OR研究の最前線 (2)

極値問題

無限次元なんて恐くない

川崎 英文

1. もぐら超特急

極値問題と聞くと、皆さんはどの様な問題を思い浮かべるでしょうか?本稿の読者であれば、直ちに線形計画問題や非線形計画問題を連想することと思います。ところが私の同僚である数学者でしたら、変分問題を思い浮かべる人が多いようです。この変分問題とは何でしょう?一松信先生の著書[6]の中に、「もぐら超特急」という変分問題の面白い例が紹介されています。**もぐら超特急**というのは、地下に掘られたトンネルの中を重力を利用して走る超特急列車のことです。ジェットコースターを思い浮かべれば分かるように、摩擦や空気抵抗が無ければ、原理的には重力のみを推力として2つの都市をもぐら超特急で移動することができます。そこで、最も速く移動するにはどのようなトンネルを掘ればよいかというのが解きたい問題です。

2つの都市を博多と横浜として、この最適化問題をモデル化してみます。博多・横浜間の直線距離 L は約860×1000mですが、この程度の距離ならば地球の丸みを無視しても構わないでしょう。始発博多を $A=(0,0)$、終点横浜を $B=(L,0)$ 、トンネルの形を関数 $x(t)$, $t\in[0,L]$、重力定数を $g=9.8$ で表すと、移動に要する時間(分)は

$$\frac{1}{60}\int_0^L\sqrt{\frac{1+\dot{x}(t)^2}{2gx(t)}}dt \tag{1}$$

で与えられます。ただし、$\dot{x}(t)$ は $x(t)$ の導関数を表します。従って、もぐら超特急問題は $x(0)=x(L)=0$ を満たす曲線のうちで、積分(1)を最小にするものを見つける問題になります。この様に、未知数がベクトル $x\in R^n$ ではなく関数 $x(t)$ である最適化問題を変分問題と呼びます。もぐら超特急問題の原形は1696年にJ. Bernouliにより提唱された最速降下問題であり、J. Bernoulli、Newton、de l'Hospitalらによって、最速降下曲線はサイクロイドであることが解明されました。

かわさき ひでふみ 九州大学大学院数理学研究科
〒812 福岡市東区箱崎 6-10-1

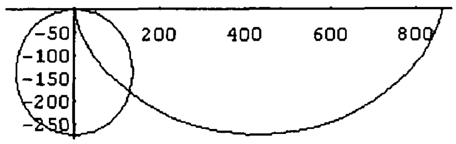

図1

サイクロイドというのは、円盤を転がしたときに円周上の1点が描く曲線のことです(図1)。もぐら超特急の場合、博多・横浜間の所要時間は $\frac{1}{60}\sqrt{\frac{2\pi L}{1000g}}=0.422\sqrt{\frac{L}{1000}}$分 $=12.4$分 です。また、円盤の半径は $\frac{L}{2\pi}=137km$ ですから、トンネルの最深部の深さは $\frac{L}{\pi}=274km$ になります。始発・終点においてサイクロイドの接線の傾きは無限大になりますので、乗客は文字通り奈落の底に落ちてゆく感覚を味わうことになるでしょう。

2. 等周問題

もぐら超特急問題に現れた制約は $x(0)=x(L)=0$ という扱いやすい条件でした。この種の制約条件は**端点条件(境界条件)**と呼ばれます。これとは異なるタイプの制約条件を持つ変分問題として、等周問題がよく知られています。もともと**等周問題**とは、与えられた長さ l の紐で囲まれる図形の面積を最大化する問題を指します。簡単のために、この問題を少し変形した次の変分問題を考える事にします。「平面上の2点 $(0,0)$, $(1,0)$ を通り、$l>1$ の長さを持ち、線分 $[0,1]$ と共に最大面積を囲む上半平面内の曲線 $x(t)$, $t\in[0,1]$ を求めよ。」曲線の長さも囲まれる面積も積分で表されますから、等周問題は端点条件 $x(0)=x(1)=0$ に加えて $\int_0^1\sqrt{1+\dot{x}(t)^2}dt=l$ を満たす曲線のうちで、積分 $\int_0^1 x(t)\,dt$ を最小化する問題になります。最適解は2点 $(0,0)$, $(1,0)$ を通る円弧になる事が知られています。

