付録２

Lesson 1 微分

微分可能

(i) $\mathbb{R}\to\mathbb{R}$の関数$f(x)$が$x=a$で微分可能であるとは，極限値

$$A=\lim_{x\to a}\frac{f(x)-f(a)}{x-a},$$

が存在することである．Aを$f(x)$の$x=a$における微係数という．

注意：$x\to a$はただxが単調にaに近づくことを意味しない．例えばxがaにその周りを振動しながら近づいてもよい．従って$y=f(x)$が折れ線で$x=a$がその節点である場合には，$f(x)$が$x=a$で微分可能ではない．

この定義は次のように言い換えることができる(現代風の定義).

(ii) $\mathbb{R}\to\mathbb{R}$の関数$f(x)$が$x=a$で微分可能であるとは，xの一次関数$P(x;a)=A(x-a)+f(a)$で

$$\lim_{x\to a}\frac{f(x)-P(x;a)}{x-a}=0,$$

を満たすものが存在することである．この時のAを$f(x)$の$x=a$における微係数という．

注意：上の条件(ii)を満たすものは一つしかないことは次のように示される．条件(ii)を満たすもうひとつの$P'(x;a)=B(x-a)+f(a)$が存在するとしても

$$0\le|A-B|=\left|\frac{f(x)-P(x;a)}{x-a}-\frac{f(x)-P'(x;a)}{x-a}\right|\le\left|\frac{f(x)-P(x;a)}{x-a}\right|+\left|\frac{f(x)-P'(x;a)}{x-a}\right|,$$

だから，$x\to a$の極限をとれば$A=B$となることから分かる．受験数学のはさみうちだ.

二つの定義(i)および(ii)は同値である．すなわち，一方の定義からもう一方の定義が従う．これは

$$\frac{f(x)-P(x;a)}{x-a}=\frac{f(x)-f(a)}{x-a}-A,$$

より明らかであろう．

<u>最良一次近似，接線</u>

$P(x;a)$を$x=a$における$f(x)$の一次近似という．定義(ii)の条件を満たす$P(x;a)$，すなわちAが$x=a$における$f(x)$の微係数である一次近似は，一次近似の中で最良の近似を$x=a$の近傍で与える．実際，この条件を満たす$P(x;a)$の誤差$e(x;a)=f(x)-P(x;a)$は

$$\lim_{x\to a}\frac{e(x;a)}{x-a}=0,$$

を満たす．これは誤差$e(x;a)$が$x=a$の近傍で$x-a$の高次の量になっていることを意味する．一般の一次近似の場合，誤差は$x=a$の近傍で$x-a$の高次の量にならず，上の極限値はゼロ以外の値になる．定義(ii)の条件を満たす$P(x;a)$に対する直線$y=P(x;a)$を曲線$y=f(x)$の$x=a$における接線という．

Lesson 2 多変数関数の微分

Lesson1 では$\mathbb{R}\to\mathbb{R}$の関数の微分を定義したが，ここでは$\mathbb{R}^N\to\mathbb{R}$の関数の微分を定義しよう．

<u>微分可能</u>[1]

$\mathbb{R}^N\to\mathbb{R}$の関数$h(\boldsymbol{x})\ (=h(x_1,\cdots,x_N))$が$\boldsymbol{x}=\boldsymbol{a}$において微分可能とは

$$Q(\boldsymbol{x};\boldsymbol{a})=\sum_{i=1}^{N}A_i(x_i-a_i)+h(\boldsymbol{a}),$$

とすると，$\boldsymbol{x}\to\boldsymbol{a}$のどんな近づけ方に対しても

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}}\frac{h(\boldsymbol{x})-Q(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x}-\boldsymbol{a}|}=0,$$

