を考えると、次のような方法で公式を書き下せば間違いにくいのではないかと思います。

s で偏微分する場合で考えてみましょう。

まず一つ目の方法として、

1 変数関数の合成関数の微分公式を繰り返す

と考えてみます。まず f が x だけの関数であるかのように考えて

$$\frac{\partial f}{\partial x}\frac{\partial \xi}{\partial s}$$

と 1 変数関数の微分公式で d を ∂ に変えたものを書きます。次に、f を y だけの関数であるかのように考えて、このとなりに

$$\frac{\partial f}{\partial y}\frac{\partial \eta}{\partial s}$$

を書きます。そして最後に真ん中に $+$ を書いて

$$\frac{\partial f}{\partial x}\frac{\partial \xi}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial \eta}{\partial s}$$

とするわけです。

もう一つの方法として、

> f のすべての変数が s に依存しているのだから、f の偏微分はすべて出てくる

という方に着目してみましょう。この場合にはまず f の偏微分を

$$\frac{\partial f}{\partial x} \qquad \frac{\partial f}{\partial y}$$

と間をあけて並べて書いてしまいます。そして、x による偏微分には x に入っている関数 ξ を偏微分したい変数 s で偏微分したものを掛け、y による偏微分には y に入っている関数 η を s で偏微分したものを掛けて足し合わせるわけです。

$$\frac{\partial f}{\partial x}\frac{\partial \xi}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial \eta}{\partial s}$$

となります。

まだ証明していないのだから納得はできないかも知れませんが、証明の前にこのように何度も書き下してみて何が起きているのか自分なりの手応えを掴んでおくことが大切だと思い、少ししつこく書いてみました。★

問題 31. $\varphi(s,t)=\bigl(\sin(st)\bigr)\bigl(\cos(s^2+t^2)\bigr)$ とする。
(1) この節で紹介した公式を使わずに、1 変数関数の合成関数の微分公式の範囲内で二つの偏導関数を計算せよ。(つまり、このゼミでこれまでやってきたように偏微分せよ。)
(2) $\varphi(s,t)$ を $f(x,y)=(\sin x)(\cos y)$ に $\xi(s,t)=st$ と $\eta(s,t)=s^2+t^2$ を合成した合成関数と見ることによりこの節で紹介した公式を使って二つの偏導関数を計算し、(1) で計算したものと一致していることを確認せよ。

7.2.5 一般の場合

このゼミでは原則として 2 変数の場合しか扱いませんが、ここまでの 4 つの場合でどのように公式が変化したかをよく見てもらえれば、f が m 変数関数で、そこに m 個の n 変数関数を入れた場合の微分公式も正しく予想できるだろうと思うので、ここに答を書いておきましょう。

$f(x_1, x_2, \ldots, x_m)$ に $x_i = \xi_i(s_1, s_2, \ldots, s_n)$ を入れてできる $s_1, s_2, \ldots, s_n$ の関数を φ とします。

$$\varphi(s_1, s_2, \ldots, s_n) = f(\xi_1(s_1, s_2, \ldots, s_n), \xi_2(s_1, s_2, \ldots, s_n), \ldots, \xi_n(s_1, \ldots, s_n))$$

です。これの s_i による偏微分は次のようになります。ただし、長くなるので「どこでの微分か」を省いた書き方で書きます。

$$\frac{\partial \varphi}{\partial s_i} = \frac{\partial f}{\partial x_1}\frac{\partial \xi_1}{\partial s_i} + \frac{\partial f}{\partial x_2}\frac{\partial \xi_2}{\partial s_i} + \cdots + \frac{\partial f}{\partial x_m}\frac{\partial \xi_m}{\partial s_i} = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}\frac{\partial \xi_j}{\partial s_i} \tag{1}$$

です。思った通りでしたか?