3. 測地線

別のタイプの制約条件を持つ変分問題として、測地線の問題があります。**測地線**とは曲面上の２点を結ぶ最短経路のことです。例えば、半径１の球面上の測地線の問題は次のように定式化されます。３次元空間内の曲線は、運動する点の軌跡と解釈出来ます。蝿が一匹ぶんぶんと飛びまわれば、ひとつの曲線ができるわけです。つまり、x_1ー座標、x_2ー座標、x_3ー座標が時刻 t と共に変化すると考えればよいのです。そこで、各座標を $x_1(t)$, $x_2(t)$ $x_3(t)$ と書くと、曲線が球面上にあるという制約条件は

$$x_1(t)^2 + x_2(t)^2 + x_3(t)^2 - 1 = 0 \ \ \forall t \tag{2}$$

と表現されます。結局、この問題は端点条件 $(x_1(0), x_2(0), x_3(0)) = a$, $(x_1(1), x_2(1), x_3(1)) = b$ に加えて**相制約(状態制約)**と呼ばれる制約条件(2)の下で、積分$\int_0^1 \sqrt{\dot{x}_1(t)^2 + \dot{x}_2(t)^2 + \dot{x}_3(t)^2}\, dt$ を最小化する問題になります。直感的にも明らかな様に、２点を通る大円が測地線です。

4. Euler-Lagrange 方程式

以上、変分問題の例を３つ挙げましたが、これらはそれぞれ次の形にまとめる事が出来ます。先ず、もぐら超特急問題は

$$\begin{array}{ll} \text{Minimize} & \int_0^T f(t, x(t), \dot{x}(t))\, dt \\ \text{subject to} & x(0) = a, \ x(T) = b \end{array} \tag{3}$$

の形をしています。問題(3)は**変分法の基本問題**と呼ばれています。等周問題は

$$\begin{array}{ll} \text{Minimize} & \int_0^T f(t, x(t), \dot{x}(t))\, dt \\ \text{subject to} & \int_0^T h(t, x(t), \dot{x}(t)\, dt = 0 \\ & x(0) = a, \ x(T) = b \end{array} \tag{4}$$

と表されます。広い意味では、この形をした変分問題を**等周問題**と呼びます。さらに測地線の問題は、$(x_1(t), x_2(t), x_3(t))$ をベクトル記号を用いて$x(t)$ と書くと、次の様に表現されます。

$$\begin{array}{ll} \text{Minimize} & \int_0^T f(t, x(t), \dot{x}(t))\, dt \\ \text{subject to} & h(t, x(t)) = 0 \ \ \forall t \\ & x(0) = a, \ x(T) = b \end{array} \tag{5}$$

それでは、変分問題を実際に解くにはどうしたらよいでしょうか?基本問題の最適解は**Euler 方程式**と呼ばれる２階微分方程式

$$\frac{d}{dt} f_{\dot{x}}(t, x(t), \dot{x}(t)) - f_x(t, x(t), \dot{x}(t)) = 0 \tag{6}$$

を解くことにより得られます。ここで、$f_{\dot{x}}$ は関数 $f(t, x, \dot{x})$ の第３変数 $\dot{x}$ に関する偏微分を表します。次に、等周問題(4)の最適解は、適当な定数 λ を選び、Lagrange 関数を$L(t, x, \dot{x}) := f(t, x, \dot{x}) + \lambda h(t, x, \dot{x})$ と定義するとき、**Euler-Lagrange 方程式**と呼ばれる２階微分方程式

$$\frac{d}{dt} L_{\dot{x}}(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \tag{7}$$

を解くことにより求まります。さらに、測地線問題(5)の最適解は、t に依存する適当な定数 $\lambda(t)$ を選び、Lagrange 関数を $L(t, x, \dot{x}) := f(t, x, \dot{x}) + \lambda(t) h(t, x)$ と定義するとき、Euler-Lagrange 方程式と呼ばれる２階微分方程式

$$\frac{d}{dt} L_{\dot{x}}(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t)) = 0 \tag{8}$$

を解くことにより求まります。

ここまで書けば、非線形計画問題

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & h_i(x) = 0 \ \ i = 1, \ldots, m \end{array} \tag{9}$$

に対するLagrange 乗数法との共通点に気付いたことと思います。Lagrange 乗数法によれば、x が非線形計画問題(9)の極小解ならば、適当な定数$\lambda_1, \ldots, \lambda_m$ を選び、Lagrange 関数を$L(x) := f(x) + \lambda_1 h_1(x) + \cdots + \lambda_m h_m(x)$ と定義するとき、$\frac{\partial L(x)}{\partial x_i} = 0 \ \ i = 1, \ldots, n$ が成立します。

