経済数学

2018

牧野圭吾

0. 集合などの記法

集合 set とは，**元 element** の集まりである．**数学的対象物**(数とは限らない)が元となりうる．ある物 a が集合Aの要素であることを$a \in A$, ある物 a が集合Aの要素でないことと$a \notin A$と記す．1 つも元を持たない集合を**空集合 empty set** といい，$\emptyset$と記す．集合の記法は 2 つある．①**外延的記法**：$A=\{a_1, a_2, \cdots, a_n\}$といったものと②**内包的記法**：$A = \{a \in \boldsymbol{R}: 1 \leq a \leq 7\}$といったものである．2 つのもの$x, y$が同じものなら$x = y$, 異なるものなら$x \neq y$と記す．

$\forall x$は任意のxについて,$\exists x$はあるxが存在して，という意味である．

2 つの集合A, Bが存在するとしよう．①$\forall x, x \in A \Rightarrow x \in B$であるとき$A$は$B$の**部分集合 subset** であるといい,$A \subset B$と記す．$A \subset B$かつ$B \subset A$ であるとき,$A = B$と記す．②$A \cap B = \{x: x \in A \; and \; x \in B\}$と記し，**共通部分 intersection** という．③$A \cup B =\{x: x \in A \; or \; x \in B\}$と記し，**和集合 union** という．④$A \setminus B = \{x: x \in A \; and \; x \notin B\}$と記し，集合の**差**という．特に$B \subset A$のとき, A を**全体集合**といい,$A \setminus B$をA^cと記し**補集合 complement** という．

元が全てそれ自身集合であるような集合(「集合の集合」)を**集合族 family of sets** という．任意の集合Xについて,Xの部分集合全体がつくる集合族を特に**冪集合 power set** という．添字の集合をIとして，各$i \in I$に対して集合X_iが定まるとする．$(X_i)_{i \in I}$を**添字付けられた集合族 indexed family of sets** という．特に$\forall i \in I, X_i \in X$となる集合族$(X_i)_{i \in I}$を部分集合族という．冪集合も部分集合族の一種であることに注意せよ．$\bigcup_{i \in I} X_i = \{x \in X, \exists i \in I, x \in X_i\}$, $\bigcap_{i \in I} X = \{x \in X, \forall i \in I, x \in X_i\}$と記す．

自然数全体の集合を$\boldsymbol{N} = \{1, 2, 3, \cdots\}$とする．整数全体の集合を$\boldsymbol{Z} = \{\cdots, -1, 0, 1, \cdots\}$とする．整数$m$と 0 でない整数$n$をもちいて$\frac{m}{n}$と表現される数を有理数といい，有理数全体の集合を$\boldsymbol{Q}$と記す．有理数に無理数を併せたものを**実数 real number** といい，実数全体の集合を$\boldsymbol{R}$と記す．実数は，本来は代数の公理，順序の公理，連続の公理を満たすものとして定義される．$\boldsymbol{R}_+ = \{x \in \boldsymbol{R}: x \geq 0\}, \boldsymbol{R}_{++} = \{x \in \boldsymbol{R}: x > 0\}$と記す．

1. 関数

定義 1.1　2 つの集合X, Yが与えられたとき,Xの各要素にYの要素を(複数でも)対応付ける規則を，集合Xから集合Yへの**対応 correspondence** という．

定義 1.2　2 つの集合X, Yが与えられたとき,Xの各要素にYの要素を唯一つ unique に対応付ける規則fを，集合Xから集合Yへの**関数 function** という．

$$f : X \to Y$$

と記す．$x \in X$について$f : x \mapsto f(x)$とも記す．Xを**定義域 domain of definition**, $\{y \in Y: \exists x \in X, f(x) = y\} \subset Y$を**値域 range** という．

定義 1.3　$x \in X, y \in Y$を満たす**順序対 ordered pair**(x, y)の全体を,XとYの**直積集合 Cartesian product** といい,$X \times Y$と記す．

定義 1.4　$f : X \to Y$について,$G_f = \{(x, f(x)): x \in X\}$を関数$f$の**グラフ graph** という．

定義 1.5　$f : X \to Y, g : Y \to Z$とする．$g \circ f :\ X \to Z$を**合成関数 composite function** という．$g \circ f : x \mapsto g(f(x))$とも記す．

定義 1.6　$f : X \to Y$とする．

(a) $x_1, x_2 \in X$ and $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$のとき,$f$は**単射 injection** である.