を満たすものが存在することである．

---

1 教科書等では全微分可能と呼ばれている．

$\mathbb{R}^N$のベクトルA_iが一意に定まることを Lesson 1 の議論と同様に示してみよう．もうひとつの$Q'(\boldsymbol{x};\boldsymbol{a}) = B_i(x_i - a_i) + h(\boldsymbol{a})$が存在すれば

$$0 \leq \frac{|\sum_{i=1}^N (A_i - B_i)(x_i - a_i)|}{|\boldsymbol{x} - \boldsymbol{a}|} = \left| \frac{h(\boldsymbol{x}) - Q(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x} - \boldsymbol{a}|} - \frac{h(\boldsymbol{x}) - Q'(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x} - \boldsymbol{a}|} \right|$$
$$\leq \left| \frac{h(\boldsymbol{x}) - Q(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x} - \boldsymbol{a}|} \right| + \left| \frac{h(\boldsymbol{x}) - Q'(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x} - \boldsymbol{a}|} \right|,$$

だから

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} \frac{|\sum_{i=1}^N (A_i - B_i)(x_i - a_i)|}{|\boldsymbol{x} - \boldsymbol{a}|} = 0.$$

これより$A_i = B_i\ (i = 1, \cdots, N)$が従う．実際，$\boldsymbol{n}$を$\mathbb{R}^N$における固定された単位ベクトルとし，$\boldsymbol{x} = \boldsymbol{a} + t\boldsymbol{n}\ (t > 0)$として$t \to 0$とすると，$\sum_{i=1}^N (A_i - B_i) n_i = 0$が上の式より従い，$\boldsymbol{n} = (1,0,\cdots,0)$，$\boldsymbol{n} = (0,1,0,\cdots,0)$，$\cdots$，$\boldsymbol{n} = (0,\cdots,1)$の場合を考えればよい．

注意：$\boldsymbol{x} \to \boldsymbol{a}$のどんな近づけ方に対しても... というのは厳密でないぞと筆者を叱る読者もいよう．そんな厳密な読者の為にこれを Weierstraß に従って言い直しておく．

任意の$\varepsilon > 0$に対しある$\delta > 0$が存在し，$0 < |\boldsymbol{x} - \boldsymbol{a}| < \delta$を満たす全てのベクトル$\boldsymbol{x}$に対して

$$\frac{|h(\boldsymbol{x}) - Q(\boldsymbol{x};\boldsymbol{a})|}{|\boldsymbol{x} - \boldsymbol{a}|} < \varepsilon,$$

を満たすものが存在する．つまりどんな小さな$\varepsilon > 0$に対しても，$\boldsymbol{x} = \boldsymbol{a}$の近傍では上の不等式が成り立つということ．

<u>偏微分可能</u>

$h(\boldsymbol{x})$が$\boldsymbol{x} = \boldsymbol{a}$で微分可能なら，$h(\boldsymbol{x})$は$\boldsymbol{x} = \boldsymbol{a}$で偏微分可能で，$Q(\boldsymbol{x};\boldsymbol{a})$の$x_i - a_i$の係数$A_i$は$\boldsymbol{x} = \boldsymbol{a}$における$h(\boldsymbol{x})$の$x_i$に関する偏微分係数に等しい．すなわち，

$$A_i = \frac{\partial h}{\partial x_i}(\boldsymbol{a}) \quad (i = 1, \cdots, N).$$

微分可能の定義では$\boldsymbol{x} \to \boldsymbol{a}$の近づけ方は任意なので，先ほどと同様に，$\boldsymbol{n}$を$\mathbb{R}^N$における固定された単位ベクトルとし$\boldsymbol{x} = \boldsymbol{a} + t\boldsymbol{n}$として$t \to 0$を考えると，

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}}\frac{h(\boldsymbol{x})-Q(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x}-\boldsymbol{a}|}=\lim_{t\to 0}\frac{h(\boldsymbol{a}+t\boldsymbol{n})-Q(\boldsymbol{x};\boldsymbol{a})}{|t|}=0.$$

すなわち

$$\lim_{t\to 0}\frac{h(\boldsymbol{a}+t\boldsymbol{n})-h(\boldsymbol{a})-t\sum_{i=1}^{N}A_i n_i}{|t|}=\lim_{t\to 0}\frac{h(\boldsymbol{a}+t\boldsymbol{n})-h(\boldsymbol{a})-t\sum_{i=1}^{N}A_i n_i}{\pm t}=0.$$