方程式と微分方程式の違いはありますが、目的関数と制約関数を Lagrange 乗数λ, $\lambda(t)$, λ_i を用いて組み合わせ Lagrange 関数を構成する手順は同じです。一見すると、等周問題の積分で表される制約条件は難しそうですが、所詮は１個の等式制約ですから１個のLagrange 乗数 λ で事足ります。それに比べて、簡単そうに見える測地線問題の相制約は、t 毎に条件が設定された無限個の等式制約ですので、無限個の Lagrange 乗数 $\lambda(t)$ が必要になります。

ここまで考察が進めば、これらの問題を統一的に取り扱えるのではないかと考えるのは自然なことです。Girsanov[5], Ioffe, Tihomirov[9], Luenberger[14], Neustadt[15]をはじめとする多くの研究者の努力により変分問題はおろか更に広いクラスの問題である最適制御問題も統一的に取り扱う事が可能になりました。非線形計画問題において未知数は n 次元 Euclid 空間の点 $x = (x_1, \ldots, x_n)$ ですが、変分問題では未知

数は関数 $x(t)$ なのです。関数の集まりは関数空間と呼ばれ、一般に無限次元空間になります。関数 $x(t)$ というと、そのグラフである曲線を思い浮かべる人が多いかと思いますが、関数空間では関数 $x(t)$ もひとつの点と見なします。R^3 の点を３個の成分を持った点と考えるのと同様に、関数を無限個の成分を持った点と考えるのです。このアイディアに基づき、関数 $x(t)$ に対して

$$F(x) := \int_0^1 f(t, x(t), \dot{x}(t))\, dt \tag{10}$$

$$H(x) := \begin{pmatrix} x(0) - a \\ x(1) - b \\ \int_0^1 f(t, x(t), \dot{x}(t))\, dt \end{pmatrix} \tag{11}$$

とおけば、等周問題(4)は次のようなスッキリした形に書き直されます。

$$\begin{array}{ll} \text{Minimize} & F(x) \\ \text{subject to} & H(x) = 0. \end{array} \tag{12}$$

同様に、不等式制約を持つ極値問題は次のように抽象化されます。

$$\boxed{\text{(P)} \quad \begin{array}{ll} \text{Minimize} & F(x) \\ \text{subject to} & G(x) \in K, \quad H(x) = 0 \end{array}}$$

ここで K はノルム空間の内部が空でない閉凸錐です。例えば、時刻 t 毎に最低限の達成目標 $l(t)$ を設定した

$$x(t) \geq l(t) \quad \forall t \tag{13}$$

という片側相制約は、凸錐 $K := \{\text{関数}x;\, x(t) \geq 0\ \forall t\}$ を用いて、$x - l \in K$ と表すことが出来ます。現在、研究の対象となっているのは、この抽象化された不等式制約 $G(x) \in K$ を持つ最適化問題に対する２次の最適性条件です。

5.　１次と２次の最適性条件

人間ドックに入ると、胃の透視や腹部超音波検診等さまざまな検査を受けます。検査の結果が全て白なら万々歳ですが、疑わしい検査結果が出た場合は精密検査をうけることになります。例えば、胃のレントゲン写真に不審な陰が見つかれば、胃カメラによる詳しい検査を受けます。最適性条件も同様に、最適解を見つけるための検査だと思えばよいのです。胃の透視が１次の最適性条件に、胃カメラが２次の最適性条件に対応します。

一般に最適性条件というのは、Lagrange 乗数法や Kuhn-Tucker 条件の様に、最適解が満たすべき（最適性必要条件）、あるいは最適解であることを保証する条件（最適性十分条）を指します。特に、１階微分のみを用いて表現された最適性条件を**１次の最適性条件**、１階微分と２階微分を用いて表わされた最適性条件を**２次の最適性条件**と呼びます。ここで１階微分と２階微分の違いは、作図における定規とコンパスの違いだと思えばよいでしょう。定規で描くことが出来るのは直線だけですが、定規とコンパスを使えばそれなりに曲線も描くことが出来ます。

ところで高校の頃、１変数関数 $f(x)$ の極値を求めるには $f'(x) = 0$ を解けばよいと習いましたが、この条件は１次の最適性(必要)条件です。さらに、大学１年の時 Taylor 展開の応用として次のことを学びました。「極小解において $f''(x) \geq 0$ が成立し、逆に $f'(x) = 0$ と $f''(x) > 0$ が成立すれば x は極小解になる。」前者が２次の最適性必要条件、後者が２次の最適性十分条件です。見た目は全然違いますが、実は Euler 方程式も１次の最適性必要条件です。

