第3章　微分法

この章では，微分について学んでいきます．いくつかのポイントを前もって示しておきます．

1. 数学史からみた微分
2. 微分の定義
3. 平均値の定理
4. 導関数
5. 高階導関数
6. テイラーの定理
7. 多変数関数の微分
8. 偏微分・偏導関数
9. 多変数関数のテイラーの定理

3.1 数学史からみた微分

すでに書きましたが，解析学の視点とは,「極限操作を通じて，数や関数の性質を調べる」ということです．極限操作とは,

- 数列の収束先，極限値を求める
- 級数（加算無限個の足し算）の収束先，極限値を求める
- 包含関係にある集合の加算無限個列を考え，その集合列の極限の集合を求める

などのことを指します.

典型的な無理数，円周率や自然対数などを有理数の級数として求める大昔の数学者のアプローチは，まさに解析学の真骨頂といえます[1]．無理数という対象に，有理数という相対的になじみの深い対象から接近しているところが，解析＝ analysis（分析）という言葉にぴったりですね.

そんな，大昔の数学者にとって，次の問題が契機となって微分法が考えられます.

[問題] 関数 $f(x)$ のグラフとして曲線が与えられているとします．このとき，各点 x でのこの曲線の傾きの大きさ・接線や法線の方程式を求めよ．

この問題を考える動機になったものとして

- 2 つの曲線が交わる角度の計算（デカルト）
- 望遠鏡 (ガリレイ) と時計 (ホイヘンス) の製作
- 関数の最大・最小を求めること (フェルマー 1638)
- 運動の速度と加速度 (ガリレイ 1638, ニュートン 1686)
- 天文学，重力の法則の検証 (ケプラー，ニュートン)

このような背景の下で，数学の全歴史の中で最大の発明ともいわれる微積分法が導入されました．ニュートン，ライプニッツの他，ベルヌーイ一族，オイラーなどの大数学者が，微積分法の発展に貢献しています.

[1] ここでの説明は，E. ハイラー/G. ワナー著 蟹江幸博訳『解析教程』シュプリンガー・フェアラーク東京,1997 年, 第 2 章の記述を参考にしています．興味のある学生諸君は，原著を読んでください.

3.2 微分の定義

微分の定義からはじめます．1 変数関数 $f:\mathbf{R}\to\mathbf{R}$ を考えます．点 $a\in\mathbf{R}$ において

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h} \tag{3.1}$$

が存在するとき，関数 f は $a\in\mathbf{R}$ において**微分可能**といいます．また (3.1) の値を関数 f の $a\in\mathbf{R}$ における**微分係数**とよび,

$$\frac{df}{dx}(a) \text{ あるいは } f'(a)$$

のように記します.

注意 12. 微分の定義について，注意しなくてはならないことがあります．それは，微分係数 $f'(a)$ は，極限を計算する前提となっている h のゼロへの近づけ方に依存してはいけないということです．正の値から段々と小さくしていってゼロに近づけても，負の値から段々と大きくしていってゼロに近づけても，*(3.1)* の値は，一意になっていることが大切なのです．例えば数列 $\{h_n\}_{n=1}^{\infty}$ を $h_n=\frac{1}{2^n}$ としようと $\{h_n\}_{n=1}^{\infty}$ を $h_n=-\frac{1}{4^n}$ はたまた $\{h_n\}_{n=1}^{\infty}$ を $h_n=\left(\frac{-1}{3}\right)^n$ としようと，上の定義における極限値としての微分係数 $f'(a)$ は等しくなくてはいけないのです.（最後の h のゼロの近づけ方は，正負の符号を入れ替わっていきます.）

例えば，$f(x)=|x|$ という関数は $0\in\mathbf{R}$ において微分可能でありません．なぜならば，h を正の値を保ちつつ近づけた極限である 1 と h を負の値を保ちつつ近づけた極限である -1 が等しくないからです.

研究課題 2. いくつかの有限個の点で微分不可能な連続関数を考えるのは簡単です．しかし，連続関数でありながら定義域の全域で微分不可能な関数もいくつか知られています．どんな関数なのか，図書館の数学の書籍やインターネットを用いてしらべてみましょう.

つぎに導関数を定義しましょう，実数直線上のすべての $x\in\mathbf{R}$ において微分可能であるとき，各点 $x\in\mathbf{R}$ に対して，微分係数 $f'(x)$ を対応させる

$$x\longmapsto f'(x)$$

を f の**導関数**とよび，f' あるいは，関数の値と区別できるときは $f'(x)$ と表します.

例 2. $f(x)=x^n,(n\in\mathbf{N})$ の $a\in\mathbf{R}$ における微分係数を求めてみましょう.

$$\begin{aligned}
\lim_{h\to 0}\frac{(a+h)^n-a^n}{h} &= \lim_{h\to 0}\frac{(a+h-a)\left((a+h)^{n-1}+(a+h)^{n-2}a+\cdots+(a+h)a^{n-2}+a^{n-1}\right)}{h}\\
&= \lim_{h\to 0}\frac{h\left((a+h)^{n-1}+(a+h)^{n-2}a+\cdots+(a+h)a^{n-2}+a^{n-1}\right)}{h}\\
&= \lim_{h\to 0}\left((a+h)^{n-1}+(a+h)^{n-2}a+\cdots+(a+h)a^{n-2}+a^{n-1}\right)\\
&= na^{n-1}
\end{aligned}$$

このことから，導関数 $f'(x)=nx^{n-1}$ もわかります.

