最適化理論-変分法と$Fréchet$微分その３-

師玉康成

目 次

第1章　$Fréchet$微分

1.1　$Fréchet$微分と$Gâteaux$微分

大分道草を食べましたが, 再び最初に扱った変分問題に戻りましょう。それは汎関数

$$f \in C^1[0,1] \mapsto \ J(f) = \int_0^1 \{f'(x)\}^2 dx$$

を条件

$$f(0) = 0,\ f(1) = 1$$

で最小化するものでした。

この問題について，$J(f)$ を関数の集合

$$E = \{f \in C^1[0,1],\ f(0) = 0,\ f(1) = 1\}$$

から $\boldsymbol{R}$ への写像

$$J \ :\ E \to \boldsymbol{R}$$

について何らかの意味の微分の概念を導入し実数値関数の極小条件

$$\left.\frac{dJ(f)}{df}\right|_{f=\bar{f}} = 0$$

のような扱いはできないだろうかという動機で，前章までに，まず定義域である関数の集合についての数学的構造を調べました。
E をその部分集合として含む $C^1[0,1]$ などは関数空間と呼ばれ，特に Banach 空間というものになっていました。
これには，要素 $f, g \in C^1[0,1]$ の加算 $f+g$ と，係数体の元 $\alpha \in \mathbf{R}$ によるスカラー倍 αf という代数的な構造（線形空間の構造）が定義され，
要素 $f \in C^1[0,1]$ にはノルム $|f|$ も定義され，さらに実数の集合のような収束の扱いも可能でした。

ここでは本題の，これら関数空間上での微分の概念を扱うことにします。

そのためには$Fréchet$微分という概念が必要になってきます。また，集合 E だけではなくそれに付随した

$$\vec{E} = \{f \in C^1[0,1],\ f(0) = f(1) = 0\}$$

も用いることになります。

1.1.1 微分の再定義

よく知られるように実数関数

$$\varphi\ :\ \boldsymbol{R} \to \boldsymbol{R}$$

の $x = \bar{x}$ での $\varphi'(\bar{x})$ は

$$\varphi'(\bar{x}) = \lim_{h \to 0} \frac{\varphi'(\bar{x}+h) - \varphi(\bar{x})}{h} \tag{1.1}$$

で定義されています。
しかし，この定義を E のような関数集合に適用することはできません。分母の

$$\varphi'(\bar{x}+h) - \varphi(\bar{x})$$

は定義できますが，これを E の元 h で割る演算できないからです。

割り算を使わない工夫が必要になります。そこで上記の微分の定義式 (**??**) を次のように書き換えてみます。

$$\begin{aligned}\varphi(\bar{x}+h) - \varphi(\bar{x}) &= \varphi'(\bar{x})h + O(h) \\ &= \varphi'(\bar{x})h + O_1(h)|h| \qquad (1.2)\end{aligned}$$

ただし，

$$O(h) = \varphi(\bar{x}+h) - \varphi(\bar{x}) - \varphi'(\bar{x})h$$

$$O_1(h) = \frac{O(h)}{|h|},$$

とします。
ここで，φ が微分可能であれば，上の微分の定義式 (**??**) の右辺の極限が左辺と一致するのですから，

$$O_1(h) \to 0\ (|h| \to 0)$$

が成立っている。

そこで，$L_{\bar{x}}$ が任意の h について

$$\varphi(\bar{x}+h)-\varphi(\bar{x})=L_{\bar{x}}\cdot +O_1(h)|h| \tag{1.3}$$
$$O_1(h)\to 0\ (|h|\to 0) \tag{1.4}$$

を充たすとき，この $L_{\bar{x}}$ を φ の $x=\bar{x}$ での微分と定義することにします。

$J(f)$ についても，見かけの形式は

$$J(\bar{f}+\Delta f)-J(\bar{f})=L_{\bar{f}}\cdot\Delta f+O_1(\Delta f)\cdot|\Delta f| \tag{1.5}$$

とするべきでしょう。このとき，和 $\bar{f}+\Delta f$, 積 $L_{\bar{f}}\cdot\Delta f$, ノルム $|\Delta f|$ はどのような意味を持つか考察しましょう。