このように、大学1年生の教科書にも載っている基本的な概念である２次の最適性条件(特に、十分条件)を、問題(P)に対して統一的な形で構成する事が我々の分野の大きな目標のひとつです。

そして、合成関数、epi-導関数、包絡線項等の研究がここ10年程の間に進展した結果、霧が晴れてようやく頂上付近が見えてきたと言った状況なのです。なお、ここに挙げた見慣れない概念は本節の後半で説明します。

ところで、これらの１次の最適性条件を導く基本的なアイディアは意外と単純です。

今、見たこともない果物が目の前にあると想像して下さい。中に空洞が空いているかもしれませんし、種が沢山はいっているかもしれません。この果物の中身を知るにはどうしたらよいでしょうか?

答えは簡単です。ナイフで切って見れば分かります。勿論切り口をひとつ見ただけでは不十分ですから、何度もナイフを入れる必要があります。かと言ってやみくもにナイフを入れたのでは、バラバラになった果物の残骸が残り後の処理に困りますから、整然とナイフを入れなければなりません。これらの要求を満たす最も単純で自然なナイフの入れ方は、メロンを切るように、調べたい箇所を通る直線に沿って放射状に

切ることです。このアイディアを極値問題(P)に当てはめると、方向 y を変えながら解の候補 $\bar{x}$ を通る直線 $\bar{x}+\theta y$ $(\theta>0)$ 上で $F(x)$, $G(x)$, $H(x)$ のグラフを切り、その切り口を観察することになります。この際、関数が $\bar{x}$ で増大するか、減少するかさえ調べれば十分です。この単純なアイディアが、変分法や最適制御問題のような複雑な問題を解く時も、そっくりそのまま使えるのです。その結果、適当な正則性の仮定の下で、問題(P)の極小解において

$$\begin{cases} F'(\bar{x})y<0, \\ G'(\bar{x})y \in clcone(K-G(\bar{x})),\ H'(\bar{x})y=0. \end{cases} \tag{14}$$

を満たす方向 $y \in X$ は存在しない事が導かれます。ここで、F' 等は Fréchet 微分、$clcone(K-G(\bar{x}))$ は $K-G(\bar{x})$ を含む最小の閉錐を表します。Fréchet 微分に馴染みの無い方は、通常の微分と思っても、本稿を読むのには差し支えありません。例えば、有限次元空間における非線形計画問題

$$\text{(NLP)}\quad \begin{array}{ll} \text{Minimize} & F(x) \\ \text{subject to} & G_1(x)\le 0,\dots,G_m(x)\le 0 \\ & H_1(x)=0,\dots,H_l(x)=0 \end{array}$$

の場合、Fréchet 微分 $F'(x)$ は勾配ベクトル $\nabla F(x)$ になります。また、K としては R^m の非正象限 R^m_- をとります。この時 (14) は

$$\nabla F(\bar{x})y<0,\ \nabla H_1(\bar{x})y=0,\dots,\ \nabla H_1(\bar{x})y=0,$$
$$\nabla G_j(\bar{x})y\le 0 \ \text{ if } \ G_j(\bar{x})=0$$

となります。つまり、(14) は、点 $\bar{x}$ からその方向に少し移動しても(1階微分を用いて判定する限りにおいては)制約条件を満たし、目的関数 F の値が確実に減少する方向を表します。条件(14)は1次の最適性必要条件の**主形式**と呼ばれます。この主形式を分離定理を用いて別表現したものが次の定理1で、**双対形式**と呼ばれます。

定理 1 適当な正則性の仮定の下で、$\bar{x}$ が極小解ならば、適当な $v^* \in K^o$ と $w^* \in W^*$ を選び、*Lagrange*関数を $L(x)=F(x)+<v^*,G(x)>+<w^*,H(x)>$ と定義する時

$$L'(\bar{x})=0,\quad <v^*,G(\bar{x})>=0 \tag{15}$$

が成立する。ただし、W^* は共役空間、K^o は極錐 $\{v^*\in V^*;\ <v^*,v>\le 0\ \forall v\in K\}$ を表す。

非線形計画問題(NLP)の場合、$W=R^l$、その共役空間は R^l、極錐 K^o は非負象限 R^m_+、$<v^*,G(x)>$ は内積 $\sum_{i=1}^m v_iG_i(x)$ になりますので、$L(x)$ は通常の Lagrange 関数 $F(x)+\sum_{i=1}^m v_iG_i(x)+\sum_{j=1}^l w_jH_j(x)$ に一致します。従って(15)はKarush-Kuhn-Tucker 条件に他なりません。また変分問題に適用すると、Euler 方程式を始めとする上述の最適性条件が全て導かれます。