(b) $y \in Y \Rightarrow \exists x, f(x) = y$のとき,$f$は**全射 surjection** である

(c) 単射かつ全射であるとき, **全単射 bijection** である.

定義 1.7 $f: X \to Y$が全単射であるとき,$\forall y \in Y, \#\{x \in X : f(x) = y\} = 1$となり, **逆関数 inverse function**

$$f^{-1}: Y \to X$$

が定まる.

2. ユークリッド空間と位相

定義 2.1 集合$\boldsymbol{R}^n$を,$\boldsymbol{R}$ の n 個の直積$\boldsymbol{R} \times \boldsymbol{R} \times \cdots \times \boldsymbol{R}$とし, **実$\boldsymbol{n}$次元数空間 real n-space** という. $\boldsymbol{R}^n$の元$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$を**$\boldsymbol{n}$次元ベクトル n-dimensional vector** という. ベクトルは順序対である. 紙幅の都合からベクトルは$x = (x_1, x_2, \cdots, x_n)$と略記される.

定義 2.2 2 つのn次元ベクトル$x = (x_1, x_2, \cdots, x_n)$と$y = (y_1, y_2, \cdots, y_n)$について,

(a) $x + y = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$とする.

(b) $\alpha x = (\alpha x_1, \alpha x_2, \cdots, \alpha x_n)$. 但し$\alpha \in \boldsymbol{R}$

(c) $x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$とする演算を**内積** inner product という.

(c)のように内積を定められた$\boldsymbol{R}^n$を**$\boldsymbol{n}$次元ユークリッド空間 n-dimensional Euclidian space** という. $x \in \boldsymbol{R}^n$の長さ(**ノルム norm**)を$\sqrt{x \cdot x} = |x|$とする. xとyのなす角をθとするとき,$x \cdot y = |x||y|\cos\theta$が成り立つ(幾何学的な意味を考察せよ).

定義 2.3 $a \in \boldsymbol{R}^n, r \in \boldsymbol{R}_{++}$について,$B(a; r) = \{x \in \boldsymbol{R}^n : |x - a| < r\}$とし, **n 次元開球 open n-ball** という.

定義 2.4 ある$a \in S \subset \boldsymbol{R}^n$について,$\exists r \in \boldsymbol{R}_{++}, B(a; r) \subset S$となるとき,$a$は$S$の**内点 interior point** であるという. 集合Sの内点の集合を**内部 interior** といい,$\operatorname{int} S$と記す.

定義 2.5 $\operatorname{int} S = S$となる集合Sを**開集合 open set** という.

定義 2.6 集合Sの補集合$S \setminus \boldsymbol{R}^n = S^c$が開集合のとき,$S$は閉集合であるという.

定義 2.7 $S \subset \boldsymbol{R}^n$と$a \in \boldsymbol{R}^n$について,$\forall \varepsilon > 0, B(a; \varepsilon) \cap S \neq \emptyset$となるとき,$a$は$S$の**接触点 adherent point** または**触点**という. 触点全体の集合を**閉包 closure** といい,$\bar{S}$と記す. 一般に$S \subset \bar{S}$.

命題 2.8 $S = \bar{S} \Leftrightarrow S$は閉集合.

定義 2.9 一次元ユークリッド空間$\boldsymbol{R}$において,$a, b \in \boldsymbol{R}$とする. $(\mathrm{a}, \mathrm{b}) = \{x \in \boldsymbol{R} : \mathrm{a} < \mathrm{x} < \mathrm{b}\}$を開区間という. $[\mathrm{a}, \mathrm{b}] = \{x \in \boldsymbol{R} : \mathrm{a} \leq \mathrm{x} \leq \mathrm{b}\}$を閉区間という.

定義 2.10 ある点$a \in \boldsymbol{R}^n$と正数$r \in \boldsymbol{R}_{++}$によって$S \subset \boldsymbol{R}^n$が$S \subset B(a; r)$とできるとき,$S$は**有界 bounded** であるという. Sが有界かつ閉集合であるとき,Sは**コンパクト compact** であるという.

