微分積分学続論　（担当：坪井　俊）

1　多変数関数のグラフ

1.1　多変数関数の増減、偏微分、方向微分、勾配ベクトル場（グラディエント・ベクトル場）

- 数直線、座標平面、座標空間はそれぞれ、座標 t, (x,y), (x,y,z) で点の位置が表される。点の位置により値が定まる関数を考える。n 次元ユークリッド空間 $\boldsymbol{R}^n$ で一般的に考えるが、$n=1, 2, 3$ に具体化しながら考えるのが良い。座標 $(x_1,\ldots,x_n)$ を位置ベクトル $\overrightarrow{x}$ で書くことにする。位置ベクトルはできるだけ列ベクトル $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ で書くようにする。

- n 次元ユークリッド空間 $\boldsymbol{R}^n$ の部分集合 U 上で定義された関数を考える。それは、座標 $(x_1,...,x_n)$ によって実数値が定まるもの（場合によっては、複素数値、ベクトル値、行列値が定まるもの）である。実数値多変数関数は $f(x_1,...,x_n)$ のように書かれる。この講義では $f(\overrightarrow{x})$ のように書くようにする。

- 多変数関数の定義域となる部分集合 U は、開集合であることが多いが、閉集合、特にある多変数関数の零点集合や不等号を用いて定義される集合であることもある。例えば、条件付き極値問題はそのような場合の極大値極小値を求める問題である。特に断らない限り定義域は**ユークリッド空間の開集合**とする。

- どういうときに多変数関数の増減を知りたいか。
  - 斜面に置かれた物体は、斜面の最大傾斜線の下向きに斜角を θ とすると、重力と応力により、$\sin\theta$ に比例する力を受けるとされる。
  - 流体の分子は、圧力勾配に比例する力を受けるとされる。
  - 一般にエネルギー保存則の成り立つ力学は、座標 $(p_1,\ldots,p_n,q_1,\ldots,q_n)$ をもつ偶数次元の空間上の、ハミルトン関数 $H(p_1,\ldots,p_n,q_1,\ldots,q_n)$ により、

    $$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}, \qquad \frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}$$

    と書かれる。H が極値をとる点では、位置も運動量も変化しない。

- 実数値多変数関数 $f(\overrightarrow{x}) = f(x_1,...,x_n)$ に対し、$\overrightarrow{x}$ の近くの点 $\overrightarrow{y}$ における値 $f(\overrightarrow{y})$ が、$f(\overrightarrow{x})$ に近い場合を考える。多変数関数の連続性については、リプシッツ連続の概念がもっとも使いやすい。微分可能性については、全微分可能性かつ微分が連続になる場合を扱うのが普通である。

- **【定義】（リプシッツ連続）**　f が U 上でリプシッツ連続とは正実数 K が存在し、任意の $\overrightarrow{x}$, $\overrightarrow{y} \in U$ に対し、$|f(\overrightarrow{x}) - f(\overrightarrow{y})| \leqq K\|\overrightarrow{x} - \overrightarrow{y}\|$ が成立することである。K はリプシッツ定数と呼ばれる。

- 【定義】(全微分可能)　$f : U \longrightarrow \boldsymbol{R}$ が $\overrightarrow{x} \in U$ で全微分可能とは、$A = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix}$ が存在し、任意の $\varepsilon > 0$ に対し、$\|\overrightarrow{x} - \overrightarrow{y}\| < \delta$ ならば、
$$|f(\overrightarrow{y}) - f(\overrightarrow{x}) - A(\overrightarrow{y} - \overrightarrow{x})| \leqq \varepsilon\|\overrightarrow{y} - \overrightarrow{x}\|$$
が成立することである。

- 【命題】(全微分可能 $\Longrightarrow$ 偏微分可能)　$f : U \longrightarrow \boldsymbol{R}$ が $\overrightarrow{x} \in U$ で全微分可能ならば、
$$a_i = (\frac{\partial f}{\partial x_i})_{(\overrightarrow{x})} = \lim_{h \to 0} \frac{f(\overrightarrow{x} + h\overrightarrow{e}_i) - f(\overrightarrow{x})}{h}$$
となる(右辺が存在し、左辺と一致)。ただし $\overrightarrow{e}_1$ は i 番目の座標の方向の単位ベクトルである。

- 方向微分　$f(\overrightarrow{x})$ が全微分可能であるとする。ベクトル $\overrightarrow{v} \in \boldsymbol{R}^n$ に対し、任意の $\varepsilon > 0$ に対し、$\|t\overrightarrow{v}\| < \delta$ ならば、$|f(\overrightarrow{x} + t\overrightarrow{v}) - f(\overrightarrow{x}) - t \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{pmatrix} \overrightarrow{v}| \leqq \varepsilon t\|\overrightarrow{v}\|$ が成立するから
$$\begin{aligned} (\frac{\mathrm{d}f(\overrightarrow{x} + t\overrightarrow{v})}{\mathrm{d}t})_{t=0} &= \sum_{i=1}^{n} (\frac{\partial f}{\partial x_i})_{(\overrightarrow{x})} v_i = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \cdots \dfrac{\partial f}{\partial x_n} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} \bullet \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \end{aligned}$$
全微分可能ならば $\overrightarrow{v}$ の方向の方向微分が上のように定まる。

- 【定義】(C^1 級)　$f : U \longrightarrow \boldsymbol{R}$ が、すべての $\overrightarrow{x} \in U$ において、すべての i に対し偏微分 $(\dfrac{\partial f}{\partial x_i})_{(\overrightarrow{x})}$ が存在し、$(\dfrac{\partial f}{\partial x_i})_{(\overrightarrow{x})}$ が $\overrightarrow{x}$ について連続のとき、f は連続微分可能であるまたは C^1 級であるという。

- 【命題】f が C^1 級ならば、全微分可能である。
【証明】　$\overrightarrow{y}^k = (y_1, \dots, y_k, x_{k+1}, \dots, x_n)$ ($\overrightarrow{y}^0 = \overrightarrow{x}$, $\overrightarrow{y}^n = \overrightarrow{y}$) とおいて、
$$\begin{aligned} f(\overrightarrow{y}) - f(\overrightarrow{x}) &= \sum_{k=1}^{n} \{f(\overrightarrow{y}^k) - f(\overrightarrow{y}^{k-1})\} \\ &= \sum_{k=1}^{n} \int_0^1 (\frac{\partial f}{\partial x_k})_{(\overrightarrow{y}^k + t(\overrightarrow{y}^k - \overrightarrow{y}^{k-1}))} (y_k - x_k)\mathrm{d}t \end{aligned}$$
である。偏微分 $(\dfrac{\partial f}{\partial x_k})_{(\overrightarrow{y})}$ は、$\overrightarrow{x}$ で連続だから、与えられた $\varepsilon > 0$ に対し、任意の $\overrightarrow{y}$ に対し、$\|\overrightarrow{y} - \overrightarrow{x}\| < \delta$ ならば、$|(\dfrac{\partial f}{\partial x_k})_{(\overrightarrow{y})} - (\dfrac{\partial f}{\partial x_k})_{(\overrightarrow{x})}| \leqq \dfrac{\varepsilon}{\sqrt{n}}$ をみたすような δ を定める。そうすると
$$\begin{aligned} &|\int_0^1 (\frac{\partial f}{\partial x_k})_{(\overrightarrow{y}^{k-1} + t(\overrightarrow{y}^k - \overrightarrow{y}^{k-1}))} (y_k - x_k)\mathrm{d}t - \int_0^1 (\frac{\partial f}{\partial x_k})_{(\overrightarrow{x})} (y_k - x_k)\mathrm{d}t| \\ &\leqq \frac{\varepsilon}{\sqrt{n}} |y_k - x_k| \end{aligned}$$

ゆえに

$$|f(\vec{y}) - f(\vec{x}) - \sum_{k=1}^{n} \big(\frac{\partial f}{\partial x_k}\big)_{(\vec{x})}(y_k - x_k)| \leqq \sum_{k=1}^{n} \frac{\varepsilon}{\sqrt{n}}|y_k - x_k| \leqq \varepsilon\|\vec{y} - \vec{x}\|$$

（平均値の定理により

$$f(\vec{y}^{\,k}) - f(\vec{y}^{\,k-1}) = \big(\frac{\partial f}{\partial x_k}\big)_{(\vec{y}^{\,k-1}+s_k(\vec{y}^{\,k}-\vec{y}^{\,k-1}))}(y_k - x_k)$$