ところが残念な事に、(15)の解が極小解になるとは限りません。極大解や鞍点でも(15)を満たすことがあります。これは1階微分という「定規」を用いることによる限界です。つまり切り口を調べる際に、定規では切り口の曲線の傾きしか分からないのです。曲線の曲り具合を読み取るには、「コンパス」に対応する2階微分を利用します。この事により、より精密な検査が可能になります。

先程1次の最適性条件を導く際、我々は真っ直ぐナイフを入れました。2次の最適性条件の先駆者である Fiacco, McCormic[3] はこの手法を押し進めて2次の最適性条件を導きましたが、Ben-Tal, Zowe[1] らは別の手法を提案しました。彼らはカーブを描きながらナイフを入れて、切り口を観察したのです。正確には **curved variation** と呼ばれる曲線 $x+\theta y+\theta^2 z$ 上で関数 $F(x)$, $G(x)$, $H(x)$ の増減を2階微分まで用いて調べました。ただ、全ての curved variation を調べる必要はありません。既に1次検査で白と判定された方向 y はパスして構いません。精密検査が必要なのは**臨界方向**と呼ばれる次の条件を満たす y です。

$$\begin{cases} F'(\bar{x})y=0, \\ G'(\bar{x})y \in clcone(K-G(\bar{x})),\ H'(\bar{x})y=0. \end{cases} \tag{16}$$

例えば、非線形計画問題で $F(x_1,x_2)=(x_1-1)^2+x_2^2$, $G(x_1,x_2)=x_1-x_2^2\le 0$, $\bar{x}=(\frac{1}{2},\frac{1}{\sqrt{2}})$ の場合、図の接線方向が臨界方向になります。

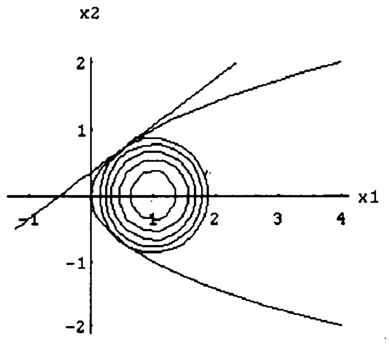

図2

実際にはBen-Talらは全ての臨界方向ではなく

$$\begin{cases} F'(\bar{x})y = 0, \\ G'(\bar{x})y \in cone(K - G(\bar{x})),\ H'(\bar{x})y = 0. \end{cases} \tag{17}$$

を満たす臨界方向に対して精密な条件を与えました。

定理 2 [1] $\bar{x}$ を (P) の極小解とすると、適当な正則性の仮定の下で、条件 *(17)* を満たす各臨界方向 y に対して、ある $v^* \in K^o$, $w^* \in W^*$ が存在して次式を満たす。

$$L'(\bar{x}) = 0, \tag{18}$$

$$L''(\bar{x})(y,y) \geq 0, \tag{19}$$

$$< v^*, G(\bar{x}) >= 0, \quad < v^*, G'(\bar{x})y >= 0. \tag{20}$$

非線形計画問題の場合、$L''(x)$ は Lagrange 関数の Hesse 行列 $(\partial^2 L/\partial x_i \partial x_j)$ に、$L''(x)(y,y)$ は $y^T L''(x)y$ になります。

一見すると、臨界方向に関する2つの条件(16)と(17)の違いはほんの僅かに思えます。しかしこの僅かな違いに大変面白い現象が潜んでいることが80年代後半に分かりました。この違いを説明する為に、5節を振り返って見ましょう。すると、1次にせよ2次にせよこれまで述べた最適性条件は全て Lagrange 関数を用いて記述されていた事に気がつきます。ところが、Ben-Tal らが検査し残した臨界方向を調べるためにはLagrange 関数だけでは不足なのです。馴染みは無いかもしれませんが、次の集合を導入することにより、それらの臨界方向にメスを入れることが可能になりました。