1.1.2 アフィン空間 $E,\vec{E}$ の構造

ここで

$$\vec{E}=\{f\in C^1[0,1];f(0)=f(1)=0\}$$

という空間を考えます。

$$\varphi_1,\varphi_2\in\vec{E}$$

に対して，

$$\varphi_1+\varphi_2\in C^1[0,1]$$
$$(\varphi_1+\varphi_2)(0)=\varphi_1(0)+\varphi_2(0)=0$$
$$(\varphi_1+\varphi_2)(1)=\varphi_1(1)+\varphi_2(1)=0$$

よって

$$\varphi_1+\varphi_2\in\vec{E}$$

また

$$\alpha\in\boldsymbol{R},\ \alpha\varphi_1\in\vec{E}$$

さらに，$\varphi\in\vec{E}$ に対して

$$\|\varphi\|_{C^1}\triangleq\|\varphi\|_C+\|\varphi'\|_C$$

とすると $\vec{E}$ は前章で扱ったノルム空間です。

E に付随した $\vec{E}$ を導入した理由は，次によります。

$$E=\{f\in C^1[0,1];f(0)=0,f(1)=1\}$$

はバナハ空間 $C^1[0,1]$ の部分集合ではありますが，これの任意の元 $f, g \in E$ については

$$\begin{aligned} &f+g \in C^1[0,1] \\ &(f+g)(0) = f(0)+g(0) = 0 \\ &(f+g)(1) = f(1)+g(1) = 2 \end{aligned}$$

となって $f+g \notin E$. また

$$\alpha \in \boldsymbol{R},\ \alpha f(1) \neq 1,\ (\alpha \neq 1 \text{ のとき})$$

であるので $\alpha f \notin E$. です。つまり E は線形空間の構造は持っていません。

　ただし

$$\begin{aligned} &f-g \in C^1[0,1] \\ &(f-g)(0) = f(0)-g(0) = 0 \\ &(f-g)(1) = f(1)-g(1) = 0 \end{aligned}$$

よって $f-g \in \vec{E}$ です。
また,

$$\bar{f} \in E,\ \varphi \in \vec{E}$$

とすると

$$\bar{f}+\varphi \in \{\bar{f}+\varphi; \bar{f} \in E, \varphi \in \vec{E}\} = E$$

です。

　ここで $g \in E$ とすると

$$g = \bar{f} + g - \bar{f},\ g - \bar{f} \in \vec{E}$$

より

$$f, g \in E,\ f = g + f - g,\ f - g \in \vec{E}$$

　「線形空間もどき」の性質があるわけです。
この空間 E を (線形空間 $\vec{E}$ をともなう) アフィン空間とよびます。
E 自身は線形空間ではないが，E の元 $f \in E$ を一つ固定して原点のように扱えば,

$$\vec{E} = \{g - f | g \in E\}$$

は線形空間になっています。また f に変動として $\bar{g} \in \vec{E}$ を加えた $f + \bar{g}$ も E の元になっています。

以上によって変分法の第1章の議論を再び持ち出すことが可能になりました。すなわち，まず

$$f \in E, \qquad \Delta f \in \vec{E}$$
$$f + \Delta f \in E$$

とおきます。E 上の写像

$$J \ : \ E \to \boldsymbol{R}$$

について，

$$J(\bar{f} + \Delta f) - J(\bar{f}) = L_{\bar{f}} \cdot \Delta f + O(\Delta f) \cdot \|\Delta f\|_C$$

であり，第1章の計算を繰り返して

$$\begin{aligned} J(\bar{f} + h\Delta f) - J(\bar{f}) &= \int_0^1 \{(\bar{f} + h\Delta f)'\}^2 dx - \int_0^1 (\bar{f}')^2 dx \\ &= \int_0^1 \{(\bar{f}')^2 + 2\bar{f}'\Delta h + h^2(\Delta \bar{f}')^2\} dx + \int_0^1 (\bar{f}')^2 dx \\ &= 2h \int_0^1 \bar{f}'(x)\Delta f'(x) dx + h \int_0^1 h(\Delta f'(x))^2 dx \end{aligned}$$

を得ます。

さてここで，$h = 1$ とすると

$$J(\bar{f} + \Delta f) - J(\bar{f}) = 2\int_0^1 \bar{f}'(x)\Delta f'(x)dx + \int_0^1 (\Delta f'(x))^2 dx$$

