微分積分学第一　V 類 T 組 第 7 回講義内容（2005 年 6 月 9 日）井上淳

1　実数とは何か

2　関数とは何か？

3　1 変数関数の微分法

4　Taylor の定理とその応用（１変数の場合）

5　多変数関数の微分法

5.1　偏微分可能性

5.2　全微分可能性

5.3　合成関数の偏微分

5.4　高階偏導関数

6　Taylor の定理とその応用（2 変数の場合）

6.1　2 変数の Taylor の定理

定理 6.1 (Taylor の定理) 長方形領域 $I \times J$ 上で $f \in C^m(I \times J)$ とすると、

$$f(x,y) = \sum_{j+k=0}^{m-1} \frac{1}{j!k!} \frac{\partial^{j+k} f}{\partial x^j \partial y^k}(a,b)(x-a)^j (y-b)^k + R_m$$

となる。ここで

$$R_m = \sum_{j+k=m} \frac{1}{j!k!}(x-a)^j (y-b)^k \frac{\partial^{j+k} f}{\partial x^j \partial y^k}(a+\theta(x-a), b+\theta(y-b)), \quad (\exists \theta \in (0,1)).$$

6.2　合成関数の微分、偏微分について

6.2.1　１変数の場合（復習）

以下、x の属している区間を I、t の属している区間を D とする。

命題 6.1 $f(x) \in C^1(I),\ \varphi(t) \in C^1(D)$ かつ $\varphi(D) \subset I$ とする。このとき、

$$\frac{d}{dt} f(\varphi(t)) = \dot{f}(\varphi(t))\dot{\varphi}(t).$$

略証：前に述べたの Landau 記号を用いると見通しがよい。微分可能だから

$$f(x+h) - f(x) - \alpha h = o(h), \quad x(t+k) - x(t) - \beta k = o(k),$$

となる $\alpha (= \dot{f}(x)$ 及び $\beta (= \dot{\varphi}(t))$ がある。故に、

$$\begin{aligned} f(x(t+k)) - f(x(t)) &= [f(x(t)) + \alpha(x(t+k) - x(t)) + o(x(t+k) - x(t))] - f(x(t)) \\ &= \alpha[\beta k + o(k)] + o(\beta k + o(k)) = \alpha\beta k + o(k). \quad \square \end{aligned}$$

6.2.2 ２変数の場合

以下、x の属している区間を I、y の属している区間を J、t の属している区間を D 或いは D_t、s の属している区間を D_s とする。

注意：Schwarz の定理によって $f \in C^2(I)$ ならば $\partial_x \partial_y f = \partial_y \partial_x f,\ f_{xy} = f_{yx}$ となる。より一般的に $f \in C^m(I)$ ならばその高階偏微分は、どの独立変数に関し何回偏微分したかには依るが偏微分する順序に依らないから $\partial_x^j \partial_y^k f\ (j + k \le m)$ 等と表記してよい。

命題 6.2 *(i)* $f(x,y) \in C^1(I \times J),\ \varphi(t), \psi(t) \in C^1(D)$ かつ $\varphi(D) \subset I,\ \psi(D) \subset J$ とする。このとき、

$$\frac{d}{dt} f(\varphi(t), \psi(t)) = f_x(\varphi(t), \psi(t))\dot{\varphi}(t) + f_y(\varphi(t), \psi(t))\dot{\psi}(t),$$

(ii) $f(x,y) \in C^2(I \times J),\ \varphi(t), \psi(t) \in C^2(D)\ \varphi(D) \subset I,\ \psi(D) \subset J$

$$\begin{aligned} \frac{d^2}{dt^2} f(\varphi(t), \psi(t)) =& [f_{xx}(\varphi(t), \psi(t))\dot{\varphi}(t) + f_{yx}(\varphi(t), \psi(t))\dot{\psi}(t)]\dot{\varphi}(t) + f_x(\varphi(t), \psi(t))\ddot{\varphi}(t) \\ &+ [f_{xy}(\varphi(t), \psi(t))\dot{\varphi}(t) + f_{yy}(\varphi(t), \psi(t))\dot{\psi}(t)]\dot{\psi}(t) + f_y(\varphi(t), \psi(t))\ddot{\psi}(t), \end{aligned}$$

(iii) $f(x,y) \in C^2(I \times J),\ u(t,s), v(t,s) \in C^2(D_t \times D_s)\ u(D_t, D_s) \subset I,\ v(D_t, D_s) \subset J$

$$\begin{aligned} \frac{\partial}{\partial t} f(u(t,s), v(t,s)) &= f_x(u(t,s), v(t,s))u_t(t,s) + f_y(u(t,s), v(t,s))v_t(t,s), \\ \frac{\partial}{\partial s} f(u(t,s), v(t,s)) &= f_x(u(t,s), v(t,s))u_s(t,s) + f_y(u(t,s), v(t,s))v_s(t,s), \end{aligned}$$