定義 1 [10] ノルム空間 V の任意の要素 u, v に対して、V の部分集合 $\{w \in V;\ \exists o(1)\ s.t.\ \theta^2 u + \theta v + w + o(1) \in K\ \ \forall \theta > 0\}$ を $K(u,v)$ と書く。ただし、$o(1)$ は $\theta \to +0$ のとき 0 に収束する V の要素とする。また、$K(G(\bar{x}), G'(\bar{x})y)$ を $K(y)$ と書く。

定理 3 [10] $\bar{x}$ を問題 (P) の正則な極小解とする。このとき $K(y)$ が空でない各臨界方向 y に対し、ある $v^* \in K^o$, $w^* \in W^*$ が存在して、*(18), (20)* に加えて

$$L''(\bar{x})(y,y) - 2\delta^*(v^*|K(y)) \geq 0, \tag{21}$$

が成立する。ただし $\delta^*(v^*|K(y))$ は支持関数 $sup\{< v^*, v >;\ v \in K(y)\}$ を表す。

定理3と従来の結果の違いは付加項 $\delta^*(v^*|K(y))$ です。定理3を導いた時点では、この項が何を意味するのかよく分かりませんでしたが、その後の研究[11][12]により、付加項が不等式制約から生成される包絡線の2階微分と密接に関係する事が明らかになりました。突然、包絡線と言ってもピンと来ないと思いますので、包絡線の簡単な例を挙げておきます。

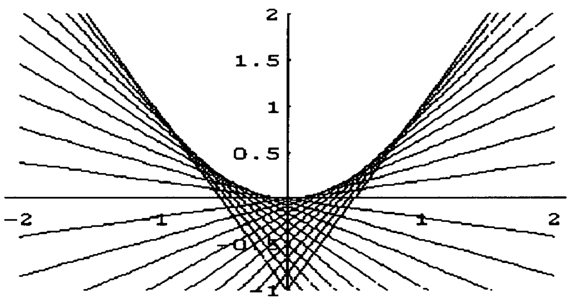

図3

パラメータ $t \in [-1,1]$ を持つ直線族 $G(x,t) = 2tx - t^2$ の max-型関数 $f(x) := \max_t G(x,t)$ を考えると、包絡線 $f(x) = x^2$, $(|x| \leq 1)$ が生成されます(図3)。ここで重要なことは、各直線の x に関する2階微分は0であるにも拘わらず、包絡線の2階微分は正になるという事です。前節の終わりに片側相制約 $x(t) \geq l(t)\ \forall t$ を挙げて説明したように、一般不等式制約 $G(x) \in K$ は有限個あるいは無限個の不等式制約を抽象化したものです。また、不等式制約というのは、「なた」を使った一刀彫りの様に、x−空間を削り取って許容領域を形造るものです。有限個の不等式制約では有限回しかなたを入れませんから、ゴツゴツとした許容領域が形造られるだけですが、無限個の不等式制約の場合は、なたの入れ方によっては滑らかな許容領域が出来上がります。つまり包絡線(包絡面)が形成されることがあり、その結果上述の2階微分のギャップが生じます。実は、このギャップを埋めるのが付加項 $\delta^*(v^*|K(y))$ なのです。この事実により、この項は**包絡線項**と呼ばれます。詳しい説明は[10][11][12]を参照して下さい。

包絡線項に関連した文献は、集合制約を持つ場合への定理3の拡張がCominetti[2], 2次の最適性十分条件はIoffe[7], Penot[16], 下半2次方向微分についてはIoffe[8]があります。特に、2次の十分条件を研究していく上で、Rockafellar[17] とIoffe[7] は重要です。Rockafellar は合成関数の形をした目的関数の最小化問題に対して、epi-導関数を導入して2次の最適性条件を論じました。また、Ioffe はRockafellar の結果と包絡

線項を融合させました。ここで言う**合成関数**とは、例えば有限個の関数のmax-型関数 $f(x) := \max_{i=1}^{m} f_i(x)$ を $F(x) := (f_1(x),\ldots,f_m(x))$ と $g(y_1,\ldots,y_m) := \max_{i=1}^{m} y_i$ を用いて $f(x) = g(F(x))$ と表したものを指します。各 $f_i(x)$ が滑らかならば、$F(x)$ も滑らかですし、$g(y)$ は区分的に線形になっています。このように、問題の難しさを２つに分解して取り扱おうというのが、Rockafellar の基本的な姿勢です。また、**epi-導関数**は epi-収束の概念を用いて定義された微分です。大雑把に言えば、通常の微分の定義式に表れる極限を epi-収束という極限で置き換えた物が epi-導関数です。パラメータ t をもつ下半連続な関数族 f_t が下半連続関数 f に **epi-収束**するとは、f_t のエピグラフ $\mathrm{epi}\, f_t := \{(x,r)\,;\, f_t(x) \le r\}$ が f のエピグラフに

$$\lim_t d(\mathrm{epi}\, f_t,\, \xi) = d(\mathrm{epi}\, f,\, \xi) \quad \forall \xi$$