となる $0 < s_k < 1$ が存在することを用いてもよい。）

- 開集合 U 上で f が C^1 級ならば、連続なベクトル場 $\begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}_{(\vec{x})}$ が定まる。これを勾配ベクトル場、グラジエントベクトル場と呼ぶ。$\mathrm{grad}(f)$, ∇f のように書く。
- f が C^1 級のとき、$Df = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{pmatrix}$ とおくと、

$$\big(\frac{\mathrm{d}f(\vec{x}+t\vec{v})}{\mathrm{d}t}\big)_{(t)} = (Df)_{(\vec{x}+t\vec{v})}\vec{v}$$

と書かれる。

- 【命題】　$f : U \longrightarrow \boldsymbol{R}$ が C^1 級ならば局所リプシッツ連続である。

【証明】　U に含まれ $\vec{x}$ の近傍を含むある有界閉集合上で $|\dfrac{\partial f}{\partial x_i}| \leqq K$ となる K が存在する。（有界閉集合上で最大値が存在する。）

$$\begin{aligned} f(\vec{x}+\vec{v}) - f(\vec{x}) &= \int_0^1 \big(\frac{\mathrm{d}f(\vec{x}+t\vec{v})}{\mathrm{d}t}\big)_{(t)}\mathrm{d}t \\ &= \sum_{i=1}^{n} \big(\frac{\partial f}{\partial x_i}\big)_{(\vec{x}+t\vec{v})} v_i \end{aligned}$$

$$|f(\vec{x}+\vec{v}) - f(\vec{x})| \leqq K\sqrt{n}\|\vec{v}\|$$

- 【演習問題】　(1)　$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \boldsymbol{R}^n$ に対し、$\dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{n}|v_i| \leqq \|\vec{v}\| \leqq \sum_{i=1}^{n}|v_i|$ を示せ。

(2)　$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \boldsymbol{R}^n$ に対し、$\vec{x} \bullet \vec{y} = \displaystyle\sum_{i=1}^{n} x_i y_i$ とするとき、$|\vec{x} \bullet \vec{y}| \leqq \|\vec{x}\| \cdot \|\vec{y}\|$ を示せ。

- 【例】 $f(\vec{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \vec{x} \bullet A\vec{x}$ とすると、

$$\mathrm{grad}(f)\vec{v} = \vec{x} \bullet ({}^tA + A)\vec{v} = ({}^tA + A)\vec{x} \bullet \vec{v}$$

A を最初から対称行列とすると、$\mathrm{grad}(f)_{(\vec{x})} = 2A\vec{x}$ となる。$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ならば、$\mathrm{grad}(f)_{(\vec{x})} = \begin{pmatrix} 2\lambda_1 x_1 \\ \vdots \\ 2\lambda_n x_n \end{pmatrix} = 2\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\vec{x}$ である。

曲線 $t \longmapsto \vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} e^{2\lambda_1 t}x_1^0 \\ \vdots \\ e^{2\lambda_n t}x_n^0 \end{pmatrix}$ に対し、

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \begin{pmatrix} \dfrac{\mathrm{d}x_1}{\mathrm{d}t}(t) \\ \vdots \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t}(t) \end{pmatrix} = \begin{pmatrix} 2\lambda_1 e^{2\lambda_1 t}x_1^0 \\ \vdots \\ 2\lambda_n e^{2\lambda_n t}x_n^0 \end{pmatrix} = 2\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\vec{x} = \mathrm{grad}(f)_{(\vec{x})}$$

となる。すなわち曲線 $\vec{x}(t)$ の微分は、$\mathrm{grad}(f)_{\vec{x}(t)}$ と一致する。

$$\vec{\varphi}(t, \vec{x}) = \vec{\varphi}_t(\vec{x}) = \begin{pmatrix} e^{2\lambda_1 t}x_1^0 \\ \vdots \\ e^{2\lambda_n t}x_n^0 \end{pmatrix}$$

で定まる $\vec{\varphi} : \boldsymbol{R} \times \boldsymbol{R}^n \longrightarrow \boldsymbol{R}^n$ を gradient vector field $\mathrm{grad}(f)$ が生成する gradient flow という。$\vec{\varphi}_s \circ \vec{\varphi}_t = \vec{\varphi}_{s+t}$ が成立する。

- 対称行列 A に対し、A の固有値を $\lambda_1, \ldots,, \lambda_n$ とすると、ある直交行列 P に対して、${}^tPAP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ となる。

- $\boldsymbol{R}^n$ の原点 $\vec{0}$ の近傍 U 上で定義された関数 $f : U \longrightarrow \boldsymbol{R}$ を考える。P を直交行列とし、P の第 j 列ベクトルを $\vec{v}_j$ とする : $P = \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix}$. 方向微分を考えると $\left(\dfrac{\mathrm{d}f(t\vec{v}_j)}{\mathrm{d}t}\right)_{t=0} = (Df)_{(\vec{0})}\vec{v}_j$ となる。点を $\vec{x} = \sum_{j=1}^{n} y_i\vec{v}_j$ と表すことは、座標変換を与えている。$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ として、$\vec{x} = P\vec{y}$ である。$g(\vec{y}) = f(P\vec{y})$ とおくと $\dfrac{\partial g}{\partial y_j} = (Df)_{(\vec{0})}\vec{v}_j$ となり、$(Dg)_{(\vec{0})} = (Df)_{(\vec{0})}P$ となる。

- 連続写像 $\vec{x} : (a, b) \longrightarrow \boldsymbol{R}^n$ を弧、道または曲線と呼ぶ (arc, path, curve)。連続とは $\vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ の各成分 $x_i(t)$ が連続であることである。各成分 $x_i(t)$ が連続微

分可能、すなわち C^1 級であるとき、$(\frac{\mathrm{d}\vec{x}}{\mathrm{d}t})_{(t)} = \begin{pmatrix} (\frac{\mathrm{d}x_1}{\mathrm{d}t})_{(t)} \\ \vdots \\ (\frac{\mathrm{d}x_n}{\mathrm{d}t})_{(t)} \end{pmatrix}$ を $\vec{x}(t)$ における**速度ベクトル**と呼ぶ。このときは $\vec{x}(t)$ を点の運動と考えている。

- C^1 級関数 $f : U \longrightarrow \boldsymbol{R}$ の値の変化は、曲線 $\vec{x} : (a, b) \longrightarrow U$ を用い、$\vec{x}$ をいろいろととって、$f \circ \vec{x}$ の変化を見るのが良い。

$$\frac{\mathrm{d}(f \circ \vec{x})}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}t} = (Df)_{(\vec{x}(t))} (\frac{\mathrm{d}\vec{x}}{\mathrm{d}t})_{(t)} = \mathrm{grad}(f)_{(\vec{x}(t))} \bullet (\frac{\mathrm{d}\vec{x}}{\mathrm{d}t})_{(t)}$$

この値は、$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \in \ker(Df)$, すなわち、$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}$ が $\mathrm{grad}(f)$ と直交するとき 0 である（瞬間的に増減しない）。また、すべての t に対し、$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \in \ker(Df)$ ならば、$f \circ \vec{x}$ の値は変化しない。従って、$\vec{x}$ は、f の等位面上にある。

$\|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\|$ が一定として $\vec{x}$ を変化させるならば、$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}$ が $\mathrm{grad}(f)$ と平行のとき、最も変化が大きい。

1.2 ニュートン力学

- ニュートンの運動方程式は、質点の運動の加速度ベクトルが、質点にかかる力に比例するというものである。

$$m(\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2})_{(t)} = \vec{F}_{(\vec{x}(t))}$$

- ポテンシャル力とは、位置の関数であるポテンシャル U があり、$\vec{F} = -\mathrm{grad}(U)$ となるもののことである。例えば、質量 $m_1, \ldots, m_k$ の k 個の質点が $\vec{x}^1, \ldots, \vec{x}^k$ にあるとき（その瞬間）、ニュートンの万有引力のポテンシャルは、単位質量に対し、$U = -\sum_{j=1}^{k} \frac{Gm_i}{\|\vec{x} - \vec{x}^j\|}$ である。

