微分積分学 -偏微分-

いろはの学習備忘録
http://168iroha.net

最終更新 2018 年 12 月 2 日

目次

第 1 章

偏微分

1.1 偏微分の定義

独立変数 $x_1, x_2, \cdots, x_n \in \mathbb{R}$ による関数 $y = f(x_1, x_2, \cdots, x_n) \in \mathbb{R}$ における微分について考える．微分の定義によれば，1 つの変数の微小変化率の極限であることから $y = f(x_1, x_2, \cdots, x_n)$ においても同様にして x_1 について

$$\lim_{h \to 0} \frac{f(x_1 + h, x_2, \cdots, x_n) - f(x_1, x_2, \cdots, x_n)}{h}$$

という 1 つの変数における微小変化率の極限を考えることができる．

これより，以下の定義が与えられる．

定義 1.

$\mathbb{R}^n$ 上で定義される独立変数 $x_1, x_2, \cdots, x_n$ による関数 $y = f(x_1, x_2, \cdots, x_n) \in \mathbb{R}$ を与えたとする．$(x_{1,0}, x_{2,0}, \cdots, x_{n,0}) \in \mathbb{R}^n$ で x_k の微小変化率の極限を

$$f_{x_k}(x_{1,0}, x_{2,0}, \cdots, x_{n,0}) = \lim_{h \to 0} \frac{f(x_{1,0}, x_{2,0}, \cdots, x_{k,0} + h, \cdots, x_{n,0}) - f(x_{1,0}, x_{2,0}, \cdots, x_{k,0}, \cdots, x_{n,0})}{h}$$

と与えたとき $f_{x_k}(x_{1,0}, x_{2,0}, \cdots, x_{n,0})$ を f の $(x_{1,0}, x_{2,0}, \cdots, x_{n,0})$ における x_k についての**偏微分係数**といい，微分係数が存在するとき f は x_k について $(x_{1,0}, x_{2,0}, \cdots, x_{n,0})$ で**偏微分可能**であるという．また，$x_1, x_2, \cdots, x_n$ の全てにおいて $(x_{1,0}, x_{2,0}, \cdots, x_{n,0})$ で偏微分可能であるとき，単に f は $(x_{1,0}, x_{2,0}, \cdots, x_{n,0})$ で**偏微分可能**であるという．

領域 $D \subseteq \mathbb{R}^n$ の任意の $(x_1, x_2, \cdots, x_n) \in D$ で偏微分可能であるとき

$$f_{x_k}(x_1, x_2, \cdots, x_n) = \lim_{h \to 0} \frac{f(x_1, x_2, \cdots, x_k + h, \cdots, x_n) - f(x_1, x_2, \cdots, x_k, \cdots, x_n)}{h}$$

によって関数 f_{x_k} を定義することができ，これを f の x_k における**偏導関数**といい，偏導関数を求める操作を**偏微分**という．これはライプニッツ記号に倣って

$$\frac{\partial y}{\partial x_k} = \frac{\partial f}{\partial x_k} = f_{x_k}$$

のようにあらわすこともある．

また，ベクトル $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}$ を用いて $f(\boldsymbol{x})$ とあらわし，ベクトルの各成分として偏微分をあらわすこともある．

偏微分はあたかも偏微分をする変数以外を固定して，その変数について微分をするという操作に等しいことから，関数が偏微分であるなら 1 変数関数の微分と等価である．また，1 変数関数の微分を偏微分と差別して**常微分**という．

つまり，偏微分可能であれば常微分において成り立つ定理がそのまま成り立つ．　これより，関数の滑らかさに関する以下の定義を与えることができる．

定義 2.

関数 f が領域 D で k 階偏微分可能かつそれが連続であるとき，f は D で C^k **級関数**であるという．特に，f が C^0 級であることは f が連続関数であることを，C^∞ 級であることは無限回数偏微分可能であることを示すクラスであり，関数の滑らかさの程度を示す．

$k \geq 1$ であれば f は**滑らかな関数**であるといい，D 内で不連続点が有限個であるとき，f は D で**区分的に滑らかな関数**であるという．

1.2　全微分

独立変数 $x_1, x_2, \cdots, x_n \in \mathbb{R}$ による関数 $y = f(x_1, x_2, \cdots, x_n) \in \mathbb{R}$ において，全ての変数を微小量変化させたときの y の変化量について考える．y の変化量を Δy，それぞれの独立変数の変化量を $\Delta x_1, \Delta x_2, \cdots, \Delta x_n$ とすることで

$$\Delta y = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2, \cdots, x_n)$$

とあらわされるが，それぞれの変数による y の微小変化量は $f_{x_k}\Delta x_k$ であるため

$$\Delta y \fallingdotseq \sum_{k=1}^{n} f_{x_k}(x_1, x_2, \cdots, x_n)\Delta x_k$$

と近似されることを考えることができる．

また，$\boldsymbol{x} = (x_1, x_2, \cdots, x_n), \Delta\boldsymbol{x} = (\Delta x_1, \Delta x_2, \cdots, \Delta x_n)$ および $\boldsymbol{f}(\boldsymbol{x}) = (f_{x_1}(\boldsymbol{x}), f_{x_2}(\boldsymbol{x}), \cdots, f_{x_n}(\boldsymbol{x}))$ とあらわすとすれば

$$\Delta y \fallingdotseq \sum_{k=1}^{n} f_{x_k}(x_1, x_2, \cdots, x_n)\Delta x_k = \boldsymbol{f}(\boldsymbol{x}) \cdot \Delta\boldsymbol{x}$$

とあらわされる．これより，常微分のイプシロン-デルタ論法を用いた形式により近似できる必要十分条件は

$$\Delta y = f(\boldsymbol{x} + \Delta\boldsymbol{x}) - f(\boldsymbol{x}) = (y_0 + \varepsilon)|\Delta\boldsymbol{x}|$$

に対して $|\Delta\boldsymbol{x}| \to 0$ で $\varepsilon \to 0$ となる定数 y_0 が存在することとあらわされる．y_0 については

$$\frac{\boldsymbol{f}(\boldsymbol{x}) \cdot \Delta\boldsymbol{x}}{|\Delta\boldsymbol{x}|}$$

に対応する．これより，$|\Delta\boldsymbol{x}| \to 0$ で $\Delta y \to \boldsymbol{f}(\boldsymbol{x}) \cdot \Delta\boldsymbol{x}$ となるには $|\Delta\boldsymbol{x}|$ よりも $\varepsilon|\Delta\boldsymbol{x}|$ の方が収束が早い，すなわち

$$\lim_{|\Delta\boldsymbol{x}| \to 0} \frac{\varepsilon|\Delta\boldsymbol{x}|}{|\Delta\boldsymbol{x}|} = 0$$

となることである．このとき，$\varepsilon|\Delta\boldsymbol{x}|$ の項は単に ε とあらわすこともできる．

これより，以下の定義を与えることができる．

定義 3.

独立変数 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ による関数 $y = f(\boldsymbol{x}) \in \mathbb{R}$ において，$\Delta\boldsymbol{x} = (\Delta x_1, \Delta x_2, \cdots, \Delta x_n)$ により $\Delta y = f(\boldsymbol{x} + \Delta\boldsymbol{x}) - f(\boldsymbol{x})$ と y の変化量を定義する．$\boldsymbol{f}(\boldsymbol{x}) = (f_{x_1}(\boldsymbol{x}), f_{x_2}(\boldsymbol{x}), \cdots, f_{x_n}(\boldsymbol{x}))$ および ε を用いて

$$\Delta y = \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{x} + \varepsilon = f_{x_1}(\boldsymbol{x})\Delta x_1 + f_{x_2}(\boldsymbol{x})\Delta x_2 + \cdots + f_{x_n}(\boldsymbol{x})\Delta x_n + \varepsilon$$

とあらわされるとき

$$\lim_{|\Delta\boldsymbol{x}| \to 0} \frac{\varepsilon}{|\Delta\boldsymbol{x}|} = 0$$

となるならば，$|\Delta\boldsymbol{x}| \to 0$ で $\Delta x_k \to 0$ および $\Delta y \to 0$ となるが，このことを微小量 $\mathrm{d}x_1, \mathrm{d}x_2, \cdots, \mathrm{d}x_n$ および $\mathrm{d}y$ を用いて

$$\mathrm{d}y = f_{x_1}(\boldsymbol{x})\mathrm{d}x_1 + f_{x_2}(\boldsymbol{x})\mathrm{d}x_2 + \cdots + f_{x_n}(\boldsymbol{x})\mathrm{d}x_n$$

とあらわすことができる．これを $y = f(\boldsymbol{x})$ の**完全微分**もしくは**全微分**という．

次に，$\boldsymbol{x}_0 \in \mathbb{R}^n$ で $f(\boldsymbol{x})$ が全微分可能であるとき，$\boldsymbol{x}_0$ における $f(\boldsymbol{x})$ の振る舞いについて考える．このとき，Δy についての式を変形することで

$$\lim_{|\Delta \boldsymbol{x}| \to 0} f(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = \lim_{|\Delta \boldsymbol{x}| \to 0} ((y_0 + \varepsilon)|\Delta \boldsymbol{x}| + f(\boldsymbol{x}_0)) = f(\boldsymbol{x}_0)$$