定義 2.11 集合Sについて,$\forall x, y \in S, \forall \alpha \in (0, 1), \alpha x + (1 - \alpha) y \in S$となるとき,$S$は**凸集合 convex set** という.

3. 行列の演算

定義 3.1 $m \times n$行列Aを,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

とする. i行目j列目にある数をa_{ij}としている. $m \times 1$行列はベクトルである.

定義 3.2　$m \times n$行列A, Bについて

$$A \pm B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \pm \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & \cdots & a_{1n} \pm b_{1n} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & \cdots & a_{2n} \pm b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} \pm b_{m1} & a_{m2} \pm b_{m2} & \cdots & a_{mn} \pm b_{mn} \end{pmatrix}$$

定義 3.3　m行n列の行列($m \times n$行列)Aと$n \times l$行列Bがあるとする.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{pmatrix}$$

$$AB = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1l} \\ c_{21} & c_{22} & \cdots & c_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{ml} \end{pmatrix}$$

このとき, $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$とする．$c_{ij}$は, Aの第i行ベクトルとBの第j列ベクトルの内積であると思えば良い．ABは$m \times l$行列である．$AB \neq BA$などころか, ABが定義できてもBAが定義できないことは多い(この定義をよく確認せよ).

定義 3.4

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

に対して$n \times m$の**転置行列 transposed matrix**A^Tを以下の様に定める.

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

定義 3.5　列数と行数が一致する$n \times n$行列を**正方行列 square matrix** という．$A^T = A$のとなる行列は**対称行列 symmetric matrix** という.

定義 3.6　Aを$n \times n$行列とする．任意の$h \in \boldsymbol{R}^n, h \neq 0$について

(a) $h^T A h < 0$ならばAは**負値定符号 negative definite** という.

(b) $h^T A h > 0$ならばAは**正値定符号 positive definite** という.

(c) $h^T A h \leq 0$ならばAは**負値半定符号 negative semidefinite** という.

(d) $h^T A h \geq 0$ならばAは**正値半定符号 positive semidefinite** という.

定義 3.7　$1, 2, \cdots, n$を並べ替える操作を**置換 permutation** といい，σで表す．$\sigma(1) = i_1, \sigma(2) = i_2, \cdots$であるとき,

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

と記す. n 文字の置換全体を$\mathrm{S_n}$と記す．文字も動かさないとき単位置換1_nという．逆変換は逆置換σ^{-1}という.

定義 3.8　2 つの置換σ, τの合成変換を**積**といい，$\tau\sigma$と記す．一般には$\tau\sigma \neq \sigma\tau$.

命題 3.8　$(\sigma\tau)\rho = \sigma(\tau\rho), 1_n\sigma = \sigma 1_n, \sigma\sigma^{-1} = \sigma^{-1}\sigma = 1_n$は明らか.

定義 3.9 n文字の置換で 2 文字のみを入れ替え他の$n-2$文字を動かさないものを**互換**という.

定理 3.10　任意の置換σは互換の積で表されるが，必要な互換の数は一意ではない．しかし，互換の個数が偶数か奇数であるかはσのみに依存して決定される．証明は齋藤（1966）p.76．偶数個の時は偶置換，奇数個の時は奇置換といい，それぞれに$\mathrm{sgn}\sigma = +1, \mathrm{sgn}\sigma = -1$とする$\mathrm{sgn}$を定義する.

定義 3.11 $n \times n$行列Aについて，行列式 determinant は以下のように定まる．

$$\det A = |A| = \sum_{\sigma \in S_n} \mathrm{sgn}\sigma \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

定義 3.12 Aを$n \times n$行列とする．**首座小行列 naturally ordered principal minor**A_kを以下のように定義する．

$$\mathrm{A_k} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

定理 3.13 Aを対称な$n \times n$行列とする．

(a) $|A_k| > 0 \ \ for\ k = 1, \cdots, n$と，$A$が正値定符号であることは同値.

(b) $(-1)^k |A_k| < 0 \ \ for\ k = 1, \cdots, n$と，$A$が負値定符号であることは同値.

