

情報数学III講義 (第2回)

平成28年10月12日

1.7 多変数の Kock-Lawvere の公理

$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\varphi : d \in D \mapsto f(\mathbf{x} + \mathbf{a}d) \in \mathbb{R}^m$ という関数に対して, 多変数の Kock-Lawvere の公理から以下が成り立つ.

$$\exists! \mathbf{b} \in \mathbb{R}^m \quad \text{s.t.} \quad \varphi(d) = \varphi(0) + \mathbf{b}d \quad (1)$$

このただひとつの \mathbf{b} を $f'(\mathbf{x}_0)(\mathbf{a})$ と書く. $f'(\mathbf{x}_0)(\mathbf{a})$ は \mathbf{x}_0 にも \mathbf{a} にも依存する. ここで, $f'(\mathbf{x}_0)(\mathbf{a})$ が以下の二つを満たした線形変換であることを示す. ただし, $\mathbf{a}_1 \in \mathbb{R}^n, \mathbf{a}_2 \in \mathbb{R}^n, \mathbf{a} \in \mathbb{R}^n, \alpha \in \mathbb{R}$ である.

$$f'(\mathbf{x}_0)(\mathbf{a}_1 + \mathbf{a}_2) = f'(\mathbf{x}_0)(\mathbf{a}_1) + f'(\mathbf{x}_0)(\mathbf{a}_2) \quad (2)$$

$$f'(\mathbf{x}_0)(\alpha \mathbf{a}) = \alpha f'(\mathbf{x}_0)(\mathbf{a}) \quad (3)$$

- (5) の証明

$$f(\mathbf{x}_0 + (\mathbf{a}_1 + \mathbf{a}_2)d) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{a}_1 + \mathbf{a}_2)d$$

であることと,

$$\begin{aligned} f(\mathbf{x}_0 + (\mathbf{a}_1 + \mathbf{a}_2)d) &= f((\mathbf{x}_0 + \mathbf{a}_1d) + \mathbf{a}_2d) \\ &= f(\mathbf{x}_0 + \mathbf{a}_1d) + f'(\mathbf{x}_0 + \mathbf{a}_1d)(\mathbf{a}_2)d \\ &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{a}_1)d + \{f'(\mathbf{x}_0) + f''(\mathbf{x}_0)(\mathbf{a}_1)d\}(\mathbf{a}_2)d \\ &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{a}_1)d + f'(\mathbf{x}_0)(\mathbf{a}_2)d + f''(\mathbf{x}_0)(\mathbf{a}_1)(\mathbf{a}_2)d^2 \\ &= f(\mathbf{x}_0) + \{f'(\mathbf{x}_0)(\mathbf{a}_1) + f'(\mathbf{x}_0)(\mathbf{a}_2)\}d \end{aligned}$$

からわかる.

- (6) の証明

$$f(\mathbf{x}_0 + (\alpha \mathbf{a})d) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\alpha \mathbf{a})d$$

であることと,

$$\begin{aligned}f(\mathbf{x}_0 + \mathbf{a}(\alpha d)) &= f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{a})(\alpha d) \\ &= f(\mathbf{x}_0) + \alpha f'(\mathbf{x}_0)(\mathbf{a})d\end{aligned}$$

からわかる.

1.8 多変数の微分

- 1変数 $y = f(x)$ の場合

$f(x)$ 上にある点を (x_0, y_0) として, x の変化量 $\Delta x = x - x_0$ と y の変化量 $\Delta y = y - y_0$ について考える. このとき Δy は $\Delta y = f(x + \Delta x) - f(x)$ と一般的に複雑な形になる. この複雑な $f(x)$ による関数を簡単にするため, 点 (x_0, y_0) における接線を用いて近似すると $\Delta y = a\Delta x$ (a は定数, $a = f'(x_0)$) のようにひとつの数で特徴付けられた比例関数で表せる.

- 2変数 $z = f(x, y)$ の場合

$z_0 = f(x_0, y_0)$ とする. x の変化量 $\Delta x = x - x_0$, y の変化量 $\Delta y = y - y_0$ に対して z の変化量 $\Delta z = z - z_0$ について考える. 関数 z を微分すると接平面が得られ, $\Delta z = a\Delta x + b\Delta y$ (a, b は定数) と表せる. a を求めるには, $f(x)$ を $\Delta y = 0$ と y の値を固定して得られる x の1変数関数 $f(x, y_0)$ とみなして x で微分すればよい.

$$a = \frac{\partial f}{\partial x}(x, y)$$

同様に, b を求めるには $\Delta x = 0$ とする.