となり，$f(\boldsymbol{x})$ が $\boldsymbol{x}_0$ で全微分可能であるならば $f(\boldsymbol{x})$ は $\boldsymbol{x}_0$ で連続ということである．これはちょうど，$f(\boldsymbol{x})$ が $\boldsymbol{x}_0$ で偏微分可能であるならば $f(\boldsymbol{x})$ は $\boldsymbol{x}_0$ で連続であることと似たような対応をすることがわかる．

ここで，全微分の十分条件についての定理を示す．

全微分の十分条件

独立変数 $x_1, x_2, \cdots, x_n \in \mathbb{R}$ による関数 $y = f(x_1, x_2, \cdots, x_n) \in \mathbb{R}$ が領域 $D \subset \mathbb{R}^n$ で C^1 級関数であるとき，全微分が可能である．

Proof.

y の変化量を Δy，それぞれの独立変数の変化量を $\Delta x_1, \Delta x_2, \cdots, \Delta x_n$ とすることで

$$\Delta y = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2, \cdots, x_n)$$

とあらわされるとする．また，x_1 に対する偏微分についてはラグランジュの平均値の定理より

$$\begin{aligned}&\frac{f(x_1 + \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n)}{\Delta x_1}\\ =&f_{x_1}(x_1 + \theta_1 \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n)\end{aligned}$$

を満たす $0 < \theta_1 < 1$ が存在し，これを Δy に代入することで

$$\begin{aligned}\Delta y =& f_{x_1}(x_1 + \theta_1 \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n)\Delta x_1\\ &+ f(x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2, \cdots, x_n)\end{aligned}$$

が得られる．同様にして，x_2 の偏微分についても

$$\begin{aligned}&\frac{f(x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2, \cdots, x_n + \Delta x_n)}{\Delta x_2}\\ =&f_{x_2}(x_1, x_2 + \theta_2 \Delta x_2, \cdots, x_n + \Delta x_n)\end{aligned}$$

を満たす $0 < \theta_2 < 1$ が存在し，これを Δy に代入することで

$$\begin{aligned}\Delta y =& f_{x_1}(x_1 + \theta_1 \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n)\Delta x_1\\ &+ f_{x_2}(x_1, x_2 + \theta_2 \Delta x_2, \cdots, x_n + \Delta x_n)\Delta x_2\\ &+ f(x_1, x_2, \cdots, x_n + \Delta x_n) - f(x_1, x_2, \cdots, x_n)\end{aligned}$$

が得られる．これを繰り返すことにより，

$$\begin{aligned}\Delta y =& f_{x_1}(x_1 + \theta_1 \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n)\Delta x_1\\ &+ f_{x_2}(x_1, x_2 + \theta_2 \Delta x_2, \cdots, x_n + \Delta x_n)\Delta x_2\\ &+ \cdots\\ &+ f_{x_{n-1}}(x_1, x_2, \cdots, x_{n-1} + \theta_{n-1}\Delta x_{n-1}, x_n + \Delta x_n)\Delta x_{n-1}\\ &+ f_{x_n}(x_1, x_2, \cdots, x_n + \theta_n \Delta x_n)\Delta x_n\end{aligned}$$

とあらわすことができることがわかる．

f は C^1 級の関数であるため $\Delta x_1, \Delta x_2, \cdots, \Delta x_n \to 0$ で例えば f_{x_1} については

$$f_{x_1}(x_1 + \theta_1 \Delta x_1, x_2 + \Delta x_2, \cdots, x_n + \Delta x_n) \to f_{x_1}(x_1, x_2, \cdots, x_n)$$

となり，これは $f(x_1, x_2, \cdots, x_n)$ の全ての偏導関数について同様のことが成り立つ．

これより，$\Delta x_1, \Delta x_2, \cdots, \Delta x_n \to 0$ で $\Delta y \to 0$ となるが，このことを微小量 $\mathrm{d}x_1, \mathrm{d}x_2, \cdots, \mathrm{d}x_n$ および $\mathrm{d}y$ を用いて

$$\mathrm{d}y = f_{x_1}(x_1, x_2, \cdots, x_n)\mathrm{d}x_1 + f_{x_2}(x_1, x_2, \cdots, x_n)\mathrm{d}x_2 + \cdots + f_{x_n}(x_1, x_2, \cdots, x_n)\mathrm{d}x_n$$

とあらわすことができ，これは全微分の定義に等しい．

よって，命題は証明された．

□

1.3　偏微分の連鎖律

独立変数 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ による領域 D で定義される関数 $y = f(\boldsymbol{x}) \in \mathbb{R}$ のそれぞれの変数の関係が，別の変数 $\boldsymbol{u} = (u_1, u_2, \cdots, u_m)$ によって与えられる，すなわち

$$x_k = x_k(\boldsymbol{u}) \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u})$$

のように合成関数として与えられたとする．なお，$x_k(\boldsymbol{u})$ は偏微分可能であるとする．

$f(\boldsymbol{x})$ が D で全微分可能であるとすれば，y の微小変化 Δy はそれぞれの独立変数の微小変化 $\Delta x_1, \Delta x_2, \cdots, \Delta x_n$ と余剰項 ε を用いて

$$\Delta y = \sum_{k=1}^{n} f_{x_k}(\boldsymbol{x}) \Delta x_k + \varepsilon$$

とあらわされる．このとき，Δy は u_k の変化のみによって生成されるとすれば，その u_l の微小変化 Δu_l で商をとることで

$$\frac{\Delta y}{\Delta u_l} = \sum_{k=1}^{n} f_{x_k}(\boldsymbol{x}) \frac{\Delta x_k}{\Delta u_l} + \frac{\varepsilon}{\Delta u_l}$$

とあらわされ，余剰項については

$$\frac{\varepsilon}{\Delta u_l} = \frac{\varepsilon}{|\boldsymbol{x}|} \cdot \frac{|\boldsymbol{x}|}{\Delta u_l} = \frac{\varepsilon}{|\boldsymbol{x}|} \cdot \sqrt{\left(\frac{\Delta x_1}{\Delta u_l}\right)^2 + \left(\frac{\Delta x_2}{\Delta u_l}\right)^2 + \cdots + \left(\frac{\Delta x_n}{\Delta u_l}\right)^2}$$

とあらわされるが，$x_k(\boldsymbol{u})$ は偏微分可能であることから，$\frac{\Delta x_l}{\Delta u_l}$ は $\Delta u_l \to 0$ の極限をとっても有限の値をとる．つまり，

$$\lim_{\Delta u_l \to 0} \frac{\varepsilon}{\Delta u_l} = 0$$

である．

これより，以下の定理を与えることができる．

偏微分の連鎖律

独立変数 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ による領域 D で定義される全微分可能な関数 $y = f(\boldsymbol{x})$ を与えたとする．このとき，別の変数 $\boldsymbol{u} = (u_1, u_2, \cdots, u_m)$ によって $\boldsymbol{x}$ の対応が与えられる，すなわち

$$x_k = x_k(\boldsymbol{u})$$

のような合成関数として与えられたとする．

このとき，$x_k(\boldsymbol{u})$ が偏微分可能であるならば y を u_l で偏微分した偏導関数は

$$\frac{\partial y}{\partial u_l} = \sum_{k=1}^{n} f_{x_k}(\boldsymbol{x}) \frac{\partial x_k}{\partial u_l} = \sum_{k=1}^{n} \frac{\partial y}{\partial x_k} \frac{\partial x_k}{\partial u_l}$$

となる．これを偏微分における**連鎖律**もしくは**チェーンルール**といい，常微分における連鎖律の拡張である．

Proof.

導入より自明である．

□

また，連鎖律は $m \times n$ 行列を用いることで

$$\begin{pmatrix} \dfrac{\partial y}{\partial u_1} \\ \dfrac{\partial y}{\partial u_2} \\ \vdots \\ \dfrac{\partial y}{\partial u_m} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x_1}{\partial u_1} & \dfrac{\partial x_2}{\partial u_1} & \cdots & \dfrac{\partial x_n}{\partial u_1} \\ \dfrac{\partial x_1}{\partial u_2} & \dfrac{\partial x_2}{\partial u_2} & \cdots & \dfrac{\partial x_n}{\partial u_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_1}{\partial u_m} & \dfrac{\partial x_2}{\partial u_m} & \cdots & \dfrac{\partial x_n}{\partial u_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial y}{\partial x_1} \\ \dfrac{\partial y}{\partial x_2} \\ \vdots \\ \dfrac{\partial y}{\partial x_n} \end{pmatrix}$$

とあらわされ，これは行列によって y の $\boldsymbol{x}$ による偏導関数が $\boldsymbol{u}$ による偏導関数に変換されることを示しており，偏微分の連鎖律を一般的にあらわした形である．また，この行列は $\mathbb{R}^n \to \mathbb{R}^m$ への写像であり，複数回の偏導関数の変換も行列の積によって与えることができる．

ここで，以下の定義を与える．

定義 4.

