ラグランジュの未定乗数法 コメンタール

(lxfonts パッケージを利用してみた.)

久島広幸 (h2@hisasima.jp)

2015 年 2 月 22 日
revised 2022 年 5 月 18 日
re-revised 2022 年 7 月 30 日

概要

ラグランジュ (Lagurange)[*1] の未定乗数法とは，ある制約条件のもとでの関数の極値を求める解法のひとつである．本稿では，通常の教科書的な定式化の他に，ベクトルの考え方を前面に出しての定式化の方法を整理してみた．またあわせて，未定乗数の持つ意味をまとめた．

1 条件付き極値問題

> $g(x,y)=x^2+y^2-1=0$ という条件があるとき，$f(x,y)=2xy$ の最大値はいくつになるか．また最小値はいくつになるか．

という問題を考えてみる[*2]．どのような方法で解けばいいか？いや，そもそも解けるのか？最大値と最小値を求めるのだから，$f(x,y)$ の振る舞いを分析して，まず，極値を求めることが王道だろう．最大値であるならばそれは極値でもある．最小値もしかり．しかしながら，その逆，つまり極値であれば最大値（最小値）である，ということはできない．停留というものがあるし[*3]．

ということで，まず極値である．極値をあたえる点であれば，その点においては関数の全微分は 0 である，という事柄を利用する．$f(x,y)$ の全微分は

$$df(x,y)=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy=2ydx+2xdy$$

であるから，点 (u,v) が極値を与える点であるとすれば，その点での全微分は 0，すなわち

$$df(x,y)\Big|_{(u,v)}=2ydx+2xdy\Big|_{(u,v)}=0 \qquad \therefore\ 2vdx+2udy=0$$

であるはずである．そして，問題から $g(x,y)=0$ という条件が与えられているので，当然

$$g(u,v)=u^2+v^2-1=0$$

でもあるはずである．この 2 式をぐっと睨んでも，前に進みそうもない．dx や dy が邪魔である．

ここで関数 $g(x,y)$ をもうひとつ嚙み砕いてみる．まず条件式から，$g(x,y)=x^2+y^2-1=0$ を満たすよう要請されていることはいいとして，よくみると $g(x,y)=0$ と $g(x,y)$ は定数であることが要請されてもい

---

*1 Joseph-Louis Lagrange, 1736 年 1 月 25 日 - 1813 年 4 月 10 日.

*2 一見仰々しく見えるが，この問題の描像はいたって簡単で，半径 1 の円上の点たちの中で $f(x,y)=2xy$ を最大／最小にする点はどの点か，というものである．

*3 本稿では，極値というものの中に停留点も含めている．つまり，極値 $=\{$ 極大値を与える点，極小値を与える点，停留点 $\}$ という言葉づかいにしている．

る．これはとりもなおさず $g(x,y)$ の全微分が 0 であることが要請されていることにほかならない．すなわち

$$dg(x,y)=\frac{\partial g}{\partial x}dx+\frac{\partial g}{\partial y}dy=2xdx+2ydy=0$$

であって，もちろん (u,v) でも全微分は 0 つまり

$$dg(x,y)\Big|_{(u,v)}=2xdx+2ydy\Big|_{(u,v)}=0 \quad \therefore\ 2udx+2vdy=0$$

ということになる．

結果をまとめると

$$2vdx+2udy=0\ , \tag{1.1}$$
$$2udx+2vdy=0\ . \tag{1.2}$$

ここから連立方程式の見立てで，(1.1) $\times\, u-$ (1.2) $\times\, v$ 計算すると

$$2(u^2-v^2)dy=0$$

dy は 0 ではない（！）ので（さらにいえば，どのような dy でも成立するので）

$$u^2-v^2=0 \quad \therefore\ u=\pm v$$

ともとまる．

一方で，ベクトルの力を借りてみる．x,y 成分 が a,b である 2 次元ベクトルを $\overrightarrow{(a,b)}$ とあらわし，内積演算を $\bullet$ であらわせば (1.1) と (1.2) は

$$\begin{aligned}(1.1) &\quad\Longleftrightarrow\quad \overrightarrow{(2v,2u)}\bullet\overrightarrow{(dx,dy)}=0\\ (1.2) &\quad\Longleftrightarrow\quad \overrightarrow{(2u,2v)}\bullet\overrightarrow{(dx,dy)}=0\end{aligned}$$