以降、変数は明示しない：

$$\begin{aligned} \frac{\partial^2}{\partial t^2} f(u(t,s), v(t,s)) &= [f_{xx}u_t + f_{yx}v_t]u_t + f_x u_{tt} + [f_{xy}u_t + f_{yy}v_t]v_t + f_y v_{tt}, \\ \frac{\partial^2}{\partial s \partial t} f(u(t,s), v(t,s)) &= [f_{xx}u_s + f_{yx}v_s]u_t + f_x u_{st} + [f_{xy}u_s + f_{yy}v_s]v_t + f_y v_{st}, \\ \frac{\partial^2}{\partial s^2} f(u(t,s), v(t,s)) &= [f_{xx}u_s + f_{yx}v_s]u_s + f_x u_{ss} + [f_{xy}v_s + f_{yy}v_s]v_s + f_y v_{ss}. \end{aligned}$$

6.3 Taylor の定理の証明

$g(t) = f(a+th, b+tk)$ とすると、この関数は t について $C^n(D)$ だから、1 変数の Taylor の定理を用いて

$$f(a+h, b+k) = g(1) = g(0) + g'(0) + \frac{1}{2!}g''(0) + \cdots + \frac{1}{(m-1)!}g^{(n-1)}(0) + \frac{1}{m!}g^{(m)}(\theta) \quad (0 < \theta < 1).$$

そこで、上に述べた合成関数の微分公式を用いる。

$$g'(t) = hf_x(a+th, b+tk) + kf_y(a+th, b+tk),$$

$$\begin{aligned} g''(t) &= h[hf_{xx} + kf_{yx}] + k[hf_{xy} + kf_{yy}] = h^2 f_{xx} + hk(f_{yx} + f_{xy}) + k^2 f_{yy} \\ &= h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy} = (h\partial_x + k\partial_y)^2 f, \\ g^{(m)}(t) &= (h\partial_x + k\partial_y)^m f = \sum_{j+\ell=m} \binom{m}{j} h^j k^\ell \partial_x^j \partial_y^\ell f. \end{aligned}$$

これより、定理の公式は直ちに導かれる。特に

$$g(t) = \sum_{\ell=0}^{m-1} \frac{1}{\ell!} g^{(\ell)}(0) t^\ell + \frac{1}{m!} g^{(m)}(\theta t) t^m$$

だから剰余項は

$$R_m = \frac{1}{m!} g^{(m)}(\theta) = \sum_{j+k=m} \frac{1}{j!k!} (x-a)^j (y-b)^k \frac{\partial^{j+k} f}{\partial x^j \partial y^k} (a + \theta(x-a), b + \theta(y-b)), \quad (\exists \theta \in (0,1))$$

となる。

6.4 極大値、極小値を求める一つの方法

命題 6.3 長方形領域 $I \times J$ で定義された関数 $f \in C^2(I \times J)$ が、領域の点 (a,b) で

$$f_x(a,b) = f_y(a,b) = 0, \quad \text{かつ} \quad J_f(a,b) = \begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix} \quad \text{が負定値行列}$$

となるとする。すると、関数 f は $(x,y) = (a,b)$ で極大値をとる。

定義 6.1 長方形領域 $I \times J$ で定義された関数 $f \in C^1(I \times J)$ に対し、$f_x(a,b) = f_y(a,b) = 0$ となる点 (a,b) を関数 $f(x,y)$ の停留点 *(stationary point)* という。

定義 6.2 $n \times n$-行列 $A = (a_{ij})$ が負定値行列であるとは、任意の $\xi = {}^t(\xi_1, \xi_2, \cdots, \xi_n) \in \mathbb{R}^n \setminus \{(0,0,\cdots,0)\}$ に対して

$$\xi \cdot A\xi = \sum_{i=1}^n \xi_i \Big(\sum_{j=1}^n a_{ij} \xi_j \Big) = \sum_{i,j=1}^n a_{ij} \xi_i \xi_j < 0$$

となることである。

特に $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ が負定値行列となる必要十分条件は

$$a < 0, \ c^2 - ab < 0$$

である。実際、任意の $\xi = {}^t(\xi_1, \xi_2)$ に対して

$$\xi_1(a\xi_1 + c\xi_2) + \xi_2(c\xi_1 + b\xi_2) = a\xi_1^2 + 2c\xi_1\xi_2 + b\xi_2 < 0$$