独立変数 $x_1, x_2, \cdots, x_n \in \mathbb{R}$ および $u_1, u_2, \cdots, u_m \in \mathbb{R}$ が与えられたとき，任意の k に対する x_k について $x_k = x_k(u_1, u_2, \cdots, u_m)$ というようにあらわされるとする．このとき，

$$J_x = \begin{pmatrix} \dfrac{\partial x_1}{\partial u_1} & \dfrac{\partial x_2}{\partial u_1} & \cdots & \dfrac{\partial x_n}{\partial u_1} \\ \dfrac{\partial x_1}{\partial u_2} & \dfrac{\partial x_2}{\partial u_2} & \cdots & \dfrac{\partial x_n}{\partial u_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_1}{\partial u_m} & \dfrac{\partial x_2}{\partial u_m} & \cdots & \dfrac{\partial x_n}{\partial u_m} \end{pmatrix}$$

と与えられる行列 J_x を**ヤコビ行列**という．$m = n$ の場合においては正方行列となるため行列式を定義することができ，ヤコビ行列の行列式 $|J_x|$ を**ヤコビ行列式**もしくは**ヤコビアン**という．また，$\boldsymbol{u} = (u_1., u_2, \cdots, u_m)$ としたとき，$\boldsymbol{u} = \boldsymbol{u}_0$ におけるヤコビ行列は $J_x(\boldsymbol{u}_0)$ とあらわす．

なお，慣習としてヤコビ行列は J_x を転置した形式で扱われるが，行列式や変換行列としての意味は変わらないため本質的に同一のものである．基本的にこの慣習に則ってヤコビ行列を用いる．

よって，ヤコビ行列を用いることで連鎖律は

$$\begin{pmatrix} \dfrac{\partial y}{\partial u_1} & \dfrac{\partial y}{\partial u_2} & \cdots & \dfrac{\partial y}{\partial u_m} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial y}{\partial x_1} & \dfrac{\partial y}{\partial x_2} & \cdots & \dfrac{\partial y}{\partial x_n} \end{pmatrix} J_x$$

とあらわすことができる．

また，独立変数 $\boldsymbol{x}' = (x'_1, x'_2, \cdots, x'_l) \in \mathbb{R}^l$ により $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u})$ ではなく $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{x}'(\boldsymbol{u}))$ というように合成関数として与えられるとすれば，これは $\mathbb{R}^m \to \mathbb{R}^l \to \mathbb{R}^n$ という写像を与えている．このとき，連鎖律より

$$\frac{\partial x_i}{\partial u_j} = \sum_{k=1}^{l} \frac{\partial x_i}{\partial x'_k} \frac{\partial x'_k}{\partial u_j}$$

とあらわされることから

$$\begin{pmatrix} \dfrac{\partial y}{\partial u_1} \\ \dfrac{\partial y}{\partial u_2} \\ \vdots \\ \dfrac{\partial y}{\partial u_m} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x'_1}{\partial u_1} & \dfrac{\partial x'_2}{\partial u_1} & \cdots & \dfrac{\partial x'_l}{\partial u_1} \\ \dfrac{\partial x'_1}{\partial u_2} & \dfrac{\partial x'_2}{\partial u_2} & \cdots & \dfrac{\partial x'_l}{\partial u_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x'_1}{\partial u_m} & \dfrac{\partial x'_2}{\partial u_m} & \cdots & \dfrac{\partial x'_l}{\partial u_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x_1}{\partial x'_1} & \dfrac{\partial x_2}{\partial x'_1} & \cdots & \dfrac{\partial x_n}{\partial x'_1} \\ \dfrac{\partial x_1}{\partial x'_2} & \dfrac{\partial x_2}{\partial x'_2} & \cdots & \dfrac{\partial x_n}{\partial x'_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_1}{\partial x'_l} & \dfrac{\partial x_2}{\partial x'_l} & \cdots & \dfrac{\partial x_n}{\partial x'_l} \end{pmatrix} \begin{pmatrix} \dfrac{\partial y}{\partial x_1} \\ \dfrac{\partial y}{\partial x_2} \\ \vdots \\ \dfrac{\partial y}{\partial x_n} \end{pmatrix}$$

となり，この行列部の転置が合成写像のヤコビ行列 $J_{x \circ x'}$ となる．

$\boldsymbol{u}_0 \in \mathbb{R}^m$ に対するヤコビ行列で考えれば，$J_{x'}$ は $\boldsymbol{u}_0$ がそのまま対応するため $J_{x'}(\boldsymbol{u}_0)$ となるが，J_x は $\boldsymbol{x} \circ \boldsymbol{x}'$ より

$\boldsymbol{x}'$ を作用させることにより与えられるため $J_x(\boldsymbol{x}'(\boldsymbol{u}_0))$ となる．つまり，合成関数におけるヤコビ行列は

$$J_{x\circ x'}(\boldsymbol{u}_0) = J_x(\boldsymbol{x}'(\boldsymbol{u}_0))J_{x'}(\boldsymbol{u}_0)$$

となる．これより，$\boldsymbol{x}$ と $\boldsymbol{x}'$ が C^1 級であるならば $J_x(\boldsymbol{x}'(\boldsymbol{u}_0))$ と $J_{x'}(\boldsymbol{u}_0)$ は C^0 級であり，行列の積は線型であることから $J_x(\boldsymbol{x}'(\boldsymbol{u}_0))J_{x'}(\boldsymbol{u}_0)$ は連続，つまりは C^0 級である．これより，$J_{x\circ x'}(\boldsymbol{u}_0)$ も C^0 級であり，$\boldsymbol{x}\circ\boldsymbol{x}'$ が C^1 であることがわかる．

よって，以下の補題を与えることができる．

補題 1. $r \geq 1$ で C^r 級写像 $f:\mathbb{R}^m \to \mathbb{R}^l$ と $g:\mathbb{R}^l \to \mathbb{R}^n$ を与えたとき，この合成写像 $g\circ f:\mathbb{R}^m \to \mathbb{R}^n$ は C^r 級写像である．

Proof.

数学的帰納法により証明を与える．

$r=1$ のときは合成写像のヤコビ行列より真であることは明らかであるため，r で命題が成り立つことを仮定して $r+1$ で成り立つことを示す．f と g が C^{r+1} 級であれば J_f と J_g は C^r 級であるため $\boldsymbol{x}\in\mathbb{R}^m$ で

$$J_{g\circ f}(\boldsymbol{x}) = J_g(f(\boldsymbol{x}))J_f(\boldsymbol{x})$$

となり，仮定より $J_g(f(\boldsymbol{x}))J_f(\boldsymbol{x})$ も C^r 級となる．つまり，$J_{g\circ f}(\boldsymbol{x})$ も C^r 級であり，$g\circ f$ は C^{r+1} 級となる．

よって，命題は証明された．

□

1.4　接超平面

常微分では関数に対する接線の方程式を考えることができるが，偏微分においても同様のことを，すなわち独立変数 $x_1, x_2, \cdots, x_n \in \mathbb{R}$ による領域 D で定義される C^1 級関数 $y = f(x_1, x_2, \cdots, x_n) \in \mathbb{R}$ に接する超平面について考える．まず，独立変数 $x\in\mathbb{R}$ による関数 $y=f(x)\in\mathbb{R}$ の x_0 における接線の方程式は

$$y - f(x_0) = f'(x_0)(x - x_0)$$

となるが，

$$-f'(x_0)(x-x_0) + (y - f(x_0)) = 0$$

とあらわすことにより，$(x_0, f(x_0))$ を通る $(-f'(x_0), 1)$ を法線ベクトルとする 2 次元における超平面の式となることがわかる．また，$z = y - f(x) = 0$ と定義された関数において

$$(-f'(x_0), 1) = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)$$

とあらわされる．

ここで，以下の定義を与える．

定義 5.

変数 $x_1, x_2, \cdots, x_n, y \in \mathbb{R}$ による方程式 $\varphi(x_1, x_2, \cdots, x_n, y) = 0$ が与えられたとき，$x_1, x_2, \cdots, x_n$ を独立変数とする関数 $f(x_1, x_2, \cdots, x_n)$ が

$$\varphi(x_1, x_2, \cdots, x_n, f(x_1, x_2, \cdots, x_n)) = 0$$

を満たすとき，$y = f(x_1, x_2, \cdots, x_n)$ を $\varphi(x_1, x_2, \cdots, x_n, y) = 0$ によって定まる**陰関数**という．

また，対称的に $y = f(x_1, x_2, \cdots, x_n)$ の形で定義される関数を**陽関数**という．

また，$f(x)$ の法線ベクトル $\boldsymbol{n}$ は形式的に

$$\boldsymbol{n} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) z$$

とあらわすことができ，z に対する作用素を考えることができる．

ここで，以下の定義を与える．

定義 6.