証明は Debreu (1952).

4. 連続

定義 4.1 $A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}^m$とする．このとき,$f$は$n$変数$m$次元ベクトル値関数という．特に$m = 1$のとき実数値関数という．

定義 4.2 $A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}^m$として,$a \in \bar{U}, b \in \boldsymbol{R}^m$とする．$\boldsymbol{x}$**が**$\boldsymbol{a}$**に近づくときの**$\boldsymbol{f(x)}$**の極限**が$b$であるとは,

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in A, |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon$$

であることを言う．これをイプシロン―デルタ論法という．

$$\lim_{x \to a} f(x) = b$$

$$f(x) \to b \ (x \to a)$$

などと記す.

定義 4.3 $A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}^m$とする．$a \in A$について

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in A, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$$

となるとき,fはaで**連続 continuous** であるという．

$$\lim_{x \to a} f(x) = f(a)$$

とも記す．任意の$a \in A$で連続なとき，単にfは連続であるという．

定義 4.4 $A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}^m, a \in \bar{A}$について,$\lim_{x \to a} f(x) = 0$であることを,$f$は$a$において**無限小**であるという．$\lim_{x \to a} f(x) = +\infty$であることを,$f$は$a$において**無限大**であるという．

定義 4.5 $A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}$とし,$A \subset T \subset \boldsymbol{R}^n, g : T \to \boldsymbol{R}$とし,$a \in A$について$\exists \varepsilon > 0, \forall x \in U, 0 < \|x - a\| < \varepsilon \Rightarrow g(x) \neq 0$とする．$f$が

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0$$

を満たすとき,aにおいてfはgに比べて**無視できる**という．aにおいてgに比べ無視できる任意の関数を一般に$o(g)(x \to a)$と記す．f, gがaにおいて共に無限小であるとき,fはgより**高位の無限小**であるという．

5. 一変数関数の微分法

定義 5.1 $\boldsymbol{R}$の開区間I上で定義される関数$f : I \to \boldsymbol{R}^n$があるとする．$t \in I$に対し，極限

$$\lim_{h \to 0} \frac{f(t + h) - f(t)}{h} = c$$

が存在するとき,fはtで**微分可能**であるといい,cをfのtにおける**微分係数 derivative** という．このとき,

$$c = f'(t) = \frac{df}{dt}(t) = Df(t) = (f(t))'$$

などと表記する．Iの各点で微分可能なとき，$I \to \boldsymbol{R}^n$である関数$t \mapsto f'(t)$をfの**導関数 derivative** という．

命題 5.2　開区間$I, f : I \to \boldsymbol{R}^n$, $t \in I, f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$が微分可能であるなら$f'(t) = \begin{pmatrix} f_1{}'(t) \\ \vdots \\ f_n{}'(t) \end{pmatrix}$.

$f'(t)$は**接ベクトル**ともいう．

定義 5.3　$f : I \to \boldsymbol{R}^n$の導関数$f'$が$I$で微分可能であるとき，$f'$の導関数である$(f')'$が定義される．これを$f$の**二階導関数 second derivative** といい，f''で表す．帰納的に，k階導関数$f^{(k)}$が定義され微分可能のとき，$f^{(k)}$の導関数として$k+1$階導関数$f^{(k)'} = f^{(k+1)}$が定義される．

定義 5.4　$k \in \boldsymbol{N}$とする．$f : I \to \boldsymbol{R}^n$が$I$で$k$階までの導関数が存在して，$f^{(k)}$が連続であるとき，$f$は$I$で$\boldsymbol{C^k}$**級($\boldsymbol{k}$回連続微分可能)**であるという．任意回数微分可能な関数をIで$\boldsymbol{C^\infty}$**級**であるという．

定理 5.5 (Taylor)　$n \in \boldsymbol{N}$とする．区間$[a, x]$(または$[x, a]$)$= I$でn回微分可能な実数値関数$f : I \to \boldsymbol{R}$について

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

となる$c \in \mathrm{int}I$が存在する．更に，

$$f(x) = f(a) + \sum_{i=1}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i + o((x-a)^n), \qquad (x \to a)$$