とあらわせて，どちらも $\overrightarrow{(dx,dy)}$ に直交していることがわかる．したがってこの左側の 2 つのベクトルは平行，つまりベクトル的には実数倍ということになるから，λ という実数定数を使って

$$\overrightarrow{(2v,2u)}=\lambda\overrightarrow{(2u,2v)} \quad \therefore\ \begin{cases}v=\lambda u\\ u=\lambda v\end{cases}$$

とあらわせる．この u と v の関係から

$$u=\lambda v=\lambda\lambda u \quad \therefore\ \lambda^2=1 \quad \therefore\ \lambda=\pm 1 \quad \therefore\ u=\pm v$$

というように[*4]，連立方程式と見立てたときと同様の結果を得る．

このようにして極値であるための条件はもとまった．あとは具体的に極値をあたえる点をもとめればよい．まず $v=u$（つまり $\lambda=1$ のとき）を条件に放り込めば

$$u^2+u^2-1=0 \quad \therefore\ u=\pm\frac{1}{\sqrt{2}}\ .$$

従って，極値をあたえる点は $(1/\sqrt{2},\ 1/\sqrt{2})$ と $(-1/\sqrt{2},\ -1/\sqrt{2})$ があげられ，そのとき $f=1$ である．

つぎに $v=-u$（つまり $\lambda=-1$ のとき）を条件に放りむのだけれども，

$$u^2+u^2-1=0 \quad \therefore\ u=\pm\frac{1}{\sqrt{2}}\ .$$

は上と同じであり．$v=-u$ というところが異なるだけであるので，極値をあたえる点は $(1/\sqrt{2},\ -1/\sqrt{2})$ と $(-1/\sqrt{2},\ 1/\sqrt{2})$ となる．そしてそのとき $f=-1$ である．

[*4] この書き方はくどすぎる．計算過程をどこまで細かく書くか，省いて良い範囲はどの程度か，脚注化するのがよいのか，などなど，個人的な悩みは尽きない．なんでも「簡単な計算により」ですますわけにもいかないだろうし．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

最後にのこるのは，最大か最小かであるが，いまここでは少しズルをして，最大・最小の判定のために極座標の力を借りることにする．$x = \cos\theta, y = \sin\theta$ と極座標で表現すると，条件式は

$$g(x, y) = x^2 + y^2 - 1 = \cos^2\theta + \sin^2\theta - 1 = 0$$

となって，三角関数の性質そのものになってしまう．つまり極座標であらわせば，自動的に条件式が満たされていることになる．そのうえで f を考えると

$$f(x, y) = 2xy = 2\cos\theta\sin\theta = \sin 2\theta$$

となり，これもまた三角関数の性質から $-1 \leq f \leq 1$ となって，先に求めた極値が最大値と最小値になっていることもわかる[*5]．

さてさてもうひとつ．極値を与える点が，極大値か極小値をあたえるものであって（つまり，局所的に最大とか最小とか），最大や最小にならない場合もあるのではないか，という疑問もこれまたごもっともである．そしてそこには，かのワイエルシュトラス（Weierstraß）[*6] の

> $g(x, y) = 0$ を与える点の集合が有界閉集合であって，かつ，$f(x, y)$ が連続である．
> $\Downarrow$
> $f(x, y)$ は $g(x, y) = 0$ という条件のもとで最大値と最小値を持つ．

という定理がある．なので，極値のどれかは最大値／最小値となっているはずである．少し安心していい．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

このようにして，ひとまず条件付き極値問題の具体的な解法を得たわけであるけれど，これを汎用化したい．それが，次節から述べる，かの有名なラグランジュの未定乗数法をである．とはいっても，内容の骨格（真髄？）はこの節で述べたことそのままである．