演習 11. 次の各関数の導関数を，定義に立ち返って求めましょう.

1. $\sin x$

2. e^x

3. $\log x$

高階導関数について説明する前に，微分についての重要な定理を挙げておきます．

定理 1. 関数 $f : \mathbf{R} \to \mathbf{R}$ が $a \in \mathbf{R}$ において微分可能であるとき f は点 a で連続となる．

(証明) 数列 $\{x_n\}_n$ が $x_n \to a$ であるとします．$h_n = x_n - a$ と定義すると，$h_n \to 0$ がわかります．このとき，

$$f(x_n) - f(a) = f(a + h_n) - f(a) \to h_n \cdot f'(a) \quad (n \to \infty)$$

から

$$f(x_n) - f(a) \to 0 \quad (n \to \infty)$$

つまり

$$f(x_n) \to f(a) \quad (n \to \infty)$$

が結論されます．(証明終わり)

微分可能ならば連続．逆は真ならずと心に刻みましょう．

次の定理も，多変数関数の微分を考えるときに重要となります．

定理 2. 関数 $f : \mathbf{R} \to \mathbf{R}$ が $a \in \mathbf{R}$ において微分可能であることの必要かつ十分条件は

$$(\exists b \in \mathbf{R}) \left[\lim_{|h| \to 0} \frac{f(a+h) - f(a) - bh}{|h|} = 0 \right] \tag{3.2}$$

演習 12. 定理 *2* を証明してください．

定理 2 は (3.2) を微分の定義にしても問題ないことを意味します．実は，多変数関数において，微分を考えるときは (3.2) のような形式で微分を考えるほうが便利なのです．

さらに，微分法の基本的な計算規則をまとめておきましょう．

定理 3 (微分法の基本規則１). 関数 $f : \mathbf{R} \to \mathbf{R}$ が $a \in \mathbf{R}$ と関数 $g : \mathbf{R} \to \mathbf{R}$ が $a \in \mathbf{R}$ が微分可能，$\lambda \in \mathbf{R}$ ならば，

1. $h : x \longmapsto f(x) + g(x)$ の導関数は $h'(x) = f'(x) + g'(x)$

2. $h : x \longmapsto \lambda f(x)$ の導関数は $h'(x) = \lambda f'(x)$

3. $h : x \longmapsto f(x)g(x)$ の導関数は $h'(x) = f'(x)g(x) + f(x)g'(x)$

4. $h : x \longmapsto \frac{f(x)}{g(x)}$ の導関数は，$(\forall x \in \mathbf{R})g(x) \neq 0$ であるとき，

$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

(証明)3 つ目の主張のみを証明し，残りは演習問題とします.

$$\begin{aligned}
h'(x) &= \lim_{h\to 0}\frac{f(x+h)g(x+h)-f(x)g(x)}{h} \\
&= \lim_{h\to 0}\left\{\frac{f(x+h)g(x+h)-f(x)g(x+h)}{h}+\frac{f(x)g(x+h)-f(x)g(x)}{h}\right\} \\
&= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\lim_{h\to 0}g(x+h)+f(x)\lim_{h\to 0}\frac{g(x+h)-g(x)}{h0} \\
&= f'(x)g(x)+f(x)g'(x)
\end{aligned}$$

(証明おわり)

次の定理は，合成関数の微分則あるいは鎖法則 (chain rule) とよばれる非常に重要な定理です.

定理 4 (微分法の基本規則２). 関数 $f:\mathbf{R}\to\mathbf{R}$ が $a\in\mathbf{R}$ $a\in\mathbf{R}$ で微分可能で，関数 $g:\mathbf{R}\to\mathbf{R}$ が $a\in\mathbf{R}$ が $f(a)\in\mathbf{R}$ で微分可能なとき

$$h : x \longmapsto g(f(x))$$

という合成関数で定義される $h(x)$ は $a\in\mathbf{R}$ で微分可能で $h'(a)=g'(f(a))f(a)$ となる.

(証明) この定理はきちんと証明しようとすると，意外とやっかいです.

$b=f(a)$，$\lambda=f'(a)$，$\mu=g'(f(a))$ と記すことにしましょう．さらに

$$\varphi(x)=f(x)-f(a)-\lambda(x-a) \tag{3.3}$$

$$\psi(y)=g(y)-g(b)-\mu(y-b) \tag{3.4}$$

$$\rho(x)=g(f(x))-g(f(a))-\mu\lambda(x-a) \tag{3.5}$$

と定義します．このとき

$$\lim_{x\to a}\frac{|\varphi(x)|}{|x-a|}=0 \tag{3.6}$$

$$\lim_{y\to b}\frac{|\psi(y)|}{|y-b|}=0 \tag{3.7}$$

が成立します.