このとき (1.5) 式の $L_{\bar{f}} \cdot \Delta f$ は

$$L_{\bar{f}} \cdot \Delta f = 2\int_0^1 \bar{f}'(x)\Delta f'(x)dx$$

で定義するべきでしょう。
$\Delta f \in \vec{E}$ の任意性を考えれば，上の式は $\vec{E}$ から $\boldsymbol{R}$ への写像

$$L_{\bar{f}} \ : \ \varphi \in \vec{E} \mapsto 2\int_0^1 \bar{f}'(x)\varphi'(x)dx \in \boldsymbol{R}$$

を定義しているものと考えらます。

1.1.3 $L_{\bar f}$ の線形性と連続性

$$L_{\bar f}\ :\ \varphi \in \vec{E} \mapsto 2\int_0^1 \bar{f}'(x)\varphi'(x)dx \in \boldsymbol{R}$$

は以下に観るように $\varphi \in \vec{E}$ についての線形性と連続性をもっています。
$\varphi_1, \varphi_2 \in \vec{E}$ とすると

$$\begin{aligned}
L_{\bar f}(\varphi_1+\varphi_2) &= 2\int_0^1 \bar{f}'(x)(\varphi_1+\varphi_2)'(x)dx \\
&= 2\int_0^1 \bar{f}'(x)(\varphi_1'(x)+\varphi_2'(x))dx \\
&= 2\int_0^1 \bar{f}'(x)\varphi_1'(x)dx + 2\int_0^1 \bar{f}'(x)\varphi_2'(x)dx \\
&= L_{\bar f}(\varphi_1) + L_{\bar f}(\varphi_2) \\
L_{\bar f}(\alpha\varphi_1) &= 2\int_0^1 \bar{f}'(x)(\alpha\varphi_1)'(x)dx \\
&= 2\int_0^1 \bar{f}'(x)\alpha\cdot\varphi_1'(x)dx \\
&= \alpha L\bar{f}(\varphi_1)
\end{aligned}$$

よって $L_{\bar f}$ は 線形です。次に

$$\begin{aligned}
\frac{1}{2}|L_{\bar f}(\varphi_1) - L_{\bar f}(\varphi_2)| &= \left|\int_0^1 \bar{f}'(x)\varphi_1'(x)dx - \int_0^1 \bar{f}'(x)\varphi_2'(x)dx\right| \\
&= \left|\int_0^1 \bar{f}'(x)(\varphi_1'(x)-\varphi_2'(x))dx\right| \\
&\le \int_0^1 |\bar{f}'(x)|\cdot|\varphi_1'(x)-\varphi_2'(x)|dx \\
&\le \int_0^1 |\bar{f}'(x)|dx\cdot\|\varphi_1(x)-\varphi_2(x)\|_{C^1} \\
&\le K_{\bar f}\|\varphi_2-\varphi_2\|_{C^1}
\end{aligned}$$

であるから任意の $\varepsilon > 0$ について

$$\delta(\varepsilon) = \frac{\varepsilon}{2K_{\bar f}}$$

と選べば,

$$\|\varphi_1-\varphi_2\|_{C^1} < \delta(\varepsilon)$$

なる全ての $\varphi_1, \varphi_2 \in C^1[0,1]$ について

$$|L_{\bar f}(\varphi_1) - L_{\bar f}(\varphi_2)| < \varepsilon$$

よって，$L_{\bar{f}}$ は連続です。最後に残余項の条件，すなわち，定義式 (1.4) の条件

$$O_1(h) \to 0 \ (|h| \to 0)$$

について調べてみます。
まず

$$\int_0^1 (\varphi'(x))^2 dx = \int_0^1 \frac{(\varphi'(x))^2}{\|\varphi\|_{C^1}} dx \cdot \|\varphi\|_{C^1}$$

です。ここで

$$\begin{aligned}\left|\int_0^1 \frac{(\varphi'(x))^2}{\|\varphi\|_{C^1}} dx\right| &\leq \frac{1}{\|\varphi\|_{C^1}} \cdot \int_0^1 |\varphi'(x)|^2 dx \\ &\leq \frac{1}{\|\varphi\|_{C^1}} \cdot \int_0^1 \|\varphi'(x)\|_{C^1}^2 dx \\ &= \frac{1}{\|\varphi\|_{C^1}} \cdot \|\varphi(x)\|_{C^1}^2 \\ &= \|\varphi\|_{C^1} \ \to \ 0 \ (\|\varphi\|_{C^1} \to 0)\end{aligned}$$