2　ラグランジュの未定乗数法の定式化 —— ベクトル編

n 変数関数 $f(x_1, x_2, \ldots, x_n)$ が，拘束条件 $g(x_1, x_2, \ldots, x_n) = 0$ のもとで極値を取るときの点 $(x_1, x_2, \ldots, x_n)$ を求める方法を導出する．なお，表記を簡便にするために，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ のように n 個の変数の組をベクトルであらわすこともある[*7]．したがって，関数の表記においても $f(\boldsymbol{x})$ のように使うこともある[*8]．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

最初に $x_1, x_2, \ldots, x_n$ という n 個の変数について，n 個の成分をもつベクトルを作り出す偏微分演算子 $\boldsymbol{\nabla}_n$ を

$$\boldsymbol{\nabla}_n = \overrightarrow{\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)}$$

と書くことにする．ベクトル解析でよく使われるナブラ演算子の自然な拡張である．右辺であらわしているものは，

[*5] だったら最初から極座標を使えばいいではないか，というのはごもっとも．あんまりいい例ではなかったのかもしれない．

[*6] Karl Theodor Wilhelm Weierstraß, 1815 年 10 月 31 日 – 1897 年 2 月 19 日．

[*7] 点の位置ベクトルであるという見立てに他ならない．

[*8] そんなことを言わずにひとつに統一しろ，というお叱りはごもっともである．この手の物言いは，ひとえに紙と鉛筆と手間の問題なのである，といってもいいように思うがどうだろうか．

ベクトルの成分表記である．n 変数関数 $F(\boldsymbol{x})$ すなわち $F(x_1, x_2, \ldots, x_n)$ に作用させれば

$$\nabla_n F = \overrightarrow{\left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n}\right)}$$

というベクトルが作られる．この形のベクトルは，通常「勾配ベクトル」と呼ばれる*9．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

さてこれから，極値の候補を見つけることになるが，拘束条件 $g(\boldsymbol{x}) = 0$ のもとで

$$\nabla_n g(\boldsymbol{x}) = \mathbf{0}$$

を満たす解 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ （通常これを特異点という）があるかどうかを確認しておく必要がある．なぜならば，この特異点では，$g(\boldsymbol{x})$ の勾配ベクトル（$\nabla_n g(\boldsymbol{x})$）がゼロベクトルになるので，解析的には扱いづらくなることが予想されるからである．なので，もし特異点があるのならば，その特異点が極値の候補になるかどうかを別途確認しておき，のちの解析的計算（偏微分など）では特異点のことを忘れてもいいような準備をしておくと安心になる．それゆえ，先に特異点を持つかどうかの吟味をしておくことが肝要になる．

とは言ってもそれはその時どきの問題の設定次第でもあるので，ここからの説明記述では，特異点の吟味をたなにあげて進めてゆくことにする*10．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ が極値であるとしよう．極値なのであるから，当然 $f(\boldsymbol{x})$ の全微分は 0 であるはずである．したがって

$$\begin{aligned} df(\boldsymbol{x}) &= \frac{\partial f(\boldsymbol{x})}{\partial x_1}dx_1 + \frac{\partial f(\boldsymbol{x})}{\partial x_2}dx_2 + \cdots + \frac{\partial f(\boldsymbol{x})}{\partial x_n}dx_n \\ &= \overrightarrow{\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \frac{\partial f(\boldsymbol{x})}{\partial x_2}, \ldots, \frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)} \bullet \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} \\ &= \nabla_n f(\boldsymbol{x}) \bullet \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} = 0 \end{aligned} \tag{2.1}$$

と，ベクトルの内積であらわせる（前節と同様に $\bullet$ はベクトルの内積演算である）．一方で拘束条件の式 $g(\boldsymbol{x}) = 0$ の全微分をとるとやはり

$$\begin{aligned} dg(\boldsymbol{x}) &= \frac{\partial g(\boldsymbol{x})}{\partial x_1}dx_1 + \frac{\partial g(\boldsymbol{x})}{\partial x_2}dx_2 + \cdots + \frac{\partial g(\boldsymbol{x})}{\partial x_n}dx_n \\ &= \overrightarrow{\left(\frac{\partial g(\boldsymbol{x})}{\partial x_1}, \frac{\partial g(\boldsymbol{x})}{\partial x_2}, \ldots, \frac{\partial g(\boldsymbol{x})}{\partial x_n}\right)} \bullet \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} \\ &= \nabla_n g(\boldsymbol{x}) \bullet \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} = 0 \end{aligned} \tag{2.2}$$