$\mathbb{R}^n$ における正規直交基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ で成分をそれぞれ $x_1, x_2, \cdots, x_n$ としたとき，

$$\nabla = \sum_{k=1}^{n} \boldsymbol{e}_k \frac{\partial}{\partial x_k}$$

と定義される作用素を**ハミルトンの演算子**といい，∇ を**ナブラ**と読む．

また，偏微分の定義よりナブラは明らかに線型作用素であることから **微分線型作用素** ということもある．

これより，$z = y - f(x_1, x_2, \cdots, x_n) = 0$ とおいたとき，∇z が超曲面 f 上の法線ベクトルであることを示せばいい．つまり，f 上に定義される曲線の接線と常に直交することを示せばいい．このとき，媒介変数 t により

$$\begin{cases} x_1 = x_1(t) \\ x_2 = x_2(t) \\ \vdots \\ x_n = x_n(t) \end{cases}$$

とあらわされるならば，$y(t) = f(x_1(t), x_2(t), \cdots, x_n(t))$ は f 上の曲線となる．この曲線を示すベクトル $\boldsymbol{r}(t)$ を

$$\boldsymbol{r}(t) = (x_1(t), x_2(t), \cdots, x_n(t), y(t))$$

と定義することで，$z = y - f(x_1(t), x_2(t), \cdots, x_n(t)) = 0$ 上の曲線となる．これより，t による導関数は連鎖律より

$$\begin{aligned} -\sum_{k=1}^{n} \frac{\partial z}{\partial x_k} \frac{\mathrm{d}x_k}{\mathrm{d}t} + \frac{\partial z}{\partial y} \frac{\mathrm{d}y}{\mathrm{d}t} &= 0 \\ \nabla z \cdot \left(\frac{\mathrm{d}x_1}{\mathrm{d}t}, \frac{\mathrm{d}x_2}{\mathrm{d}t}, \cdots, \frac{\mathrm{d}x_n}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}\right) &= 0 \\ \nabla z \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} &= 0 \end{aligned}$$

となり，$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ は $\boldsymbol{r}(t)$ の接ベクトルであることから，これは接ベクトルと ∇z が常に直交することを示している．つまり，∇z は f 上の法線ベクトルである．

ここで，接超平面が存在することの意味を考えれば，f は C^1 級であるため，$\boldsymbol{x}_0 \in D$ における接超平面は $\boldsymbol{x}_0$ の近傍で f に近似されることがわかる．この近似された式は，$\boldsymbol{x}$ に対するナブラを $\nabla_{\boldsymbol{x}}$ とすることで

$$\begin{aligned} \nabla z(\boldsymbol{x}_0, f(\boldsymbol{x}_0)) \cdot (\boldsymbol{x} - \boldsymbol{x}_0, f(\boldsymbol{x}) - f(\boldsymbol{x}_0)) &= 0 \\ -\nabla_{\boldsymbol{x}} f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + (f(\boldsymbol{x}) - f(\boldsymbol{x}_0)) &= 0 \\ f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \nabla_{\boldsymbol{x}} f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) \end{aligned}$$

とあらわすこともできる．これは元は接超平面の方程式であるが，$f(\boldsymbol{x})$ についての式であると考えることで，$\boldsymbol{x}_0$ の近傍で局所的に近似できる式であるといえる．

同様にして，$U \subset \mathbb{R}^n$ で定義される C^1 級関数 $g: D \to \mathbb{R}^m$ が $\boldsymbol{x} \in U$ で $\boldsymbol{y} = g(\boldsymbol{x}) \in \mathbb{R}^m$ とあらわされる場合について考える．$\boldsymbol{y}$ を $\boldsymbol{y} = (y_1, y_2, \cdots, y_m)$ とあらわせば，$\boldsymbol{y}$ の各成分について f と同様のことが考えられるため $\boldsymbol{x}_0 \in U$ での近似は $g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x}))$ とあらわすことで

$$y_k = g_k(\boldsymbol{x}_0) + \nabla_{\boldsymbol{x}} g_k(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)$$

と与えられ，$\boldsymbol{y}$ の全ての成分をまとめれば

$$\begin{aligned}
g(\boldsymbol{x}) =& g(\boldsymbol{x}_0) + \begin{pmatrix} \nabla_{\boldsymbol{x}} g_1(\boldsymbol{x}_0) \\ \nabla_{\boldsymbol{x}} g_2(\boldsymbol{x}_0) \\ \vdots \\ \nabla_{\boldsymbol{x}} g_m(\boldsymbol{x}_0) \end{pmatrix} (\boldsymbol{x} - \boldsymbol{x}_0) \\
=& g(\boldsymbol{x}_0) + \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial g_1}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_1}{\partial x_n}(\boldsymbol{x}_0) \\ \frac{\partial g_2}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial g_2}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_2}{\partial x_n}(\boldsymbol{x}_0) \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial g_m}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_m}{\partial x_n}(\boldsymbol{x}_0) \end{pmatrix} (\boldsymbol{x} - \boldsymbol{x}_0) \\
=& g(\boldsymbol{x}_0) + J_g(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)
\end{aligned}$$

とあらわされる．

ここで，このことについての定義を与える．

定義 7.

独立変数 $\boldsymbol{x} \in \mathbb{R}^n$ による領域 $D \subset \mathbb{R}^n$ で定義される C^1 級写像 $f : D \to \mathbb{R}^m$ は $\boldsymbol{x}_0 \in D$ で ε を用いて

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + J_f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + \varepsilon$$

とあらわされ，$\boldsymbol{x} \to \boldsymbol{x}_0$ で $|\varepsilon| \to 0$ となることから $\boldsymbol{x}$ が十分に $\boldsymbol{x}_0$ に近いならば ε を無視して十分な近似として与えることができる．また，これは $\mathbb{R}^2$ における $\boldsymbol{x}_0$ での接線の方程式の一般形の 1 つであると考えることができる．

このことを $\boldsymbol{x}_0$ を中心とした **1 次近似**もしくは**線型近似**という．

ここで，接超平面について再び考えれば，$y = f(\boldsymbol{x})$ は z による定数倍を除いた唯一の陰関数であることから，より一般の z が複数の陰関数をもつ場合についても同様のことが適用されると考えることができる．

そこで，以下の補題および定理を与える．

補題 2.　領域 $D \subset \mathbb{R}^n$ で定義される C^1 級写像 $f : D \to \mathbb{R}^m$ において，$\boldsymbol{x} \in D$ で $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x}))$ とあらわされるとする．凸な領域 $U \subset D$ で

$$\left| \frac{\partial f_i}{\partial x_j} \right| \leq K$$

が成り立つとして，$\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{x}' \in U$ を与えれば，リプシッツ定数についての関係式

$$|f(\boldsymbol{x} + \boldsymbol{x}') - f(\boldsymbol{x})| \leq \sqrt{mn} K |\boldsymbol{x}'|$$

が成り立つ．

Proof.

任意の f_i は C^1 級であることから全微分可能であり，$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ のようにあらわすこととすれば

$$f_i(\boldsymbol{x} + \boldsymbol{x}') - f_i(\boldsymbol{x}) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} x'_j + \varepsilon$$

が成り立つ．$|\boldsymbol{x}'| \to 0$ で $\varepsilon \to 0$ となることから，独立変数 t を導入して $\boldsymbol{x} + t\boldsymbol{x}' \in U$ とあらわすことで $t \to 0$ で $\varepsilon \to 0$ となるため

$$\frac{\mathrm{d} f_i}{\mathrm{d} t}(\boldsymbol{x} + t\boldsymbol{x}') = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\boldsymbol{x} + t\boldsymbol{x}') x'_j$$

が成り立つ．これより，

$$\begin{aligned} f_i(\boldsymbol{x}+\boldsymbol{x}') - f_i(\boldsymbol{x}) &= [f_i(\boldsymbol{x}+t\boldsymbol{x}')]_0^1 \\ &= \int_0^1 \left(\frac{\mathrm{d}f_i}{\mathrm{d}t}(\boldsymbol{x}+t\boldsymbol{x}') \right) \mathrm{d}t \\ &= \int_0^1 \left(\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}+t\boldsymbol{x}') x'_j \right) \mathrm{d}t \end{aligned}$$

となり，絶対値で評価をすることにより

$$\left| \int_0^1 \left(\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}+t\boldsymbol{x}') x'_j \right) \mathrm{d}t \right| \leq \int_0^1 \left(\sum_{j=1}^{n} K|x'_j| \right) \mathrm{d}t = K \sum_{j=1}^{n} |x'_j| \leq K\sqrt{n}|\boldsymbol{x}'|$$

が成り立ち

$$|f_i(\boldsymbol{x}+\boldsymbol{x}') - f_i(\boldsymbol{x})| \leq K\sqrt{n}|\boldsymbol{x}'| \ \Rightarrow \ |f(\boldsymbol{x}+\boldsymbol{x}') - f(\boldsymbol{x})| \leq \sqrt{mn}K|\boldsymbol{x}'|$$

が得られる．

よって，命題は証明された．

□

補題 3.　一般線型群 $GL_n(\mathbb{R})$ に対して，これの逆元を対応させる写像 $\varphi : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ は C^∞ 級である．

Proof.