実際、

$$\begin{aligned} \frac{\partial}{\partial x_i} \frac{1}{\|\vec{x} - \vec{x}^j\|} &= \frac{\partial}{\partial x_i} \{\sum_{i=1}^{3} (x_i - x_i^{(j)})^2\}^{-\frac{1}{2}} \\ &= \{\sum_{i=1}^{3} (x_i - x_i^{(j)})^2\}^{-\frac{3}{2}} (-\frac{1}{2}) 2(x_i - x_i^{(j)}) \\ &= -\frac{x_i - x_i^{(j)}}{\|\vec{x} - \vec{x}^j\|^3} \end{aligned}$$

だから、$-\mathrm{grad}(-\frac{1}{\|\vec{x} - \vec{x}^j\|}) = -\frac{\vec{x} - \vec{x}^j}{\|\vec{x} - \vec{x}^j\|^3}$ であり、

$$\vec{F} = -\mathrm{grad}(U) = -\sum Gm_j \frac{\vec{x} - \vec{x}^j}{\|\vec{x} - \vec{x}^j\|^3}$$

となる。

この計算は、ある瞬間に系に影響を与えない単位質量の質点が $\vec{x}$ にあったときのポテンシャルを与えている。

- $\vec{x}^1$, $\vec{x}^2$ にある質量 m_1, m_2 2つの質点の運動を記述する問題は2体問題と呼ばれる。それぞれに働く重力は次のようになる。

$$\vec{F}_{(\vec{x}^1)} = -Gm_1m_2\frac{\vec{x}^1-\vec{x}^2}{\|\vec{x}^1-\vec{x}^2\|^3} = -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x}-\vec{x}^2\|})_{(\vec{x}^1)}$$
$$\vec{F}_{(\vec{x}^2)} = -Gm_1m_2\frac{\vec{x}^2-\vec{x}^1}{\|\vec{x}^2-\vec{x}^1\|^3} = -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x}-\vec{x}^1\|})_{(\vec{x}^2)}$$

$\vec{q} = \begin{pmatrix}\vec{x}^1\\ \vec{x}^2\end{pmatrix}$ とおけば、$\vec{F}_{(\vec{q})} = -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x}^2-\vec{x}^1\|})_{(\vec{q})}$ である。運動方程式は、

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{pmatrix}m_1\vec{x}^1\\ m_2\vec{x}^2\end{pmatrix} = -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x}^2-\vec{x}^1\|})_{(\vec{q})}$$

3体問題については、同様に $\vec{q} = \begin{pmatrix}\vec{x}^1\\ \vec{x}^2\\ \vec{x}^3\end{pmatrix}$ とおけば、$\vec{F}_{(\vec{q})} = -\mathrm{grad}(-\sum_{j<k}\frac{Gm_jm_k}{\|\vec{x}^j-\vec{x}^k\|})_{(\vec{q})}$ である。

- $\vec{p} = \begin{pmatrix}\vec{p}^1\\ \vec{p}^2\\ \vec{p}^3\end{pmatrix} = \begin{pmatrix}m_1\frac{\mathrm{d}\vec{x}^1}{\mathrm{d}t}\\ m_2\frac{\mathrm{d}\vec{x}^2}{\mathrm{d}t}\\ m_3\frac{\mathrm{d}\vec{x}^3}{\mathrm{d}t}\end{pmatrix}$ とおき、$H = \sum_{j=1}^{3}\frac{\|\vec{p}^j\|^2}{2m_j} - \sum_{j<k}\frac{Gm_jm_k}{\|\vec{x}^j-\vec{x}^k\|}$ とおけば、

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -\mathrm{grad}_{\vec{q}}H \qquad \frac{\mathrm{d}\vec{q}}{\mathrm{d}t} = \mathrm{grad}_{\vec{p}}H$$

となる。ハミルトン正準方程式と呼ばれる。

紛らわしいが、$\mathrm{grad}_{\vec{q}}$ あるいは $\mathrm{grad}_{\vec{p}}$ は、$\vec{q}$ の関数とみて、あるいは $\vec{p}$ の関数とみて、勾配を計算するという意味である。$\begin{pmatrix}-\mathrm{grad}_{\vec{q}}\\ \mathrm{grad}_{\vec{p}}\end{pmatrix}$ を skew-gradient あるいは symplectic gradient と呼んで、

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\vec{p}\\ \vec{q}\end{pmatrix} = \text{s-grad}(H)$$

とも書かれる。

1.3 参考：制限三体問題のポテンシャル

- 2体問題において、$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{m_1\vec{x}^1+m_2\vec{x}^2}{m_1+m_2} = \vec{0}$ であるから、質点 m_1, m_2 の重心 $\vec{x}^0 = \frac{m_1\vec{x}^1+m_2\vec{x}^2}{m_1+m_2}$ は、$\frac{\mathrm{d}^2\vec{x}^0}{\mathrm{d}t^2} = \vec{0}$ を満たし、等速直線運動をする。

$\vec{x}^1 - \vec{x}^0 = \dfrac{m_1\vec{x}^1 + m_2\vec{x}^1 - m_1\vec{x}^1 - m_2\vec{x}^2}{m_1+m_2} = \dfrac{m_2}{m_1+m_2}(\vec{x}^1 - \vec{x}^2)$ だから、運動方程式

$$m_1\frac{\mathrm{d}^2\vec{x}^1}{\mathrm{d}t^2} = -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x} - \vec{x}^2\|})_{(\vec{x}^1)}$$

から、

$$\begin{aligned} m_1\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\vec{x}^1 - \vec{x}^0) = m_1\frac{\mathrm{d}^2\vec{x}^1}{\mathrm{d}t^2} &= -\mathrm{grad}(-\frac{Gm_1m_2}{\|\vec{x} - \vec{x}^2\|})_{(\vec{x}^1)} \\ &= -\mathrm{grad}(-\frac{Gm_1{m_2}^2}{(m_1+m_2)\|\vec{x} - \vec{x}^0\|})_{(\vec{x}^1)} \end{aligned}$$

が得られる。$\vec{x}^2 - \vec{x}^0$ についても同様である。つまり、重心に対する相対位置 $\vec{x}^1 - \vec{x}^0$, $\vec{x}^2 - \vec{x}^0$ は、次の運動方程式を満たす。

$$\begin{aligned} \frac{\mathrm{d}^2(\vec{x}^1 - \vec{x}^0)}{\mathrm{d}t^2} &= -\frac{G{m_2}^2}{m_1+m_2}\frac{\vec{x}^1 - \vec{x}^0}{\|\vec{x}^1 - \vec{x}^0\|^3} \\ \frac{\mathrm{d}^2(\vec{x}^2 - \vec{x}^0)}{\mathrm{d}t^2} &= -\frac{G{m_1}^2}{m_1+m_2}\frac{\vec{x}^2 - \vec{x}^0}{\|\vec{x}^2 - \vec{x}^0\|^3} \end{aligned}$$

を満たす。これは、$\vec{x}^0$ に質量 $\dfrac{{m_2}^2}{m_1+m_2}$ あるいは $\dfrac{{m_1}^2}{m_1+m_2}$ の不動の質点があるときの $\vec{x}^1$, $\vec{x}^2$ にある質点の運動方程式である。

- この不動の質点があるとき、一般には、$\vec{x}^0$ を焦点とする 2 次曲線を軌道とする解となる。特に、不動の質点の周りの円運動をする解が存在する。すなわち、$\vec{x}^0 = \vec{0}$ として、$\vec{x}(t) = R_{\omega t}\vec{x}(0)$ ($R_{\omega t} = \begin{pmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{pmatrix}$ は ωt の回転、$\vec{x}(0) = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ 0 \end{pmatrix}$) とすると、

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2} = -\omega^2\vec{x}$$

となる。

$$\omega^2 = \frac{G{m_1}^2}{m_1+m_2}\frac{1}{\|\vec{x}^2\|^2} = \frac{G{m_2}^2}{m_1+m_2}\frac{1}{\|\vec{x}^1\|^2}$$