結局，次式を得る．

$$\lim_{t\to 0}\frac{h(\boldsymbol{a}+t\boldsymbol{n})-h(\boldsymbol{a})}{t}=\sum_{i=1}^{N}A_i n_i.$$

後は$\boldsymbol{n}=(1,0,\cdots,0)$，$\boldsymbol{n}=(0,1,0,\cdots,0)$，$\cdots$，$\boldsymbol{n}=(0,\cdots,1)$の場合を考えればよい．左辺は偏微分の定義そのものになる．

上の結果より直ちに方向微分可能性が言える．

<u>方向微分</u>

$h(\boldsymbol{x})$が$\boldsymbol{x}=\boldsymbol{a}$で微分可能なら，$h(\boldsymbol{x})$は$\boldsymbol{x}=\boldsymbol{a}$で方向微分可能．すなわち，

$$\lim_{t\to 0}\frac{h(\boldsymbol{a}+t\boldsymbol{n})-h(\boldsymbol{a})}{t}=\sum_{i=1}^{N}n_i\frac{\partial h}{\partial x_i}(\boldsymbol{a}).$$

<u>最良一次近似</u>

$Q(\boldsymbol{x};\boldsymbol{a})$を$\boldsymbol{x}=\boldsymbol{a}$における$h(\boldsymbol{x})$の一次近似という．微分可能の条件を満たす$Q(\boldsymbol{x};\boldsymbol{a})$は$\boldsymbol{x}=\boldsymbol{a}$の近傍における$h(\boldsymbol{x})$の一次近似のうち最良の近似を$\boldsymbol{x}=\boldsymbol{a}$の近傍で与える．実際，この条件を満たす$Q(\boldsymbol{x};\boldsymbol{a})$の誤差$e(\boldsymbol{x};\boldsymbol{a})=h(\boldsymbol{x})-Q(\boldsymbol{x};\boldsymbol{a})$は

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}}\frac{e(\boldsymbol{x};\boldsymbol{a})}{|\boldsymbol{x}-\boldsymbol{a}|}=0,$$

を満たし，これは誤差$e(\boldsymbol{x};\boldsymbol{a})$が$\boldsymbol{x}=\boldsymbol{a}$の近傍で$|\boldsymbol{x}-\boldsymbol{a}|$の高次の量であることを意味する．微分可能の条件を満たさない一般の一次近似の場合，誤差は$\boldsymbol{x}=\boldsymbol{a}$の近傍で$|\boldsymbol{x}-\boldsymbol{a}|$の高次の量にならない．

<u>微分可能な条件</u>

一変数関数の場合と同様に，多変数関数の場合でも微分可能なのは最良一次近似が存在することであったが，どんな場合にこれが存在するのか？つまり微分可能な条件は何かという疑問が自然にわいてくるだろう．微分可能ならばすべてのx_iに関して(一階)偏微分可能なことは上で見た．しかしこの逆は成り立たない．すなわちすべてのx_iに関して(一階)偏微分可能でも微分可能であるとは言えない．例えばもしその一階の偏導関数がすべて連続ならば，微分可能である[2]．また微分可能でも偏導関数が連続とは限らない．以下では関数は十分な滑らかさを持ち，(高階の)偏導関数も全て連続として議論を進める．

<u>全微分</u>

関数$h(\boldsymbol{x})$が微分可能なときに，$h(\boldsymbol{x})$の「全微分」と称する

$$dh = \frac{\partial h}{\partial x_1}dx_1 + \cdots + \frac{\partial h}{\partial x_N}dx_N,$$