C^1 級写像 $f : \mathbb{R} \to GL_n(\mathbb{R})$ と $g : \mathbb{R} \to GL_n(\mathbb{R})$ を与えたとき，積の導関数の導出より $x \in \mathbb{R}$ で

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$$

が成り立つことがわかる．

これと同様にして，$x \in GL_n(\mathbb{R})$ の (i,j) 成分を $x_{i,j}$ とあらわせば

$$\frac{\partial}{\partial x_{i,j}}(\varphi(x)x) = \varphi_{x_{i,j}}(x)x + \varphi(x) = \varphi_{x_{i,j}}(x)x + x^{-1}$$

となるが，単位行列 $I \in GL_n(\mathbb{R})$ を用いて $\varphi(x)x = I$ であることから

$$\varphi_{x_{i,j}}(x)x + x^{-1} = 0 \ \Leftrightarrow \ \varphi_{x_{i,j}}(x) = -x^{-1}x^{-1}$$

となる．

このとき，φ を $x_{i,j}$ による n 階偏導関数を $\varphi^{(n)}_{x_{i,j}}$ とあらわすとすれば

$$\varphi^{(n)}_{x_{i,j}}(x) = (-1)^n n!(x^{-1})^{n+1}$$

となると考えられる．$n=1$ のときは明らかであるため，n で成り立つことを仮定して $n+1$ で成り立つことを示す．$n=1$ のときと同様にして $\varphi^{(n)}_{x_{i,j}} \frac{n!}{x}^{n+1}$ を $x_{i,j}$ で偏微分すれば

$$\begin{aligned} \frac{\partial}{\partial x_{i,j}}(\varphi^{(n)}_{x_{i,j}} x^{n+1}) =& \frac{1}{n!}\varphi^{(n+1)}_{x_{i,j}}(x)x^{n+1} + \frac{1}{n!}\varphi^{(n)}_{x_{i,j}}(n+1)!x^n \\ =& \frac{1}{n!}\varphi^{(n+1)}_{x_{i,j}}(x)x^{n+1} + (-1)^n(n+1)x^{-1} \end{aligned}$$

となり，$\varphi^{(n)}_{x_{i,j}} \frac{1}{n!} x^{n+1} = (-1)^n I$ であることから

$$\frac{1}{n!}\varphi^{(n+1)}_{x_{i,j}}(x)x^{n+1} + (-1)^n(n+1)x^{-1} = 0 \ \Leftrightarrow \ \varphi^{(n+1)}_{x_{i,j}}(x) = (-1)^{n+1}(n+1)!(x^{-1})^{(n+1)+1}$$

となり，数学的帰納法により φ を $x_{i,j}$ による n 階偏導関数についての式が成り立つことが示された．

これにより，φ は無限回数偏微分可能であり，$x^{-1} \in GL_n(\mathbb{R})$ より，それは連続であるため φ は $GL_n(\mathbb{R})$ 全域で

C^∞ 級である．

よって，命題は証明された．

□

逆写像定理

$r \geq 1$ で領域 $D \subset \mathbb{R}^n$ で定義される C^r 級写像 $f : D \to \mathbb{R}^n$ において，$\boldsymbol{x}_0 \in D$ でヤコビ行列 $J_f(\boldsymbol{x}_0)$ が正則であるとき，$\boldsymbol{x}_0$ 近傍の領域 U で f は逆写像 $g : V \to U$ をもつ．

これを**逆写像定理**という．

Proof.

まず，U を $-\boldsymbol{x}_0$ だけ平行移動させた領域 U_0 を考えれば，$\boldsymbol{0} \in U_0$ であり，$f_0(\boldsymbol{0}) = \boldsymbol{0}$ となるような C^r 級写像 f_0 を定義して可換図式

$$\begin{tikzcd} U \arrow[r, "f"] & V \\ U_0 \arrow[u, "+\boldsymbol{x}_0"] \arrow[r, "f_0"'] & V_0 \arrow[u, "\phi"'] \end{tikzcd}$$

を満たすように V_0 および ϕ を定義する．このとき，f は $\boldsymbol{x} \in U$ で $\boldsymbol{x} - \boldsymbol{x}_0 \in U_0$ となることから

$$f(\boldsymbol{x}) = \phi(f_0(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{x}_0)$$

となる．また，f は C^r 級であることから U で十分に 1 次近似が可能であり，$\boldsymbol{x}_0$ を中心とした 1 次近似は ϕ は ε を用いることで

$$\phi(\boldsymbol{x}) = f(\boldsymbol{x}_0) + J_f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + \varepsilon$$

とあらわされ，f_0 も同様にして C^r 級であることから U_0 で十分に 1 次近似が可能であり，$\boldsymbol{0}$ を中心とした 1 次近似は ε' を用いることで

$$f_0(\boldsymbol{x} - \boldsymbol{x}_0) = f_0(\boldsymbol{0}) + J_{f_0}(\boldsymbol{0})(\boldsymbol{x} - \boldsymbol{x}_0 - \boldsymbol{0}) + \varepsilon' = \boldsymbol{x} - \boldsymbol{x}_0 + \varepsilon'$$

となる．これより，f を ϕ と f_0 の定義に基づいて計算すると

$$\begin{aligned} f(\boldsymbol{x}) =& \phi(\boldsymbol{x} - \boldsymbol{x}_0 + \varepsilon' + \boldsymbol{x}_0) \\ =& \phi(\boldsymbol{x} + \varepsilon') \\ =& f(\boldsymbol{x}_0) + J_f(\boldsymbol{x}_0)(\boldsymbol{x} + \varepsilon' - \boldsymbol{x}_0) + \varepsilon \\ =& f(\boldsymbol{x}_0) + J_f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + \varepsilon + J_f(\boldsymbol{x}_0)\varepsilon' \end{aligned}$$

となり，f の 1 次近似に一致することがわかる．そのため，以後 ε および ε_0 を無視するとする．

また，ϕ は線型かつ $J_f(\boldsymbol{x}_0)$ の正則性より逆写像 ϕ^{-1} が存在し，$\boldsymbol{y} = \phi(\boldsymbol{x})$ で

$$\begin{aligned} \boldsymbol{y} =& f(\boldsymbol{x}_0) + J_f(\boldsymbol{x}_0)(\phi^{-1}(\boldsymbol{y}) - \boldsymbol{x}_0) \\ \phi^{-1}(\boldsymbol{y}) - \boldsymbol{x}_0 =& J_f^{-1}(\boldsymbol{x}_0)(\boldsymbol{y} - f(\boldsymbol{x}_0)) \\ \phi^{-1}(\boldsymbol{y}) =& \boldsymbol{x}_0 + J_f^{-1}(\boldsymbol{x}_0)(\boldsymbol{y} - f(\boldsymbol{x}_0)) \end{aligned}$$

となり，$J_f^{-1}(\boldsymbol{x}_0)$ は ϕ^{-1} の $\boldsymbol{x}_0$ におけるヤコビ行列といえる．

これより，ϕ^{-1} が存在することで f_0 は

$$f_0(\boldsymbol{x} - \boldsymbol{x}_0) = (\phi^{-1} \circ f)(\boldsymbol{x}) - \boldsymbol{x}_0$$

とあらわされ，$J_{f_0}(\boldsymbol{0})$ については合成写像のヤコビ行列の関係より単位行列 I を用いて

$$J_{f_0}(\boldsymbol{0}) = J_{\phi^{-1} \circ f}(\boldsymbol{x}_0) = J_{\phi^{-1}}(f(\boldsymbol{x}_0))J_f(\boldsymbol{x}_0) = J_f^{-1}(\boldsymbol{x}_0)J_f(\boldsymbol{x}_0) = I$$

となる．

ここで，f_0 が U_0 で逆写像 $g_0 : V_0 \to U_0$ が存在すると仮定する．このとき，$\boldsymbol{y} = f(\boldsymbol{x})$ とすることで

$$\begin{aligned}\boldsymbol{y} =& \phi(f_0(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{x}_0)\\ \phi^{-1}(\boldsymbol{y}) =& f_0(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{x}_0\\ \boldsymbol{x} - \boldsymbol{x}_0 =& g_0(\phi^{-1}(\boldsymbol{y}) - \boldsymbol{x}_0)\\ \boldsymbol{x} =& g_0(\phi^{-1}(\boldsymbol{y}) - \boldsymbol{x}_0) + \boldsymbol{x}_0\end{aligned}$$

というように $\boldsymbol{x}$ について解くことができ，これが f の逆写像 g となる．つまり，f_0 が U_0 で逆写像が存在すれば f は U で逆写像 $g : V \to U$ をもつということである．

さらに，ϕ と ϕ^{-1} は明らかに C^r 級であるため，g_0 が C^r 級となるならば g は ϕ^{-1} と g_0 の合成によってあらわされるため C^r 級となる．以上より，f_0 が C^r 級で $f_0(\boldsymbol{0}) = \boldsymbol{0}$ かつ $J_{f_0}(\boldsymbol{0}) = I$ を満たすとき，U_0 で C^r 級逆写像 g_0 をもつならば，f は U で逆写像をもつということがわかる．これより，f_0 の C^r 級逆写像 g_0 が存在することを示す．

$\boldsymbol{x} \in U_0$ で $\boldsymbol{y} = f_0(\boldsymbol{x})$ とあらわすとしたとき，ある $\boldsymbol{0}$ に近い $\boldsymbol{y}$ に対して，f_0 で $\boldsymbol{x}$ が定まるということは

$$f_0(\boldsymbol{x}) = \boldsymbol{y}$$

が一意に成り立つことであるが，そのために写像

$$h(\boldsymbol{x}) = (h_n(\boldsymbol{x}), h_n(\boldsymbol{x}), \cdots, h_n(\boldsymbol{x})) = \boldsymbol{y} - f_0(\boldsymbol{x})$$

