

合成関数の微分

著者：関勝寿 公開日：2019年7月1日 ソース：<https://sekika.github.io/2019/07/01/derivative-chain/>

合成関数の微分の簡略化された説明と正確な証明についてまとめる。

y が u の関数 $f(u)$ であらわれ u が x の関数 $g(x)$ であらわれるとき、合成関数 $f \circ g$ を

$$y = (f \circ g)(x) = f(g(x))$$

と書くことができる。たとえば、 $y = \log(u)$ で $u = x + 2$ のとき、 $f(u) = \log(u), g(x) = x + 2$ であるから

$$y = f(g(x)) = f(x + 2) = \log(x + 2)$$

となる。 f と g が微分可能であれば

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

が成り立つ。すなわち

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

となる。これが「合成関数の微分公式」すなわち微分法における連鎖律である。

この式は、 dx, du, dy がそれぞれ「ある数」とであるとみなせば、成り立つのが当然であるように見える。そして、実際にそのように簡略化して説明されることも多いだろう。

1. 簡略化された説明

$y = f(x)$ について、 x と y の有限微小変化量をそれぞれ Δx と Δy とし、 $\Delta x \rightarrow 0$ の極限としての無限小変化量を dx と dy とすることで、

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$$

のように導関数が定義される。ここで、 Δx に対する $u = g(x)$ の変化量を Δu とし、 Δu に対する $y = f(u)$ の変化量を Δy とすれば、 Δx を定めれば Δu と Δy が順次定まる。そして、 $\Delta x \neq 0, \Delta u \neq 0$ のとき

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}$$

が成立する。これは、 $\Delta x, \Delta u, \Delta y$ がそれぞれ「ある数」であることから明らかである。 $u = g(x)$ が微分可能、つまり連続関数であることから、 $\Delta x \rightarrow 0$ とすると $\Delta u \rightarrow 0$ となる。したがって、上式で $\Delta x \rightarrow 0$ の極限を取ると

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x} \right) = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x}$$

となることから、

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

が導かれる。

以上の説明は Wikipedia の連鎖律では「間違った証明」とであるとされている。なぜ間違っているとされているのかといえば、上記の説明の中で $\Delta x \neq 0, \Delta u \neq 0$ としているが、 $\Delta x \neq 0$ であっても $\Delta u = 0$ であるとは限らないためである。実際に $g(x)$ が定数関数であればそうなるし、定数関数でなくても、微分可能な関数 $g(x)$ のある x において、任意の $\epsilon > 0$ に対して $0 < \|\Delta x\| < \epsilon, \Delta u = 0$ となるような Δx が存在するようなものを考えることは可能である。ただし、高校の数学の授業でそのよう

な関数を扱うことはまずないであろう。

とはいえ、高校の数学では連続関数や微分可能性の $\epsilon - \delta$ 論法による厳密な定義までは学習しないことから、高校生に教えるときにそこまでの厳密性がないことをもって「間違った証明である」とするのも、かえって混乱させてしまう可能性もある。正確な説明は無理だから「 dx, du, dy はそれぞれが数であるかのごとく分数のように計算できるけど、なぜそうなのかは考えなくてもいいよ」としてしまいうのも一つの方法ではあるが、そこまでブラックボックスにしてしまうのも味気ない。やはり「理由があってそうなっている」ということは伝えたいとすれば、このような簡略化された説明は認められても良いと考える。実際になかけんの数学ノートには、と書かれていて、高校の教科書でも簡略化された説明がなされているようである。

正確さを期するのであれば「これはあまり正確な証明ではないけれど、正確な証明は大学で勉強してね」とするのが妥協点であろうか。ここでは「証明」とまでは言えないことから「簡略化された説明」とした。

2. 正確な証明

上記の簡略化された説明を修正して正しい証明とする方法として、Wikipedia の連鎖律では関数 Q を導入している。この証明方法よりは、合成関数の微分の証明（伶俐玲瓏）に書かれているように、 Δu による割り算を避けて証明する方が簡明であると感じる（やっていることは本質的に同じではあるが）。少しだけ書き直す。

- $u = g(x)$ において、 x の増分 Δx に対する u の増分を Δu
- $y = f(x)$ において、 x の増分 Δu に対する y の増分を Δy とする。ある x における $\Delta x \neq 0$ に対する ϵ_1 を

$$\Delta u = [g'(x) + \epsilon_1]\Delta x$$

によって定めると、

$$\epsilon_1 = \frac{\Delta u}{\Delta x} - g'(x)$$

であり、 $g'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x}$ より $\Delta x \rightarrow 0$ のとき $\epsilon_1 \rightarrow 0$ かつ $\Delta u \rightarrow 0$ である。次に、ある u における Δu に対する ϵ_2 を

$$\Delta y = [f'(u) + \epsilon_2]\Delta u \quad (\text{ただし、}\Delta u = 0 \text{ のとき } \epsilon_2 = 0)$$

$= [f'(u) + \epsilon_2][g'(x) + \epsilon_1]\Delta x$ (上の式を代入)
によって定めると、 $\Delta u \rightarrow 0$ のとき $\epsilon_2 \rightarrow 0$ である。そして、 $\Delta x \rightarrow 0$ のとき $\Delta u \rightarrow 0$ であるから $\epsilon_2 \rightarrow 0$ である。したがって、

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} [f'(u) + \epsilon_2][g'(x) + \epsilon_1] \\ &= f'(u)g'(x) \\ &= \frac{dy}{du} \cdot \frac{du}{dx} \end{aligned}$$

となり、合成関数の微分公式が導かれた。 Δy の式において「 $\Delta u = 0$ のとき $\epsilon_2 = 0$ 」としているところがミソである。