によって定義される量をよく目にする．物理や工学の教科書だけではなく，厳密さを重んずるはずの数学の教科書でもだ．だから関数$h(\boldsymbol{x})$の$\boldsymbol{x}=\boldsymbol{a}$における全微分$dh$は$h(\boldsymbol{a}+d\boldsymbol{x})-h(\boldsymbol{a})$の主要部と説明されても，その主要部になにか厳密なものが潜んでいるように感じるかもしれない．しかしdhはそんな厳めしいものではない．誤差$e(\boldsymbol{x};\boldsymbol{a})$の定義式を

$$h(\boldsymbol{x}) - h(\boldsymbol{a}) = \sum_{i=1}^{N}(x_i - a_i)\frac{\partial h}{\partial x_i}(\boldsymbol{a}) + e(\boldsymbol{x};\boldsymbol{a}),$$

のように書きなおし，さらに$|\boldsymbol{x}-\boldsymbol{a}|$の近傍で考えていることを意識して
$\boldsymbol{x}=\boldsymbol{a}+d\boldsymbol{x}$と書きなおし，

$$h(\boldsymbol{a}+d\boldsymbol{x}) - h(\boldsymbol{a}) = \sum_{i=1}^{N}\frac{\partial h}{\partial x_i}(\boldsymbol{a})dx_i + e(\boldsymbol{a}+d\boldsymbol{x};\boldsymbol{a}), \qquad (1)$$

この式の右辺第 1 項を$h(\boldsymbol{a}+d\boldsymbol{x})-h(\boldsymbol{a})$の主要部，あるいは関数$h(\boldsymbol{x})$の$\boldsymbol{x}=\boldsymbol{a}$における全微分$dh$と呼ぶだけのことである．全微分$dh$とは，$h(\boldsymbol{a}+d\boldsymbol{x})-h(\boldsymbol{a})$という微小変化量の$|d\boldsymbol{x}|$の高次の量の誤差を伴う近似以外のなにものでもない．つまり$h(\boldsymbol{x})$を$\boldsymbol{x}=\boldsymbol{a}$における最良一次近似したときの変動量である．

<u>接平面</u>

$x_{N+1} = h(\boldsymbol{x})$は$\mathbb{R}^{N+1}$における(超)曲面を表す．$Q(\boldsymbol{x};\boldsymbol{a})$を$h(\boldsymbol{x})$の$\boldsymbol{x}=\boldsymbol{a}$における微分可能な条件を満たす一次式とする．

$$x_{N+1} = Q(\boldsymbol{x};\boldsymbol{a}) = \frac{\partial h}{\partial x_i}(\boldsymbol{a})(x_i - a_i) + h(\boldsymbol{a}),$$

---

[2] すべてというのは過大であり，詳細は解析概論等の教科書を参照されたい．

を$x_{N+1}=h(\boldsymbol{x})$の$\boldsymbol{x}=\boldsymbol{a}$における接平面という．

<u>勾配</u>

上の接平面は$\mathbb{R}^{N+1}$で考えたが，ここでは再び$\mathbb{R}^N$に戻る．式(1)で見たように，微小変位$d\boldsymbol{x}$に対する変化量$h(\boldsymbol{a}+d\boldsymbol{x})-h(\boldsymbol{a})$は，$|d\boldsymbol{x}|$の高次の量を無視すれば，関数$h$の偏微分係数からなるベクトル$\nabla h\equiv\left(\frac{\partial h}{\partial x_1},\cdots,\frac{\partial f}{\partial x_N}\right)$と微小変位ベクトル$d\boldsymbol{x}\equiv(dx_1,dx_2,\cdots,dx_N)$との内積で表される． $d\boldsymbol{x}$，∇hをそれぞれ$|d\boldsymbol{x}|\boldsymbol{n}$，$|\nabla h|\boldsymbol{g}$と単位ベクトル$\boldsymbol{n}$および$\boldsymbol{g}$を用いて表すと，Schwarz の不等式[3]より

$$|dh|=|\nabla h||d\boldsymbol{x}||\,\boldsymbol{n}\cdot\boldsymbol{g}|\le|\nabla h||d\boldsymbol{x}|,$$

を得る．等号が成り立つ時，すなわち$|dh|$が最大となるのは，$\boldsymbol{n}\parallel\boldsymbol{g}$の場合である．つまり$\nabla h$の方向に変位すれば$h$が最も急に変化する．そして$|\nabla h|$はその方向における$h$の変化の割合になっている．これらのことから，$\nabla h$は$h$の勾配と呼ばれている．