7.3 行列を使った記法

「一般の場合」の微分公式(1) のように $\sum$ を使ってまとめられた式を見ると、何かを思い出しませんか? そう、**行列とベクトルの掛け算**です。実際、

$$\varphi_i = \frac{\partial \varphi}{\partial s_i}, \qquad f_j = \frac{\partial f}{\partial x_j}, \qquad x_{ij} = \frac{\partial x_j}{\partial s_i}$$

とおくと、式(1) は

$$\begin{aligned}
\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{bmatrix} &= \begin{bmatrix} \sum_{j=1}^{m} f_j x_{1j} \\ \sum_{j=1}^{m} f_j x_{2j} \\ \vdots \\ \sum_{j=1}^{m} f_j x_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} x_{1j} f_j \\ \sum_{j=1}^{m} x_{2j} f_j \\ \vdots \\ \sum_{j=1}^{m} x_{mj} f_j \end{bmatrix} \\
&= \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}
\end{aligned}$$

とまとめられます。

このまとめ方でもちろんよいのですが、第 5.2 節「全微分可能性と方向微分・偏微分」で $f(x, y)$ の二つの偏微分を**理由があって**縦に並べずに**横に**並べたので、それに合わせておいた方がよいでしょう。つまり、上のまとめ方で行と列を取り替えるわけです。

$$[\varphi_1, \varphi_2, \cdots, \varphi_n] = [f_1, f_2, \ldots, f_m] \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix}$$

とするわけです。第 5.2 節で紹介した記号では

$$d\varphi = [\varphi_1, \varphi_2, \cdots, \varphi_n], \qquad df = [f_1, f_2, \cdots, f_m]$$

となります。合成関数の微分公式は

$$d\varphi = df \begin{bmatrix} \frac{\partial \xi_1}{\partial s_1} & \frac{\partial \xi_1}{\partial s_2} & \cdots & \frac{\partial \xi_1}{\partial s_m} \\ \frac{\partial \xi_2}{\partial s_1} & \frac{\partial \xi_2}{\partial s_2} & \cdots & \frac{\partial \xi_2}{\partial s_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \xi_n}{\partial s_1} & \frac{\partial \xi_n}{\partial s_2} & \cdots & \frac{\partial \xi_n}{\partial s_m} \end{bmatrix}$$

というふうに

> 横ベクトル df に「中に入れた方の関数たち」の偏微分で作った行列を
> 右から掛けると $d\varphi$ になる

という形でまとめることができます。

1 変数と 2 変数の場合で具体的に書いてみましょう。

$\varphi(s) = f \circ \xi(s) = f(\xi(s))$ のときは

$$\frac{d\varphi}{ds} = \frac{df}{dx}\frac{d\xi}{ds}$$

ですが、それぞれを成分が一つしかない横ベクトルと 1 行 1 列の行列と思えば

$$\left[\frac{d\varphi}{ds}\right] = \left[\frac{df}{dx}\right]\left[\frac{d\xi}{ds}\right]$$

というふうに上で説明した行列とベクトルの掛け算の形になります。

$\varphi(s,t) = f(\xi(s,t))$ のときは

$$\left[\frac{\partial \varphi}{\partial s}, \frac{\partial \varphi}{\partial t}\right] = \left[\frac{df}{dx}\right]\left[\frac{\partial \xi}{\partial s}, \frac{\partial \xi}{\partial t}\right]$$

です。

$\varphi(s) = f(\xi(s), \eta(s))$ のときは

$$\left[\frac{d\varphi}{ds}\right] = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]\begin{bmatrix} \frac{d\xi}{ds} \\ \frac{d\eta}{ds} \end{bmatrix}$$

となります。

最後に $\varphi(s,t) = f(\xi(s,t), \eta(s,t))$ のときは

$$\left[\frac{\partial \varphi}{\partial s}, \frac{\partial \varphi}{\partial t}\right] = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]\begin{bmatrix} \frac{\partial \xi}{\partial s} & \frac{\partial \xi}{\partial t} \\ \frac{\partial \eta}{\partial s} & \frac{\partial \eta}{\partial t} \end{bmatrix}$$

となります。一貫して

関数たちは縦に、変数たちは横に並べる

という形になっているのです。

このまとめ方はいかにもご都合主義、上手く利用できたから利用したに過ぎないように見えるでしょう。ところが、このようにまとめてみることで、あるいはこのようにまとめることができるということの意味を積極的に探ってみることで、「微分とは何か」ということが明らかになるのです。ただし、その話はこのゼミのレベルを超えていますので、扱わないことにします。

