多変数の微分積分学1 第9回

桂田 祐史

2013年6月17日

この授業用のWWWページは
http://www.math.meiji.ac.jp/~mk/lecture/tahensuu1-2013/

1 今日は…

風邪引いて2コマ続けてやるとダウンしそうな予感がしたので、1コマで我慢することにした。

出欠は紙で取る。

問8の解説はプリント配布で勘弁してもらう。

問8は、公式に代入するだけなので、見方によっては安直な問題である(数学的理解をあまり要求していない、採点は楽で、先生によっては「学生を甘やかしている」というかもしれない)。一方で、過去の試験の採点の経験からすると、出来の悪いところである。

少し脱線するけれど、某大学では、演習時間の半分は、ノート整理で、演習はその整理したノートだけを見て解くようにさせているとか。なかなか面白いアイディアだ。ノート取れていない部分があるとか、散逸させているとか、書き間違えているとか、そもそも一度も見返してないとか。なるべく記憶の新しいうちに自分のノートを読み返すこと。

今日は定理 7.1 の証明から。

7 合成関数の微分法

非常に重要であるが、一つの難所である。簡単のようでいて難しいところがある。

- 公式そのものは、1変数の場合の素直な拡張とみなせるので、覚えるのは簡単である。
- 具体的な合成関数を微分するのは簡単、というか、ひょっとすると不要なのかもしれない。使わずに解ける問題が実は多い。(要は、多変数関数でも、偏微分を計算すれば良いので、結局は1変数関数の問題になってしまう場合が多い。)
- 具体的な関数でない場合にとても重要なものが多く(例えば微分方程式を変数変換する)、しかも計算が非常に面倒というのが少なくない。
- 具体的な計算問題のドリルをこなしてマスターする、という学校数学の勉強法は通用しない。

7.1 定理と証明

1 変数の場合

$y = f(x)$, $z = g(y)$ がいずれも微分可能であれば、合成関数 $z = g(f(x))$ も微分可能で、

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

つまり

$$(g \circ f)'(a) = g'(b) f'(a) \quad (\text{ただし } b = f(a))$$

である。

これは多変数の場合にも、自然に拡張できる。右辺をヤコビ行列の積とみなせばよい。

定理 7.1 (合成関数の微分法, chain rule (連鎖律)) Ω, D はそれぞれ $\mathbf{R}^n$, $\mathbf{R}^m$ の開集合で、$a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$, $g\colon D \to \mathbf{R}^\ell$, $f(\Omega) \subset D$, $b = f(a)$, f は a で微分可能、g は b で微分可能ならば、$g \circ f$ は a で微分可能で、

$$(g \circ f)'(a) = g'(b) f'(a).$$

$y = f(x)$, $z = g(y)$ と書けば、上式の (i, j) 成分は

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \frac{\partial y_k}{\partial x_j} \quad (1 \le i \le \ell,\ 1 \le j \le n).$$

証明の前に、1 次関数だったら当たり前、ということを見よう。$f(x) = Ax + b$, $g(y) = Cy + d$ とすると、

$$g \circ f(x) = g(f(x)) = g(Ax + b) = C(Ax + b) + d = CAx + (Cb + d).$$

これから、

$$(g \circ f)'(x) = CA = g'(y) f'(x).$$

余談 7.1 (証明もどき) 高校数学の某教科書には、合成関数の微分法に次のような「証明」がついていた。

「証明」

$$\frac{dz}{dx} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \left(\frac{\Delta z}{\Delta y} \frac{\Delta y}{\Delta x} \right).$$

z は y の関数として微分可能なので、連続であるから、$\Delta x \to 0$ のとき、$\Delta y \to 0$ である。ゆえに

$$\frac{dz}{dx} = \lim_{\Delta y \to 0} \frac{\Delta z}{\Delta y} \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{dz}{dy} \frac{dy}{dx}.$$

しかし $\Delta x \neq 0$ であっても $\Delta y \neq 0$ とは限らないので、

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \frac{\Delta y}{\Delta x}$$

は一般に成り立つ式ではない。つまり上の「証明」には大穴がある。こういう杜撰なことを許せば、上の定理の証明は簡単に出来るのだが…以下に紹介する証明は、補助関数の導入によって、その問題をていねいに回避するものであると言える。■

証明　f が a で全微分可能であるから、

(1) $$\varepsilon_1(x) := f(x) - f(a) - f'(a)(x-a) \quad (x \in \Omega)$$

とおくと

(2) $$\lim_{x\to a} \frac{\varepsilon_1(x)}{\|x-a\|} = 0.$$

また g が $b = f(a)$ で全微分可能であるから、

(3) $$\varepsilon_2(y) := g(y) - g(b) - g'(b)(y-b) \quad (y \in D)$$

とおくと

(4) $$\lim_{y\to b} \frac{\varepsilon_2(y)}{\|y-b\|} = 0.$$

(3) から得られる

$$g(y) - g(b) = g'(b)(y-b) + \varepsilon_2(y) \quad (y \in D)$$

に $y = f(x)$ を代入して ($b = f(a)$ に注意して)

$$\begin{aligned}(g\circ f)(x) - (g\circ f)(a) &= g(f(x)) - g(f(a)) \\ &= g'(b)(f(x) - f(a)) + \varepsilon_2(f(x)) \quad (x \in \Omega).\end{aligned}$$