が成立する。

- 万有引力定数 G の次元（単位）は、(力) = (質量)(長さ)(時間)$^{-2}$ = G(質量)^2(長さ)$^{-2}$ だから、G は、(質量)$^{-1}$(長さ)^3(時間)$^{-2}$ の次元を持つ。最後の等式の次元は、(角速度)^2 = (時間)$^{-2}$ = (質量)$^{-1}$(長さ)^3(時間)$^{-2}$(質量)(長さ)$^{-2}$ = (長さ)(時間)$^{-2}$ である。まず、長さの単位を変えて、$\|\vec{x}^2 - \vec{x}^1\| = 1$ とする。質量の比を $m_2/m_1 = \mu/(1-\mu)$ とすると、$\|\vec{x}^1\| = \mu$, $\|\vec{x}^2\| = 1-\mu$ である。時間の単位を変えて角速度を 1 にする。このとき、$1 = G(m_1+m_2)$ であり、$Gm_1 = 1-\mu$, $Gm_2 = \mu$ である。

- 円運動をする解は、

$$\begin{aligned} \vec{x}^1(t) &= -\mu R_t\vec{e}_1 \\ \vec{x}^2(t) &= (1-\mu)R_t\vec{e}_1 \end{aligned}$$

ここで、$\overrightarrow{x}$ にある質点 m は、$-\mathrm{grad}(-\frac{Gmm_1}{\|\overrightarrow{x}-\overrightarrow{x}^1(t)\|}-\frac{Gmm_2}{\|\overrightarrow{x}-\overrightarrow{x}^2(t)\|})$ の力を受けるので、運動方程式は、

$$\frac{\mathrm{d}^2\overrightarrow{x}}{\mathrm{d}t^2}=-\mathrm{grad}(-\frac{1-\mu}{\|\overrightarrow{x}-\overrightarrow{x}^1(t)\|}-\frac{\mu}{\|\overrightarrow{x}-\overrightarrow{x}^2(t)\|})$$

- $R_t\overrightarrow{y}(t)=\overrightarrow{x}(t)$ という座標で円運動をする解を書き直すと、$\overrightarrow{y}^1(t)=-\mu\overrightarrow{e}_1$, $\overrightarrow{y}^2(t)=(1-\mu)\overrightarrow{e}_1$ である。この 2 点は「遠心力」と釣り合っている。$J=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}$ として、

$$JR_t\overrightarrow{y}(t)+R_t\big(\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}\big)_{(t)}=\big(\frac{\mathrm{d}\overrightarrow{x}}{\mathrm{d}t}\big)_{(t)},$$
$$J^2R_t\overrightarrow{y}(t)+2JR_t\big(\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}\big)_{(t)}+R_t\big(\frac{\mathrm{d}^2\overrightarrow{y}}{\mathrm{d}t^2}\big)_{(t)}=\big(\frac{\mathrm{d}^2\overrightarrow{x}}{\mathrm{d}t^2}\big)_{(t)}$$

だから、

$$\begin{aligned}\big(\frac{\mathrm{d}^2\overrightarrow{y}}{\mathrm{d}t^2}\big)_{(t)}&=R_{-t}\big(\frac{\mathrm{d}^2\overrightarrow{x}}{\mathrm{d}t^2}\big)_{(t)}-J^2\overrightarrow{y}(t)-2J\big(\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}\big)_{(t)}\\&=-\mathrm{grad}(-\frac{1-\mu}{\|\overrightarrow{y}-\overrightarrow{y}^1(t)\|}-\frac{\mu}{\|\overrightarrow{y}-\overrightarrow{y}^2(t)\|})+\begin{pmatrix}y_1(t)\\y_2(t)\\0\end{pmatrix}-2J\big(\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}\big)_{(t)}\\&=-\mathrm{grad}(-\frac{1-\mu}{\|\overrightarrow{y}-\mu\overrightarrow{e}_1\|}-\frac{\mu}{\|\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1\|}-\frac{{y_1}^2+{y_2}^2}{2})-2J\big(\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}\big)_{(t)}\end{aligned}$$

- ここで、$\frac{\mathrm{d}\overrightarrow{y}}{\mathrm{d}t}=\overrightarrow{0}$ かつ $\frac{\mathrm{d}^2\overrightarrow{y}}{\mathrm{d}t^2}=\overrightarrow{0}$ を満たす解（$\overrightarrow{y}$ の座標で静止した状態が続く点）を探すと、

$$\mathrm{grad}(\frac{1-\mu}{\|\overrightarrow{y}+\mu\overrightarrow{e}_1\|}+\frac{\mu}{\|\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1\|}+\frac{{y_1}^2+{y_2}^2}{2})=\overrightarrow{0}$$

となる点である。このような点が 5 つあり、ラグランジュ点と呼ばれる。

- この式を書き直すと、

$$-\frac{(1-\mu)(\overrightarrow{y}+\mu\overrightarrow{e}_1)}{\|\overrightarrow{y}+\mu\overrightarrow{e}_1\|^3}-\frac{\mu(\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1)}{\|\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1\|^3}+\overrightarrow{y}=\overrightarrow{0}$$

- $\mu\overrightarrow{e}_1$, $(1-\mu)\overrightarrow{e}_1$ を 1 辺とする正三角形の頂点が、この式を満たしていることが作図により容易にわかる。

- 残りの 3 点は、$\overrightarrow{e}_1$ 軸上にあることが、$-\frac{(1-\mu)(\overrightarrow{y}+\mu\overrightarrow{e}_1)}{\|\overrightarrow{y}+\mu\overrightarrow{e}_1\|^3}-\frac{\mu(\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1)}{\|\overrightarrow{y}-(1-\mu)\overrightarrow{e}_1\|^3}+\overrightarrow{y}$ の成分表示からわかる。

成分表示は

$$\begin{pmatrix} -\dfrac{(1-\mu)(y_1+\mu)}{((y_1+\mu)^2+{y_2}^2+{y_3}^2)^{3/2}}-\dfrac{\mu(y_1-(1-\mu))}{((y_1-(1-\mu))^2+{y_2}^2+{y_3}^2)^{3/2}}+y_1 \\ -\dfrac{(1-\mu)y_2}{((y_1+\mu)^2+{y_2}^2+{y_3}^2)^{3/2}}-\dfrac{\mu y_2}{((y_1-(1-\mu))^2+{y_2}^2+{y_3}^2)^{3/2}}+y_2 \\ -\dfrac{(1-\mu)y_3}{((y_1+\mu)^2+{y_2}^2+{y_3}^2)^{3/2}}-\dfrac{\mu y_3}{((y_1-(1-\mu))^2+{y_2}^2+{y_3}^2)^{3/2}} \end{pmatrix}$$

この y_3 成分が 0 になるのは、$y_3=0$ のときだけである。従って、y_1y_2 平面上で考えると、y_1, y_2 成分の表示は

$$\begin{pmatrix} -\dfrac{(1-\mu)(y_1+\mu)}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\dfrac{\mu(y_1-(1-\mu))}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+y_1 \\ -\dfrac{(1-\mu)y_2}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\dfrac{\mu y_2}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+y_2 \end{pmatrix}$$

y_2 成分が 0 になるのは、$y_2=0$ または、

$$-\frac{1-\mu}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\frac{\mu}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+1=0$$

のときである。$(-\mu,0)$, $(1-\mu,0)$ への距離がともに 1 未満の点では、

$$\frac{1-\mu}{((y_1+\mu)^2+{y_2}^2)^{3/2}}>1-\mu,\quad \frac{\mu}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}>\mu$$

であり、

$$-\frac{1-\mu}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\frac{\mu}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+1<0$$

となる。y_1 軸上の区間 $[-\mu,1-\mu]$ からの距離が増加すると、この値は単調に増加し、0 をとる。(y_1,y_2) が

$$-\frac{1-\mu}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\frac{\mu}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+1=0$$

をみたすとき、y_1 成分は、

$$\begin{aligned} &-\frac{(1-\mu)(y_1+\mu)}{((y_1+\mu)^2+{y_2}^2)^{3/2}}-\frac{\mu(y_1-(1-\mu))}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+y_1 \\ =&-\frac{(1-\mu)\{(y_1+\mu)-(y_1-(1-\mu))\}}{((y_1+\mu)^2+{y_2}^2)^{3/2}}+y_1-(y_1-(1-\mu)) \\ =&-\frac{(1-\mu)}{((y_1+\mu)^2+{y_2}^2)^{3/2}}+(1-\mu) \\ =&-\frac{\mu}{((y_1-(1-\mu))^2+{y_2}^2)^{3/2}}+\mu \end{aligned}$$