を定義する．このとき，$J_h(\boldsymbol{0})$ は零行列であり，f_0 の $\boldsymbol{0}$ を中心とした 1 次近似により恒等写像に近い写像であることがわかる．そのため，h の偏導関数は十分に $\boldsymbol{0}$ に近く，

$$\forall\varepsilon_0 > 0, \exists\delta_0 > 0 \ s.t.\ \varepsilon_0 \leq \frac{1}{2n}, |\boldsymbol{x}| < \delta_0, \left|\frac{\partial h_i}{\partial x_j}\right| < \varepsilon_0$$

を満たすように δ_0 を選択すれば ε_0 をとることができる．$\boldsymbol{x} + \boldsymbol{x}' \in U_0$ となるような任意の $\boldsymbol{x}' \in U_0$ を与えれば，h のシプリッツ定数について，補題より

$$|h(\boldsymbol{x} + \boldsymbol{x}') - h(\boldsymbol{x})| < \varepsilon n|\boldsymbol{x}'|$$

が成り立つ．

ここで，$\boldsymbol{x}_1 \in U_0$ による $h(\boldsymbol{x}_1)$ を考えれば，これは十分に $\boldsymbol{0}$ に近いが，$h(\boldsymbol{x}_1) = \boldsymbol{0}$ となるとは限らない．そこで，これにより生じた誤差を補正するように

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 - h(\boldsymbol{x}_1)$$

を与え，これを繰り返すことにより

$$\begin{cases}\boldsymbol{x}_0 = \boldsymbol{0}\\ \boldsymbol{x}_1 = \boldsymbol{y}\\ \boldsymbol{x}_{k+1} = \boldsymbol{x}_k - (f_0(\boldsymbol{x}_k) - \boldsymbol{y}) \quad (k \geq 2)\end{cases}$$

といった漸化式による数列 $(\boldsymbol{x}_n)$ を与えれば，これは収束すると考えることができる．

このとき，$|\boldsymbol{y}| < \frac{\delta_0}{2}$ ならば全ての $\boldsymbol{x_k}$ で $|\boldsymbol{x}| < \delta_0$ となることを数学的帰納法により確かめる．$k = 0$ および $k = 1$ のときは仮定より自明であるため，$k - 1$ および k で成り立つことを仮定して $k + 1$ で成り立つことを示す．$\boldsymbol{x}_{k+1}$ と $\boldsymbol{x}_k$ の差について

$$\begin{aligned}|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k| =& |(\boldsymbol{x}_k - (f(\boldsymbol{x}_k) - \boldsymbol{y})) - (\boldsymbol{x}_{k-1} - (f(\boldsymbol{x}_{k-1}) - \boldsymbol{y}))|\\ =& |(\boldsymbol{x}_k - f(\boldsymbol{x}_k)) - (\boldsymbol{x}_{k-1} - f(\boldsymbol{x}_{k-1}))|\\ =& |h(\boldsymbol{x}_k) - h(\boldsymbol{x}_{k-1})|\\ <& \varepsilon n|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}|\\ <& \frac{1}{2}|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}|\end{aligned}$$

といった関係式が得られる．よって，これを繰り返すことにより

$$|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k| < \frac{1}{2^k}|\boldsymbol{x}_1 - \boldsymbol{x}_0| = \frac{1}{2^k}|\boldsymbol{y}|$$

となる．また，$\boldsymbol{x}_{k+1}$ について漸化式を展開すると

$$\boldsymbol{x}_{k+1} = -\sum_{i=0}^{k}(f_0(\boldsymbol{x}_i) - \boldsymbol{y})$$

となり，$f_0(\boldsymbol{x}_i) - \boldsymbol{y} = -(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)$ となることから

$$\boldsymbol{x}_{k+1} = \sum_{i=0}^{k}(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)$$

が得られる．

よって，三角不等式により

$$|\boldsymbol{x}_{k+1}| \leq \sum_{i=0}^{k}|\boldsymbol{x}_{i+1} - \boldsymbol{x}_i| < \sum_{i=0}^{k}\frac{1}{2^i}|\boldsymbol{y}|$$

となるが，無限等比級数の関係で

$$\sum_{i=0}^{\infty}\frac{1}{2^i} = \frac{1}{1-\frac{1}{2}} = 2$$

となることを用いれば

$$|\boldsymbol{x}_{k+1}| < 2|\boldsymbol{y}| < \delta_0$$

となる．よって，任意の $\boldsymbol{x}_k$ は $|\boldsymbol{x}_k| < \delta_0$ を満たし，任意の $\boldsymbol{x}_k$ はリプシッツ定数についての関係を満たすことがわかる．つまり，$(\boldsymbol{x}_n)$ はコーシー列であり，このとき $(\boldsymbol{x}_n)$ はある $\boldsymbol{x}_\infty$ に収束し，元の漸化式に対して

$$\boldsymbol{x}_\infty = \boldsymbol{x}_\infty - (f_0(\boldsymbol{x}_\infty) - \boldsymbol{y}) \ \Leftrightarrow\ \boldsymbol{y} = f_0(\boldsymbol{x}_\infty)$$

となり，$|\boldsymbol{x}| < \delta_0$ を満たす $\boldsymbol{x}$ は $|\boldsymbol{y}| < \frac{\delta_0}{2}$ を満たす $\boldsymbol{y}$ に対して一意に定まることがわかる．

ここで，$|\boldsymbol{y}| < \frac{\delta_0}{2}$ を示す領域 V_0 で定義される写像 $g_0'(\boldsymbol{y}) = \boldsymbol{x}$ を与えれば，g_0' はリプシッツ連続である．このとき，任意の $\boldsymbol{x}_1, \boldsymbol{x}_2 \in U_0$ で $\boldsymbol{y}_1 = f_0(\boldsymbol{x}_1)$ および $\boldsymbol{y}_2 = f_0(\boldsymbol{x}_2)$ が $|\boldsymbol{y}_1| < \frac{\delta_0}{2}$ および $|\boldsymbol{y}_2| < \frac{\delta_0}{2}$ を満たすとして，全微分の導出に則れば，ε_0' を用いて

$$\begin{aligned} f_0(\boldsymbol{x}_2) - f(\boldsymbol{x}_1) =& J_{f_0}(\boldsymbol{x}_1)(\boldsymbol{x}_2 - \boldsymbol{x}_1) + \varepsilon_0'|\boldsymbol{x}_2 - \boldsymbol{x}_1| \\ \boldsymbol{y}_2 - \boldsymbol{y}_1 =& J_{f_0}(g_0'(\boldsymbol{y}_1))(g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)) + \varepsilon_0'|g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)| \end{aligned}$$

とあらわすことができる．

また，$h(\boldsymbol{x})$ において $\boldsymbol{y} = \boldsymbol{x}$ とすれば $J_h = I - J_{f_0}$ であり，J_h の各成分の絶対値は ε_0 未満であるため，J_h^k の各成分の絶対値は $n^{k-1}\varepsilon_0^k$ 未満となる．ここで，$n^{k-1}\varepsilon_0^k$ とは

$$n^{k-1}\varepsilon_0^k \leq \frac{n^{k-1}}{(2n)^k} = \frac{1}{2^k n}$$

となるため J_h^k の各成分は 1 未満となることがわかる．つまり，

$$\sum_{k=0}^{\infty} J_h^k$$

は各成分において絶対収束する．このとき，$(I - J_h)^{-1}$ に収束するため

$$\begin{aligned} \sum_{k=0}^{\infty} J_h^k =& \sum_{k=0}^{\infty}(I - J_{f_0})^k \\ (I - J_h)^{-1} =& (I - (I - J_{f_0}))^{-1} \\ J_{f_0}^{-1} =& (I - J_h)^{-1} \end{aligned}$$

となり，J_h の各成分の絶対値が ε_0 未満，すなわち $|\boldsymbol{x}| < \delta_0$ であるならば J_{f_0} の逆行列が存在することが示される．

よって，$g_0'(V_0)$ において J_{f_0} は正則であり

$$g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1) = J_{f_0}^{-1}(g_0'(\boldsymbol{y}_1))(\boldsymbol{y}_2 - \boldsymbol{y}_1) - J_{f_0}^{-1}(g_0'(\boldsymbol{y}_1))\varepsilon_0'|g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)|$$

というように $g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)$ について変形できる．ここで，ε_0' を含む項について考えれば

$$\frac{|g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)|}{|\boldsymbol{y}_2 - \boldsymbol{y}_1|} < \delta_0 \cdot \frac{2}{\delta_0} = 2$$

となることから

$$\begin{aligned}\lim_{\boldsymbol{y}_2 \to \boldsymbol{y}_1} \varepsilon_0' |g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)| &= \lim_{\boldsymbol{y}_2 \to \boldsymbol{y}_1} \varepsilon_0' \frac{|g_0'(\boldsymbol{y}_2) - g_0'(\boldsymbol{y}_1)|}{|\boldsymbol{y}_2 - \boldsymbol{y}_1|} |\boldsymbol{y}_2 - \boldsymbol{y}_1| \\ &< \lim_{\boldsymbol{y}_2 \to \boldsymbol{y}_1} 2\varepsilon_0' |\boldsymbol{y}_2 - \boldsymbol{y}_1| \\ &= \boldsymbol{0}\end{aligned}$$