この式の右辺に、(1) から得られる

$$f(x) - f(a) = f'(a)(x-a) + \varepsilon_1(x)$$

を代入して

$$\begin{aligned}(g\circ f)(x) - (g\circ f)(a) &= g'(b)\left[f'(a)(x-a) + \varepsilon_1(x)\right] + \varepsilon_2(f(x)) \\ &= g'(b)f'(a)(x-a) + g'(b)\varepsilon_1(x) + \varepsilon_2(f(x)).\end{aligned}$$

ゆえに

(5) $$\lim_{x\to a} \frac{g'(b)\varepsilon_1(x)}{\|x-a\|} = 0$$

と

(6) $$\lim_{x\to a} \frac{\varepsilon_2(f(x))}{\|x-a\|} = 0$$

を証明できれば、

$$\lim_{x\to a} \frac{g'(b)\varepsilon_1(x) + \varepsilon_2(f(x))}{\|x-a\|} = 0$$

となるので、$g\circ f$ が a で全微分可能で $(g\circ f)'(a) = g'(b)f'(a)$ であることが結論できる。

(5) は、(2) から

$$\left\|\frac{g'(b)\varepsilon_1(x)}{\|x-a\|}\right\| \le \|g'(b)\| \frac{\|\varepsilon_1(x)\|}{\|x-a\|} \to 0 \quad (x \to a)$$

が成り立つので明らかである。

(6) を証明するために、補助関数 M を導入する。

$$M(y) := \begin{cases} \dfrac{\varepsilon_2(y)}{\|y-b\|} & (y \in D \setminus \{b\}) \\ 0 & (y = b) \end{cases}$$

とおくと ($\varepsilon_2(b) = 0$ に注意して)

$$M\colon D \to \mathbf{R}^m, \quad \lim_{y \to b} M(y) = 0, \quad \varepsilon_2(y) = \|y-b\|\, M(y) \quad (y \in D).$$

特に $\varepsilon_2\left(f(x)\right) = \|f(x)-b\|\, M(f(x))$ である。さて、

$$f(x) - b = f(x) - f(a) = f'(a)(x-a) + \varepsilon_1(x)$$

より

$$\|f(x)-b\| = \|f'(a)(x-a) + \varepsilon_1(x)\| \le \|f'(a)\|\, \|x-a\| + \|\varepsilon_1(x)\|.$$

ゆえに

$$\frac{\|\varepsilon_2(f(x))\|}{\|x-a\|} = \frac{\|f(x)-b\|\, \|M\left(f(x)\right)\|}{\|x-a\|} \le \left(\|f'(a)\| + \frac{\|\varepsilon_1(x)\|}{\|x-a\|}\right) \|M(f(x))\|$$
$$\to (\|f'(a)\| + 0) \cdot 0 = 0 \quad (x \to a).$$

ゆえに (6) が示された。■

2 点ほど指摘しておく。

- 行列のノルムを導入しておいたのが、見通しを良くすることに役立っている。
- 上で補助関数 M を導入する理由は、
  $$\frac{\|\varepsilon_2(f(x))\|}{\|x-a\|} = \frac{\|f(x)-b\|}{\|x-a\|} \cdot \frac{\|\varepsilon_2(f(x))\|}{\|f(x)-b\|} \le C \frac{\|\varepsilon_2(f(x))\|}{\|f(x)-b\|} \to C \cdot 0 = 0 \quad (x \to a)$$
  という素朴な「証明」の穴をふさぐためである。$x \ne a$ であっても、$f(x) \ne b$ であるとは限らないので、分母 $\|f(x)-b\|$ が 0 となりうることに注意する。

7.2 例

7.2.1 曲線に沿った微分

例 7.2 (chain rule の簡単な例: 曲線に沿った微分) (1) $z = xy$, $x = \varphi(t)$, $y = \psi(t)$ とするとき、

$$\frac{dz}{dt} = \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} = y\frac{dx}{dt} + x\frac{dy}{dt} = \psi(t)\varphi'(t) + \varphi(t)\psi'(t).$$

別解として

$$\frac{dz}{dt} = \frac{d}{dt}\left(\varphi(t)\psi(t)\right) = \varphi'(t)\psi(t) + \varphi(t)\psi'(t).$$

(こうやってしまうと、1 変数関数についての定理だけで済む。)