Lesson 3 陰関数定理 1

$h(x_1,x_2)$を$\mathbb{R}^2\to\mathbb{R}$の十分滑らかな関数とする．$h(x_1,x_2)=0$を満たす$\mathbb{R}^2$の点集合（等位集合と呼ばれる）がどのようなものかを考えよう．$\mathbb{R}^2$の問題なので，頭の中で図を描く思考実験で何とかなるだろう．

$\mathbb{R}\to\mathbb{R}$の関数のグラフは$\mathbb{R}^2$上で描かれるのと同様に，$\mathbb{R}^2\to\mathbb{R}$の関数は$\mathbb{R}^3$におけるグラフとして描けるだろう．$h(x_1,x_2)=0$は$x_3=h(x_1,x_2)$と$x_3=0$の共通集合だから$\mathbb{R}^3$における曲面と平面の交わりで，$x_1x_2$平面($\mathbb{R}^2$)上の曲線になるはずで，

---

[3] $\boldsymbol{x}\neq\boldsymbol{0}$，$\boldsymbol{y}\neq\boldsymbol{0}$に対して$(\boldsymbol{x}\cdot\boldsymbol{x})(\boldsymbol{y}\cdot\boldsymbol{y})\ge(\boldsymbol{x}\cdot\boldsymbol{y})^2$が成り立つ．等号は$\boldsymbol{x}\parallel\boldsymbol{y}$の場合．
$\because$ $s\in\mathbb{R}$として二次関数$q(s)$を$q(s)=(s\boldsymbol{x}+\boldsymbol{y})\cdot(s\boldsymbol{x}+\boldsymbol{y})$と定義する．内積の性質(i)より$\forall s\in\mathbb{R}$に対して$q(s)\ge 0$．一方，性質(ii)および(iii)より$q(s)=(\boldsymbol{x}\cdot\boldsymbol{x})s^2+2(\boldsymbol{x}\cdot\boldsymbol{y})s+\boldsymbol{y}\cdot\boldsymbol{y}$,だから二次方程式の判別式の条件から不等式を得る．判別式がゼロ（不等式で等号が成り立つ場合）ならば，$\exists s^*\in\mathbb{R}$に対して$q(s^*)=0$．性質(i)より$s^*\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{0}$．これより$\boldsymbol{x}\parallel\boldsymbol{y}$．逆に$\boldsymbol{x}\parallel\boldsymbol{y}$ならば，$\exists s^*\in\mathbb{R}$に対して$q(s^*)=0$が成り立ち，$q(s)\ge 0$だから判別式がゼロで不等式$(10-1)$の等号が従う．■

$x_1 = \varphi_1(x_2)$あるいは$x_2 = \varphi_2(x_1)$という形で表されるはずだ． $h(\varphi_1(x_2), x_2) = 0$あるいは$h(x_1, \varphi_2(x_1)) = 0$という関係を微分してやれば，合成関数の微分の公式が使えて$h_{x_1}\varphi_1{}' + h_{x_2} = 0$と$h_{x_1} + h_{x_2}\varphi_2{}' = 0$と計算できるから, 結局$\varphi_1{}' = -\frac{h_{x_2}}{h_{x_1}}$, $\varphi_2{}' = -\frac{h_{x_1}}{h_{x_2}}$という公式も得られる. もちろん前者では$h_{x_1} \neq 0$, 後者では$h_{x_2} \neq 0$でなければばならない.