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

これを偏微分と呼ぶ. a を求める偏微分のイメージは, 図1のような風呂敷状のグラフを xy 平面と平行に切断して得られる断面での微分といえる.

1.9 様々な写像

これまで \mathbb{R} への関数を見てきたが, 他にも \mathbb{R}^2 や \mathbb{R}^3 への関数, たとえば

$$\mathbb{R} \longrightarrow \mathbb{R}^2 \text{ (平面上の運動)}$$

$$\mathbb{R} \longrightarrow \mathbb{R}^3 \text{ (空間上の運動)}$$

などが考えられる. 例えば \mathbb{R} を時間, \mathbb{R}^2 を平面上の点, \mathbb{R}^3 を空間内の点とすれば, 上の二つは平面上の運動, 空間内での運動を表す関数と考えられる.

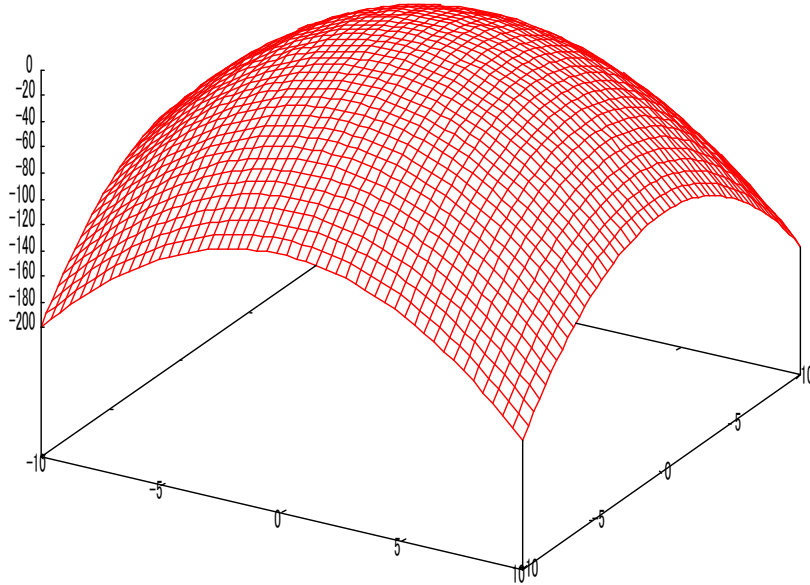


図 1: $z = f(x, y)$ のグラフ

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $z = f(x, y)$ の場合

$$\begin{aligned} \Delta z &= \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \end{aligned}$$

と多変数の微分を行列の積で表すことができる。

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ の場合

f を $f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ の 2 つに分けて $f(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix}$ と置けるとする。

$z_1 = f_1(x, y)$, $z_2 = f_2(x, y)$ とすると以下のように表せ、

$$\begin{aligned} \Delta z_1 &= \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ \Delta z_2 &= \begin{pmatrix} \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \end{aligned}$$

これらをまとめることで,

$$\begin{pmatrix} \Delta z_1 \\ \Delta z_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

と行列の積で表せる.

以上をまとめると,

- $\mathbb{R} \rightarrow \mathbb{R}$ の関数を微分すると比例関係が出る.
- $\mathbb{R}^2 \rightarrow \mathbb{R}$ の関数を偏微分すると行列 $\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}$ が出る.
- $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ の関数を偏微分すると $\begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}$ が出る.

このように, 線形写像を行列の形で考えることが出来る.

1.10 合成関数の微分

$\mathbb{R} \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$ のとき, 合成関数 $g \circ f$ の微分は以下で表される.

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

多変数の合成関数の微分を行うために, これを一般化する. $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m \xrightarrow{g} \mathbb{R}^l$ のとき, 合成関数 $g \circ f$ の微分は以下で表される.

$$(g \circ f)'(\mathbf{x}) = \underbrace{g'(f(\mathbf{x}))}_{l \times m} \cdot \underbrace{f'(\mathbf{x})}_{m \times n}$$

定理

関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, g: \mathbb{R}^m \rightarrow \mathbb{R}^l, g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^l$ に対して $\mathbf{x}_0 \in \mathbb{R}^n$ としたとき, 以下の式が成り立つ.