(2) $f(x,y) = \sqrt{x^2+y^2}$, $x = \varphi(t)$, $y = \psi(t)$ とするとき、

$$\frac{d}{dt} f\left(\varphi(t), \psi(t)\right) = f_x x_t + f_y y_t = \frac{x}{\sqrt{x^2+y^2}}\varphi'(t) + \frac{y}{\sqrt{x^2+y^2}}\psi'(t) = \frac{\varphi(t)\varphi'(t) + \psi(t)\psi'(t)}{\sqrt{\varphi(t)^2 + \psi(t)^2}}.$$

これも別解として

$$\frac{d}{dt} f\left(\varphi(t), \psi(t)\right) = \frac{d}{dt}\sqrt{\varphi(t)^2 + \psi(t)^2} = \frac{\dfrac{d}{dt}\left(\varphi(t)^2 + \psi(t)^2\right)}{2\sqrt{\varphi(t)^2 + \psi(t)^2}} = \frac{\varphi'(t)\varphi(t) + \psi'(t)\psi(t)}{\sqrt{\varphi(t)^2 + \psi(t)^2}}.$$

(こうやってしまうと、1 変数関数についての定理だけで済む。) ■

学生との Q&A から　上の例 ($z = xy$, $x = \varphi(t)$, $y = \psi(t)$) で

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

という chain rule を使ったが、これはどうやって出せるか (定理の「あてはめ」ですね)、という基本的 (ということは重要) な質問をされた。

z が 2 つの変数 x と y の関数である、つまり $z = z(x, y)$ であることを見て取るのが第 1 ステップである。すると、

$$\frac{\partial z}{\partial \boxed{}} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial \boxed{}} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \boxed{}}$$

となる。この $\boxed{}$ に t を入れて出来上がり。

もし u が 3 変数 x, y, z の関数、つまり $u = u(x, y, z)$ であれば、

$$\frac{\partial u}{\partial \boxed{}} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \boxed{}} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \boxed{}} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \boxed{}}$$

となる。

z_i が $y_1, \ldots, y_m$ の関数、つまり $z_i = z_i(y_1, \ldots, y_m)$ であれば、

$$\frac{\partial z_i}{\partial \boxed{}} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k}\frac{\partial y_k}{\partial \boxed{}}.$$

(次の例は次回に回しました。)

例 7.3 (波動方程式) $f\colon \mathbf{R} \to \mathbf{R}$, $c \in \mathbf{R}$ とするとき、$u\colon \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ を $u(x, t) := f(x - ct)$ で定めるとき、u_x, u_{xx}, u_t, u_{tt} を f を用いて表し、$\dfrac{1}{c^2}u_{tt}(x, t) = u_{xx}(x, t)$ を満たすことを示せ。

f の変数を ξ と書き ($f(\xi)$)、$\xi = x - ct$ とすると、

$$u_x(x, t) = f_\xi \xi_x = f'(\xi) \cdot 1 = f'(\xi),$$

$$u_{xx}(x, t) = (f')_\xi \xi_x = f''(\xi) \cdot 1 = f''(\xi),$$

$$u_t(x, t) = f_\xi \xi_t = f'(\xi) \cdot (-c) = -cf'(\xi) = -cf'(x - ct),$$

$$u_{tt}(x, t) = -c(f')_\xi \xi_t = -cf''(\xi) \cdot (-c) = (-c)^2 f''(\xi) = c^2 f''(x - ct)$$

であるから、

$$\frac{1}{c^2}u_{tt}(x, t) = u_{xx}(x, t)$$

が成り立つ。■

7.2.2　空間極座標

重要な変数変換に、極座標変換がある。平面極座標は良く知っているだろうから、ここでは 3 次元空間の極座標を紹介する。古い講義ノート (`http://www.math.meiji.ac.jp/~mk/tahensuu1/tahensuu1-2011.pdf`) の付録に n 次元極座標の紹介がある。

空間に、互いに直交する座標軸 x 軸, y 軸, z 軸を取って座標を入れた xyz 座標系で、(x, y, z) という座標を持つ点を P とする。

- P の原点からの距離を r $(r \geq 0)$

- $\overrightarrow{\mathrm{OP}}$ が z 軸の正方向となす角を θ $(0 \le \theta \le \pi)$

- P を xy 平面に正射影した点を P' として、動径 $\overrightarrow{\mathrm{OP}'}$ を x 軸の正の部分から反時計回りに測った角を ϕ $(0 \le \phi < 2\pi)$

とすると

$$\begin{cases} x = r \sin\theta \cos\phi \\ y = r \sin\theta \sin\phi \\ z = r \cos\theta \end{cases}$$

が成り立つ。このとき、r, θ, ϕ を P の**空間極座標** (**3次元極座標**あるいは**球 (面) 座標**, **spherical coordinate**) と呼ぶ。