となる．つまり，拘束条件を満たす点は，すべてこの (2.2) を満たさねばならないのである．その上で (2.1) と比較すると，おのおの内積が 0 なのであるから

$$\left.\begin{array}{l} \nabla_n f(\boldsymbol{x}) \text{ と } \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} \text{ は直交している．} \\ \nabla_n g(\boldsymbol{x}) \text{ と } \overrightarrow{(dx_1, dx_2, \ldots, dx_n)} \text{ は直交している．} \end{array}\right\} \Longrightarrow \nabla_n f(\boldsymbol{x}) \text{ と } \nabla_n g(\boldsymbol{x}) \text{ は平行}^{*11}\text{である}$$

*9 ついでに，もうすこしベクトル味をだしてみよう．x_i 方向の単位ベクトルを $\hat{\boldsymbol{e}}_i$ とし，これらの内積は正規直行関係 $\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j = \delta_{ij}$ であるとすれば，

$$\nabla_n = \hat{\boldsymbol{e}}_1\frac{\partial}{\partial x_1} + \hat{\boldsymbol{e}}_2\frac{\partial}{\partial x_2} + \cdots + \hat{\boldsymbol{e}}_n\frac{\partial}{\partial x_n} \quad \text{であり} \quad \nabla_n F = \hat{\boldsymbol{e}}_1\frac{\partial F}{\partial x_1} + \hat{\boldsymbol{e}}_2\frac{\partial F}{\partial x_2} + \cdots + \hat{\boldsymbol{e}}_n\frac{\partial F}{\partial x_n}$$

と書ける．ここでは利用しないけれども，この表式は代数的操作を行う場合にいろいろと都合がいい．

*10 さらにもう少し言えば，特異点の吟味の他にも，そもそもとして $f(\boldsymbol{x})$，$g(\boldsymbol{x})$ がともに C^1-級の関数であることも必要になる．ただ，筆者は多変数関数の C^1-級がよくわかってはいない．たとえば 2 変数関数 $f(x, y) = xy + T$（T は定数）は C^1-級だと思うのだが，実は自信がない．偏微分可能だし，偏導関数は連続であるので問題ないと思うのだが．

といえることになる．平行であるのだから，λ 倍と書ける（いままでのところ，実数以外の数の出番はないので，λ は実数であるとする）．すなわち 極値を与える $\boldsymbol{x}$ においては

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \text{ が極値を与える} \implies \boldsymbol{\nabla}_n f(\boldsymbol{x}) = \lambda \boldsymbol{\nabla}_n g(\boldsymbol{x}) \tag{2.3}$$

という関係がなりたつ．この右辺の必要条件を噛み砕けば

$$\overrightarrow{\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \frac{\partial f(\boldsymbol{x})}{\partial x_2}, \ldots, \frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)} = \lambda \overrightarrow{\left(\frac{\partial g(\boldsymbol{x})}{\partial x_1}, \frac{\partial g(\boldsymbol{x})}{\partial x_2}, \ldots, \frac{\partial g(\boldsymbol{x})}{\partial x_n}\right)} \tag{2.4}$$

すなわち

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = \lambda \frac{\partial g(\boldsymbol{x})}{\partial x_i} \quad (i = 1, 2, \ldots, n) \tag{2.5}$$

となる．この偏微分の関係であらわされる n 個の連立方程式から λ を求めあげ，そのうえで拘束条件を当てはめて，極値となる点を求めれば良い[*12]．なお，(2.4)（もちろん (2.5) も）は必要条件であるから，これが成り立つ点は極値を与える点の候補であるだけで，必ずしも極値になっていない場合もあるということは忘れてはいけない[*13]．

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

さて残るは，これらの極値の候補が，はたして最大値や最小値になっているかどうか，である．そしてそれは解きたい問題個別に検討するしかない．もちろん，先のワイエルシュトラスの定理のもとでの検討になるのだが，あの定理は「大体の場合，最大値と最小値はあるよ」と言っているに等しいので，その追い風を受けて検討を進めれば良い．