すなわち、$(-\mu,0)$, $(1-\mu,0)$ への距離がともに 1 の点でのみ 0 になる。これが正三角形解である。

1.4　多変数関数のグラフ、多変数関数のグラフの接平面

- $\boldsymbol{R}^n$ の開集合 U 上の多変数関数 $f : U \longrightarrow \boldsymbol{R}$ に対しては、そのグラフ

$$\{(\vec{x}, f(\vec{x})) \in \boldsymbol{R}^n \times \boldsymbol{R} \bigm| \vec{x} \in U\}$$

  を考えると、($n = 1, 2$ に対しては）視覚的にも理解しやすい。

- 曲線 $\vec{x}(t)$ $(t \in (a, b))$ をいろいろととり、関数 $f \circ \vec{x}$ の増減について考えることは基本である。

- f について、極大点、極小点が定義される。
  f の極大点、極小点においては、$DF = \mathbf{0}$ となる。

- 多変数関数のグラフの $(\vec{x}, f(\vec{x}))$ における接平面は、

$$\{(\vec{x}, f(\vec{x})) + (\vec{v}, (Df)_{(\vec{x})}\vec{v}) \bigm| \vec{v} \in \boldsymbol{R}^n\}$$

  で定義される。この接平面（接超平面）は、$(\vec{x}, f(\vec{x}))$ を起点とするベクトル
  $\begin{pmatrix} \vec{e}_1 \\ (Df)_{(\vec{x})}\vec{e}_1 \end{pmatrix}, \dots, \begin{pmatrix} \vec{e}_n \\ (Df)_{(\vec{x})}\vec{e}_n \end{pmatrix}$ で張られる。このベクトル空間は

$$\ker\begin{pmatrix} (Df)_{(\vec{x})} & -1 \end{pmatrix} = \{\begin{pmatrix} \vec{v} \\ w \end{pmatrix} \in \boldsymbol{R}^{n+1} \bigm| (Df)_{(\vec{x})}\vec{v} - w = 0\}$$

  とも書かれる。

- ベクトル値関数 $\vec{F} : U \longrightarrow \boldsymbol{R}^m$ に対して同様のグラフおよびその接平面（接空間）を考えることができる。

- ベクトル値関数 $\vec{F} : U \longrightarrow \boldsymbol{R}^m$ のグラフとは、

$$\{(\vec{x}, \vec{F}(\vec{x})) \in \boldsymbol{R}^n \times \boldsymbol{R}^m \bigm| \vec{x} \in U\}$$

- ベクトル値関数 $\vec{F} : U \longrightarrow \boldsymbol{R}^m$ のグラフの $(\vec{x}, \vec{F}(\vec{x}))$ における接平面（接空間）は、

$$\{(\vec{x}, \vec{F}(\vec{x})) + (\vec{v}, (D\vec{F})_{(\vec{x})}\vec{v}) \bigm| \vec{v} \in \boldsymbol{R}^n\}$$

  で定義される。この接平面（接空間）は、$(\vec{x}, \vec{F}(\vec{x}))$ を起点とするベクトル
  $\begin{pmatrix} \vec{e}_1 \\ (D\vec{F})_{(\vec{x})}\vec{e}_1 \end{pmatrix}, \dots, \begin{pmatrix} \vec{e}_n \\ (D\vec{F})_{(\vec{x})}\vec{e}_n \end{pmatrix}$ で張られる。このベクトル空間は

$$\ker\begin{pmatrix} (D\vec{F})_{(\vec{x})} & -I \end{pmatrix} = \{\begin{pmatrix} \vec{v} \\ \vec{w} \end{pmatrix} \in \boldsymbol{R}^{n+m} \bigm| (D\vec{F})_{(\vec{x})}\vec{v} - \vec{w} = 0\}$$

  とも書かれる。

- 接空間は $(\vec{x}, \vec{F}(\vec{x}))$ を通過する曲線の速度ベクトルの集合と一致する。

1.5 2階微分のなす行列、接平面とグラフの位置関係、凸関数

- 多変数関数のグラフ $\{(\vec{x}, f(\vec{x})) \mid \vec{x} \in U\}$ とその $(\vec{x}, f(\vec{x}))$ における接平面

$$\{(\vec{x}, f(\vec{x})) + (\vec{v}, (Df)_{(\vec{x})}\vec{v}) \mid \vec{v} \in \boldsymbol{R}^n\}$$

の位置関係を考えるとき、それらが $(\vec{x}, f(\vec{x}))$ の近傍で $(\vec{x}, f(\vec{x}))$ 以外の共通部分を持つかどうかということが重要である。

- U の各点 $\vec{x}$ において $(\vec{x}, f(\vec{x}))$ の近傍で $(\vec{x}, f(\vec{x}))$ 以外の共通部分を持たないとき、f のグラフは凸であるとよぶ。

1.6 2次超曲面 その1

- f が2次の多項式

$$f(x_1, \ldots, x_n) = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

であるとする。$(x_1^0, \ldots, x_n^0, f(x_1^0, \ldots, x_n^0))$ における接平面は

$$\Big(\frac{\partial f}{\partial x_j}\Big)_{(\vec{x}^0)} = a_i + \sum_{j=1}^{n} (a_{ij} + a_{ji}) x_j^0$$

だから

$$f(x_1^0, \ldots, x_n^0) + \sum_{i=1}^{n} \{a_i + \sum_{j=1}^{n} (a_{ij} + a_{ji}) x_j^0\}(x_i - x_i^0)$$

のグラフとなる。

$$\begin{aligned} & f(x_1, \ldots, x_n) - \big(f(x_1^0, \ldots, x_n^0) + \sum_{i=1}^{n} \{a_i + \sum_{j=1}^{n} (a_{ij} + a_{ji}) x_j^0\}(x_i - x_i^0)\big) \\ = & \sum_{i,j=1}^{n} b_{ij}(x_i - x_i^0)(x_j - x_j^0) \end{aligned}$$

のように書かれる。a_{ij} は対称とする。

$$\begin{aligned} & a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i,j=1}^{n} a_{ij} x_i x_j \\ & -\big(a_0 + \sum_{i=1}^{n} a_i x_i^0 + \sum_{i,j=1}^{n} a_{ij} x_i^0 x_j^0 + \sum_{i=1}^{n} \{a_i + \sum_{j=1}^{n} 2a_{ij} x_j^0\}(x_i - x_i^0)\big) \\ = & \sum_{i,j=1}^{n} a_{ij} x_i x_j - \sum_{i,j=1}^{n} a_{ij} x_i^0 x_j^0 - \sum_{i=1}^{n} \sum_{j=1}^{n} 2a_{ij} x_j^0 (x_i - x_i^0) \\ = & \sum_{i,j=1}^{n} (a_{ij} x_i x_j - 2a_{ij} x_j^0 x_i + a_{ij} x_i^0 x_j^0) \\ = & \sum_{i,j=1}^{n} a_{ij} (x_i - x_i^0)(x_j - x_j^0) \end{aligned}$$

- $A = \big(a_{ij}\big)_{i,j=1,\dots,n}$ を対称行列とし、$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ に対し、$\displaystyle\sum_{i,j=1}^{n} a_{ij}v_iv_j = {}^t\vec{v}A\vec{v}$ であるが、A は、${}^t\vec{v}A\vec{v} \geqq 0$ かつ $\vec{v} = \vec{0} \Longleftrightarrow {}^t\vec{v}A\vec{v} = 0$ が成立するとき、正定値、A は、${}^t\vec{v}A\vec{v} \leqq 0$ かつ $\vec{v} = \vec{0} \Longleftrightarrow {}^t\vec{v}A\vec{v} = 0$ が成立するとき、負定値呼ばれる。これは、A の固有値がすべて正、すべて負であることと同値である。
- ２次関数については、２次同次成分を対称行列 A で表すとき、A が正定値ならば、グラフは下に凸、負定値ならばグラフは上に凸となる。

1.7 C^2 級関数

- 方向微分 $\dfrac{\mathrm{d}}{\mathrm{d}t}f(\vec{x} + t\vec{v})$ を考えると、その方向の f の変化が分かったが、$\vec{v}$ の方向に下または上に凸であるかどうかは、２階微分 $\dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\mathrm{d}}{\mathrm{d}t}f(\vec{x} + t\vec{v})$ を計算できればわかる。
- 偏微分関数がさらに連続微分可能である場合を考える。
- 【定義】　$f : U \longrightarrow \boldsymbol{R}$ が C^2 級とは、$\dfrac{\partial f}{\partial x_i}$ が、C^1 級であることである。
- このとき、$\dfrac{\partial}{\partial x_j}\dfrac{\partial f}{\partial x_i}$ は $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ と書かれる。
（この微分の順序の書き方には最近逆の説明をする人も増えている。）
- $f(x,y) = xy\dfrac{x^2-y^2}{x^2+y^2}$　$(f(0,0) = 0)$ について、$\dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y} \neq \dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}$
- 【定理】　f が C^2 級ならば $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$ である。（これは微分積分学の講義で証明しているはず。）