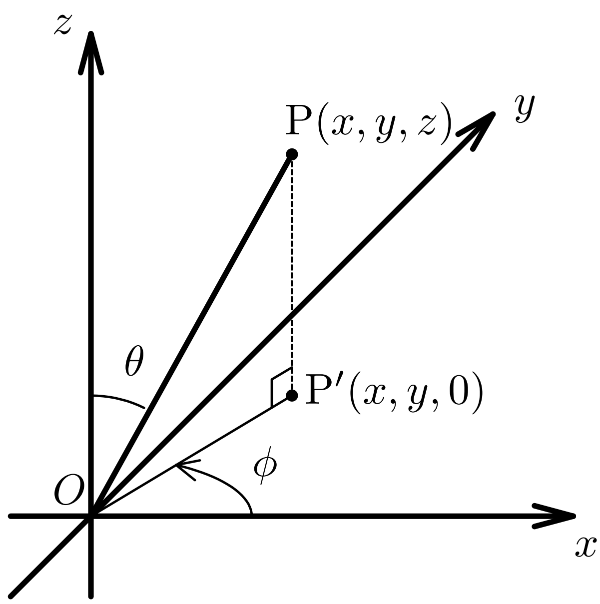

図 1: 球座標の θ, ϕ

θ と ϕ の測り方の違いに注意

注意 7.4 (極座標とのつきあい方) 3次元以上の極座標の式には、色々なバリエーションがある。x 軸の正方向から測った角度を θ としたり ($x = r\cos\theta$ となる)、地球儀のように緯度・経度形式にしたり。ここに紹介した式を丸暗記するよりは、導出の仕方を理解して、自分で導出が出来るように練習しておくのが望ましい。■

問　地球の表面上にある2点の緯度経度が分かっているときに、その2点間の (表面に沿っての最短の) 道のりの長さの求め方を説明せよ (ただし地球は球であると考える)。具体的な問題が欲しければ、アレクサンドリアの図書館 (北緯 $31°12' = 31.20°$, 東経 $29°55' = 29.91°$) と生田キャンパス (北緯 $35°37' = 35.61°$, 東経 $139°33' = 139.55°$) の地球表面に沿った道のりの長さを求めよ。(答: 昔、某学生が計算して、9555km くらい、とのことでした。正しいかどうか保証しません。)

例 7.5 (変数の極座標変換、2変数1階の場合) 関数の変数をデカルト座標から極座標に「変換」する、つまり関数 $f = f(x, y)$ が与えられたとき、

$$g(r, \theta) := f(x, y), \quad x = r\cos\theta, \quad y = r\sin\theta$$

で定義される関数 g を考える — と分かりやすくなることが (非常にしばしば) ある。

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

であるから、g は $\varphi(r,\theta):=\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$ と f との合成関数である。

$$g_r=f_x x_r+f_y y_r,\quad g_\theta=f_x x_\theta+f_y y_\theta$$

であり (これは $(g_r\ g_\theta)=(f_x\ f_y)\begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}$ と書けるので、確かに $(f\circ\varphi)'=f'\varphi'$ が成り立っている)、

$$x_r=\cos\theta,\quad y_r=\sin\theta,\quad x_\theta=-r\sin\theta,\quad y_\theta=r\cos\theta$$

であるから、

$$g_r=f_x\cos\theta+f_y\sin\theta,\quad g_\theta=-f_x r\sin\theta+f_y r\cos\theta.$$

実は g の2階偏導関数を f の偏導関数で表したり、逆に f の2階偏導関数を g の偏導関数で表すことが重要であるが、それはまた後で紹介する。■

余談 7.2 上の例は簡単だが、f を具体的に与えていないため、本質的に2変数の例となっている (つまり1変数関数の合成関数の微分法だけでは得られない)。

1変数関数から「組み立てられない[1]」という意味で本質的に多変数の関数の具体例はほとんど知られていない (皆無というわけではないらしいが、1変数関数が豊富な具体例を持っているのとは対照的である)。従って、具体的な計算問題で、多変数関数の合成関数の微分法の定理が必要となるものを出題することは難しい。これは、本質的に多変数でない関数は応用上現れないということとは違う。応用上現れる関数のほとんどは具体的な式では書けない、ということである。■

7.3 高階の導関数

要するに偏導関数を計算すれば良いので、既に述べたことと積の微分法くらいで計算はどんどん出来る。

試しに1変数では、

$$\frac{dz}{dx}=\frac{dz}{dy}\frac{dy}{dx}$$

より

$$\begin{aligned}\frac{d^2z}{dx^2}&=\frac{d}{dx}\left(\frac{dz}{dy}\frac{dy}{dx}\right)=\left(\frac{d}{dx}\frac{dz}{dy}\right)\frac{dy}{dx}+\frac{dz}{dy}\left(\frac{d}{dx}\frac{dy}{dx}\right)\\&=\frac{d^2z}{dy^2}\frac{dy}{dx}\cdot\frac{dy}{dx}+\frac{dz}{dy}\frac{d^2y}{dx^2}=\frac{d^2z}{dy^2}\left(\frac{dy}{dx}\right)^2+\frac{dz}{dy}\frac{d^2y}{dx^2}.\end{aligned}$$

