担当教員: 柳田 伸太郎　研究室: A441　E-mail:yanagida@math.nagoya-u.ac.jp

ラグランジュの未定乗数法

作成日: 12/11/2016 Version: 0.1　配布日: 12/15/2016

前回 (第 9 セット) では主に $\mathbb{R}^n$ の開集合上で定義された関数 f の最大最小問題について復習した。今回は別の写像 (または関数)g を用いて $g(x) = c$ と定義される部分集合上での関数 f の最大最小問題を考察する。

Lagrange の未定乗数法

条件付きの最大最小問題のうち比較的簡単なものは高校数学でも扱われている。例えば

関数 $f(x, y) = x^2 + y^2 + 3xy$ の曲線 $x^2 + y^2 = 1$ 上での最大最小を求めよ。

という問題なら、極座標 $(x, y) = (\cos\theta, y = \sin\theta)$ を用いて $f(\cos\theta, \sin\theta) = 1 + \frac{3}{2}\sin(2\theta)$ となるから最大値 $5/2$, 最小値 $-1/2$ とすぐにわかる。では次の問題はどうだろうか。

問題 1. 関数 $f(x, y, z) = x^3 + y^3 + 3xy$ の曲線 $x^2 + y^2 = 1$ での最大値と最小値を求めよ。

極座標を用いて 1 変数関数の極値問題を解いても解答は得られるが、計算は最初の問題の場合ほど楽ではない。このような問題に有効なのが Lagrange の未定乗数法である。

$\mathbb{R}^n$ 内の開集合 U 上で定義された関数 $f : U \to \mathbb{R}$ と写像 $g : U \to \mathbb{R}^m$ を考える。これらは U 上で C^1 級と仮定する。そして $S := g^{-1}(0) = \{x \in U \mid g(x) = 0\}$ とおく。この時 $a \in S$ が関数 f の S 上での極値点であるとは、a を中心とする開球 $U(a; \varepsilon) = \{x \in \mathbb{R}^n \mid |x - a| < \varepsilon\}$ が存在して、f が点 a で $S \cap U(a; \varepsilon)$ 上の最大値または最小値を取るときを言う。また最大値、最小値に応じて f の a での値を極大値、極小値と呼ぶ。このような極値点を求めるための、次の定理で与えられる手法を Lagrange の未定乗数法と言う。

定理 1. $U \subset \mathbb{R}^n, f, g, S = g^{-1}(0)$ を上と同様のものとする。$a \in S$ が f の S 上での極値点であるとする。このとき次のいずれかが成り立つ。

(1) $U \times \mathbb{R}^m$ 上の関数 $\Phi(x, \lambda) := f(x) - \langle \lambda, g(x) \rangle$ に対して、ある $\lambda^o \in \mathbb{R}^m$ が存在して次が成り立つ。

$$\frac{\partial \Phi}{\partial x_j}(a, \lambda^o) = 0, \quad \frac{\partial \Phi}{\partial \lambda_i}(a, \lambda^o) = 0 \qquad (j = 1, \ldots, n,\ i = 1, \ldots, m).$$

但し $\langle \lambda, g(x) \rangle$ は $\lambda \in \mathbb{R}^m$ と $g(x) \in \mathbb{R}^m$ の Euclid 内積。

(2) $\operatorname{rank}(g'(a)) < m$.

この定理の使い方を $m = 1$ の場合に説明しよう。

まず $g'(x) \neq 0$ $(x \in U)$ と仮定しよう。もし $a \in S = g^{-1}(0)$ が f の S 上での極値点であれば、定理の (2) はこの仮定から成り立たないので定理の (1) が成り立つ。つまり $U \times \mathbb{R} \subset \mathbb{R}^{n+1}$ 上の関数 $\Phi(x, \lambda) = f(x) - \lambda g(x)$ に対して、定理から

$$\frac{\partial \Phi}{\partial x_j}(a, \lambda^o) = 0, \quad \frac{\partial \Phi}{\partial \lambda}(a, \lambda^o) = 0 \qquad (j = 1, \ldots, n)$$

担当教員: 柳田 伸太郎　研究室: A441　E-mail:yanagida@math.nagoya-u.ac.jp

が成り立つような $\lambda^o \in \mathbb{R}$ が存在する。逆にこの方程式を満たす $(a, \lambda^o) \in U \times \mathbb{R}$ があれば、$\frac{\partial \Phi}{\partial \lambda}(x, \lambda) = -g(x)$ より $g(a) = 0$、つまり $a \in S$ である。従って f の S 上での極大値、極小値を求めるためには上記の方程式の解 (a, λ^o) を全て求め、それらの周りでの f の挙動を調べれば良い。但しこれらの点の中には極値点でないもの存在しうることに注意。

また、例えば $S = g^{-1}(0)$ が有界閉集合であれば、f は必ず S 上で最大値と最小値を取る。最大値や最小値を与える点は極値点だから、上のようにして得られた候補の中に最大値・最小値を与える点が必ず存在する。なお S が有界閉集合でない場合には f は最大値や最小値を取るとは限らないので注意すること。