3　ラグランジュの未定乗数法の定式化 —— 教科書編

通常の教科書での説明では，まずいきなり

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

という新たな関数 L が与えられて，さらにそのときには

$$\frac{\partial L}{\partial x_i} = 0 \quad (i = 1, 2, \ldots n)\ , \quad \frac{\partial L}{\partial \lambda} = 0$$

となるので，これを解けば良い，ということになっている．いや本当のところはもっときちんと丁寧に説明されているだろうが，「公式」としてはこのような物言いがされているのが実際のところではあるまいか．しかしこの「公式」の正当性（保証）はどっから来るのだろうか．

我々は，前節での方法論，つまり勾配ベクトルとベクトルの直交関係を使う方法論から (2.3) という結果が求まることを知っている．それゆえ (2.3) を若干変形することによって

$$\boldsymbol{\nabla}_n f(\boldsymbol{x}) = \lambda \boldsymbol{\nabla}_n g(\boldsymbol{x}) \iff \boldsymbol{\nabla}_n \left(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\right) = \boldsymbol{0} \tag{3.1}$$

がもとめられる．ここで λ も変数であるというひとつの飛躍を施す．そのために

$$\boldsymbol{\nabla}_{n,\lambda} = \overrightarrow{\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial \lambda}\right)}$$

[*11] n 次元のベクトルが「平行」である，ということを幾何的にイメージするのはなかなか難しい．本末を転倒させて，「平行」とは「実数倍であらわすことができること」と考えた方が，話は早い．

[*12] 1 節で使った例を確認しておこう．$\partial f/\partial x = 2y$，$\partial f/\partial y = 2x$ であり $\partial g/\partial x = 2x$，$\partial g/\partial y = 2y$ であるから

$$\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \quad \therefore\ 2y = \lambda 2x \quad \therefore\ y = \lambda x\ ,$$
$$\frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y} \quad \therefore\ 2x = \lambda 2y \quad \therefore\ x = \lambda y\ .$$

この結果を拘束条件 $g(x, y) = x^2 + y^2 - 1 = 0$ にあてはめれば $(\lambda y)^2 + (\lambda x)^2 - 1 = 0$ すなわち $\lambda^2 = 1$ であるから $\lambda = \pm 1$ となり同様の結果が導かれる（ま，そりゃそうだ）．

[*13] 求めた点にたいして，極値かどうかの個別の吟味が必要になるが，大体の場合は見ただけでわかるのではないだろうか．いやもしかして，本稿では「極値」に停留点を含めたので，特異点を除けば，これは必要条件ではなく必要十分条件なのかもしれないなぁ．

という記号を用意して，あらためて (3.1) を

$$\boldsymbol{\nabla}_{n,\lambda} L(\boldsymbol{x}, \lambda) = \boldsymbol{\nabla}_{n,\lambda} \left(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\right) = \boldsymbol{0}$$

と記せば，ここから自動的に

$$\frac{\partial L}{\partial x_i} = 0 \quad (i = 1, 2, \ldots n)\ , \qquad \frac{\partial L}{\partial \lambda} = 0$$

が導出されてくる．これが上の「公式」の正当性を保証する理路であると考えて良い．そして，L の形式から，$\partial L / \partial \lambda = 0$ は拘束条件そのものになっているのである[*14]．

* *

視点をすこし変えてみよう．拘束条件 $g(\boldsymbol{x}) = 0$ の全微分をとると

$$dg(\boldsymbol{x}) = \frac{\partial g(\boldsymbol{x})}{\partial x_1}dx_1 + \frac{\partial g(\boldsymbol{x})}{\partial x_2}dx_2 + \cdots + \frac{\partial g(\boldsymbol{x})}{\partial x_n}dx_n = \sum_{i=1}^{n} \frac{\partial g(\boldsymbol{x})}{\partial x_i}dx_i = 0 \tag{3.2}$$

であり，これは常に成り立っていることが前提されている．$f(\boldsymbol{x})$ の全微分は

$$df(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})}{\partial x_1}dx_1 + \frac{\partial f(\boldsymbol{x})}{\partial x_2}dx_2 + \cdots + \frac{\partial f(\boldsymbol{x})}{\partial x_n}dx_n = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x})}{\partial x_i}dx_i$$