の意味で収束することを指します。ただし $d(\mathrm{epi}\, f, \xi)$ は集合 $\mathrm{epi}\, f$ と点 ξ の距離を表します。例えば、パラメータ $t \in R$ をもつ関数族

$$f_t(x) := \begin{cases} 0 & \text{if } x = t, \\ 1 & \text{if } x \ne t. \end{cases}$$

について、$t \to +0$ のとき f_t は f_0 に epi-収束します。しかし、各点収束の意味では $f(x) \equiv 1$ に収束します。この例が端的に示すように、epi-収束は、各点収束が取りこぼしていた値(この例では、各関数の最小点 0)を拾い上げる事が出来ます。epi-収束については Rockafellar, Wets[18] を参考にするとよいでしょう。

6. 無限次元空間の心得

極値問題を(P)の形に抽象化することの利点のひとつは n 次元 Euclid 空間 R^n との共通点が明確になる事です。しかし何もかも R^n と同じと言うわけには行きません。本節では、特に注意すべき事項をピックアップしておきます。詳しくは、関数解析の入門書や[5][9][13][14][15][19]を参照して下さい。

(A) R^n では閉球 $\{x;\, \| x \| \le r\}$ はコンパクトですが、無限次元ノルム空間の閉球はコンパクトにはなりません。その結果、連続関数が閉球上で最大値、最小値を持たないことがあります。また、閉凸集合 C と点 $a \notin C$ に対して、a に最も近い C の点が存在しないことがあります。さらに、アルゴリズムの収束性を証明する際よく利用する論法「生成された点列 x_k は有界であるから、適当な部分列を選ぶことにより、その極限は条件…を満たす。」は使えません。有界点列だからと言って、収束する部分列を選べるとは限らないからです。

(B) 無限次元空間でも R^n でも様々な距離を考えることが出来ますが、R^n の距離は本質的には皆同じです。しかし、無限次元の場合には本質的に異なる距離が何種類か存在します。次の２つの関数の距離は近いでしょうか、それとも遠いでしょうか?

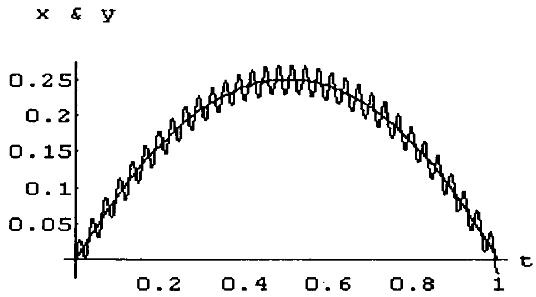

図4

「どちらとも考えられる。」と言うのが正解です。距離 $d(x,y) := \max_t |x(t) - y(t)|$ で測ると近く、距離 $d_1(x,y) := \max_t |x(t) - y(t)| + \max_t |\dot{x}(t) - \dot{y}(t)|$ で測ると遠くなります。それ故、極小解（その点の近くでは最小解）を議論する際はどちらの距離で考えているのかを明確にしなければなりません。変分法では、距離 d に関する極小解を**強極値**(C-極値)、距離 d_1 に関する極小解を**弱極値**(D-極値)と呼びます。後者は関数の値だけでなく導関数の値まで近いとき初めて２つの関数は近いと考えますから、近い関数が少ないことになります。そこで一番良いと言っても説得力に欠けるので、弱極値と呼ばれるわけです。

(C) ほとんどの関数空間では角度を考える事が出来ません。角度を考えることの出来るノルム空間を内積空間、完備な内積空間を Hilbert 空間と呼びます。Hilbert 空間は関数空間の中では最も良い性質を持つ空間ですから、そこで議論できれば一般には都合の良い事が多いのですが、残念な事に次の２つの理由により不等式制約には向いていません。Hilbert 空間の代表例である $L^2 := \{x;\, \int x(t)^2\, dt < \infty\}$ の閉凸錐 $\{x \in L^2\,;\, x(t) \ge 0 \ \ \forall t(\text{正確には a.e. t})\}$ は内点を持ちません。また、片側相制約を持つ変分問題は Lagrange 乗数が特異性を持つため、本質的に Hilbert 空間では議論が無理です。さらに(B)で挙げた２つの距離を入れた連続関数、連続微分可能関数の空間はどちらも内