偏微分方程式論からの有名な例を2つ紹介する。変数変換 (独立変数の変換は、要するに合成関数である！) をして「見方を変える」ことが重要なテクニックである。具体的に分からない関数 (なにしろ未知関数だから！) の合成関数の、高階の偏導関数の計算が必要になるのは仕方がない。

[1]大学1年生が知っているような1変数関数と、和・差・積・商、合成、テンソル積を取ることで得られる関数の意味。

例 7.6 (変数の極座標変換、2 変数 2 階の場合 (偏微分方程式で重要な Laplacian の極座標表示))
C^2 級の関数 $f\colon (x,y)\mapsto f(x,y)$ があるとき、

$$x=r\cos\theta,\quad y=r\sin\theta,\quad g(r,\theta):=f(x,y),$$

すなわち、

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

このとき

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

が成り立つ。chain rule と積の微分法により、

$$\begin{aligned}
g_r&=f_x x_r+f_y y_r,\\
g_{rr}&=(f_{xx}x_r+f_{xy}y_r)x_r+f_x x_{rr}+(f_{yx}x_r+f_{yy}y_r)y_r+f_y y_{rr}\\
&=f_{xx}x_r^2+(f_{xy}+f_{yx})\,x_r y_r+f_{yy}y_r^2+f_x x_{rr}+f_y y_{rr},\\
g_{\theta\theta}&=f_{xx}x_\theta^2+(f_{xy}+f_{yx})\,x_\theta y_\theta+f_{yy}y_\theta^2+f_x x_{\theta\theta}+f_y y_{\theta\theta}.
\end{aligned}$$

$x_r=\cos\theta$, $y_r=\sin\theta$, $x_{rr}=0$, $y_{rr}=0$, $x_\theta=-r\sin\theta$, $y_\theta=r\cos\theta$, $x_{\theta\theta}=-r\cos\theta$, $y_{\theta\theta}=-r\sin\theta$ であるから、

$$\begin{aligned}
g_{rr}&=f_{xx}\cos^2\theta+(f_{xy}+f_{yx})\cos\theta\sin\theta+f_{yy}\sin^2\theta,\\
\frac{1}{r}g_r&=\frac{f_x\cos\theta}{r}+\frac{f_y\sin\theta}{r},\\
\frac{1}{r^2}f_{\theta\theta}&=\frac{1}{r^2}\left(f_{xx}r^2\cos^2\theta-(f_{xy}+f_{yx})\,r^2\sin\theta\cos\theta+f_{yy}r^2\cos^2\theta-f_x r\cos\theta-f_y r\sin\theta\right)\\
&=f_{xx}\cos^2\theta-(f_{xy}+f_{yx})\cos\theta\sin\theta+f_{yy}\sin^2\theta-\frac{f_x\cos\theta}{r}-\frac{f_y\sin\theta}{r}.
\end{aligned}$$

ゆえに

$$g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta}=f_{xx}+f_{yy}.\blacksquare$$

ちなみに 3 変数バージョンは

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=\frac{\partial^2 g}{\partial r^2}+\frac{2}{r}\frac{\partial g}{\partial r}+\frac{1}{r^2}\left(\frac{\partial^2 g}{\partial\theta^2}+\frac{1}{\tan\theta}\frac{\partial g}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2 g}{\partial\phi^2}\right)$$

となり、工夫なしに馬鹿正直に計算すると、1 時間半以上かかる (桂田先生調査)。

問 9　C^2 級の関数 $u\colon \mathbf{R}^2\ni(x,t)\mapsto u(x,t)\in\mathbf{R}$ と正定数 c があるとき、

$$\xi=x-ct,\quad \eta=x+ct,\quad v(\xi,\eta)=u(x,t),\quad \text{すなわち}\quad v(\xi,\eta):=u\left(\frac{\xi+\eta}{2},\frac{\eta-\xi}{2c}\right)$$

とおく。このとき次式を証明せよ (左辺、右辺どちらから始めても良い、余裕あれば両方)。

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x^2}=-4\frac{\partial^2 v}{\partial\xi\partial\eta}.$$

余談 7.3 (ギリシャ文字に親しもう) 上の問のようにギリシャ文字を用いると苦戦する学生が一定の数いるので、あるとき、ギリシャ文字を避けて授業したことがある。そうすると確かにその部分はスムーズに進行するのだが、ともすると学生がギリシャ文字に慣れる機会を奪ってしまいかねないことに気づいた。今ではネットで「ギリシャ文字 書き方」のように検索すると、ギリシャ文字の筆順まで調べられるご時世なので、ギリシャ文字を避けずに授業をする気になった。■