となるための条件は、上に述べたものである。

命題 6.3 の証明: さて、$f_x(a,b)=0,\ f_y(a,b)=0$ だから

$$f(x,y)=f(a,b)+R_2,$$

$$\begin{aligned}R_2&=\frac{1}{2}\big[f_{xx}(\cdots)(x-a)^2+2f_{xy}(\cdots)(x-a)(y-b)+f_{yy}(\cdots)(y-b)^2\big]\\&=\frac{1}{2}(x-a,y-b)\cdot J_f(\cdots)\begin{pmatrix}x-a\\y-b\end{pmatrix},\quad J_f(\cdots)=\begin{pmatrix}f_{xx}(\cdots)&f_{yx}(\cdots)\\f_{xy}(\cdots)&f_{yy}(\cdots)\end{pmatrix}\end{aligned}$$

但し、記法上 $(\cdots)=(a+\theta(x-a),b+\theta(y-b))$ とした。$f\in C^2(I\times J)$ であるから、$J_f(a,b)$ は負定値行列という仮定より、$(\cdots)$ が (a,b) の十分近くにあるとき $J_f(\cdots)$ も負定値行列。あとは、１変数の場合と同様の議論で、(x,y) が (a,b) の十分近くにあるとき

$$f(x,y)\le f(a,b)\quad 即ち、f(x,y)\ は\ (a,b)\ で極大.\quad\square$$

注意：$J_f(a,b)$ が負定値でも正定値でもない場合、点 (a,b) が極値を与える点かどうか、一般的な判定はできない。

定義 6.3 一般に点 (a,b) は、二つのベクトル $(x_1,y_1)\,(x_2,y_2)\neq 0$ が存在して t に関する関数 $g(t)=f(a+tx_1,b+ty_1)$ が $t=0$ で極小、s に関する関数 $h(s)=f(a+sx_2,b+sy_2)$ が $s=0$ で極大となるとき、関数 $f(x,y)$ の峠点という。

例えば、関数 $f(x,y)=x^2-y^2$ に対する停留点 $(0,0)$ は、$f(x,0)=x^2$ が $x=0$ で極小値、$f(0,y)=-y^2$ が $y=0$ で極大値を持つので、峠点である。

$$\begin{pmatrix}f_{xx}(0,0)&f_{xy}(0,0)\\f_{yx}(0,0)&f_{yy}(0,0)\end{pmatrix}=\begin{pmatrix}2&0\\0&-2\end{pmatrix}.$$

より一般に、関数 $f(x,y)$ に対する停留点 $(a,b),\ f_x(a,b)=0,\ f_y(a,b)=0,$ で

$$\begin{aligned}\det\begin{pmatrix}f_{xx}(a,b)-\lambda&f_{xy}(a,b)\\f_{yx}(a,b)&f_{yy}(a,b)-\lambda\end{pmatrix}&=\det(H_f(a,b)-\lambda\mathbb{I}_2)\\&=\lambda^2-(f_{xx}(a,b)+f_{yy}(a,b))\lambda+f_{xx}(a,b)f_{yy}(a,b)-f_{xy}(a,b)f_{yx}(a,b)=0\end{aligned}$$

を満たす実根[1] λ_1,λ_2 の性質を調べる。これらは停留点 (a,b) での Hessian（ヘッセ行列、ヘッシアン）$H_f(a,b)$ に対応する固有値[2]になり、以下のように分類される。

(i) $\lambda_1>0,\ \lambda_2>0$ ならば停留点 (a,b) は極小点、

(ii) $\lambda_1<0,\ \lambda_2<0$ ならば停留点 (a,b) は極大点、

(iii) $\lambda_1>0,\ \lambda_2<0$ ならば停留点 (a,b) は峠点、

(iv) $\lambda_1\lambda_2=0$ ならばその停留点 (a,b) の近辺の関数の挙動はこの情報だけでは不明。

注意：記法上 $(\cdots)=(a+\theta(x-a),b+\theta(y-b))$ と書いたが、講義中には

$$J_f(\cdots)=\begin{pmatrix}f_{xx}(\cdots)&f_{yx}(\cdots)\\f_{xy}(\cdots)&f_{yy}(\cdots)\end{pmatrix}\quad ではなく\quad \begin{pmatrix}f_{xx}&f_{yx}\\f_{xy}&f_{yy}\end{pmatrix}(\cdots)$$

と書いて、意味が分らないと言う質問があった。この表記で分かってくれただろうか？

[1]実数値解のこと。この式は虚根（虚数値解というのか？）を持たないことはすぐ分かる
[2]行列とその固有値については線形代数の講義で説明される

注意：$f_x(a,b)=f_y(a,b)=0$ かつ

$$\begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix} \quad \text{が負定値行列}$$

ならば、関数 f は $(x,y)=(a,b)$ で極大値をとる、がすっきり分らないという質問もあった。次回もう一度説明する予定である。

=====================================

メモ：出席者が大分減ってきた感じがするが、視力がないので、人数を数えられない。大体７０名前後か？更に、幾分蒸し暑かったせいか、まぶたを静かに閉じたままの姿も数多く見受けられた。