よって

$$O_1(\varphi) = \int_0^1 \frac{(\varphi'(x))^2}{\|\varphi\|_{C^1}} dx$$

は定義式 (1.1) の条件を充たしています。
以上をまとめれば，任意の $\varphi \in \vec{E}$ に対して

$$J(\bar{f} + \varphi) - J(\bar{f}) = L_{\bar{f}} \cdot \varphi + O_1(\varphi) \|\varphi\|_{C^1}$$

が成立っています。
$L_{\bar{f}}$ は 線形かつ連続であり，$O_1(\varphi) \to 0 \ (\|\varphi\|_{C^1} \to 0)$

1.2 $Fréchet$微分

1.2.1 $Fréchet$微分の定義

E, F を $\vec{E}, \vec{F}$ を伴ったノルムアフィン空間とします。$\vec{E}$ はノルム線形空間で
写像

$$J \ : \ E \to F$$

定義 1.2.1 J が $f \in E$ で $Fréchet$微分可能であるとは次が成立つことをいいます。

f に依存する連続線形写像 (作用素)

$$L_f \ : \ \vec{E} \to \vec{F}$$

が存在して，任意の $\varphi \in \vec{E}$ に対して

$$J(f+\varphi) - J(f) = L_f(\varphi) + O(\varphi)\|\varphi\|_{\vec{F}}$$

であり，

$$\|O(\varphi)\|_{\vec{F}} \to 0, \qquad (\|\varphi\|_{\vec{E}} \to 0)$$

この L_f は一意に存在します。
実際,

$$J(f+\varphi) - J(f) = L_f(\varphi) + O(\varphi)\|\varphi\|_{\vec{E}}$$

$L_{f_1}(\varphi), L_{f_2}(\varphi)$ が同じ条件を満たすとすると：$L_{f_1}(\varphi) = L_{f_2}(\varphi)$ です。これを観るには,

$$\begin{array}{lrcl} \|\varphi\|_{\vec{E}} \to 0 \quad \text{のとき} & \|O_1(\varphi)\|_{\vec{F}} & \to & 0 \\ & \|O_2(\varphi)\|_{\vec{F}} & \to & 0 \end{array}$$

$$\begin{aligned} J(f+\varphi) - J(f) &= L_{f_1}(\varphi) + O_1(\varphi)\|\varphi\|_{\vec{E}} \\ &= L_{f_2}(\varphi) + O_2(\varphi)\|\varphi\|_{\vec{E}} \end{aligned}$$

$$L_{f_1}(\varphi) - L_{f_2}(\varphi) = O_1(\varphi)\|\varphi\|_{\vec{E}} - O_2(\varphi)\|\varphi\|_{\vec{E}}$$

$\|L_{f_1}(\varphi) - L_{f_2}(\varphi)\|_{\vec{F}} \to 0, \qquad \|\varphi\|_{\vec{E}} \to 0$

$$\lim_{h\to 0} \frac{J(f+h\varphi) - J(f)}{h} = \int_0^1 \frac{f'(x) - g'(x)}{\sqrt{1+(f'(x))^2}} dx \tag{1.6}$$

$$\begin{aligned} J(f+h\varphi) - J(f) &= L_{f_1}(h\varphi) + O_1(h\varphi)\|h\varphi\|_{\vec{E}} \\ &= hL_{f_1}(\varphi) + |h|O_1(h\varphi)\|\varphi\|_{\vec{E}} \end{aligned}$$

$$\begin{aligned} (1.6)\ \text{の左辺} &= \lim_{h\to 0} \frac{hL_{f_1}(\varphi) + |h|O_1(h\varphi)\|\varphi\|_{\vec{E}}}{h} \\ &= L_{f_1}(\varphi) + \lim_{h\to 0} \frac{|h|O_1(h\varphi)\|\varphi\|_{\vec{E}}}{h} \\ &= L_{f_1}(\varphi) = L_{f_2}(\varphi) \end{aligned}$$