【証明】　平均値の定理により、次をみたす $\beta \in (0,1)$, $\alpha \in (0,1)$ が存在する。

$$
\begin{aligned}
&\{f(x+h,y+k) - f(x,y+k)\} - \{f(x+h,y) - f(x,y)\} \\
&= k\{\big(\frac{\partial f}{\partial y}\big)_{(x+h,y+\beta k)} - \big(\frac{\partial f}{\partial y}\big)_{(x,y+\beta k)}\} \\
&= hk\big(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\big)_{(x+\alpha h,y+\beta k)}
\end{aligned}
$$

同様に、次をみたす $\alpha' \in (0,1)$, $\beta' \in (0,1)$ が存在する。

$$
\begin{aligned}
&\{f(x+h,y+k) - f(x+h,y)\} - \{f(x,y+k) - f(x,y)\} \\
&= h\{\big(\frac{\partial f}{\partial x}\big)_{(x+\alpha' h,y+k)} - \big(\frac{\partial f}{\partial x}\big)_{(x+\alpha' h,y)}\} \\
&= hk\big(\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\big)_{(x+\alpha' h,y+\beta' k)}
\end{aligned}
$$

$\dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}$, $\dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}$ が連続だから、任意の $\varepsilon > 0$ に対し、$\delta > 0$ で、$\sqrt{h^2+k^2} \leqq \delta$ ならば、$|\big(\dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}\big)_{(x+h,y+k)} - \big(\dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}\big)_{(x,y)}| \leqq \varepsilon$, $|\big(\dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}\big)_{(x+h,y+k)} - \big(\dfrac{\partial}{\partial y}\dfrac{\partial f}{\partial x}\big)_{(x,y)}| \leqq \varepsilon$

となるものが存在する。

$$
\begin{aligned}
&hk|\big(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\big)_{(x,y)}-\big(\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\big)_{(x,y)}|\\
&=hk|\big(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\big)_{(x,y)}-\big(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\big)_{(x+\alpha h,y+\beta k)}+\big(\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\big)_{(x+\alpha' h,y+\beta' k)}-\big(\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\big)_{(x,y)}|\\
&\leqq 2hk\varepsilon
\end{aligned}
$$

だから、$\big(\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\big)_{(x,y)}=\big(\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\big)_{(x,y)}$.

- f が C^2 級ならば、

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}t}f(\overrightarrow{x}+t\overrightarrow{v})&=\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n}\big(\frac{\partial f}{\partial x_i}\big)_{(\overrightarrow{x}+t\overrightarrow{v})}v_i\\
&=\sum_{i=1}^{n}\sum_{j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x}+t\overrightarrow{v})}v_iv_j
\end{aligned}
$$

任意の $\overrightarrow{x}\in U$ に対し、$\sum_{i,j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x})}v_iv_j>0$ ならば、$\overrightarrow{v}$ の方向の直線上で下に凸である。

- 任意の $\overrightarrow{v}\neq\overrightarrow{0}$ に対し、>0 ならば、f のグラフは下に凸というのは、C^2 ならば正しいのだが、議論としては少し乱暴である。

  実際、少し様子は異なるが René Thom による次のような例がある。

  $\boldsymbol{R}^2$ 上の C^∞ 級関数 f で、$\overrightarrow{0}$ を通る任意の直線上で、$\overrightarrow{0}$ のある近傍上で $f=0$（つまり、任意の $\overrightarrow{v}$ に対し、$\boldsymbol{R}\ni t\longmapsto f(t\overrightarrow{v})$ は $t=0$ の近傍で 0 だが、f は 0 の近傍で 0 ではない。

  f は、$y\leqq x^2$, $y\geqq 2x^2$ で $f=0$ とし、$x^2<y<2x^2$ で $f>0$ となるように定義する。例えば、$(0,\infty)$ で正となる $\boldsymbol{R}$ 上の C^∞ 級関数 ρ を用いて、
  $f(x,y)=(\rho(x)+\rho(-x))\rho(y-x^2)\rho(2x^2-y)$ とすればよい。

- f が C^2 級ならば、任意の $\varepsilon>0$ に対し、$\delta>0$ で $\|v\|\leqq\delta$ ならば、
  $|\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x}+\overrightarrow{v})}-\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x})}|\leqq\varepsilon$ となる δ が取れる。このとき、

$$
\begin{aligned}
&|\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Bigg(f(\overrightarrow{x}+t\overrightarrow{v})-\Bigg(f(\overrightarrow{x})+t\sum_{i=1}^{n}\big(\frac{\partial f}{\partial x_i}\big)_{(\overrightarrow{x})}v_i+\frac{t^2}{2}\sum_{i,j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x})}v_iv_j\Bigg)\Bigg)|\\
&=|\sum_{i,j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x}+t\overrightarrow{v})}v_iv_j-\sum_{i,j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x})}v_iv_j|\\
&\leqq\varepsilon(\sum_{i=1}^{n}|v_i|)^2\leqq n\|\overrightarrow{v}\|^2\varepsilon
\end{aligned}
$$

$$
|f(\overrightarrow{x}+\overrightarrow{v})-\Bigg(f(\overrightarrow{x})+\sum_{i=1}^{n}\big(\frac{\partial f}{\partial x_i}\big)_{(\overrightarrow{x})}v_i+\frac{1}{2}\sum_{i,j=1}^{n}\big(\frac{\partial^2 f}{\partial x_i\partial x_j}\big)_{(\overrightarrow{x})}v_iv_j\Bigg)|\leqq\frac{1}{2}n\|\overrightarrow{v}\|^2\varepsilon
$$

- f が U 上 C^2 級で $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{(\vec{x})}$ が正定値ならば、f のグラフは、下に凸である。

 実際、

$$f(\vec{x}+\vec{v}) - \left(f(\vec{x}) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{(\vec{x})} v_i\right) \geqq \frac{1}{2} \sum_{i,j=1}^{n} \{\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{(\vec{x})} - n\varepsilon\delta_{ij}\} v_i v_j$$

 は $\|\vec{v}\| \leqq \delta$ ならば $\geqq 0$ である。なぜならば $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j} - n\varepsilon\delta_{ij}\right)_{i,j=1,\dots,n}$ は $n\varepsilon$ が $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\dots,n}$ の最小固有値よりも小さいならば正定値である。

- 【例】 $z = x^3 - y^2$ について、$\begin{pmatrix} \dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 & -2y \end{pmatrix}$

 $\begin{pmatrix} \dfrac{\partial^2 z}{\partial x^2} & \dfrac{\partial^2 z}{\partial x \partial y} \\ \dfrac{\partial^2 z}{\partial x \partial y} & \dfrac{\partial^2 z}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & 0 \\ 0 & -2 \end{pmatrix}$ であり、$x < 0$ で上に凸、$x > 0$ は、接平面はグラフと非自明な交わりを持つ。

- 【例】 $z = x^3 - 3xy^2 = \mathrm{Re}(x + y\sqrt{-1})^3$.