示すべきことは

$$\lim_{x\to a}\frac{|\rho(x)|}{|x-a|}=0$$

です.

$$\begin{aligned}
\rho(x) &= g(f(x))-g(b)-\mu\lambda(x-a) \\
&= g(f(x))-g(b)-\mu(f(x)-f(a)-\varphi(x)) \\
&= \{g(f(x))-g(b)-\mu(f(x)-f(a))\}+\mu(\varphi(x)) \\
&= \psi(f(x))+\mu(\varphi(x))
\end{aligned}$$

となりますから，つぎの二つの式を示せば十分です.

$$\lim_{x\to a}\frac{|\psi(f(x))|}{|x-a|}=0 \tag{3.8}$$

$$\lim_{x \to a} \frac{|\mu\varphi(x)|}{|x-a|} = 0 \tag{3.9}$$

ここで (3.9) は (3.6) と線形写像の性質から導かれます．また (3.7) により

$$(\forall \varepsilon > 0)(\exists \delta > 0)\ |f(x) - b| < \delta \Longrightarrow |\psi(f(x))| < \varepsilon |f(x) - b|$$

ここで適当な $\delta_1 > 0$ をとると，この条件は $|x - a| < \delta_1$ に対して成り立ちます．結局

$$\begin{aligned} |\psi(f(x))| &< \varepsilon |f(x) - b| \\ &= \varepsilon |\varphi(x) - \lambda(x - a)| \\ &\leq \varepsilon |\varphi(x)| + \varepsilon |\lambda||x - a| \end{aligned}$$

となる不等式が得られます．式 (3.8) はこれから出ます.（証明おわり）

3.3 平均値の定理

微分による解析の中心的な道具となるテイラーの定理を導出する準備となる平均値の定理を証明しましょう．平均値の定理は，関数の挙動を調べる便利な道具でもあります.

最初に次の定理を証明します．内容は，直感的に自然なものです.

関数 $f(x)$ が $x = c$ において極大値をとるとは,

$$(\exists \delta > 0)(\forall x \in (c - \delta, c + \delta))\ f(x) \geq f(c)$$

をみたすことを言います．$x = c$ において極小値をとるとは

$$(\exists \delta > 0)(\forall x \in (c - \delta, c + \delta))\ f(x) \leq f(c)$$

となることをいいます．極大値か極小値をとることを単に極値をとるといいます.

定理 5. $a < b$ とし開区間 (a, b) 上で関数 $f(x)$ が微分可能とする．このとき，$a < c < b$ となる c において f が極値をとるとき $f'(c) = 0$ となる.

(証明) $x = c$ で極大とします.（極小の場合も同様です.）このとき

$$(\exists \delta > 0)(\forall h)\ |h| < \delta \Longrightarrow f(c) \leq f(c + h)$$

$h > 0$ のとき

$$\frac{f(c + h) - f(c)}{h} \leq 0$$

より，$h > 0$ のまま $h \to 0$ とすると

$$f'(c) \leq 0$$

他方, $h < 0$ のとき

$$\frac{f(c + h) - f(c)}{h} \geq 0$$

より，$h < 0$ のまま $h \to 0$ とすると

$$f'(c) \geq 0$$

これら 2 つの等号つき不等式が同時に成立することにより，$f'(c) = 0$ がえられます.（証明おわり）

定理 6 (Rolle の定理). 関数 $f : \mathbf{R} \to \mathbf{R}$ を微分可能とする．$a < b$ となる実数 a, b について $f(a) = f(b)$ であるとき，$a < c < b$ となる実数 c が存在して $f'(c) = 0$

(証明) 閉区間 $[a, b]$ 上で，f は連続なので，ワイエルストラスの定理より，最大値 $f(M_0)$ と最小値 $f(m_0)$ をもつ．

[$f(M_0) = f(m_0) = f(a) = f(b)$ **の場合**] 明らかに

$$(\forall x \in [a, b])\ f(x) = f(a)$$

となります．

$$(\forall x \in (a, b))\ f'(x) = 0$$

が成立するので，任意の $x \in (a, b)$ がもとめる c となります．

[$f(M_0) > f(a) = f(b)$ **の場合**] $M_0 \in (a, b)$ となる．このとき，$x = M_0$ は極大点であるから，定理 5 より

$$f'(M_0) = 0$$

をえる．M_0 がもとめる c となります．

[$f(m_0) < f(a) = f(b)$ **の場合**] $m_0 \in (a, b)$ となる．このとき，$x = M_0$ は極小点であるから，定理 5 より

$$f'(m_0) = 0$$

をえる．m_0 がもとめる c となります．(証明おわり)

定理 7 (平均値の定理). 関数 $f : \mathbf{R} \to \mathbf{R}$ は開区間 $(a, b) \neq \emptyset$ 上で微分可能，閉区間 $[a, b]$ 上で連続とする．このとき

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

(証明) 次の関数を考える．

$$\varphi(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

$\varphi(a) = f(a) = \varphi(b)$ が成立します．そこで，この $\varphi(x)$ に Rolle の定理を用いると，もとめる $c \in (a, b)$ がえられます．(証明おわり)

微分法の解析上の応用で，よく使われる以下の定理は，実は平均値の定理に基づきます．

定理 8 (単調性の判定).