となる．$x - a = h$とすれば，

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{f^{(i)}(a)}{i!}h^i + o(h^n), \qquad (h \to 0)$$

となる．

最適化で特に重要なのは，

$$f(a+h) = f(a) + f'(a)h + o(h) \qquad (\text{一次近似公式})$$

$$f(a+h) = f(a) + f'(a)h + \frac{f''(a)}{2}h^2 + o(h^2) \qquad (\text{二次近似公式})$$

の 2 つである．

6. 多変数関数の微分

定義 6.1　$A \subset \boldsymbol{R}^n, f : A \to \boldsymbol{R}^m, a \in A$について，任意の$n$次元ベクトル$e \in \boldsymbol{R}^n$に対し実変数$t \in \boldsymbol{R}$の関数

$$g(t) = f(a + te)$$

がある正数$\varepsilon > 0$によって定まる近傍$(-\varepsilon, \varepsilon)$で定義されているとしよう．この関数$g(t)$が$t = 0$で微分可能なら，関数$f$は$a$において$\boldsymbol{e}$**方向に微分可能**であるといい，$g'(0)$を$f$の$a$における$\boldsymbol{e}$**方向の微分係数**といい，

$$g'(0) = D_e f(a) = \frac{\partial f}{\partial e}(a)$$

と記す．

命題 6.2　$D_e f(a) = \begin{pmatrix} D_e f_1(a) \\ \vdots \\ D_e f_m(a) \end{pmatrix}$が命題 5.2 と同様に成り立つ．

定義 6.3　$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$を$\boldsymbol{R}^n$の自然基底という．$f : A \to \boldsymbol{R}^m$が$a \in A$において$e_i$方向に微分可能であるとき，つまり極限

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

が存在するとき,fはaにおいて第i座標x_iについて**偏微分可能**であるという．$D_{e_i}f(a)$を偏微分係数といい,

$$D_{e_i}f(a) = \frac{\partial f}{\partial x_i}(a) = f_{x_i}(a) = D_i f(a)$$

等と記す．Aの各点で微分可能なとき，関数$\frac{\partial f}{\partial x_i}: a \mapsto \frac{\partial f}{\partial x_i}(a)$を$f$の**偏導関数**という．

定義 6.4　偏導関数$\frac{\partial f}{\partial x_i}: a \mapsto \frac{\partial f}{\partial x_i}(a)$が第$j$座標$x_j$に関し偏微分可能ならば，二階偏導関数

$$\begin{cases} \dfrac{\partial^2 f}{\partial x_j \partial x_i}, & (i \neq j) \\ \dfrac{\partial^2 f}{\partial x_i^2}(a), & (i = j) \end{cases}$$

が定義される．$f_{ij}(x), f_{ii}(x)$とも記す。同様に高階偏導関数も定義される．/

定義 6.5　$k \in \boldsymbol{N}$とする．開集合$U \subset \boldsymbol{R}^n, f : U \to \boldsymbol{R}^m$が$k$階までの全ての偏導関数が存在して,$U$上連続であるとき,$f$は$A$で$\boldsymbol{C^k}$**級($\boldsymbol{k}$回連続微分可能)**であるという．任意回数微分可能な関数を$\boldsymbol{C^\infty}$**級**であるという．

定理 6.6 (Young)　$f : U \to \boldsymbol{R}^m$が$C^k$級ならば,$f$の$k$階までの全ての偏導関数は偏微分の順序によらない．

定義 6.7　開集合$U \subset \boldsymbol{R}^n$で定義される$f : U \to \boldsymbol{R}^m$が$x \in U$で**微分可能**であるとは,$h \in \boldsymbol{R}^n$とし，ある$m \times n$行列$M$が存在して

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Mh}{|h|} = 0$$

となることである．$f(x+h) - f(x) = Mh + o(|h|)$と同値である．$M$を微分係数といい,$\mathrm{M} = f'(x)$などと記す．

命題 6.8　$f'(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$. ヤコビ行列 **Jacobian matrix** と呼ばれる．