上のように$h(x_1, x_2) = 0$が曲線であることは直感的に理解できが，これを厳密に保証するのが微分積分学の鬼門の 1 つである陰関数定理である．その証明には様々な微分積分学のテクニックを駆使したものや，高級な現代数学の証明まである．じっくり勉強することを強く勧めて，ここではその定理を簡単に紹介しよう．

<u>陰関数定理１</u>

$h(\boldsymbol{x})$を$\mathbb{R}^N \to \mathbb{R}$の滑らかな関数とする．$\mathbb{R}^N$内の一点$\boldsymbol{x}^0$において$h(\boldsymbol{x}^0) = 0$およびある独立変数$x_p$ $(p \in \{1, \cdots, N\})$に対して

$$\frac{\partial h}{\partial x_p}(\boldsymbol{x}^0) \neq 0, \tag{2}$$

ならば，$\boldsymbol{x}^0$の近傍で滑らかな関数$\varphi(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N)$が一意に存在し，その等位集合は曲線

$$x_p = \varphi\left(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N\right), \tag{3}$$

で表される．すなわち

$$x_p^0 = \varphi\left(x_1^0, \cdots, x_{p-1}^0, x_{p+1}^0, \cdots, x_N^0\right), \tag{4}$$

$$h\left(x_1, \cdots, x_{p-1}, \varphi\left(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N\right), x_{p+1}, \cdots, x_N\right) = 0. \tag{5}$$

式(2)をx_m $(m \neq p)$で微分すればチェーンルールによって次式を得る．

$$\frac{\partial \varphi}{\partial x_m}\left(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N\right)$$

$$= -\frac{\frac{\partial h}{\partial x_m}\left(x_1, \cdots, x_{p-1}, \varphi\left(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N\right), x_{p+1}, \cdots, x_N\right)}{\frac{\partial h}{\partial x_p}\left(x_1, \cdots, x_{p-1}, \varphi\left(x_1, \cdots, x_{p-1}, x_{p+1}, \cdots, x_N\right), x_{p+1}, \cdots, x_N\right)}$$

$$(m \neq p). \qquad (6)$$

先に直感的に導出した結果は上の定理の二変数の場合に一致する．

上の定理を関数の引数の入れ替えを行って見やすい形に書き直しておく．x_pをz，それ以外の引数を$N-1$個の成分を持つベクトル$\boldsymbol{y}$の成分$y_1, \cdots, y_{N-1}$として表し，関数Hを

$$H(\boldsymbol{y}, z) = h\left(x_1, \cdots, x_{p-1}, x_p, x_{p+1}, \cdots, x_N\right),$$

によって定義すると，定理１は次のように言い換えられる．

<u>定理１’</u>

$H(\boldsymbol{y}, z)$を$\mathbb{R}^{N-1} \times \mathbb{R} \to \mathbb{R}$の滑らかな関数とする．$\mathbb{R}^{N-1} \times \mathbb{R}$内の一点$(\boldsymbol{y}^0, z^0)$において$H(\boldsymbol{y}^0, z^0) = 0$および

$$\frac{\partial H}{\partial y}(\boldsymbol{y}^0, z^0) \neq 0,$$

ならば，$\boldsymbol{y}^0$の近傍において滑らかな$\mathbb{R}^{N-1} \to \mathbb{R}$の関数$\varphi(\boldsymbol{y})$が一意に存在し，

$$z^0 = \varphi(\boldsymbol{y}^0),$$
$$H\big(\boldsymbol{y}, \varphi(\boldsymbol{y})\big) = 0.$$

さらに

$$\frac{\partial \varphi}{\partial y_i}(\boldsymbol{y}) = -\frac{\dfrac{\partial H}{\partial y_i}\big(\boldsymbol{y}, \varphi(\boldsymbol{y})\big)}{\dfrac{\partial H}{\partial z}\big(\boldsymbol{y}, \varphi(\boldsymbol{y})\big)} \quad (i = 1, \cdots, N-1). \qquad (7)$$

<u>超曲面の接平面，法線ベクトル</u>

式(3)は$\mathbb{R}^N$における(超)曲面を表す．点$\left(x_1^0, \cdots, x_p^0, \cdots, x_N^0\right)$におけるその接平面は

$$x_p - x_p^0 = \sum_{i \neq p} \left(\frac{\partial \varphi}{\partial x_i}\left(x_1^0, \cdots, x_{p-1}^0, x_{p+1}^0, \cdots, x_N^0\right) \right) (x_i - x_i^0),$$