問題 2. $g(x, y) := x - y + 1,\ f(x, y) := 2y^3 - x^3 - 6x^2 - 6x$ について考える。

(1) $g' = 0$ となる点は存在しないことを確かめよ。

(2) $\Phi(x, y, \lambda) := f(x, y) - \lambda g(x, y)$ とする。$\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial \lambda}$ を計算せよ。

(3) 点 ${}^t(x, y, \lambda) = {}^t(0, 1, -6)$ は $\frac{\partial \Phi}{\partial x} = \frac{\partial \Phi}{\partial y} = \frac{\partial \Phi}{\partial \lambda} = 0$ を満たすが ${}^t(x, y) = {}^t(0, 1)$ は $f(x, y)$ の極値点ではないことを示せ。

問題 3. 次の関数の $S := \{{}^t(x, y, z) \in \mathbb{R}^3\,;\, x^2 + y^2 + z^2 = 1\}$ での最大値、最小値を求めよ。

(1) $f(x, y, z) = xy + yz + zx.$　　(2) $f(x, y, z) = z^2 + xy.$

問題 4. (プリント Y009 の問題 20) 上面の開いた直方体の箱を作る。表面積が 12 となるもので体積の最大のものがどのような箱か、Lagrange の未定乗数法を用いて考察せよ。

問題 5. 次の集合 S 上での関数 f の極値点および最大値・最小値を求めよ。

(1) $f(x, y, z) = x^2 + y^2 + z^2,\ S = \{{}^t(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = d\}$. 但し $a, b, c, d \in \mathbb{R}$ で $(a, b, c) \neq (0, 0, 0)$.

(2) $f(x, y) = x^2 + y^2,\ S = \{{}^t(x, y) \in \mathbb{R}^2 \mid x^4 + y^4 + 3xy = 2\}$.

問題 6. a, b, c, d を正の実数とする。条件 $x + y + z = d,\ x, y, z > 0$ のもとで関数 $f(x, y, z) = x^a y^b z^c$ の最大値を求めよ。

問題 7. $\mathbb{R}^3$ 内の楕円面 $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ に内接する直方体で各辺が x 軸, y 軸, z 軸に平行なもののうち、体積が最大なものを求めよ。ただし a, b, c は正の実数とする。

今までは g が関数の場合の極大極小問題を考えたが、g が写像の場合も扱っておこう。

問題 8. 条件 $x + y - z = 2,\ -x + y + z = 3$ のもとで関数 $f(x, y, z) = x^2 + y^2 + z^2$ の最小値を求めよ。

問題 9. 原点中心の半径 1 の球面と平面 $x + 2y + 3z = 0$ の共通部分に出来る曲線上の点と、点 ${}^t(0, 1, 0)$ との距離の最小値と最大値を求めよ。

問題 10. $\mathbb{R}$ の区間 I で定義された関数 f は、任意の $x, y \in I$ と任意の $t \in [0,1]$ に対して $f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$ を満たすとき凸関数と呼ばれる。

(1) 区間 I 上の凸関数 f と任意の $a_1, \ldots, a_n \in I$, $t_1, \ldots, t_n \geq 0$ に対して $t_1+\cdots+t_n=1$ ならば $t_1a_1+\cdots+t_na_n \in I$ であり、かつ $f(t_1a_1+\cdots+a_nt_n) \leq t_1f(a_1)+\cdots+t_nf(a_n)$ が成り立つことを示せ。

(2) $f(x)=x\log x$ は $[0,\infty)$ で凸関数であることを示せ。但し $f(0)=0$ とする。

(3) $U := \{x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0\ (i=1,\ldots,n)\}$ 上の関数 $H(x) := -\sum_{i=1}^n x_i \log x_i$ (但し $0\log 0 := 0$ とする) を考える。部分集合 $S := \{x \in U \mid x_1+\cdots+x_n=1\}$ 上での $H(x)$ の最大値とそれを与える点を求めよ。

Lagrange の未定乗数法の幾何学的な意味

ここでは Lagrange の未定乗数法の幾何学的な意味を問題形式で考察する。以後 U を $\mathbb{R}^n$ 内の開集合とし、f と g を U 上の C^1 級実数値関数とする。また $c \in \mathbb{R}$ として $S := g^{-1}(c) = \{x \in U \mid g(x)=c\}$ が曲面だと仮定する。つまり $\nabla g(x) \neq 0$ が任意の $x \in S$ に対して成り立つとする。曲面 S の点 $x \in S$ における接空間とは $\nabla g(x)$ と直交し x を通る平面のことであった。

問題 11. 任意の $a \in S$ を取る。このとき $\nabla g(a)$ と直交する任意のベクトル $u \in \mathbb{R}^n$ に対して、$\varepsilon > 0$ を十分小さく取ると、ある曲線 $c : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ で $c(0)=a$, $c(t) \in S$ $(t \in (-\varepsilon,\varepsilon))$, $c'(0)=u$ となるものが存在することを示せ。(ヒント: 陰関数定理)