定理 6.9　$f'(x)$がC^1級$\Leftrightarrow$ fは微分可能で$f'(x)$は連続

定理 6.10（多変数の Taylor）$U \subset \boldsymbol{R}^n, f : U \to \boldsymbol{R}$が$C^k$級とする．二点$x, x+h$を結ぶ線分が$U$に含まれているとすれば,$0 < \theta < 1$となる実数$\theta$が存在して

$$f(x+h) = f(x) + \sum_{m=1}^{k-1} \frac{1}{m!}(d^m f)_x(h) + \frac{1}{k!}(d^k f)_{x+\theta h}(h)$$

$$(d^m f)_x(h) = \sum_{1 \le i_1, \cdots, i_m \le n} \frac{\partial^m f}{\partial x_{i_1} \cdots \partial x_{i_m}}(x) h_{i_1} h_{i_2} \cdots h_{i_m}$$

最適化で特に重要なのは,

$$f(x+h) = f(x) + Df(x+\theta h)h$$

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2}h^T D^2 f(x+\theta h)h$$

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2}h^T D^2 f(x)h + o(|h|^2)$$

である．$Df(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$は$\nabla f(x)$とも記す．$D^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$はヘッセ行列 **Hessian matrix** といい, Young の定理(定理 6.6)よりfがC^2級なら$D^2 f(x)$は対称行列となる．

7. 微分公式

一変数実数値関数に関しては，以下のことが知られている．$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$をネイピア数という．$a, \alpha$を定数$, a > 0, a \neq 1$として

$$(x^\alpha)' = \alpha x^{\alpha - 1}, (e^x)' = e^x, (\log_e x)' = (\log x)' = \frac{1}{x}, (\log_a x)' = \frac{1}{x \log_e a}$$

開集合$U \subset \boldsymbol{R}^n$上で定義される$f, g : U \to \boldsymbol{R}^m$が微分可能なとき，

$$\big(f(x) \pm g(x)\big)' = f'(x) \pm g'(x)$$

$$\alpha \in \boldsymbol{R}, \big(\alpha f(x)\big)' = \alpha f'(x)$$

$m = 1$のとき(実数値関数であるとき)は，

$$\big(f(x) \cdot g(x)\big)' = f'(x)g(x) + f(x)g'(x)$$

g(x) $\neq 0$のとき，

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{g(x)f'(x) - f(x)g'(x)}{\big(g(x)\big)^2}$$

合成関数について考える．開集合$U \subset \boldsymbol{R}^m, W \subset \boldsymbol{R}^n$について$f : U \to \boldsymbol{R}^n$，$g : W \to \boldsymbol{R}^l, f(U) \subset W$とする．両関数が微分可能なとき，

$$\Big(g\big(f(x)\big)\Big)' = g'(f(x))f'(x)$$

しばしば**連鎖律 chain rule** という．

定理 7.1 (Euler)　$f : \boldsymbol{R}^n \to \boldsymbol{R}$が定数λと$t \in \boldsymbol{Z}$について

$$f(\lambda x) = \lambda^t f(x)$$

となっているとする．このとき,fは**t次同次hoogeneous of degree t**であるという．すると，

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) x_i = t f(x)$$

が成立する．

以降,Uは凸集合かつ開集合であるとし,$f : U \to \boldsymbol{R}$がC^2級であるとする．但し，指示や特別な文脈がある場合はその限りでない．また，最適化に関する条件は殆どが Taylor の定理(定理 5.5，定理 6.10)から証明されるが，ここでは紙幅の都合から扱わない. Sundaram(1996)参照.

8. 制約式のない最適化

制約式のない最適化問題は

$$\max_{x \in U} f(x)$$

というように定式化される.

定義 8.1　$f(x) \geq f(y)\ for\ all\ y \in U \cap B(x; r)$となる$r > 0$が存在するとき,$x \in U$は$f$の **local maximum** であるという．$f(x) > f(y)$としたとき，**strict local maximum** という．Uでfが最大値をとるような点を **global maximum** という. Minimum も同様に定義される.

定理 8.2 (FOC)　制約式のない場合を考える．$x^* \in U$が local maximum なら$Df(x^*) = 0$となる．これを最適化の**一階条件 First Order Condition** という.