と表される．これを式(6)を用いて書きう直すと

$$\sum_{i=1}^{N} \left(\frac{\partial h}{\partial x_i}(\boldsymbol{x}^0) \right) (x_i - x_i^0) = 0. \qquad (8)$$

式(8)はこの接平面上のベクトル(接ベクトル)$\boldsymbol{x} - \boldsymbol{x}^0$が勾配$\nabla h$と直交していることを意味する．すなわち$\nabla h$は接平面の法線方向にある．

Lesson 4 陰関数定 2

陰関数定理は多変数のベクトル値関数の場合に拡張できる．拡張された定理はラグランジュの未定乗数法の理論的基盤を与えるものである．

定理２

$\boldsymbol{y} \in \mathbb{R}^n$，$\boldsymbol{z} \in \mathbb{R}^m$とし$\boldsymbol{H}(\boldsymbol{y}, \boldsymbol{z})$を$\mathbb{R}^n \times \mathbb{R}^m \Rightarrow \mathbb{R}^m$の滑らかなベクトル値関数とする．

$$\boldsymbol{H}(\boldsymbol{y}^0, \boldsymbol{z}^0) = 0,$$

かつ行列

$$D_z\boldsymbol{H}(\boldsymbol{y}, \boldsymbol{z}) = \begin{pmatrix} \frac{\partial H_1}{\partial z_1}(\boldsymbol{y}, \boldsymbol{z}) & \cdots & \frac{\partial H_1}{\partial z_m}(\boldsymbol{y}, \boldsymbol{z}) \\ \vdots & \ddots & \vdots \\ \frac{\partial H_m}{\partial z_1}(\boldsymbol{y}, \boldsymbol{z}) & \cdots & \frac{\partial H_m}{\partial z_m}(\boldsymbol{y}, \boldsymbol{z}) \end{pmatrix},$$

が$(\boldsymbol{y}^0, \boldsymbol{z}^0)$で正則ならば, $\boldsymbol{y}^0$の近傍に滑らかな$\mathbb{R}^n \longrightarrow \mathbb{R}^m$のベクトル値関数$\boldsymbol{\varphi}(\boldsymbol{y})$が一意に存在し

$$\boldsymbol{z}^0 = \boldsymbol{\varphi}(\boldsymbol{y}^0),$$
$$\boldsymbol{H}\big(\boldsymbol{y}, \boldsymbol{\varphi}(\boldsymbol{y})\big) = \boldsymbol{0}.$$

さらに

$$D\boldsymbol{\varphi}(\boldsymbol{y}) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}) & \cdots & \frac{\partial \varphi_1}{\partial y_n}(\boldsymbol{y}) \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_m}{\partial y_1}(\boldsymbol{y}) & \cdots & \frac{\partial \varphi_m}{\partial y_n}(\boldsymbol{y}) \end{pmatrix},$$

$$D_y\boldsymbol{H}(\boldsymbol{y}, \boldsymbol{z}) = \begin{pmatrix} \frac{\partial H_1}{\partial y_1}(\boldsymbol{y}, \boldsymbol{z}) & \cdots & \frac{\partial H_1}{\partial y_n}(\boldsymbol{y}, \boldsymbol{z}) \\ \vdots & \ddots & \vdots \\ \frac{\partial H_m}{\partial y_1}(\boldsymbol{y}, \boldsymbol{z}) & \cdots & \frac{\partial H_m}{\partial y_n}(\boldsymbol{y}, \boldsymbol{z}) \end{pmatrix},$$

とすると

$$D\boldsymbol{\varphi}(\boldsymbol{y}) = -\Big(D_z\boldsymbol{H}\big(\boldsymbol{y}, \boldsymbol{\varphi}(\boldsymbol{y})\big)\Big)^{-1} D_y\boldsymbol{H}\big(\boldsymbol{y}, \boldsymbol{\varphi}(\boldsymbol{y})\big).$$

これはチェーンルールによる合成関数の微分に他ならない．