となることから $\boldsymbol{y}_2 \to \boldsymbol{y}_1$ で ε_0' を含む項は $\boldsymbol{0}$ に収束し，$g_0'(\boldsymbol{y})$ は全微分可能であり，g_0' は連続，つまりは C^0 級であることがわかる．

$GL_n(\mathbb{R})$ の任意の元の乗法逆元を対応させる写像 $l : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ を与えたとき，写像の対応関係を示すと

$$\begin{array}{ccccccc} V_0 & \xrightarrow{g_0'} & U_0 & \xrightarrow{J_{f_0}} & GL_n(\mathbb{R}) & \xrightarrow{l} & GL_n(\mathbb{R}) \\ \Cup & & \Cup & & \Cup & & \Cup \\ \boldsymbol{y} & \longmapsto & g_0'(\boldsymbol{y}) & \longmapsto & J_{f_0}(g_0'(\boldsymbol{y})) & \longmapsto & J_{f_0}^{-1}(g_0'(\boldsymbol{y})) \end{array}$$

となり，このような対応によって $J_{f_0}^{-1}(g_0'(\boldsymbol{y}))$ が与えられ，このとき $J_{g_0'} = J_{f_0}^{-1}$ である．補題より l は C^∞ 級であり，J_{f_0} は C^{r-1} 級であることから，この合成は C^r 級であり，g_0' が C^0 級であることから $J_{f_0}^{-1} = J_{g_0'}$ も C^0 級である．つまり，g_0' は C^1 級であるといえる．同様にして，$n < r$ で g_0' が C^n 級であるとすれば $J_{f_0}^{-1} = J_{g_0'}$ も C^n 級であり，g_0' が C^{n+1} 級であるといえる．以上より，数学的帰納法から g_0' は C^r 級である．

また，g_0' の定義より $g_0' \circ f_0 = \mathrm{id}_{U_0}$ かつ $f_0 \circ g_0' = \mathrm{id}_{V_0}$ が成り立ち，g_0' は f_0 の逆写像である．つまり，$g_0 = g_0'$ ということである．

よって，命題は証明された．

□

これより，以下の系が与えられる．

系 1.　領域 $D \subset \mathbb{R}^n$ で定義される C^1 級写像 $f : D \to \mathbb{R}^n$ において，$\boldsymbol{x}_0 \in D$ でヤコビ行列 $J_f(\boldsymbol{x}_0)$ が正則であるとき，f の逆写像 f^{-1} により

$$J_{f^{-1}}(f(\boldsymbol{x}_0)) = (J_f(\boldsymbol{x}_0))^{-1}$$

となる．すなわち，逆写像のヤコビ行列は元の写像のヤコビ行列の逆行列となる．

Proof.

逆写像定理より $\boldsymbol{x}_0$ 近傍で f の逆写像 f^{-1} が存在する．$f^{-1} \circ f = \mathrm{id}_D$ より，合成写像 $f^{-1} \circ f$ のヤコビ行列は単位行列 I を用いることにより

$$\begin{aligned} J_{f^{-1} \circ f}(\boldsymbol{x}_0) &= J_{f^{-1}}(f(\boldsymbol{x}_0)) J_f(\boldsymbol{x}_0) \\ I &= J_{f^{-1}}(f(\boldsymbol{x}_0)) J_f(\boldsymbol{x}_0) \\ J_{f^{-1}}(f(\boldsymbol{x}_0)) &= (J_f(\boldsymbol{x}_0))^{-1} \end{aligned}$$

となり，等式が得られる．

よって，命題は証明された．

□

陰関数定理

独立変数 $x_1, x_2, \cdots, x_n, y \in \mathbb{R}$ による領域 D で定義される $r \geq 1$ による C^r 級関数 $z = \varphi(x_1, x_2, \cdots, x_n, y) \in \mathbb{R}$ が $\boldsymbol{p} \in D$ で

$$z(\boldsymbol{p}) = 0, \quad \frac{\partial z}{\partial y}(\boldsymbol{p}) \neq 0$$

が成り立つとき，$\boldsymbol{p}$ の近傍で $z = 0$ による陰関数 $y = f(x_1, x_2, \cdots, x_n)$ が一意に存在し，これは C^r 級関数である．

これを，**陰関数定理**という．

Proof.

$D \subset \mathbb{R}^n \times \mathbb{R}$ とすることで $(\boldsymbol{x}, y) \in D$ のようにあらわしたとき，$g : \mathbb{R}^{n+1} \supset D \to \mathbb{R}^{n+1}$ を

$$g(\boldsymbol{x}, y) = \begin{pmatrix} \boldsymbol{x} \\ \varphi(\boldsymbol{x}, y) \end{pmatrix}$$

とすれば g は C^r 級であり，$\boldsymbol{p} = (\boldsymbol{p}_x, p_y) \in \mathbb{R}^n \times \mathbb{R}$ とあらわすことにより

$$g(\boldsymbol{p}_x, p_y) = \begin{pmatrix} \boldsymbol{p}_x \\ \varphi(\boldsymbol{p}_x, p_y) \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_x \\ 0 \end{pmatrix}$$

が成り立つ．

また，このヤコビ行列をとれば

$$J_g(\boldsymbol{p}_x, \varphi(\boldsymbol{p}_x, p_y)) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \varphi_{x_1}(\boldsymbol{p}_x, p_y) & \varphi_{x_2}(\boldsymbol{p}_x, p_y) & \cdots & \varphi_{x_n}(\boldsymbol{p}_x, p_y) & \varphi_y(\boldsymbol{p}_x, p_y) \end{pmatrix}$$

となり，単位行列 I と零ベクトル $\mathbf{0}$ および $\boldsymbol{x}$ に対するナブラ $\nabla_{\boldsymbol{x}}$ を用いることにより

$$J_g(\boldsymbol{p}_x, \varphi(\boldsymbol{p}_x, p_y)) = \begin{pmatrix} I & \mathbf{0} \\ \nabla_{\boldsymbol{x}}\varphi(\boldsymbol{p}_x, p_y) & \varphi_y(\boldsymbol{p}_x, p_y) \end{pmatrix}$$

とあらわすことができる．この行列式をとれば

$$|J_g(\boldsymbol{p}_x, \varphi(\boldsymbol{p}_x, p_y))| = I\varphi_y(\boldsymbol{p}_x, p_y) - \nabla_{\boldsymbol{x}}\varphi(\boldsymbol{p}_x, p_y) \cdot \mathbf{0} = \varphi_y(\boldsymbol{p}_x, p_y) = \frac{\partial z}{\partial y}(\boldsymbol{p}) \neq 0$$

となり，逆写像定理が適用可能である．

このとき，$\boldsymbol{p}$ 近傍の領域 $U \subset D$ と $g(\boldsymbol{p}_x, p_y) = (\boldsymbol{p}_x, 0)$ 近傍の領域 $V = g(U)$ が存在して

$$g|_U : U \to V$$

は C^r 級逆写像 $h : V \to U$ をもつ．また，$h(\boldsymbol{x}, y) = (\eta(\boldsymbol{x}, y), \psi(\boldsymbol{x}, y))$ とおけば任意の $(\boldsymbol{x}, y) \in V$ で

$$(\boldsymbol{x}, y) = (g \circ h)(\boldsymbol{x}, y) = g(\eta(\boldsymbol{x}, y), \psi(\boldsymbol{x}, y)) = (\eta(\boldsymbol{x}, y), \varphi(\eta(\boldsymbol{x}, y), \psi(\boldsymbol{x}, y)))$$

となり，$\boldsymbol{x} = \eta(\boldsymbol{x}, y)$ である．つまり，$h(\boldsymbol{x}, y) = (\boldsymbol{x}, \psi(\boldsymbol{x}, y))$ であり，$(\boldsymbol{x}, y) \in V$ ならば $(\boldsymbol{x}, \psi(\boldsymbol{x}, y)) \in U$ であることがわかる．

射影作用素 $\pi_y : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ を $\pi_y(\boldsymbol{x}, y) = y$ を与えれば，$\psi = \pi_y \circ h$ とあらわされるため，ψ は C^r 級である．また，$\varphi = \pi_y \circ g$ となるため，任意の $(\boldsymbol{x}, y) \in V$ で

$$\varphi(\boldsymbol{x}, \psi(\boldsymbol{x}, y)) = \varphi(h(\boldsymbol{x}, y)) = (\pi_y \circ g \circ h)(\boldsymbol{x}, y) = \pi_y(\boldsymbol{x}, y) = y$$

が成り立つ．$\boldsymbol{p}_x$ 近傍の領域 $U_x \subset \mathbb{R}^n$ および p_y 近傍の領域 $U_y \subset \mathbb{R}$ を

$$U_x \times U_y \subset U, \quad U_x \times \{0\} \subset V$$

を満たすようにとり，写像 $\phi : U_x \to \mathbb{R}$ を $\phi(\boldsymbol{x}) = \psi(\boldsymbol{x}, 0) \in \psi(U_x \times \{0\})$ と与えれば，ψ は C^1 級であるため ϕ も C^r 級であり

$$\phi(\boldsymbol{p}_x) = \psi(\boldsymbol{p}_x, 0) = (\pi_y \circ h)(\boldsymbol{p}_x, 0) = (\pi_y \circ h \circ g)(\boldsymbol{p}_x, p_y) = \pi_y(\boldsymbol{p}_x, p_y) = p_y$$