である．そして，前節で見た勾配ベクトルどうしの関係から

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = \lambda \frac{\partial g(\boldsymbol{x})}{\partial x_i} \tag{3.3}$$

が得られるので，これを $f(\boldsymbol{x})$ の全微分にあてはめれば

$$df(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x})}{\partial x_i}dx_i = \sum_{i=1}^{n} \lambda \frac{\partial g(\boldsymbol{x})}{\partial x_i}dx_i = \lambda \sum_{i=1}^{n} \frac{\partial g(\boldsymbol{x})}{\partial x_i}dx_i$$

となり，(3.2) を用いれば

$$= \lambda\, dg(\boldsymbol{x}) = 0$$

となる．つまり，(3.2) すなわち拘束条件の全微分が 0 であり，(3.3) が成り立てば，$f(\boldsymbol{x})$ が極値をとることが示される．そしてこの (3.3) は

$$\begin{aligned} \frac{\partial f(\boldsymbol{x})}{\partial x_i} = \lambda \frac{\partial g(\boldsymbol{x})}{\partial x_i} \quad &\Longleftrightarrow \quad \frac{\partial f(\boldsymbol{x})}{\partial x_i} - \lambda \frac{\partial g(\boldsymbol{x})}{\partial x_i} = 0 \\ &\Longleftrightarrow \quad \frac{\partial}{\partial x_i}\Big(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\Big) = 0 \end{aligned}$$

と変形できるので，$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$ とすると

$$\frac{\partial L(\boldsymbol{x}, \lambda)}{\partial x_i} = \frac{\partial}{\partial x_i}\Big(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\Big) = 0$$

となり，この節の冒頭で述べた「公式」が導かれるのである (さらにシステマチックな「公式」にするために，λ をも変数と見立て，拘束条件を偏微分化したというカラクリなのであった)．さらにもう一つ進めて，$L(\boldsymbol{x}, \lambda)$ の全微分を計算すれば

$$\begin{aligned} dL(\boldsymbol{x}, \lambda) &= \frac{\partial L(\boldsymbol{x})}{\partial x_1}dx_1 + \frac{\partial L(\boldsymbol{x})}{\partial x_2}dx_2 + \cdots + \frac{\partial L(\boldsymbol{x})}{\partial x_n}dx_n + \frac{\partial L(\boldsymbol{x})}{\partial \lambda}d\lambda \\ &= \sum_{i=1}^{n} \frac{\partial L(\boldsymbol{x})}{\partial x_i}dx_i + \frac{\partial L(\boldsymbol{x})}{\partial \lambda}d\lambda = 0 \end{aligned}$$

[*14] ここでも 1 節で使った例を確認しておこう．$L = f(x, y) - \lambda g(x, y) = 2xy - \lambda(x^2 + y^2 - 1)$ とすれば

$$\frac{\partial L}{\partial x} = 2y - \lambda 2x\ , \quad \frac{\partial L}{\partial y} = 2x - \lambda 2y\ , \quad \frac{\partial L}{\partial \lambda} = -(x^2 + y^2 - 1)$$

であり，これらがすべて 0 であるということから同様の結果が導かれる．

となる．すなわち，拘束条件 $g(\boldsymbol{x})$ のもとで $f(\boldsymbol{x})$ が極値をとるのならば，$L(\boldsymbol{x},\lambda)=f(\boldsymbol{x})-\lambda g(\boldsymbol{x})$ の全微分も 0 となるのである．

4　拘束条件が複数ある場合

いままでの議論は，拘束条件が複数あった場合にもそのまま適用できる（もちろん各々は異なる拘束条件であるとする）．ただし，独立変数の個数 n 以上の拘束条件がある場合には，単なる連立方程式になり，その連立方程式を解くことによって解がもとめられるだろう．したがって，$m<n$ という独立変数の個数より少ない m 個の拘束条件を考えることにする．それらを $g_j(\boldsymbol{x})=0$ $(j=1,2,\ldots,m)$ とあらわし，この拘束条件各々に対して全微分をとれば，

$$dg_j(\boldsymbol{x})=\sum_{i=1}^{n}\frac{\partial g_j(\boldsymbol{x})}{\partial x_i}dx_i=\boldsymbol{\nabla}_n g_j(\boldsymbol{x})\bullet\overrightarrow{(dx_1,dx_2,\ldots,dx_n)}=0 \tag{4.1}$$