$$(\forall x \in (A, B))\ f'(x) > 0$$

ならば，$f(x)$ は狭義の単調増加関数．また

$$(\forall x \in (A, B))\ f'(x) < 0$$

ならば，$f(x)$ は狭義の単調減少関数．

(証明) $A < a < b < B$ とします．平均値の定理から

$$(\exists c \in (a,b)) f(b) - f(a) = f'(c)(b-a)$$

$(a,b) \subset (A,B)$ ですから

$$f(b) - f(a) > 0 \Longleftrightarrow f'(c) > 0$$

(証明おわり)

定理 9 (定数関数の判定).

$$(\forall x \in (A,B)) \ f'(x) = 0$$

ならば，$f(x)$ は区間 (A,B) で一定値をとる．

(証明) 演習とします．

演習 13. 定理 *9* を定理 *8* にならって証明してください．

例 3. 区間 $(0,\infty)$ において，不等式

$$x > \sin x$$

が成立することを示してみましょう．

$F(x) = x - \sin x$ とおき，定義域を $[0,\infty)$ とします．F の連続性，微分可能性は OK ですね．$x > 0$ に対して $F'(x) = 1 - \cos x > 0$ より区間 $(0,\infty)$ において単調増加関数．任意の $x \in (0,\pi/2]$ に対して平均値の定理より $c \in (0,x)$ が存在して

$$F(x) - F(0) = F'(c)(x-0) = F'(c)x > 0$$

ところが $F(0) = 0$ なので区間 $(0,\pi/2]$ において $F(x) > 0$ がいえました．区間 $(\pi/2,\infty)$ については単調増加性より $F(x) \geq F(\pi/2) > 0$ がいえるので，区間全体 $(0,\infty)$ について $F(x) >$ がいえました．これは，もとめる不等式をしめします．

演習 14. 平均値の定理を使って，以下の不等式の成立を示してください．

1. $\sin x > x - \frac{x^3}{6} \quad (x > 0)$

2. $\sin x + \tan x > 2x \quad (0 < x < \pi/2)$

3. $\cos x > 1 - \frac{x^2}{2} \quad (x \neq 0)$

4. $1 - \frac{x^2}{2} + \frac{x^4}{24} > \cos x \quad (x \neq 0)$

3.4 高階導関数

高校でも習いますが，微分可能な関数 $y = f(x)$ が与えられたとき，$a \in \mathbf{R}$ における $f'(a)$ は，曲線上の点 $(a, f(a))$ における接線の傾きを示します．

導関数 $f'(x)$ は，この接線の傾きが x の変化に対応して，どのように変わるかを表します．この $f'(x)$ が x に関して微分可能なとき得られる 2 階の導関数 $f''(x)$ は何を意味するのでしょうか．

ニュートン (1665) とヨハン・ベルヌーイ (1691) が，この幾何学的な意味を考えたといわれています．

もし区間 (a,b) において，$f''(x)>0$ ならば $f'(x)$ はこの区間において増加しますから，$a<x_1<x_2<b$ とすると，$f'(x_1)<f'(x_2)$ を意味します．これは，区間において接線の傾きがより急になることに他なりません．この場合，関数 f は下に凸，あるいは単に凸であるといいます．同様に，$f''(x)<0$ のとき，凹とよびます．

演習 15. オイラー *(1755)* が考えた次の関数

$$f(x)=\frac{x}{1+x^2}$$

の *1* 階の導関数 $f'(x)$，*2* 階の導関数 $f''(x)$ を調べることで，$f(x)$ の最大値・最小値，変曲点をあきらかにしなさい．

関数の形状に関しての大域的な性質である凸性・凹性については，最適化に関する次章で包括的にあつかうことにして，話を進めます．

微分可能な関数 $f(x)$ の 1 階の導関数 $f'(x)$ が，さらに微分可能なとき，2 階の導関数 $f''(x)$ が得られます．さらに，2 階の導関数が微分可能なとき，3 階の導関数 $f^{(3)}(x)$，4 階の導関数 $f^{(4)}(x)$ というように，可能な限り定義することができる．

関数 $f(x)$ が定義域において n 階微分可能で，さらに n 階導関数 $f^{(n)}(x)$ が連続関数であるとき，n 階連続微分可能とよび，$f\in C^n$ のように書き，f は C^n 級だといいます．また連続関数を C^0 級とよぶこともあります．

演習 16. 次の関数は，何級の関数でしょうか．

1. $f(x)=e^x$

2. $f(x)=\cos x$

3. $f(x)=x^2$

4. $f(x)=|x|$

5. $f(x)=x\times|x|$

3.5 テイラーの定理

関数 $y=f(x)$ が与えられたとき，この関数の性質を調べるのに，1 階や 2 階の導関数の符号に注目する．その場合，わかることは増加関数か減少関数，凸か凹かといった，定性的な情報です．増加関数といってもピンからキリまであるし，凸関数についても事情は同様です．