7.4 逆関数の微分法

1 変数関数の場合の

$$\frac{dx}{dy}=\frac{1}{\dfrac{dy}{dx}}$$

に相当する定理がある。

定理 7.7 (逆関数の微分法) U と V は $\mathbf{R}^n$ の開集合で、$\varphi\colon U\to V$ は全単射、$a\in U$, $b=\varphi(a)$, φ は a で、φ^{-1} は b で全微分可能であるならば、

$$\left(\varphi^{-1}\right)'(b)=\left(\varphi'(a)\right)^{-1}.$$

(左辺の $^{-1}$ は逆関数を表し、右辺の $^{-1}$ は逆行列を表す。)

注意: 上の定理では微分可能な逆関数の存在を仮定しているが、実は $\det\varphi'(a)\neq 0$ という条件が成り立てば、局所的に微分可能な逆関数の存在が導かれるという、**逆関数定理**がある (非常に重要!)。それはこの講義科目の終盤に解説する予定である。

証明 逆関数の定義により、

$$\varphi^{-1}\left(\varphi(x)\right)=x\quad(x\in U).$$

φ は a で、φ^{-1} は $b=f(a)$ で全微分可能であるから、合成関数の微分法より、

$$\left(\varphi^{-1}\right)'(b)\varphi'(a)=I\quad(I\text{ は }n\text{ 次の単位行列})$$

が成り立つ。ゆえに $\left(\varphi^{-1}\right)'(b)=\varphi'(a)^{-1}$. ■

問 $\varphi(x)=x$ $(x\in\mathbf{R}^n)$ とするとき、$\varphi'(x)=I$ であることを示せ。

(解法 1) 一般に「$f(x)=Ax+b$ ならば $f'(x)=A$」であるから、$\varphi(x)=x=Ix+0$ より、$\varphi'(x)=I$.

(解法 2) $\varphi=\begin{pmatrix}\varphi_1\\ \vdots\\ \varphi_n\end{pmatrix}$ とおくと、$\varphi_i(x)=x_i$ $(i=1,\dots,n)$.

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

であるから、$\varphi'(x)=(\delta_{ij})=I$. ■

次の例で、逆関数の微分法を使って $\left(\varphi^{-1}\right)'$ を計算しているが、微分可能性を証明するには、本当は逆関数の定理のお世話になる必要があるだろう。

例 7.8 (極座標変換の逆変換のヤコビ行列 (とても重要)) $\varphi\colon (0,\infty)\times(0,2\pi)\ni\begin{pmatrix}r\\ \theta\end{pmatrix}\mapsto\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}r\cos\theta\\ r\sin\theta\end{pmatrix}\in\mathbf{R}^2$ は明らかに C^1 級であるから (実は C^∞ 級である)、全微分可能である。

$$\varphi'(r,\theta)=\begin{pmatrix}x_r & x_\theta\\ y_r & y_\theta\end{pmatrix}=\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}.$$

逆関数の微分法によって、$\begin{pmatrix} r \\ \theta \end{pmatrix} = \varphi^{-1}(x,y)$ の全微分係数は、

$$\begin{aligned}
\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix} &= \left(\varphi^{-1}\right)'(x,y) = \varphi'(r,\theta)^{-1} = \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} \\
&= \frac{1}{\cos\theta\cdot r\cos\theta-(-r\sin\theta)\sin\theta}\begin{pmatrix} r\cos\theta & +r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix}.
\end{aligned}$$

ゆえに[2]

$$r_x = \cos\theta,\quad r_y = \sin\theta,\quad \theta_x = -\frac{\sin\theta}{r},\quad \theta_y = \frac{\cos\theta}{r}.\blacksquare$$

(**以下はやや脱線** — 無視してよろしい) この結果を逆関数の微分法を使わずに求めてみよう。$r=\sqrt{x^2+y^2}$ であるから、r_x と r_y は比較的簡単に得られる。

$$\begin{aligned}
r_x &= \frac{\partial}{\partial x}\left(x^2+y^2\right)^{1/2} = \frac{1}{2}\left(x^2+y^2\right)^{-1/2}\cdot\frac{\partial}{\partial x}(x^2+y^2) = \frac{x}{\sqrt{x^2+y^2}}, \\
r_y &= \frac{\partial}{\partial y}\left(x^2+y^2\right)^{1/2} = \frac{1}{2}\left(x^2+y^2\right)^{-1/2}\cdot\frac{\partial}{\partial y}(x^2+y^2) = \frac{y}{\sqrt{x^2+y^2}}.
\end{aligned}$$