したがって、一意性が示されました。

1.2.2 $G\hat{a}teaux$微分

ここで，$Fr\acute{e}chet$微分と第一章で導出した

$$\frac{dJ(f+h\varphi)}{dh}\Big|_{h=0}$$

との関係について述べておきます。この微分は Δf に沿った微分または $G\hat{a}teaux$微分と呼ばれます。L_f を J の f での$Fr\acute{e}chet$微分とすると，

$$\begin{aligned} J(f+h\varphi)-J(f) &= L_f(h\varphi)+o_1(h\varphi) \\ &= hL_f(\varphi)+o_1(h\varphi) \end{aligned}$$

であり，両辺を h で割って，

$$fracJ(f+h\varphi)-J(f)h = L_f(\varphi)+\frac{o_1(h\varphi)}{h}$$

を得る。
さらに，

$$\frac{J(f+h\varphi)-J(f)}{h} \to \frac{dJ(f+h\varphi)}{dh}\Big|_{h=0} \quad (h\to 0)$$

$$\frac{o_1(h\varphi)}{h} \to 0 \quad (h\to 0)$$

ゆえ，

$$\frac{dJ(f+h\varphi)}{dh}\Big|_{h=0} = L_f(\varphi)$$

を得る。これは$Fr\acute{e}chet$微分可能ならば$G\hat{a}teaux$微分可能であり，以下の関係式が成立っていることを示しています。

$$L_f(\varphi) = \frac{dJ(\bar{f}+h\varphi)}{dh} \quad \Big|_{h=0}$$

問題 2 の$Fr\acute{e}chet$微分

汎関数

$$J(f) = \int_0^1 \sqrt{1+(f'(x))^2}dx$$

$$\begin{aligned} f\in E &= \{f; f\in C^1[0,1], f(0)=0, f(1)=1\} \\ \vec{E} &= \{\varphi; \varphi\in C^1[0,1], \varphi(0)=\varphi(1)=0\} \\ &= \{f-g; f,g\in E\} \end{aligned}$$

の$Fréchet$微分を求めてみましょう。

[導出]

$$\begin{aligned} J(f+\varphi)-J(f) &= \int_0^1 \sqrt{1+(f'+\varphi')^2}dx - \int_0^1 \sqrt{1+(f')^2}dx \\ &= \int_0^1 \{\sqrt{1+(f'+\varphi')^2} - \sqrt{1+(f')^2}\}dx \qquad (1.7) \end{aligned}$$

ここで φ' を固定して S を次のようにします。

$$\varphi' \overset{S}{\longmapsto} S(\varphi') = \sqrt{1+(f'+\varphi')^2}$$

そして $\varphi' = 0$ のまわりでテイラー展開を行ないます。

$$\begin{aligned} S(\varphi') &= S(0) + S'(0)\varphi' + R_2 \\ S'(\varphi') &= \frac{f'+\varphi'}{\sqrt{1+(f'+\varphi')^2}} \\ S''(\varphi') &= \frac{1}{\{1+(f'+\varphi')^2\}^{\frac{2}{3}}} \\ R_2 &= \frac{S''(\theta_{\varphi'} \cdot \varphi') \cdot (\varphi')^2}{2} \qquad (0 < \theta_{\varphi'} < 1) \end{aligned}$$

以上より

$$\begin{aligned} S(\varphi') &= \sqrt{1+(f'+\varphi')^2} \\ &= \sqrt{1+(f')^2} + \frac{f' \cdot \varphi'}{\sqrt{1+(f')^2}} + \frac{(\varphi')^2}{2\{1+(f'+\varphi')^2\}^{\frac{2}{3}}} \end{aligned}$$

$$\therefore \sqrt{1+(f'+\varphi')^2} - \sqrt{1+(f')^2} = \frac{f' \cdot \varphi'}{\sqrt{1+(f')^2}} + \frac{(\varphi')^2}{2\{1+(f'+\varphi')^2\}^{\frac{2}{3}}}$$

これらから (1.7) は次のように変形できます。

$$\int_0^1 \{\sqrt{1+(f'+\varphi')^2} - \sqrt{1+(f')^2}\}dx = \int_0^1 \frac{f' \cdot \varphi'}{\sqrt{1+(f')^2}}dx + \int_0^1 \frac{(\varphi')^2}{2\{1+(f'+\varphi')^2\}^{\frac{2}{3}}}dx$$