7.4 注意:「連鎖律」という用語について

合成関数の微分公式のことを「**連鎖律**」と呼ぶことが多いのですが、実は、教科書によって(つまり人によって)この言葉で指している公式が違っているのです。そのせいか岩波の数学辞典には連鎖律という言葉が載っていません。

連鎖律は英語で言うと chain rule となります。(そのまんまですが。)「連鎖」とか「chain」という言葉の意味するところは「連鎖反応」などと同じだと思います。とすると、上で紹介した微分公式(1) を連鎖律と呼ぶ人の心は、

> s で偏微分しただけなのに、f のすべての変数が s とつながっているために、f のすべての偏微分が出てきてしまう

というところに「連鎖」の気持ちを感じているのだと思います。だから、このように連鎖律という言葉を使う人は 1 変数関数の合成関数の微分公式のことは連鎖律とは呼ばないようです。

一方、1 変数関数の微分公式

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

は

> 全体を微分することが一つ一つの関数の微分を引き起こしている

と見ることができます。二つの関数の合成だとわかりにくいですが、例えば、

$$(f_4 \circ f_3 \circ f_2 \circ f_1)' = (f_4' \circ f_3 \circ f_2 \circ f_1) \cdot (f_3' \circ f_2 \circ f_1) \cdot (f_2' \circ f_1) \cdot f_1'$$

と関数を増やして書いてみると、いかにも「連鎖反応」が起きているふうに見えないでしょうか? これは 1 変数関数の場合だけの公式ですが、前節でのように行列を使ったまとめ方を微分だと見なすと、この公式は多変数でもそのまま成り立つことが結論されます。この形まで話を整理してはじめて「連鎖律」と呼ぶのだと考える人もいます。その場合は 1 変数関数の合成関数の微分公式も連鎖律と呼ぶことになるでしょう。

おそらく前者の使い方の方が普通だと思いますが、後者の使い方をする人も結構いますし、そういう教科書もかなりあります。そういうわけなので、このプリントでは連鎖律という言葉は使わないことにしました。

問題 28 の解答

$\varphi(s)$ は $f(x) = \sin x$ に $\xi(s) = e^s$ を合成した関数ですので、

$$\varphi'(s) = f'(\xi(s)) \cdot \xi'(s) = \cos(e^s) \cdot e^s = e^s \cos(e^s)$$

となります。 $\square$

問題 29 の解答

(1) s での偏微分は t を定数と見て s だけを変数とする 1 変数関数として微分することでした。よって、1 変数関数の合成関数の微分公式により、

$$\frac{\partial \varphi}{\partial s}(s,t) = \cos(s^2 t) \cdot 2st = 2st \cos(s^2 t)$$

となります。同様に、s を定数と見て t の関数として微分することで、

$$\frac{\partial \varphi}{\partial t}(s,t) = \cos(s^2 t) \cdot s^2 = s^2 \cos(s^2 t)$$

となります。
(2) $\varphi(s,t) = f \circ \xi(s,t)$, $f(x) = \sin x$, $\xi(s,t) = s^2 t$ ですので、

$$\frac{\partial \varphi}{\partial s}(s,t) = \frac{df}{dx}(\xi(s,t)) \cdot \frac{\partial \xi}{\partial s}(s,t) = \cos(s^2 t) \cdot 2st = 2st \cos(s^2 t)$$

および、

$$\frac{\partial \varphi}{\partial t}(s,t) = \frac{df}{dx}(\xi(s,t)) \cdot \frac{\partial \xi}{\partial t}(s,t) = \cos(s^2 t) \cdot s^2 = s^2 \cos(s^2 t)$$

となります。このように、「1 変数関数に 2 変数関数を合成した場合」は計算結果のみならず計算過程まで (1) と全く同じになります。 $\square$

問題 30 の解答

(1) 積の微分法により、

$$\varphi'(s) = (e^s)' \sin s + e^s (\sin s)' = e^s \sin s + e^s \cos s = e^s (\sin s + \cos s)$$