である．したがって，考えているすべての j の場合について

$$\left.\begin{array}{l}\boldsymbol{\nabla}_n f(\boldsymbol{x})\text{ と }\overrightarrow{(dx_1,dx_2,\ldots,dx_n)}\text{ は直交している．}\\ \boldsymbol{\nabla}_n g_j(\boldsymbol{x})\text{ と }\overrightarrow{(dx_1,dx_2,\ldots,dx_n)}\text{ は直交している．}\end{array}\right\}\Longrightarrow\boldsymbol{\nabla}_n f(\boldsymbol{x})\text{ と }\boldsymbol{\nabla}_n g_j(\boldsymbol{x})\text{ は平行である}$$

ということになる．平行であることから，各係数を η_j とすれば

$$\begin{aligned}\boldsymbol{\nabla}_n f(\boldsymbol{x})&=\eta_1\boldsymbol{\nabla}_n g_1(\boldsymbol{x})\\ \boldsymbol{\nabla}_n f(\boldsymbol{x})&=\eta_2\boldsymbol{\nabla}_n g_2(\boldsymbol{x})\\ &\cdots\cdots\\ \boldsymbol{\nabla}_n f(\boldsymbol{x})&=\eta_j\boldsymbol{\nabla}_n g_j(\boldsymbol{x})\\ &\cdots\cdots\\ \boldsymbol{\nabla}_n f(\boldsymbol{x})&=\eta_m\boldsymbol{\nabla}_n g_m(\boldsymbol{x})\end{aligned}$$

とあらわせて，辺々加えたのちに両辺を m でわって，$\lambda_j=\eta_j/m$ とすることにより

$$\boldsymbol{\nabla}_n f(\boldsymbol{x})=\lambda_1\boldsymbol{\nabla}_n g_1(\boldsymbol{x})+\lambda_2\boldsymbol{\nabla}_n g_2(\boldsymbol{x})+\cdots+\lambda_m\boldsymbol{\nabla}_n g_m(\boldsymbol{x})=\sum_{j=1}^{m}\lambda_j\boldsymbol{\nabla}_n g_j(\boldsymbol{x}) \tag{4.2}$$

すなわち

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i}=\sum_{j=1}^{m}\lambda_j\frac{\partial g_j(\boldsymbol{x})}{\partial x_i} \tag{4.3}$$

ということになる．この結果と j 個の拘束条件を組み合わせれば，λ_j を求めることができ，それを使って極値をあたえる候補の点がもとまる．

ここで

$$L(\boldsymbol{x},\lambda_1,\lambda_2,\ldots,\lambda_m)=f(\boldsymbol{x})-\sum_{j=1}^{m}\lambda_j g_j(\boldsymbol{x})$$

とし，かつ，λ_j すべてを変数とみなしすことにすれば，(4.2) つまり (4.3) から

$$\frac{\partial L}{\partial x_i}=0\quad(i=1,2,\ldots,n)$$

であり，拘束条件すべてが

$$\frac{\partial L}{\partial \lambda_j}=0\quad(j=1,2,\ldots,m)$$

とあらわせて，いわゆる「公式」をを得ることができる．

ちなみに，このときの $f(\boldsymbol{x})$ の全微分を計算すると（途中で和の順番を入れ替え，最後に (4.1) を用いれば）

$$
\begin{aligned}
df(\boldsymbol{x}) &= \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x})}{\partial x_i} dx_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\boldsymbol{x})}{\partial x_i} \right) dx_i = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j \frac{\partial g_j(\boldsymbol{x})}{\partial x_i} dx_i \\
&= \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} \frac{\partial g_j(\boldsymbol{x})}{\partial x_i} dx_i = \sum_{j=1}^{m} \lambda_j dg_j(\boldsymbol{x}) = 0
\end{aligned}
$$