定理 8.3 (SOC)　制約式のないとき，以下の条件を**二階条件 Second Order Condition** という．

(A) fが local maximum をxで持つ⇒ $D^2f(x)$が負値半定符号.

(a) fが local minimum をxで持つ⇒ $D^2f(x)$が正値半定符号.

(B) $Df(x)=0$かつ$D^2f(x)$が負値定符号⇒ xは strict local maximum.

(b) $Df(x)=0$かつ$D^2f(x)$が正値定符号⇒ xは strict local minimum.

9. 制約式のある最適化問題

$h_i\colon \boldsymbol{R}^n\to\boldsymbol{R},\ i\in\{1,\cdots,l\}$について,$h_i(x)\geq 0, i\in\{1,\cdots,l\}$が不等式制約であるとしよう．このとき，最適化問題は,

$$\max_{x\in U} f(x)$$
$$subject\ to\ x\in B=\{x\in U\colon h_i(x)\geq 0, i\in\{1,\cdots,l\}\}$$

というように定式化される.

定理 9.1 (KKT, FOC)　不等号制約式のある最適化問題を考える．$f\colon U\to\boldsymbol{R}, h_i\colon\boldsymbol{R}^n\to\boldsymbol{R},\ i\in\{1,\cdots,l\}$が$C^1$級であるとする．$x^*$が集合

$$B=\{x\in U\colon h_i(x)\geq 0, i\in\{1,\cdots,l\}\}$$

上のfの local maximum であるとする．さらに,$h_i(x)=0$をみたす添え字の全体を$\mathrm{E}\subset\{1,\cdots,l\}$として ,rank$Dh_E(x^*)=\#E$となるとき，以下の条件を満たす$\lambda^*=(\lambda_1^*,\cdots,\lambda_l^*)\in\boldsymbol{R}^n$が存在する.

$$[\mathrm{KT}-1]\quad \lambda_j^*\geq 0\ and\ \lambda_j^*h_j(x^*)=0\ for\ j=1,\cdots,l$$

$$[\mathrm{KT}-2]\quad Df(x^*)+\sum_{j=1}^{l}\lambda_j^*Dh_j(x^*)=0$$

これらの条件は，制約式のある最適化問題の FOC となっている.

10. 凹/凸関数と最適化

定義 10.1　$f: U\to\boldsymbol{R}$について

(a) subgraph sub$f=\{(x,y)\in U\times\boldsymbol{R}\colon f(x)\geq y\}$が凸集合であるとき，**凹関数 concave function** であるという.

(b) epigraph epi$f=\{(x,y)\in U\times\boldsymbol{R}\colon f(x)\leq y\}$が凸集合であるとき，**凸関数 convex function** であるという.

命題 10.2　任意の$x,y\in U$と任意の$\lambda\in(0,1)$について,

(A) fが凹関数である．⇔ $f(\lambda x+(1-\lambda)y)\geq\lambda f(x)+(1-\lambda)f(y)$

(a) fが**狭義凹関数 strictly concave** である．⇔ $f(\lambda x+(1-\lambda)y)>\lambda f(x)+(1-\lambda)f(y)$

(B) fが凸関数である．⇔ $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$

(b) fが**狭義凸関数 strictly convex** である．⇔ $f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y)$

定理 10.3　$f: U\to\boldsymbol{R}$が凹/凸関数で,(ルベーク測度が 0 となる集合上以外では),C^1級である.

定理 10.4　$f: U\to\boldsymbol{R}$について

(A) fが凹関数である．⇔任意の$x\in U$についてヘッセ行列$D^2f(x)$が負値半定符号.

(a) fが狭義凹関数である．⇐任意の$x\in U$についてヘッセ行列$D^2f(x)$が負値定符号.

(B) fが凸関数である．⇔任意の$x\in U$についてヘッセ行列$D^2f(x)$が正値半定符号.

(b) fが狭義凸関数である．⇐任意の$x\in U$についてヘッセ行列$D^2f(x)$が正値定符号.