 $\begin{pmatrix} \dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 - 3y^2 & -6xy \end{pmatrix}$

 $\begin{pmatrix} \dfrac{\partial^2 z}{\partial x^2} & \dfrac{\partial^2 z}{\partial x \partial y} \\ \dfrac{\partial^2 z}{\partial x \partial y} & \dfrac{\partial^2 z}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix}$

 $\dfrac{\partial^2 z}{\partial x^2} + \dfrac{\partial^2 z}{\partial y^2} = 0$ であり、接平面はグラフと非自明な交わりを持つ。

- どうすれば正定値、負定値がわかるか？

 ２変数の場合は、トレースと行列式。

 ３変数の場合はトレースと行列式では完全にはわからない。

 しかし、トレースが零で行列式が零でなければ、定値ではないことがわかり、凸ではないことがわかる。

- トレースが零、すなわち $\displaystyle\sum_{i=1}^{n} \frac{\partial^2 f}{\partial {x_i}^2} = 0$ を満たす f を調和関数と呼ぶ。

 複素関数の実部、虚部は調和関数となる。

 領域上の調和関数については、定数関数でない限り、最大値、最小値を領域の内点で取ることはないから、有界な領域上の調和関数が境界に連続に拡張しているならば、最大値、最小値は境界上でとる。

 複素関数 $f(z)$ について、$\mathrm{Re}(\log f) = \log |f(z)|$ であることを考えると、領域の内点では $|f(z)|$ は最大値をとらない。

1.8 ヘッセ行列、極大極小の判定

- 【定義】 $f : U \longrightarrow \boldsymbol{R}$ が C^2 級であり、$(Df)_{(\vec{x})} = \mathbf{0}$ のとき、$H = \left((\dfrac{\partial^2 f}{\partial x_i \partial x_j})_{(\vec{x})} \right)_{i,j=1,\dots,n}$ を f の $\vec{x}$ におけるヘッシアン（ヘッセ行列）とよぶ。

- 凸関数に関しての考察から、次の定理が従う。

- 【定理】 $f : U \longrightarrow \boldsymbol{R}$ が C^2 級であり、$(Df)_{(\vec{x})} = \mathbf{0}$ のとき、f の $\vec{x}$ におけるヘッシアン $H = \left((\dfrac{\partial^2 f}{\partial x_i \partial x_j})_{(\vec{x})} \right)_{i,j=1,\dots,n}$ が正定値ならば、f は $\vec{x}$ で極小値をとり、負定値ならば、f は $\vec{x}$ で極大値をとる。ヘッセ行列が正および負の固有値をもつ（特に行列式が負）ならば f は $\vec{x}$ で極大でも極小でもなく、$\{\vec{x} + \vec{v} \mid f(\vec{x} + \vec{v}) = f(\vec{x})\}$ は $\vec{x}$ 以外の点を含む。

- 【例】 $f(x, y) = x^3 + 2xy^2 - x.$

  $Df = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 + 2y^2 - 1 & 4xy \end{pmatrix}$

  $\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x^2} \end{pmatrix} = \begin{pmatrix} 6x & 4y \\ 4y & 4x \end{pmatrix}$ トレースは $10x$, 行列式は、$24x^2 - 16y^2$

  $Df = 0$ となる $\begin{pmatrix} \dfrac{1}{\sqrt{3}} \\ 0 \end{pmatrix}$ で極小、$\begin{pmatrix} -\dfrac{1}{\sqrt{3}} \\ 0 \end{pmatrix}$ で極大、$\begin{pmatrix} 0 \\ \pm\dfrac{1}{\sqrt{2}} \end{pmatrix}$ は鞍点。

- 【例】 $f = (x_1, x_2, x_3) = \cos x_1 + \cos x_2 + \cos x_3.$

  $Df = \begin{pmatrix} -\sin x_1 & -\sin x_2 & -\sin x_3 \end{pmatrix}.$

  $\begin{pmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_3} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial {x_2}^2} & \dfrac{\partial^2 f}{\partial x_2 \partial x_3} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_3} & \dfrac{\partial^2 f}{\partial x_2 \partial x_3} & \dfrac{\partial^2 f}{\partial {x_3}^2} \end{pmatrix} = \begin{pmatrix} -\cos x_1 & 0 & 0 \\ 0 & -\cos x_2 & 0 \\ 0 & 0 & -\cos x_3 \end{pmatrix}$

  $Df = 0$ となるのは、各座標が $\mathrm{mod}\ 2\pi$ で、0 または π になるときである。$\mathrm{mod}\ 2\pi$ で 8 点あるが、$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ で極大、$\begin{pmatrix} \pi \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pi \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \pi \end{pmatrix}, \begin{pmatrix} \pi \\ \pi \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \pi \\ \pi \end{pmatrix}, \begin{pmatrix} \pi \\ 0 \\ \pi \end{pmatrix}$ では極大でも極小でもなく、$\begin{pmatrix} \pi \\ \pi \\ \pi \end{pmatrix}$ で極小となる。

- 【例】 $f(x_1, x_2, x_3) = x_1 x_2 x_3$

  $DF = 0$ となる点の集合は、x_1 軸、x_2 軸、x_3 軸の和集合である。

  これらの点は極大でも極小でもない。

1.9 最小二乗法

- あるアフィン関数 $\vec{f} : \boldsymbol{R}^n \longrightarrow \boldsymbol{R}^m,\ \vec{x} \longmapsto \vec{y} = A\vec{x} + \vec{b}$ があり、$\vec{x}$ と $\vec{y}$ の関係を表しているとする。($\vec{y}$ は $\vec{x}$ の関数の 1 次結合である場合にすぐに拡張できる。) このとき、データ $\vec{x}^1, \ldots, \vec{x}^N, \vec{y}^1, \ldots, \vec{y}^N$ から、$A, \vec{b}$ を定める問題を考える。

- 1 次連立方程式、$\vec{y}^i = A\vec{x}^i + \vec{b}$ の変数は、$m \times n + m$ だから、$N \geqq m(n+1)$ でないと $A, \vec{b}$ を定めることはできないが、様々な誤差もあれば、仮定の誤りもあるので、通常は解がないことが予想される。

- これをうまく近似する $A, \vec{b}$ を定めるとなると、誤差を評価する関数を定め、その総和を最小にするのが良いが、そこで誤差として距離の 2 乗をとるものが最小二乗法である。(ほかに距離をとるというような考え方もありうるが、計算は難しくなる。)

- $\vec{f}(\vec{x}) = A\vec{x} + \vec{b}$ で

$$F(A, \vec{b}) = \sum_{i=1}^{N} \|A\vec{x}^i + \vec{b} - \vec{y}^i\|^2 = \sum_{i=1}^{N}\sum_{k=1}^{m}(\sum_{j=1}^{n} a_{kj}x_j^{(i)} + b_k - y_k^{(i)})^2$$

  を最小にする a_{kj}, b_k を求める。

- これは複雑そうだが、a_{kj}, b_k について 2 次式である。以前の議論により、F の 2 階微分のなす行列が正定値ならば、最小値を $\dfrac{\partial F}{\partial a_{kj}} = 0,\ \dfrac{\partial F}{\partial b_k} = 0$ となる点で取る。

- この問題は、$\vec{x}^i$ と $\vec{y}^i$ を用いて A の k 行目と b_k を決める問題が、$k = 1, \ldots, m$ について、独立に存在していると考えられる。

- 従って、A が行ベクトル ${}^t\vec{a},\ f(\vec{x}) = {}^t\vec{a}\,\vec{x} + b$ で

$$F(\vec{a}, b) = \sum_{i=1}^{N} \|{}^t\vec{a}\,\vec{x}^i + b - y^{(i)}\|^2 = \sum_{i=1}^{N}(\sum_{j=1}^{n} a_j x_j^{(i)} + b - y^{(i)})^2$$

  を最小にする a_j, b を求められれば良い。

-

$$\begin{aligned}
\frac{\partial F}{\partial a_J} &= 2\sum_{i=1}^{N}(\sum_{j=1}^{n} a_j x_j^{(i)} + b - y^{(i)})x_J^{(i)} \\
\frac{\partial F}{\partial b} &= 2\sum_{i=1}^{N}(\sum_{j=1}^{n} a_j x_j^{(i)} + b - y^{(i)})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 F}{\partial a_J \partial a_K} &= 2\sum_{i=1}^{N} x_J^{(i)} x_K^{(i)} \\
\frac{\partial^2 F}{\partial b \partial a_K} &= 2\sum_{i=1}^{N} x_K^{(i)} \\
\frac{\partial^2 F}{\partial b^2} &= 2N
\end{aligned}$$

$n\times N$ 行列 $X=\begin{pmatrix} x_1^{(1)} & \cdots & x_1^{(N)} \\ \vdots & \ddots & \vdots \\ x_n^{(1)} & \cdots & x_n^{(N)} \end{pmatrix}$, $\hat{X}=\begin{pmatrix} x_1^{(1)} & \cdots & x_1^{(N)} \\ \vdots & \ddots & \vdots \\ x_n^{(1)} & \cdots & x_n^{(N)} \\ 1 & \cdots & 1 \end{pmatrix}$ について、

$$\hat{X}\,{}^t\hat{X}=\begin{pmatrix} \sum x_1^{(i)}x_1^{(i)} & \cdots & \sum x_1^{(i)}x_n^{(i)} & \sum x_1^{(i)} \\ \vdots & \ddots & \vdots & \vdots \\ \sum x_n^{(i)}x_1^{(i)} & \cdots & \sum x_n^{(i)}x_n^{(i)} & \sum x_n^{(i)} \\ \sum x_1^{(i)} & \cdots & \sum x_n^{(i)} & N \end{pmatrix}$$