積空間にはなりません。これらの事実は、数理計画法が不等式制約を主な研究対象としていることを考えると、重く受け止めなければなりません。しかし多くの論文において空間の設定が軽視されているのが現状です。また、ノルム空間より広い空間で極値問題を議論している論文も沢山あります。最も一般的な設定は位相線型空間です。しかしその場合でも内容はノルム空間の場合と全く同じで、どちらの空間で記述するかは著者の趣味の問題であることが多いようです。位相線型空間を持ち出す必然性は、例えば弱位相や*−弱位相を考える時に生じますが、それらの位相の特性を利用した議論はあまりなされていないようです。

当たり前の事ですが、抽象的な問題を一般的に議論するという事は広範な問題をいっきに処理してしまう事を意味します。一般的な設定が思わぬ難問を含んでしまっているかもしれません。通り一遍の議論で難問がスラスラと解ければ苦労はしません。従って一般論を展開する時は、無意味な事をやっていないかを多くの例でチェックしながら注意深く研究を進めなければなりません。某大学(九州大学ではありません)の大学院入試の面接で、コンパクト多様体の研究を希望する学生に「コンパクト多様体の例をあげなさい。」と質問したところ、その学生は答えられず「もっと本質的な質問をして下さい。」と捨て台詞を吐いて退席したそうです。

最後になりましたが、本稿を執筆する機会を与えて下さり、さらに原稿に目を通してコメントを下さいました編集委員の方々、とりわけ水野真治先生に厚くお礼申し上げます。

参考文献

[1] A. Ben-Tal and J. Zowe, "A unified theory of first and second order conditions for extremum problems in topological vector spaces" , *Math. Program. Study* 19 (1982) 39–76.

[2] R Cominetti, "Metric regularity, tangent sets, and second-order optimality conditions", *Appl. Math. Optim.* 29 (1990) 265–287.

[3] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming : Sequential Unconstrained Minimization Techniques,* John Wiley and Sons (1968).

[4] I.M. Gelfand and S.V. Fomin, (関根訳), 変分法, 総合図書 (1970).

[5] I.V. Girsanov, *Lectures on Mathematical Theory of Extremum Problems*, Springer, New York (1972).

[6] 一松信, 偏微分と極値問題, 現代数学社 (1971).

[7] A. Ioffe, "On some recent developments in the theory of second order optimality conditions", in *S. Dolezki (ed.) Optimization,* Lecture Notes in Math. 1405, Springer, New York (1989) 55–68.

[8] A. Ioffe, "Variational analysis of a composite function: A formula for the lower second order epiderivative", *JMAA* 160 (1991) 379–405.

[9] A. Ioffe and V. Tihomirov, *"Theory of Extremal Problems"*, North-Holland, Amsterdam (1979).

[10] H. Kawasaki, "An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems", *Math. Program.* 41 (1988) 73–96.

[11] H. Kawasaki, "The upper and lower second-order directional derivatives of a sup-type function," *Math. Program.* 41 (1988) 327–339.

[12] H. Kawasaki, "Second order necessary optimality conditions for minimizing a sup-type function", *Math. Program.* 49 (1991) 213–229.

[13] 熊田禎宣, 木谷忍, 計画のための最適化数学, 井上書院 (1988).

[14] D.G. Luenberger (増淵、嘉納訳), 関数解析による最適化理論, コロナ社 (1973).

[15] L.W. Neustadt, "A general theory of extremals", *J. Computer and System Sciences*, 3 (1969) 57–92.

[16] J.P. Penot, "Optimality conditions in mathematical programming and composite optimization", *Math. Program.* 67 (1994) 225–245.

[17] R.T. Rockafellar, "First-and second-order epi-differentiability in nonlinear programming", *Trans. Amer. Math. Soc.* 307 (1988) 75–108.

[18] R.T. Rockafellar and R.J.-B. Wets, "Variational systems, an introduction", in Multifunctions and integrands (ed. G. Salinetti), Lecture Notes in Math. 1091 (1984) 1–54, Springer-Verlag.

[19] 志水清孝, 最適制御の理論と計算法, コロナ社 (1994).