上記の問題の曲線 $c(t)$ を取ると、前回のプリント Y009 の問題 8 と 9 より $\left.\frac{d}{dt}\right|_{t=0} f(c(t)) = \langle u, \nabla f(a) \rangle$ となることに注意しておく。

問題 12. 点 $a \in S$ が関数 f の S 上での極値点であったとする。このとき $\nabla g(a)$ と直交する任意のベクトル $u \in \mathbb{R}^n$ に対して $\langle u, \nabla f(a) \rangle = 0$ が成り立つことを示せ。

つまり $\nabla f(a)$ は $\nabla g(a)$ と直交するベクトル達と直交する。これから次が分かる。

問題 13. 上の状況で、ある $\lambda^o \in \mathbb{R}$ が存在して $\nabla f(a) = \lambda^o \nabla g(a)$ となることを示せ。

これは定理 1 の主張に他ならない。つまり Lagrange の未定乗数法は 「曲面 $S = g^{-1}(c)$ 上での f の極値点 a において、$\nabla f(a)$ は S の点 a における接空間に直交する」ということを主張していることが分かった。

その他の問題

Lagrange の未定乗数法は用いる話題ではないが、1 変数多項式の族で直交多項式と呼ばれるものをここで紹介しておく。物理学や工学では微分方程式の解に現れることもあって、学部 2 年生の後期に習う重要な関数族である。

担当教員: 柳田 伸太郎　研究室: A441　E-mail:yanagida@math.nagoya-u.ac.jp

問題 14. (プリント Y006 問題 2 参照) 実係数の n 次多項式 $p(t)$ で t^n の係数が 1 となるものに対して、積分

$$I=\int_0^\infty p(t)^2 e^{-t}\,dt$$

を考える。t^n の係数が 1 となるような実係数 n 次多項式 p を動かしたとき、この積分の値が最小になるような多項式 p を以下の手順で求める。

(1) 任意の非負の整数 l に対して積分 $\int_0^\infty t^l e^{-t}\,dt$ を求めよ。

(2) $n-1$ 次実係数多項式 $q(t)$ を $q(t)=q(x,t)=x_1+x_2t+\cdots+x_nt^{n-1}$ と書く。但し $x={}^t(x_1,\dots,x_n)\in\mathbb{R}^n$. このとき $\mathbb{R}^n$ 上の関数

$$J(x):=\int_0^\infty q(x,t)^2e^{-t}\,dt,\quad x={}^t(x_1,\dots,x_n)\in\mathbb{R}^n$$

に対して固有値が全て正のある n 次実対称行列 A で $J(x)=\langle Ax,x\rangle$ となるものが存在することを示せ。
(ヒント: n 次実対称行列 A に対して $\langle Ax,x\rangle\geq 0$ が任意の $x\in\mathbb{R}^n$ に対して成り立ち、しかも $\langle Ax,x\rangle=0 \iff x=0$ の場合、A の固有値は全て正の実数となる。)

(3) 行列 A を設問 (2) で求めたものとする。$x={}^t(x_1,\dots,x_n)\in\mathbb{R}^n$ の関数 I を

$$I(x):=\int_0^\infty (q(x,t)+t^n)^2e^{-t}\,dt$$

とおく。このときある $a={}^t(a_1,\dots,a_n)\in\mathbb{R}^n$, $c\in\mathbb{R}$ を用いて $I(x)=\langle Ax,x\rangle+2\langle x,a\rangle+c$ の形に書けることを示せ。

(4) 設問 (3) で求めた $I(x)$ の表示をもとに、$I(x)$ の極値点はただ一つであること、そしてそこでの値は最小値であることを示せ。また最小値を与える点 $x^o={}^t(x_1^o,\dots,x_n^o)\in\mathbb{R}^n$ と最小値 $I(x^o)$ を a,c,A を用いて書き表せ。

(5) 一方 $I(x)$ の積分表示式をもとに、$I(x)$ を偏微分してそれが 0 であるという方程式をたてることにより、次の方程式を満たす多項式 $p(x,t)=q(x,t)+t^n$ はただ一つ、$x=x^o$ のときであることを示せ。

$$\int_0^\infty p(x,t)t^{j-1}e^{-t}\,dt=0,\quad (j=1,\dots,n).$$

(6) 関数 $e^t\frac{d^n}{dt^n}(e^{-t}t^n)$ は t についての n 次実係数多項式であることを示せ。また設問 (4),(5) で得られた多項式 $p(x^o,t)=q(x^o,t)+t^n$ はこの多項式の $(-1)^n$ 倍であることを示せ。

つまり積分 I を最小にする多項式は $(-1)^n e^t\frac{d^n}{dt^n}(e^{-t}t^n)$ であることが分かった。関数

$$L_n(t):=\frac{1}{n!}e^t\frac{d^n}{dt^n}(e^{-t}t^n)$$

は Laguerre 多項式と呼ばれ、上で触れた直交多項式と呼ばれる特殊関数の一つである。他にも様々な直交多項式が存在するが、上記のような積分による特徴付けが知られている。