$\hat{X}$ の $n+1$ 個の行ベクトルが 1 次独立であることと、$\hat{X}\,{}^t\hat{X}$ が正定値であることは同値である。

このとき、$\dfrac{\partial F}{\partial a_J}=0,\ \dfrac{\partial F}{\partial b}=0$ を満たすように、$\vec{a}$, b を定めればよい。

$$\frac{1}{2}\frac{\partial F}{\partial a_J}=\sum_{i=1}^{N}(\sum_{j=1}^{n}a_jx_j^{(i)}+b-y^{(i)})x_J^{(i)}=\sum_{j=1}^{n}a_j\sum_{i=1}^{N}x_j^{(i)}x_J^{(i)}+b\sum_{i=1}^{N}x_J^{(i)}-\sum_{i=1}^{N}y^{(i)}x_J^{(i)}=0$$
$$\frac{1}{2}\frac{\partial F}{\partial b}=\sum_{i=1}^{N}(\sum_{j=1}^{n}a_jx_j^{(i)}+b-y^{(i)})=\sum_{j=1}^{n}a_j\sum_{i=1}^{N}x_j^{(i)}+Nb-\sum_{i=1}^{N}y^{(i)}=0$$

第 2 式は、平均を $\overline{x}_j,\overline{y}$ として、$\displaystyle\sum_{j=1}^{n}a_j\overline{x}_j+b-\overline{y}=0$

第 2 式が 0 のとき、第 1 式は

$$\begin{aligned}&\sum_{i=1}^{N}(\sum_{j=1}^{n}a_jx_j^{(i)}+b-y^{(i)})(x_J^{(i)}-\overline{x}_J)\\ =&\sum_{j=1}^{n}a_j\sum_{i=1}^{N}x_j^{(i)}(x_J^{(i)}-\overline{x}_J)+b\sum_{i=1}^{N}(x_J^{(i)}-\overline{x}_J)-\sum_{i=1}^{N}y^{(i)}(x_J^{(i)}-\overline{x}_J)=0\end{aligned}$$

ここで第 2 式に $\displaystyle\sum_{i=1}^{N}(x_J^{(i)}-\overline{x}_J)$ をかけたものを引くと、

$$\sum_{j=1}^{n}a_j\sum_{i=1}^{N}(x_j^{(i)}-\overline{x}_j)(x_J^{(i)}-\overline{x}_J)-\sum_{i=1}^{N}(y^{(i)}-\overline{y})(x_J^{(i)}-\overline{x}_J)=0$$

統計において、$\displaystyle\frac{1}{N}\sum_{i=1}^{N}(y^{(i)}-\overline{y})^2$ を確率変数 y の分散 variance、$\displaystyle\frac{1}{N}\sum_{i=1}^{N}(x^{(i)}-\overline{x})(y^{(i)}-\overline{y})$ を確率変数 x, y の共分散 covariance とよぶ。共分散行列を用いた 1 次方程式から a_j が定まり、この a_j を用いて、平均を代入したものから、b が求まる。

1.10　多変数のテーラー展開、多項式、２次曲面

- 【定義】　正整数 r に対し、$f : U \longrightarrow \boldsymbol{R}$ の r 階までの偏微分が存在し連続であるとき、f は C^r 級であるという。任意の正整数 r に対し、C^r 級のとき、C^∞ 級であるという。

- ベクトル $\vec{v}$ に対し、$\vec{v}$ 方向の r 階微分を考えると

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r} f(\vec{x} + t\vec{v}) = \sum_{i_1,\ldots,i_r=1}^{n} \Big(\frac{\partial^r f}{\partial x_{i_1} \cdots \partial x_{i_r}}\Big)_{(\vec{x}+t\vec{v})} v_{i_1} \cdots v_{i_r}$$

である。$\dfrac{\partial^r f}{\partial x_{i_1} \cdots \partial x_{i_r}}$ は微分の順序によらないので、$\dfrac{\partial^r f}{\partial x_{i_1} \cdots \partial x_{i_r}} = \dfrac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}$ $(r = r_1 + \cdots + r_n)$ と書かれる。

$\{1, \ldots, n\}$ から r 個の場所に 1 を r_1 個、…、n を r_n 個並べる順列の個数は、

$$\binom{r}{r_1}\binom{r - r_1}{r_2} \cdots \binom{r - r_1 - \cdots - r_{n-1}}{r_n} = \frac{r!}{r_1! \cdots r_n!}$$

だから、

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r} f(\vec{x} + t\vec{v}) = \sum_{r=r_1+\cdots+r_n} \frac{r!}{r_1! \cdots r_n!} \frac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}} {v_1}^{r_1} \cdots {v_n}^{r_n}$$

- 多変数関数のテーラー展開は、剰余項を書くことにあまり意味がないので次の形の定理とする。

【定理】　$\|\vec{v}\| \leqq \delta$ ならば、$\Big|\Big(\dfrac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x}+\vec{v})} - \Big(\dfrac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x})}\Big| \leqq \varepsilon$ とすると

$$\Big|f(\vec{x}+\vec{v}) - \sum_{r=0}^{r} \sum_{r=r_1+\cdots+r_n} \frac{1}{r_1! \cdots r_n!} \Big(\frac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x})} {v_1}^{r_1} \cdots {v_n}^{r_n}\Big| \leqq \frac{n^{r/2}}{r!} \|\vec{v}\|^r \varepsilon$$

【証明】

$$\Big|\frac{\mathrm{d}^r}{\mathrm{d}t^r}\Bigg(f(\vec{x} + t\vec{v}) - \sum_{s=0}^{r} \sum_{s=r_1+\cdots+r_n} \frac{t^s}{r_1! \cdots r_n!} \Big(\frac{\partial^s f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x})} {v_1}^{r_1} \cdots {v_n}^{r_n}\Bigg)\Big|$$
$$\leqq \Big| \sum_{r=r_1+\cdots+r_n} \frac{r!}{r_1! \cdots r_n!} \Big(\frac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x}+t\vec{v})} {v_1}^{r_1} \cdots {v_n}^{r_n}$$
$$- \sum_{r=r_1+\cdots+r_n} \frac{r!}{r_1! \cdots r_n!} \Big(\frac{\partial^r f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x})} {v_1}^{r_1} \cdots {v_n}^{r_n}\Big|$$
$$\leqq \varepsilon\Big(\sum_{i=1}^{n} |v_i|\Big)^r \leqq n^{r/2} \|\vec{v}\|^r \varepsilon$$

これを、積分して証明される。

- r 次のテーラー展開は、$\vec{x}$ の近傍で最もよく $f(\vec{x} + \vec{v})$ を近似する $\vec{v}$ の成分についての r 次多項式である。これは、$\vec{v}$ の成分についての s 次多項式 P と

$$\sum_{s=0}^{s} \sum_{s=r_1+\cdots+r_n} \frac{t^s}{r_1! \cdots r_n!} \Big(\frac{\partial^s f}{\partial {x_1}^{r_1} \cdots \partial {x_n}^{r_n}}\Big)_{(\vec{x})} {v_1}^{r_1} \cdots {v_n}^{r_n}$$

の差に s 次同次の項があれば、この差はある $\vec{v}$, $t \in [0,1]$ に対し、$c\,t^s$ の形であり、

$$
\begin{aligned}
&|f(\vec{x}+t\vec{v})-P(t\vec{v})| \geqq |c-n^{s/2}\varepsilon|\,|t|^s \\
\geqq\, & n^{s/2}\varepsilon|t|^s \\
\geqq\, & |f(\vec{x}+t\vec{v})-\sum_{s=0}^{s}\sum_{s=r_1+\cdots+r_n}\frac{t^s}{r_1!\cdots r_n!}\Big(\frac{\partial^s f}{\partial {x_1}^{r_1}\cdots\partial {x_n}^{r_n}}\Big)_{(\vec{x})}{v_1}^{r_1}\cdots{v_n}^{r_n}|
\end{aligned}
$$