が成り立つ．また，ψ の定義より ϕ は C^r 級逆写像 ϕ^{-1} をもち，$W = U_x \cap \phi^{-1}(U_y)$ を与えれば $\phi(W) \subset U_y$ であるため，$\boldsymbol{x} \in W$ とすれば $\phi(\boldsymbol{x}) \in \phi(W)$ であり

$$(\boldsymbol{x}, \phi(\boldsymbol{x})) \in W \times U_y \subset U_x \times U_y \subset U$$

となるため，$\varphi(\boldsymbol{x}, \phi(\boldsymbol{x})) = \varphi(\boldsymbol{x}, \psi(\boldsymbol{x}, 0)) = 0$ となる．逆に，$(\boldsymbol{x}, y) \in W \times U_y$ で $\varphi(\boldsymbol{x}, y) = 0$ とすれば

$$g(\boldsymbol{x}, y) = (\boldsymbol{x}, \varphi(\boldsymbol{x}, y)) = (\boldsymbol{x}, 0) = (\boldsymbol{x}, \varphi(\boldsymbol{x}, \phi(\boldsymbol{x}))) = g(\boldsymbol{x}, \phi(\boldsymbol{x}))$$

となり，$g|_U$ は全単射であるため $y = \phi(\boldsymbol{x})$ である．また，ψ が g により一意に定まるため，$\phi(\boldsymbol{x}) = \psi(\boldsymbol{x}, 0)$ より ϕ も一意に定まる．

よって，命題は証明された．

□

また，陰関数の偏導関数について以下の補題が与えられる．

補題 4.　独立変数 $x_1, x_2, \cdots, x_n, y \in \mathbb{R}$ による C^1 級の方程式 $\varphi(x_1, x_2, \cdots, x_n, y) = 0$ を与えたとする．このとき，$\varphi(x_1, x_2, \cdots, x_n, y) = 0$ による陰関数 $y = f(x_1, x_2, \cdots, x_n)$ の偏導関数は

$$\nabla y = -\frac{1}{\varphi_y}(\varphi_{x_1}, \varphi_{x_2}, \cdots, \varphi_{x_n})$$

と与えられる．

Proof.

陰関数定理より $f(x_1, x_2, \cdots, x_n)$ は C^1 級であるため，$\varphi(x_1, x_2, \cdots, x_n, y) = 0$ を x_k で偏微分すると連鎖律より

$$\begin{aligned}\varphi_{x_k} =& \sum_{l=1}^{n} \varphi_{x_l} \frac{\partial x_l}{\partial x_k} + \varphi_y \frac{\partial y}{\partial x} \\ =& \varphi_{x_k} + \varphi_y \frac{\partial y}{\partial x_k} \\ =& 0\end{aligned}$$

となり，

$$\frac{\partial y}{\partial x_k} = -\frac{\varphi_{x_k}}{\varphi_y}$$

が得られる．これを全ての $x_1, x_2, \cdots, x_n$ に対する偏導関数をまとめることで

$$\nabla y = -\frac{1}{\varphi_y}(\varphi_{x_1}, \varphi_{x_2}, \cdots, \varphi_{x_n})$$

となる．

よって，命題は証明された．

□

これより，グラフに接する超平面についての以下の定理を与えることができる．

グラフの接超平面

独立変数 $x_1, x_2, \cdots, x_n, y \in \mathbb{R}$ による領域 D で定義される方程式 $z = \varphi(x_1, x_2, \cdots, x_n, y) = 0$ において，$\boldsymbol{p} \in D$ で接超平面が存在する十分条件は，z が C^1 級であることである．

このとき，超接平面の方程式は $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ および $\boldsymbol{p} = (\boldsymbol{p}_x, p_y) \in \mathbb{R}^n \times \mathbb{R}$ とあらわすとすれば

$$\nabla z \cdot (\boldsymbol{x} - \boldsymbol{p}_x) = 0$$

である．

Proof.

陰関数定理より z が C^1 級であるならば z による C^1 級の陰関数が存在し，その陰関数の法線ベクトルは ∇z で与えられる．つまり，超接平面の方程式は

$$\nabla z \cdot (\boldsymbol{x} - \boldsymbol{p}_x) = 0$$

で与えられる．

よって，命題は証明された．

□

1.5　高次偏導関数

独立変数 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ による領域 D で定義される関数 $y = f(\boldsymbol{x}) \in \mathbb{R}$ を複数回偏微分することについて考える．このとき，関数に対する変数は複数あるため偏微分をする順序をも考えなければならない．

ここで，m 回だけ任意の変数に対して偏微分をすることを考えたとき，独立変数に対する添字を示す $\mathbb{N}$ による添字集合 $I = \{i_1, i_2, \cdots, i_m\}$ を与えられたとき，添字付けられた変数 $x_1, x_2, \cdots, x_n$ を添字集合 I により f を偏微分するということは

$$\left(\frac{\partial}{\partial x_{i_m}} \cdots \left(\frac{\partial}{\partial x_{i_2}}\left(\frac{\partial f}{\partial x_{i_1}}\right)\right)\right) = \frac{\partial^m f}{\partial x_{i_m} \cdots \partial x_{i_2} \partial x_{i_1}}$$

であると定義する．また，微分演算子を省略することで

$$\frac{\partial^n f}{\partial x_{i_m} \cdots \partial x_{i_2} \partial x_{i_1}} = f_{x_{i_1} x_{i_2} \cdots x_{i_m}}$$

とあらわされるとする．

このとき，n 階偏導関数の組み合わせは n^m 通り存在するが，仮定として f は D で n 回偏微分可能である必要がある．ここで，偏微分の順序に対して偏導関数が一致する，すなわち偏微分の順序が交換可能である条件について考える．

まず，$f^0 = f$ とおき，整列された添字集合 $N = \{1, 2, \cdots, m\}$ を与えたとする．$i, j \in N$ で $i \neq j$ を満たすとき $f^0_{x_{i_i} x_{i_j}} = f^0_{x_{i_j} x_{i_i}}$ とすれば，これを f^1 とおくとする．そして，$k, l \in N \setminus \{i, j\}$ で $k \neq l$ を満たすとき $f^1_{x_{i_k} x_{i_l}} = f^1_{x_{i_l} x_{i_k}}$ とすれば f^2 とおくといった操作を繰り返すことにより，偏微分の順序が交換可能かを考えることができる．

つまり，偏微分が交換可能かを考えるには 2 階導関数が一致するかを調べるだけで十分である．

そこで，以下の定理を与える．

偏微分の順序交換の十分条件

> 独立変数 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ による領域 D で定義される関数 $y = f(\boldsymbol{x}) \in \mathbb{R}$ が $i \neq j$ で $f_{x_i x_j} = f_{x_j x_i}$ となる十分条件は $f_{x_i x_j}$ および $f_{x_j x_i}$ が存在して，ともに連続であることである．

Proof.

x_i と x_j 以外の独立変数は考察対象でないため，$\boldsymbol{x} \in D$ を $(x_i, x_j) \in D$ というようにあらわされるとする．

$(x_i + h, x_j + k) \in D$ となるように任意の $h, k \in \mathbb{R}$ をとり，これによる関数 $F(h, k)$ を

$$F(h, k) = f(x_i + h, x_j + k) - f(x_i + h, x_j) - f(x_i, x_j + k) + f(x_i, x_j)$$

と定義する．ここで，平均値の定理より $0 < \theta_{i_1}, \theta_{j_1} < 1$ を用いて

$$\begin{aligned} F(h, k) =& (f(x_i + h, x_j + k) - f(x_i + h, x_j)) - (f(x_i, x_j + k) - f(x_i, x_j)) \\ =& (f_{x_i}(x_i + \theta_{i_1} h, x_j + k) - f_{x_i}(x_i + \theta_{i_1} h, x_j))h \\ =& f_{x_i x_j}(x_i + \theta_{i_1} h, x_j + \theta_{j_1} k)hk \end{aligned}$$

となり，同様にして $0 < \theta_{i_2}, \theta_{j_2} < 1$ を用いて

$$\begin{aligned} F(h, k) =& (f(x_i + h, x_j + k) - f(x_i, x_j + k)) - (f(x_i + h, x_j) - f(x_i, x_j)) \\ =& (f_{x_j}(x_i + h, x_j + \theta_{i_2} k) - f_{x_i}(x_i, x_j + \theta_{i_2} k))k \\ =& f_{x_j x_i}(x_i + \theta_{i_2} h, x_j + \theta_{j_2} k)hk \end{aligned}$$

が得られる．$f_{x_i x_j}$ と $f_{x_j x_i}$ は連続であるため $h, k \to 0$ とすることで

$$\begin{aligned}\lim_{h,k\to 0} f_{x_i x_j}(x_i + \theta_{i_1} h, x_j + \theta_{j_1} k)hk &= \lim_{h,k\to 0} f_{x_j x_i}(x_i + \theta_{i_2} h, x_j + \theta_{j_2} k)hk \\ f_{x_i x_j} &= f_{x_j x_i}\end{aligned}$$

が成り立つ．

　よって，命題は証明された．

□