数量的な情報を含む形で，さらによい「精度」で，局所的な挙動を調べる手段はないかということで考えられたのが，関数をべき級数の形で展開するという接近法です．

何度も強調してきましたが，古典的な解析学の発想は,「数学の言葉で定式化できる対象 (数や関数) を，極限操作，収束判定を用いて調べる」ということです．

円に内接する正 n 多角形の面積 (n 個の 2 等辺 3 角形の面積) $\longrightarrow$ 円の面積

ということと，

　　べき関数の級数 $\longrightarrow$ 対象の関数

を対応させて考えてみましょう．分析したい対象を，性質がよくわかっているものの無限個の「和」としてとらえようという態度が共通しているのです[2]．

ついでに言えば，内接正多角形の面積は各の数 n の単調増加数列をなしていて，円の面積という上限をもつという性質ために，この接近法の正しさが保証されるのです．

関数 $f(x)$ と点 x_0 が与えられたときに，

$$\alpha(x) = \alpha_0 + \alpha_1(x - x_0) + \alpha_2(x - x_0)^2 + \alpha_3(x - x_0)^3 + \cdots$$

をとれば $\alpha(x)$ は何階でも微分可能だとおもわれますが，各階の導関数について

$$\alpha^{(i)}(x_0) = f^{(i)}(x_0), \quad (i = 0, 1, 2, 3, \cdots) \tag{3.10}$$

であれば，$\alpha(x)$ は $f(x)$ のよい近似になることが，x_0 のごく近傍に限れば，いえそうだというわけです．

演習 17. 条件 *(3.10)* を用いて，各 α_i，$(i = 0, 1, 2, 3, \cdots)$ をもとめてみましょう．

高階微分可能関数 $f(x)$ を，剰余項という項を含む有限個のべき関数の多項式に表したものをテイラー多項式，$a = 0$ としてべき関数の級数に展開したものをマクローリン展開といいます．

ベルヌーイのアイデアは

$$f(x) = f(a) + \int_a^x 1 \cdot f'(t)dt$$

と書いて，$u'(t) = 1, v(t) = f'(t)$ とおいて部分積分することです．ただし，$u(t) = t - x$ とおくことがポイントです．

$$\begin{aligned} f(x) &= f(a) - [(x-t)f'(t)]_a^x + \int_a^x (x-t)f''(t)dt \\ &= f(a) + (x-a)f'(a) + \int_a^x (x-t)f''(t)dt \end{aligned}$$

をもとめ，次のステップでは $u(t) = -(x-t)^2/2!, v(t) = f''(t)$ とおいて部分積分すると

$$f(x) = f(a) + (x-a)f'(a) + \frac{f''(a)}{2!}(x-a)^2 + \int_a^x \frac{(x-t)^2}{2!} f'''(t)dt$$

となります．これを続けると

$$f(x) = \sum_{i=0}^{k} \frac{f^{(i)}(a)}{i!}(x-a)^i + \int_a^x \frac{f^{(k+1)}}{k!}(x-t)^k dt \tag{3.11}$$

を得ます．最後の項が剰余項です．実際には，ベルヌーイによるテイラー多項式では，剰余項を具体的に計算することがやっかいな場合もあり，実用性は，これから示すテイラーの定理より劣ります．

[2]数学史的には，18 世紀の前半にテイラーやマクローリンといった人が考えたことになっていて，級数展開に彼らの名前がつけられています．実際には，もっと多くの人が同じ発想に達していたようです．実際，ヨハン・ベルヌーイは 1694 年に部分積分の繰り返し適用により，剰余項を含むテイラー多項式をテイラーよりもはやく導出しています．(後述)

演習 18. 上の *(3.11)* を $f(x)=e^x, a=0$ に適用して k 項まで展開してみましょう．また，剰余項を具体的に評価してみましょう．*(*最初は $k=2$ くらいで考えてみるとわかります．*)*

定理 10 (テイラーの定理). 関数 $f:\mathbf{R}\to\mathbf{R}$ を $f\in C^m$，$a\in\mathbf{R}$ とします．任意の $x\neq a$ に対して，a と x を端点とする開区間 (x,a)（あるいは (a,x)）に属する点 c が存在して

$$f(x)=f(a)+\sum_{i=1}^{m-1}\frac{f^{(i)}(a)}{i!}(x-a)^i+\frac{f^{(m)}(c)}{m!}(x-a)^m$$

定理の証明の前に，証明に使う一般化された平均値の定理を示しておきます.

定理 11 (一般化された平均値の定理). 関数 $f:\mathbf{R}\to\mathbf{R}$ と $g:\mathbf{R}\to\mathbf{R}$ を $f,g\in C^1$，$a<b$ とします．$c\in(a,b)$ が存在して

$$f'(c)(g(b)-g(a))=g'(c)(f(b)-f(a))$$

(証明) $h(x)=f(x)(g(b)-g(a))-g(x)(f(b)-f(a))$ とおいて，定理 6 を適用します.（証明おわり）

(テイラーの定理の証明) 一般性を失うことなく $a<x$ として証明します．$g(x)=(x-a)^m$ と定義します.