これらがそれぞれ $\cos\theta$, $\sin\theta$ に等しいことは容易に分かる[3]。θ の導関数の方は少し難しい。割と多くのテキストに

$$\theta = \tan^{-1}\frac{y}{x}$$

と書かれているが[4]、これは $\pmod{\pi}$ でしか正しくない式である。本当は、$\tan^{-1}\dfrac{y}{x}$ が主値 ($(-\pi/2,\pi/2)$ 内の値) を意味するとして、

$$\theta = \begin{cases}
\tan^{-1}\dfrac{y}{x} & ((x,y)\text{ が第 1 象限内あるいは }x\text{ 軸の正の部分上の点}) \\
\tan^{-1}\dfrac{y}{x}+\pi & (x<0,\text{ 言い換えると }(x,y)\text{ が第 2,3 象限内あるいは }x\text{ 軸の負の部分上の点}) \\
\tan^{-1}\dfrac{y}{x}+2\pi & ((x,y)\text{ が第 4 象限内の点}) \\
\dfrac{\pi}{2} & (x=0\text{ かつ }y>0) \\
\dfrac{3\pi}{2} & (x=0\text{ かつ }y<0)
\end{cases}$$

[2]こういう計算をするとき、ヤコビ行列の成分の並べ方を間違えると、とんでもない結果になってしまうことに注意しよう。

[3]例えば $\dfrac{x}{\sqrt{x^2+y^2}} = \dfrac{r\cos\theta}{\sqrt{(r\cos\theta)^2+(r\sin\theta)^2}} = \dfrac{r\cos\theta}{r} = \cos\theta.$

[4]本当に困ったことである。C 言語のプログラムで、デカルト座標を極座標に直すには、`r=sqrt(x*x+y*y); theta=atan2(y,x);` のようにする。`theta=atan(y/x);` ではマズイ — というのは常識的なことなのだが、数学書の方が旧態依然のままなのは情けない。

となるはずである。

$$\frac{\partial}{\partial x}\tan^{-1}\frac{y}{x}=\frac{1}{1+(y/x)^2}\frac{\partial}{\partial x}\left(\frac{y}{x}\right)=\frac{1}{1+(y/x)^2}\cdot\left(-\frac{y}{x^2}\right)=-\frac{y}{x^2+y^2},$$
$$\frac{\partial}{\partial y}\tan^{-1}\frac{y}{x}=\frac{1}{1+(y/x)^2}\frac{\partial}{\partial y}\left(\frac{y}{x}\right)=\frac{1}{1+(y/x)^2}\cdot\frac{1}{x}=\frac{x}{x^2+y^2}$$

であることから、

$$\theta_x=-\frac{y}{x^2+y^2},\quad \theta_y=\frac{x}{x^2+y^2}$$

が導けるが、少々面倒である (例えば y 軸上でどうすれば良いか分かりますか？[5])。これがそれぞれ $-\frac{\sin\theta}{r}$, $\frac{\cos\theta}{r}$ に等しいことは容易に確められる。■

例 7.9 (Laplacian の極座標表示 (再び)) C^2 級の関数 $f\colon (x,y)\mapsto f(x,y)$ があるとき、

$$x=r\cos\theta,\quad y=r\sin\theta,\quad g(r,\theta):=f(x,y)$$

で g を定める。すなわち、

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

このとき

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

が成り立つ。上では、右辺を変形していって左辺になることで示した。

応用上は、 $\triangle f=f_{xx}+f_{yy}$ が先にあって、これを g とその偏導関数で表したいので、以下のように左辺を変形していって右辺になることを示す方が望ましい。まず x, y についての 1 階偏導関数

$$f_x=g_r r_x+g_\theta\theta_x=g_r\cos\theta-g_\theta\frac{\sin\theta}{r},$$
$$f_y=g_r r_y+g_\theta\theta_y=g_r\sin\theta+g_\theta\frac{\cos\theta}{r}$$

から、

$$\frac{\partial}{\partial x}=\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta},\quad \frac{\partial}{\partial y}=\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\frac{\partial}{\partial\theta}.$$

以下は (面倒ではあるが、機械的計算で)

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

---

[5]例えば、$x=0$, $y>0$ の範囲では、偏微分の定義から $\theta_y=0=\frac{x}{x^2+y^2}$ となることは容易に分かるが、$\theta_x=-\frac{y}{x^2+y^2}$ となることは、偏導関数の右側からの極限と、左側からの極限がともに $-\frac{y}{x^2+y^2}$ であることを確め、「$f\colon (a,b)\to\mathbf{R}$ が連続で、$c\in(a,b)$ 以外では微分可能で、$\lim\limits_{\substack{x\ne c\\ x\to c}} f'(x)=D$ であれば、f は c で微分可能で、$f'(c)=D$.」という定理を用いる。

同様に

$$f_{yy}=g_{rr}\sin^2\theta+\frac{2g_{r\theta}\cos\theta\sin\theta}{r}+\frac{g_{\theta\theta}\cos^2\theta}{r^2}+\frac{g_r\cos^2\theta}{r}-\frac{2g_\theta\cos\theta\sin\theta}{r^2}.$$

ゆえに

$$f_{xx}+f_{yy}=g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta}.\blacksquare$$