定理 10.5　fは凹関数とする．制約式のないときのfの local maximum は global maximum である．また,

global maximum となる点の集合は，空集合であるかと凸集合である．とくに狭義凹関数に関しては，空集合か一元集合となる．

定理 10.6　fは凹関数とする．xが制約のないときの maximum となることと,$Df(x)=0$は同値である．

定理 10.7（KKT）　$\max_{x\in U} f(x)\ subject\ to\ x\in B=\{x\in U: h_i(x)\geq 0, i\in\{1,\cdots,l\}\}$という 9 節で考えた問題について,$f:U\to \boldsymbol{R}, h_i:U\to\boldsymbol{R}, i\in\{1,\cdots,l\}$が全て凹関数であるとする．$\mathrm{h_i}(\bar{x})>0\ for\ all\ i\in\{1,\cdots,l\}$となる$\bar{x}$が$U$にあると仮定する．定理 9.1 の[$\mathrm{KT}-1,2$]を満たす$x^*$と$\lambda^*=(\lambda_1^*,\cdots,\lambda_l^*)\in\boldsymbol{R}^n$が存在すると,$x^*$は最適化問題の解である．

11. 準凹/凸関数と最適化

定義 11.1　$f: U\to\boldsymbol{R}$について，任意の$a\in\boldsymbol{R}$に対し

(a) 上位集合 upper-contour set $U_f(a)=\{x\in U: f(x)\geq a\}$が凸集合であるとき，**準凹関数 quasi-concave function** であるという．

(b) 下位集合 lower-contour set $L_f(a)=\{x\in U: f(x)\leq a\}$が凸集合であるとき，**準凸関数 quasi-convex function** であるという．

命題 11.2　任意の$x,y\in U$と任意の$\lambda\in(0,1)$について，

(A) fが準凹関数である．$\Leftrightarrow f(\lambda x+(1-\lambda)y)\geq \min\{f(x),f(y)\}$

(a) fが狭義準凹関数である．$\Leftrightarrow f(\lambda x+(1-\lambda)y)> \min\{f(x),f(y)\}$

(B) fが準凸関数である．$\Leftrightarrow f(\lambda x+(1-\lambda)y)\leq \max\{f(x),f(y)\}$

(b) fが狭義準凸関数である．$\Leftrightarrow f(\lambda x+(1-\lambda)y)< \max\{f(x),f(y)\}$

命題 11.3　$\mathrm{C_k}(x)=\begin{pmatrix} 0 & f_1(x) & \cdots & f_k(x) \\ f_1(x) & f_{11}(x) & \cdots & f_{1k}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_k(x) & f_{k1}(x) & \cdots & f_{kk}(x) \end{pmatrix}, for\ k=1,\cdots,n$とする．

(a) fが準凹関数なら，$(-1)^{\mathrm{k}}|C_k(x)|\geq 0\ for\ k=1,\cdots,n.$

(b) $(-1)^{\mathrm{k}}|C_k(x)|>0\ for\ k=1,\cdots,n$なら，$f$は準凹関数.

定理 11.4　fが狭義準凹関数であるとする．そのとき，制約式のないときのfの local maximum は global maximum になる．また, global maximum となる点は空集合か一元集合である．

定理 11.5（KKT）　$f:U\to\boldsymbol{R}, h_i:U\to\boldsymbol{R}, i\in\{1,\cdots,l\}$が全て準凹関数であるとする．定理 9.1 の[$\mathrm{KT}-1,2$]を満たす$x^*$と$\lambda^*=(\lambda_1^*,\cdots,\lambda_l^*)\in\boldsymbol{R}^n$が存在すると,$x^*$は次の 2 条件のいずれかを満たせば最適化問題の解である．

[$\mathrm{QC}-1$]　$Df(x^*)\neq 0$

[$\mathrm{QC}-2$]　fが凹関数

参考文献

Debreu, G. (1952) Definite and Semidefinite Quadratic Forms, *Econometica* 20 (2), pp.295-300

Sundaram, R. K. (1996), *A First Course in Optimization Theory*. Cambridge University Press.

神谷和也・浦井憲（1996),『経済学のための数学入門』東京大学出版会.

齋藤正彦（1966),『線型代数入門』東京大学出版会.

杉浦光夫（1980),『解析入門 I』東京大学出版会.

二階堂副包（1960),『現代経済学の数学的方法―位相数学による分析入門』.

松坂和夫（1968),『集合・位相入門』岩波書店.