$$\phi(t)=f(t)+\sum_{i=1}^{m-1}\frac{f^{(i)}(a)}{i!}(x-a)^i$$

$$\eta(t)=g(t)+\sum_{i=1}^{m-1}\frac{g^{(i)}(a)}{i!}(x-a)^i$$

とおきます．この二つに一般化された平均値の定理 11 を適用すると，$c\in(a<x)$ が存在して

$$\phi'(c)(\eta(b)-\eta(a))=\eta'(c)(\phi(b)-\phi(a))$$

となりますが，$\phi(x)=f(x),\eta(x)=g(x)$ ですから

$$\phi'(c)(g(x)-\eta(a))=\eta'(c)(f(x)-\phi(a)) \tag{3.12}$$

一方

$$\begin{aligned}\phi'(t) &= f'(t)+\sum_{i=1}^{m-1}\left\{\frac{f^{(i+1)}(t)}{i!}(x-t)^i-\frac{f^{(i)}(t)}{(i-1)!}(x-t)^{i-1}\right\}\\ &= \frac{f^{(m)}(t)}{(m-1)!}(x-t)^{m-1} \qquad (3.13)\end{aligned}$$

同様にして

$$\eta'(t)=\frac{g^{(m)}(t)}{(m-1)!}(x-t)^{m-1} \tag{3.14}$$

(3.12),(3.13),(3.14) より

$$\frac{(x-c)^{m-1}}{(m-1)!}f^{(m)}(c)(g(x)-\eta(a)) = \frac{(x-c)^{m-1}}{(m-1)!}g^{(m)}(c)(f(x)-\phi(a)) \tag{3.15}$$

ところが,

$$g(a)=0,\ (\forall k\in\{1,2,\cdots,m-1\})g^{(k)}(a)=1$$

より $\eta(a)=0$ がわかります.

$$(\forall t\in\mathbf{R})\ g^{(m)}(t)=m!$$

が成立することを使って，(3.15) を整理すると

$$f(x)=\phi(a)+\frac{f^{(m)}(c)}{m!}(x-a)^m$$

が得られますが，これはテイラー多項式そのものです.（証明おわり）

演習 19. *1.* $f(x)=\cos x$ に対して，$a=0,m=2$ として，定理 *1* を適用してみてください.

2. $f(x)=\log x$ に対して，$a=1,m=2$ として，定理 *1* を適用してみてください.

3. $f(x)=e^{-x^2}$ に対して，$a=0,m=5$ として，定理 *1* を適用してみてください.

つぎの形のテイラーの定理もよく使われます.

系 1. 関数 $f:\mathbf{R}\to\mathbf{R}$ を $f\in C^m$，$a\in\mathbf{R}$ とします．任意の $x\neq a$ に対して，$\theta\in(0,1)$ が存在して

$$f(x)=f(a)+\sum_{i=1}^{m-1}\frac{f^{(i)}(a)}{i!}(x-a)^i+\frac{f^{(m)}(a+\theta(x-a))}{m!}(x-a)^m$$

3.6 多変数関数の微分

多変数関数

$$f:\mathbf{R}^n\longmapsto\mathbf{R}$$

の $\boldsymbol{a}\in\mathbf{R}^n$ における微分 $\boldsymbol{b}\in\mathbf{R}^n$ を，定理 2 にならって,

$$\lim_{||\boldsymbol{h}||\to 0}\frac{f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})-\boldsymbol{b}\cdot\boldsymbol{h}}{||\boldsymbol{h}||}=0 \tag{3.16}$$

をみたす n 次元ベクトルとして定義します．ここで $\boldsymbol{b}\cdot\boldsymbol{h}=\sum_{i=1}^n b_ih_i$ という内積を $\mathbf{R}^n\longmapsto\mathbf{R}$ という線形写像と考えてください．そうすると，1 次元の場合の微分係数との形式的な類似性が自然に理解できます.

式 (3.18) の極限が存在するような $\boldsymbol{b}$ が見つかるとき，f は $\boldsymbol{a}$ において微分可能，あるいは全微分可能といいます．さらにすべての $\boldsymbol{a}\in\mathbf{R}^n$ で (全) 微分可能であるとき，単に (全) 微分可能であるといい．$\boldsymbol{x}$ に対して微分 $\boldsymbol{b}$ を対応させる写像を f の導関数といいます．導関数が連続であるとき，1 階連続微分可能といい f は C^1 級とよびます.

多変数関数の微分を，上のように定義するとき，1 変数関数の場合の自然な拡張になるのはよいのですが，例えば関数 $f(x_1, x_2) = x_1^{x_2}$ が与えられたとき，点 $(a, b) \in \mathbf{R}^2$ における微分を，定義のみに基づいて計算するのは大変です.（うそだと思うならやってみましょう.）

どうして大変かというと，微分の定義において $\boldsymbol{h}$ のとり方が，1 次元の場合にくらべてはるかに自由度が高いからです.（1 次元の場合は，実数直線上を a に向かって増やすか，減らすかだけでした.）

そこで偏微分を考えて，微分をもとめる手がかりにするのです.

3.7 偏微分・偏導関数

多変数関数 $f(x_1, x_2, \cdots, x_n)$ を考えるとき，$\boldsymbol{a} = (a_1\ a_2\ \cdots\ a_n) \in \mathbf{R}^n$ における i 変数についての偏微分を

$$\lim_{h \to 0} \frac{f(a_1, a_2, \cdots, a_{i-1}, a_i + h, a_{i+1}, \cdots, a_n) - f(a_1, a_2, \cdots, a_n)}{h} \tag{3.17}$$

という極限値で定義します．この極限値が存在するとき，f は i 変数に関して偏微分可能であるといいます．また，この極限値を

$$\frac{\partial f}{\partial x_1}(\boldsymbol{a})$$

のように書きます．また，すべての $\boldsymbol{a} \in \mathbf{R}^n$ について微分可能であるとき

$$\frac{\partial f}{\partial x_i} : \mathbf{R}^n \longmapsto \mathbf{R}$$

を (i 変数に関する) 偏導関数とよびます.