ここで

$$\begin{aligned} R_3 &= \frac{1}{2\{1+(f'+\varphi')^2\}^{\frac{2}{3}}} \\ R_4 &= (\varphi')^2 \end{aligned}$$

とおくと

$$R_2 = R_3 R_4$$

$$\begin{aligned}
|R_3| &< \frac{1}{2} = K \\
|R_4| &= |\varphi'(x)|^2 \leq \|\varphi'\|_C^2 \\
&\leq \|\varphi\|_{C^1}^2 \qquad (\because \|\varphi\|_{C^1} = \|\varphi\|_C + \|\varphi'\|_C)
\end{aligned}$$

$$\therefore |R_2| \leq K\|\varphi\|_{C^1}^2$$

$$\int_0^1 \|\varphi\|_{C^1}^2 dx$$

したがって

$$O(\varphi) \triangleq \frac{1}{\|\varphi\|_{C^1}} \int_0^1 R_2 dx$$

$O(\varphi)$ をこのように定義すると

$$\frac{1}{\|\varphi\|_{C^1}} \int_0^1 R_2 dx \leq K\|\varphi\|_{C_1}$$

であるので、$\|\varphi\|_{C_1} \to 0$ のとき $\|O(\varphi)\| \to 0$

$$\therefore J(f+\varphi) - J(f) = \int_0^1 \frac{f'(x)\varphi'(x)}{\sqrt{1+\{f'(x)\}^2}} dx + O(\varphi)\|\varphi\|_{C^1}$$

次に線型性を示します。

$$\vec{E} \ni \varphi \xmapsto{L_f} \int_0^1 \frac{f'(x)\varphi'(x)}{\sqrt{1+(f'(x))^2}} dx \in \boldsymbol{R}$$

$\forall \varphi_1, \varphi_2 \in \vec{E}$ に対して、

$$\begin{aligned}
L_f(\alpha\varphi_1 + \beta\varphi_2) &= \int_0^1 \frac{f'(x)(\alpha\varphi_1(x) + \beta\varphi_2(x))'}{\sqrt{1+(f'(x))^2}} dx \\
&= \int_0^1 \frac{f'(x)(\alpha\varphi_1(x)' + \beta\varphi_2(x)')}{\sqrt{1+(f'(x))^2}} dx \\
&= \alpha \int_0^1 \frac{f'(x){\varphi_1}'(x)}{\sqrt{1+(f'(x))^2}} dx + \beta \int_0^1 \frac{f'(x){\varphi_2}'(x)}{\sqrt{1+(f'(x))^2}} dx \\
&= \alpha L_f(\varphi_1) + \beta L_f(\varphi_2)
\end{aligned}$$

以上より線型性が示されました。次に連続性を示します。

$$\forall \varepsilon > 0, \quad \exists \delta(\varepsilon) = \frac{\varepsilon}{K_f}, \quad \|\varphi_1 - \varphi_2\|_{C^1} < \delta(\varepsilon)$$

$$\Rightarrow |L_f(\varphi_1) - L_f(\varphi_2)| < \varepsilon$$

これを示す

$$
\begin{aligned}
|L_f(\varphi_1) - L_f(\varphi_2)| &= \left| \int_0^1 \frac{f'(x)(\varphi_1(x) - \varphi_2(x))'}{\sqrt{1+(f'(x))^2}} dx \right| \qquad (1.8) \\
&\leq \int_0^1 \left| \frac{f'(x)}{\sqrt{1+(f'(x))^2}} \right| \cdot |\varphi_1'(x) - \varphi_2'(x)| dx \\
&\leq \int_0^1 \left| \frac{f'(x)}{\sqrt{1+(f'(x))^2}} \right| dx \cdot \|\varphi_1' - \varphi_2'\|_C \\
&\leq \int_0^1 \left| \frac{f'(x)}{\sqrt{1+(f'(x))^2}} \right| dx \cdot \|\varphi_1 - \varphi_2\|_{C^1} \\
&\leq K_f \|\varphi_1 - \varphi_2\|_{C^1} < \varepsilon
\end{aligned}
$$

以上より連続性が示されました。