$$(g \circ f)'(\mathbf{x}_0) = g'(f(\mathbf{x}_0)) \cdot f'(\mathbf{x}_0)$$

この定理は以下のように証明できる. $\mathbf{a} \in \mathbb{R}^n, d \in D$ とする

$$\begin{aligned} (g \circ f)(\mathbf{x}_0 + \mathbf{a}d) &= g(f(\mathbf{x}_0 + \mathbf{a}d)) \\ &= g(f(\mathbf{x}_0) + f'(\mathbf{x}_0)(\mathbf{a})d) \\ &= g(f(\mathbf{x}_0)) + g'(f(\mathbf{x}_0))(f'(\mathbf{x}_0)(\mathbf{a})d) \\ &= g(f(\mathbf{x}_0)) + (g'(f(\mathbf{x}_0)) \circ f'(\mathbf{x}_0))(\mathbf{a})d \end{aligned}$$

例えば、1変数の場合は 1×1 の行列の積と考えればよい。

また、 $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$ の場合は $g \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x_1 \ x_2) \rightarrow (y_1 \ y_2) \rightarrow z$ と表され、微分は以下のようなになる。

$$g'(f(\mathbf{x})) = \left(\frac{\partial z}{\partial y_1} f(\mathbf{x}) \quad \frac{\partial z}{\partial y_2} f(\mathbf{x}) \right) \quad 1 \times 2 \text{ の行列}$$

$$f'(\mathbf{x}) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1}(x_1, x_2) & \frac{\partial y_1}{\partial x_2}(x_1, x_2) \\ \frac{\partial y_2}{\partial x_1}(x_1, x_2) & \frac{\partial y_2}{\partial x_2}(x_1, x_2) \end{pmatrix} \quad 2 \times 2 \text{ の行列}$$

最終的に、行列を掛け算すると以下のように表せる。

$$\begin{aligned} \begin{pmatrix} \frac{\partial z}{\partial x_1} \\ \frac{\partial z}{\partial x_2} \end{pmatrix} &= \begin{pmatrix} \frac{\partial z}{\partial y_1} & \frac{\partial z}{\partial y_2} \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial z}{\partial y_1} \frac{\partial y_1}{\partial x_1} + \frac{\partial z}{\partial y_2} \frac{\partial y_2}{\partial x_1} \\ \frac{\partial z}{\partial y_1} \frac{\partial y_1}{\partial x_2} + \frac{\partial z}{\partial y_2} \frac{\partial y_2}{\partial x_2} \end{pmatrix} \end{aligned}$$

このように、線形写像の合成を行列の掛け算で記述することが出来る。

§ 行列表現

関数 $f : \mathbb{R}^n \rightarrow \mathbb{R}$ に対して微分 f' は線型写像 $f' : \mathbb{R}^n \rightarrow \mathbb{R}$ と考えることができる。いま、 $\mathbf{e}_1 = (1 \ 0 \ \dots \ 0)^T$ として $f'(\mathbf{x})(\mathbf{e}_1)$ について考える。公理より以下のようなになる。

$$f(\mathbf{x} + \mathbf{e}_1 d) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{e}_1) d$$

x_2 から x_n は動かしていないので、 x_1 で偏微分していることになる。よって、以下のように表せる。

$$f(\mathbf{x} + \mathbf{e}_1 d) = f(\mathbf{x}) + \frac{\partial f}{\partial x_1}(\mathbf{x}) d$$

次に、 $\mathbf{a} = (a_1 \ a_2 \ \dots \ a_n)^T = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_n \mathbf{e}_n$ について考える。これは線型写像なのでそれぞれに分けられることを思い出すと、以下のように整理

できる.

$$\begin{aligned} f'(\mathbf{x})(\mathbf{a}) &= f'(\mathbf{x})(a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n) \\ &= a \frac{f'(\mathbf{x})(\mathbf{e}_1)}{\frac{\partial f}{\partial x_1}(\mathbf{x})} + a_2 \frac{f'(\mathbf{x})(\mathbf{e}_2)}{\frac{\partial f}{\partial x_2}(\mathbf{x})} + \cdots + a_n \frac{f'(\mathbf{x})(\mathbf{e}_n)}{\frac{\partial f}{\partial x_n}(\mathbf{x})} \\ &= \left(\frac{\partial f}{\partial x_1}(\mathbf{x}) \quad \frac{\partial f}{\partial x_2}(\mathbf{x}) \quad \cdots \quad \frac{\partial f}{\partial x_n}(\mathbf{x}) \right) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \end{aligned}$$

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ のとき, $f': \mathbb{R}^n \rightarrow \mathbb{R}$ は $1 \times n$ 行列で表せる.