さらに，すべての i について偏導関数が連続関数であるとき f を C^1 級ということにします.

次の定理が成立します.

定理 12. 多変数関数 $f(x_1, x_2, \cdots, x_n)$ が $\boldsymbol{a} \in \mathbf{R}^n$ において *1* 階微分可能であるとき，f は $\boldsymbol{a} \in \mathbf{R}^n$ において，任意の x_i について偏微分が可能で，微分 $\boldsymbol{b}$ は

$$b_i = \frac{\partial f}{\partial x_i}(\boldsymbol{a}),\ (i = 1, 2, \cdots, n)$$

とあらわされる.

(証明) 多変数関数の微分の定義式 (3.18) にあらわれる $\boldsymbol{h} = (0, \cdots, 0, h_i, 0, \cdots, 0) \in \mathbf{R}^n$ という，第 i 要素のみが非ゼロのものを考えて，微分の第 i 要素を評価すると，それが偏微分に等しいことがわかる.（証明おわり）

注意 13. 定理 *12* があるために，多変数関数の微分の計算の手間が軽減されるのです．細かいことを言うようですが，定理 *12* は,「全微分可能ならば偏微分可能」を主張しています.「偏微分可能ならば全微分可能」というほうは，無条件には成立しません．ただし，経済学部であつかうようなほとんどの関数では，反例にあたるものに遭遇することはありません.

偏微分からなるベクトル

$$\left(\frac{\partial f}{\partial x_1}(\boldsymbol{a})\ \frac{\partial f}{\partial x_2}(\boldsymbol{a})\ \cdots\ \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \right)$$

を勾配ベクトル (gradient vector) などとよび，

$$\frac{\partial f}{\partial \boldsymbol{x}}(\boldsymbol{a}),\quad \nabla f(\boldsymbol{a})$$

などと記します．

多変数関数の微分についても，１変数関数の場合同様の基本的な計算規則が従います．

定理 13 (微分法の基本規則１)． 関数 $f:\mathbf{R}^n\to\mathbf{R}$ が $a\in\mathbf{R}$ と関数 $g:\mathbf{R}^n\to\mathbf{R}$ が $a\in\mathbf{R}$ が微分可能，$\lambda\in\mathbf{R}$ ならば，

1. $h:\boldsymbol{x}\longmapsto f(\boldsymbol{x})+g(\boldsymbol{x})$ の導関数は $\nabla h(\boldsymbol{x})=\nabla f(\boldsymbol{x})+\nabla g(\boldsymbol{x})$
2. $h:\boldsymbol{x}\longmapsto \lambda f(\boldsymbol{x})$ の導関数は $\nabla h(\boldsymbol{x})=\lambda\nabla f(\boldsymbol{x})$

高階の偏導関数を考えることができます．たとえば，$\frac{\partial f}{\partial x_i}$ の k 変数に関する偏導関数を考えることができて，実際にこの偏導関数が構成できるとき 2 階偏微分可能といいます．

$$\frac{\partial^2 f}{\partial x_i\partial x_k}(\boldsymbol{x})$$

と記します．この偏導関数を ik 要素とする n 次正方行列を $D^2f(\boldsymbol{x})$ あるいは $\nabla^2 f(\boldsymbol{x})$ と記し，ヘッセ行列とよびます．

注意 14． $\frac{\partial^2 f}{\partial x_i\partial x_k}$ と $\frac{\partial^2 f}{\partial x_k\partial x_i}$ は $i\neq k$ のとき，本来は異なったものです．ただし，ある条件の下で等しくなります．

$\boldsymbol{a}$ において，$\frac{\partial^2 f}{\partial x_i\partial x_k}$ が連続関数であるとき $\boldsymbol{a}$ において 2 階連続微分可能といい，すべての $\boldsymbol{a}\in\mathbf{R}^n$ で 2 階連続微分可能であるとき C^2 級であるといいます．

証明抜きで以下の定理をあげておきます．

定理 14． 多変数関数 $f(x_1,x_2,\cdots,x_n)$ が *2* 階連続偏微分可能であるとき

$$\frac{\partial^2 f}{\partial x_i\partial x_k}=\frac{\partial^2 f}{\partial x_k\partial x_i},\quad (i=1,2,\cdots,n),(k=1,2,\cdots,n)$$