8 平均値の定理、Taylor の定理

合成関数の微分法を用いると、多変数関数版平均値の定理・Taylor の定理が得られる。これらは例えば極値問題の解析に利用できる。

問題意識 Ω が $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$, $a\in\Omega$ とするとき、($\|h\|$ が十分小さい h に対して $a+h\in\Omega$ となるわけだが) $f(a+h)-f(a)$ はどうなるか？以前 $f(a+h)-f(a)\fallingdotseq f'(a)h=\nabla f(a)\cdot h$ ということを書いたが、$\fallingdotseq$ は数学的な主張ではない。そこをちゃんとしたい。

注意 f の値域の次元 (これまで m と書いてきたもの) は、1 とする。その理由は後述する。

アイディア

解くためには、次のアイディア一発で十分である。

$$F(t):=f(a+th)\quad(t\in[0,1])$$

とおくと、$F(0)=f(a)$, $F(1)=f(a+h)$ であるから、

$$f(a+h)-f(a)=F(1)-F(0).$$

この F は 1 変数関数であるから、1 変数関数の平均値定理、Taylor の定理が使える。

F の導関数 F' はどうなるか？$\varphi(t):=a+th$ とおくと、F は f と φ の合成関数である: $F=f\circ\varphi$. また、$\varphi'(t)=h$. 実際、

$$\varphi(t)=a+th=\begin{pmatrix}a_1+th_1\\a_2+th_2\\\vdots\\a_n+th_n\end{pmatrix}$$

であるから、

$$\varphi'(t)=\begin{pmatrix}(a_1+th_1)'\\(a_2+th_2)'\\\vdots\\(a_n+th_n)'\end{pmatrix}=\begin{pmatrix}h_1\\h_2\\\vdots\\h_n\end{pmatrix}=h.$$

ゆえに合成関数の微分法から、

$$F'(t)=\frac{d}{dt}f(a+th)=f'(\varphi(t))\varphi'(t)=f'(a+th)h=\nabla f(a+th)\cdot h.$$

特に

$$F'(0)=\left.\frac{d}{dt}f(a+th)\right|_{t=0}=f'(a)h=\nabla f(a)\cdot h.$$

この量を、a における、f の h 方向の**方向微分係数**と呼び、$\dfrac{\partial f}{\partial h}(a)$ で表す。

記号の約束: 方向微分係数

$$\frac{\partial f}{\partial h}(a) := \frac{d}{dt}f(a+th)\bigg|_{t=0} = f'(a)h = \nabla f(a)\cdot h.$$

平均値の定理

定理 8.1 (多変数関数の平均値の定理) Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ は全微分可能、$a, b \in \Omega$, $a \neq b$, $[a,b] \subset \Omega$ とするとき、

$$\exists c \in (a,b) \quad \text{s.t.} \quad f(b) - f(a) = f'(c)(b-a).$$

ただし

$$[a,b] := \{(1-t)a + tb; t \in [0,1]\},$$
$$(a,b) := \{(1-t)a + tb; t \in (0,1)\}.$$

注意 8.2 $n \geq 2$ のとき、1 変数の場合に良く登場する式

$$\frac{f(b)-f(a)}{b-a} = f'(c)$$

はナンセンスである (ベクトル $b-a$ で割れない)。■

証明 $h := b - a$, $F(t) := f(a+th)$ $(t \in \mathbf{R})$ とおくと、

$$f(b) - f(a) = F(1) - F(0), \quad F'(t) = f'(a+th)h.$$

1 変数関数の平均値の定理から、

$$\exists \theta \in (0,1) \quad \text{s.t.} \quad F(1) - F(0) = F'(\theta)\cdot 1.$$

ゆえに

$$f(b) - f(a) = f'(a+\theta h)h.$$

$c := a + \theta h$ とおくと、$c \in (a,b)$ で、$f(b) - f(a) = f'(c)(b-a)$. ■

平均値の定理をベクトル値関数に拡張することは出来ない。実数値関数の場合の定理の証明を振り返り、さかのぼると、実数値関数の最大値の存在まで行き着く。この証明を一般化するのが難しいことは理解できるであろう。ここでは一つの反例を示す。

例 8.3 $f\colon \mathbf{R}^2 \ni x \mapsto \begin{pmatrix}\cos x\\ \sin x\end{pmatrix} \in \mathbf{R}$ は明らかに C^∞ 級であり、全微分可能である。$a = 0$, $b = 2\pi$ とするとき、

$$f(b) - f(a) = \begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} = 0.$$

一方

$$f'(x) = \begin{pmatrix}(\cos x)'\\(\sin x)'\end{pmatrix} = \begin{pmatrix}-\sin x\\ \cos x\end{pmatrix}$$

で、このノルムは $\sqrt{\sin^2 x + \cos^2 x} = 1$ であるから、$f'(x) \neq 0$. 特に $\forall c \in (a, b)$ に対して、$f'(c)(b-a) \neq 0$. ゆえに $f(b) - f(a)$ と $f'(c)(b-a)$ は等しくなり得ない。■