となります。
(2) $\varphi(s) = f(\xi(s), \eta(s))$, $f(x,y) = xy$, $\xi(s) = e^s$, $\eta(s) = \sin s$ ですので、

$$\begin{aligned}\frac{d\varphi}{ds}(s) &= \frac{\partial f}{\partial x}(\xi(s), \eta(s)) \frac{d\xi}{ds}(s) + \frac{\partial f}{\partial y}(\xi(s), \eta(s)) \frac{d\eta}{ds}(s) \\ &= \eta(s)(e^s)' + \xi(s)(\sin s)' = (\sin s) \cdot e^s + e^s \cdot \cos s = e^s(\sin s + \cos s)\end{aligned}$$

となり、(1) の結果と一致します。 $\square$

問題 31 の解答

(1) t を定数と見なして s の 1 変数関数として微分すると、1 変数関数の積の微分公式と合成関数の微分公式により、

$$\begin{aligned}\frac{\partial\varphi}{\partial s}(s,t) &= \Big(\big(\cos(st)\big)t\Big)\,\big(\cos(s^2+t^2)\big) + \big(\sin(st)\big)\,\Big(\big(-\sin(s^2+t^2)\big)(2s)\Big)\\ &= t\big(\cos(st)\big)\big(\cos(s^2+t^2)\big) - 2s\big(\sin(st)\big)\big(\sin(s^2+t^2)\big)\end{aligned}$$

となります。同様に、s を定数と見なして t の 1 変数関数として微分すると、

$$\begin{aligned}\frac{\partial\varphi}{\partial t}(s,t) &= \Big(\big(\cos(st)\big)s\Big)\,\big(\cos(s^2+t^2)\big) + \big(\sin(st)\big)\,\Big(\big(-\sin(s^2+t^2)\big)(2t)\Big)\\ &= s\big(\cos(st)\big)\big(\cos(s^2+t^2)\big) - 2t\big(\sin(st)\big)\big(\sin(s^2+t^2)\big)\end{aligned}$$

となります。
(2) s による偏微分についての合成関数の微分公式により、

$$\begin{aligned}\frac{\partial\varphi}{\partial s}(s,t) &= \frac{\partial f}{\partial x}(\xi(s,t),\eta(s,t))\frac{\partial\xi}{\partial s}(s,t) + \frac{\partial f}{\partial y}(\xi(s,t),\eta(s,t))\frac{\partial\eta}{\partial s}(s,t)\\ &= \big(\cos(st)\big)\big(\cos(s^2+t^2)\big)t + \big(\sin(st)\big)\big(-\sin(s^2+t^2)\big)(2s)\\ &= t\big(\cos(st)\big)\big(\cos(s^2+t^2)\big) - 2s\big(\sin(st)\big)\big(\sin(s^2+t^2)\big)\end{aligned}$$

となります。同様に、t による偏微分についての合成関数の微分公式により、

$$\begin{aligned}\frac{\partial\varphi}{\partial t}(s,t) &= \frac{\partial f}{\partial x}(\xi(s,t),\eta(s,t))\frac{\partial\xi}{\partial t}(s,t) + \frac{\partial f}{\partial y}(\xi(s,t),\eta(s,t))\frac{\partial\eta}{\partial t}(s,t)\\ &= \big(\cos(st)\big)\big(\cos(s^2+t^2)\big)s + \big(\sin(st)\big)\big(-\sin(s^2+t^2)\big)(2t)\\ &= s\big(\cos(st)\big)\big(\cos(s^2+t^2)\big) - 2t\big(\sin(st)\big)\big(\sin(s^2+t^2)\big)\end{aligned}$$

となります。これらは (1) の計算結果と一致しています。 □

「違う計算なのに答が同じになって不思議」と感じたでしょうか、それとも「少し順番が違うだけで結局同じ計算をしているのだから一致して当然」と感じたでしょうか。もしも不思議と感じたなら、(1) の計算の途中と (2) の計算の途中をよく見比べて対応関係を見つけようとしてみて下さい。(1) で個別的に行った計算が実は (2) のように式の具体的な形によらない手順になっているのだということをキチンと証明しようというのが次の節の目標です。