これにより，C^2 級のときヘッセ行列は対称行列になります．

演習 20． 次の *2* 変数関数の，*1* 階の偏導関数，*2* 階の偏導関数をすべて求めなさい．

1. $\cos x_1\sin x_2$
2. $\frac{x^1x^2}{x^1+x^2}$
3. $x_1^{x_2}$

多変数関数 $f:\mathbf{R}^n\to\mathbf{R}$ をこれまで考えてきましたが，こうした多変数関数を m 個，

$$(f_1(\boldsymbol{x})\ f_2(\boldsymbol{x})\ \cdots\ f_m(\boldsymbol{x}))$$

と並べることで m 次元ベクトル値写像 $\boldsymbol{f}:\mathbf{R}^n\to\mathbf{R}^m$ を構成することができます．(全) 微分に関しても，多変数関数の場合と同様

$$\lim_{||\boldsymbol{h}||\to 0}\frac{||\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-B\boldsymbol{h}||}{||\boldsymbol{h}||}=0 \tag{3.18}$$

をみたす $m \times n$ 行列 B として定義します．ここで $B\boldsymbol{h} = \sum_{i=1}^{n} b_i h_i$ という内積を $\mathbf{R}^n \longmapsto \mathbf{R}^m$ という線形写像と考えてください．そうすると，多変数関数の場合の微分との形式的な類似性が自然に理解できます.

また，連続微分についてもこれまで同様に定義できて，C^1 級のとき $\boldsymbol{a}$ における微分 B は

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\boldsymbol{a}) \end{pmatrix}$$

と表せます．この B をヤコビ行列とよび $D\boldsymbol{f}(\boldsymbol{a})$ と記します.

注意 15. $\boldsymbol{f}$ が *1* 次元ベクトル値 $\boldsymbol{f}(\boldsymbol{x}) = (f(\boldsymbol{x}))$ のとき，つまり事実上，実数値関数 $f(\boldsymbol{x})$ のとき，$D\boldsymbol{f}(\boldsymbol{x})$ と $\nabla f(\boldsymbol{x})$ は同じものとかんがえてください.

定理 15 (多変数関数の鎖法則). *2* つの C^1 級の写像 $\boldsymbol{f} : \mathbf{R}^n \to \mathbf{R}^m$ と $\boldsymbol{g} : \mathbf{R}^m \to \mathbf{R}^\ell$ を考えるこのとき，$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))$ という合成写像として $\boldsymbol{h} : \mathbf{R}^n \to \mathbf{R}^\ell$ を構成し，$\boldsymbol{a} \in \mathbf{R}^n$ として $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{a})$ と記す．すると

$$D\boldsymbol{h}(\boldsymbol{a}) = D\boldsymbol{g}(\boldsymbol{b})D\boldsymbol{f}(\boldsymbol{a})$$

上式の右辺は，$\ell \times m$ 行列と $m \times n$ 行列の積です.

3.8 多変数関数のテイラーの定理

さて，1 変数のテイラーの定理を多変数関数の場合に拡張することができます．3 次の項以上を含む展開の部分を「きれい」に表現するにはテンソルの知識がいるので，1 次と 2 次の項を含むものに限定して紹介します．経済学の応用上では，以下に示すバージョンで十分です.

定理 16. 多変数関数 $f : \mathbf{R}^n \to \mathbf{R}$ を C^1 級とします．$\boldsymbol{a} \in \mathbf{R}^n$ とします．このとき，任意の $\boldsymbol{x} \neq \boldsymbol{a}$ に対して $\theta \in (0, 1)$ が存在して，

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a} + \theta(\boldsymbol{x} - \boldsymbol{a})) \cdot (\boldsymbol{x} - \boldsymbol{a})$$

さらに，f が C^2 級のとき，任意の $\boldsymbol{x} \neq \boldsymbol{a}$ に対して $\theta \in (0, 1)$ が存在して，

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}\, {}^t(\boldsymbol{x} - \boldsymbol{a}) \nabla^2 f(\boldsymbol{a} + \theta(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a})$$

(証明) $g(t) = f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a}))$ とすると，定理 15 によって g は C^1 級です．微分を計算すると

$$g'(t) = \nabla f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})) \cdot (\boldsymbol{x} - \boldsymbol{a}) \tag{3.19}$$

一方，テイラーの定理を $t = 0$ の近傍で g に適用すると，$\theta \in (0, 1)$ が存在して

$$g(t) = g(0) + g'(\theta t)t$$

なので

$$g(t) = g(0) + t\nabla f(\boldsymbol{a} + \theta t(\boldsymbol{x} - \boldsymbol{a})) \cdot (\boldsymbol{x} - \boldsymbol{a})$$

$t = 1$ とおいて，g の定義にもどると

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a} + \theta(\boldsymbol{x} - \boldsymbol{a})) \cdot (\boldsymbol{x} - \boldsymbol{a})$$

が得られます.

2 階のテイラーの定理についても同様に,

$$g''(t) = {}^t(\boldsymbol{x} - \boldsymbol{a})\nabla^2 f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}) \tag{3.20}$$

とテイラーの定理を g に $t = 0$ の近傍に適用してえられる

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\theta t)t^2$$

から

$$g(t) = g(0) + t\nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}{}^t(\boldsymbol{x} - \boldsymbol{a})\nabla^2 f(\boldsymbol{a} + \theta t(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a})\ t^2$$

となります．$t = 1$ とおいて g の定義にもどると

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}\ {}^t(\boldsymbol{x} - \boldsymbol{a})\nabla^2 f(\boldsymbol{a} + \theta(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a})$$

がえられます.（証明おわり）