となり，極値をとっていることが確認できる．

* *

ひとつ例題をやっておく．

> $f(x, y, z) = x + y + z$ は，$g_1(x, y, z) = x^2 + y^2 - 1 = 0$ と $g_2(x, y, z) = y^2 + z^2 - 1 = 0$ という条件のもとではどのような極値を取るか？

まず勾配ベクトルを計算すると

$$
\begin{aligned}
\boldsymbol{\nabla} f &= \overrightarrow{(1, 1, 1)} \\
\boldsymbol{\nabla} g_1 &= \overrightarrow{(2x, 2y, 0)} \\
\boldsymbol{\nabla} g_2 &= \overrightarrow{(0, 2y, 2z)}
\end{aligned}
$$

であるから，f が極値をとる場合には

$$
\left.
\begin{aligned}
\frac{\partial f}{\partial x} &= \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x} \qquad \therefore\ 1 = \lambda_1 2x \\
\frac{\partial f}{\partial y} &= \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y} \qquad \therefore\ 1 = \lambda_1 2y + \lambda_2 2y \\
\frac{\partial f}{\partial z} &= \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z} \qquad \therefore\ 1 = \lambda_2 2z
\end{aligned}
\right\} \tag{4.4}
$$

公式をつかう方法だと

$$
\begin{aligned}
L(x, y, z, \lambda) &= f(x, y, z) - \lambda_1 g_1(x, y, z) - \lambda_2 g_2(x, y, z) \\
&= x + y + z - \lambda_1(x^2 + y^2 - 1) - \lambda_2(y^2 + z^2 - 1)
\end{aligned}
$$

であるから

$$
\left.
\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \qquad \therefore\ 1 - \lambda_1 2x = 0 \\
\frac{\partial L}{\partial y} &= 0 \qquad \therefore\ 1 - \lambda_1 2y - \lambda_2 2y = 0 \\
\frac{\partial L}{\partial z} &= 0 \qquad \therefore\ 1 - \lambda_2 2z = 0
\end{aligned}
\right\} \tag{4.5}
$$

となって，当然のごとく両者の結果 (4.4) と (4.5) は同じ結果になっている．また λ_1, λ_2 による偏微分の条件から

$$
\left.
\begin{aligned}
\frac{\partial L}{\partial \lambda_1} &= 0 \qquad \therefore\ x^2 + y^2 - 1 = 0 \\
\frac{\partial L}{\partial \lambda_2} &= 0 \qquad \therefore\ y^2 + z^2 - 1 = 0
\end{aligned}
\right\}
$$

となって，これは拘束条件の式に他ならないことも見て取れる．あとはこの連立方程式を解くだけである．まず，x, y, z を λ_1, λ_2 であらわせば，

$$x = \frac{1}{2\lambda_1}\,, \quad y = \frac{1}{2(\lambda_1 + \lambda_2)}\,, \quad z = \frac{1}{2\lambda_2}$$

である．そして 2 つの条件式を引き算すると $x^2 - z^2 = 0 \iff (x - z)(x + z) = 0$ であるから

$$\left(\frac{1}{2\lambda_1} - \frac{1}{2\lambda_2}\right)\left(\frac{1}{2\lambda_1} + \frac{1}{2\lambda_2}\right) = \frac{\lambda_2 - \lambda_1}{2\lambda_1\lambda_2}\,\frac{\lambda_2 + \lambda_1}{2\lambda_1\lambda_2} = 0$$

つまり，$\lambda_1 = \lambda_2$ かまたは $\lambda_1 = -\lambda_2$ でなければならないことになる．しかし，$\lambda_1 = -\lambda_2$ のときは y の分母が 0 になってしまい不適切であるから[*15] $\lambda_1 = \lambda_2$ でなければならないことになる．それゆえ $y = 1/(4\lambda_1)$ がえられ，これを拘束条件の式にいれると

$$x^2 + y^2 - 1 = \left(\frac{1}{2\lambda_1}\right)^2 + \left(\frac{1}{4\lambda_1}\right)^2 - 1 = 0 \quad \therefore\ \lambda_1 = \pm\frac{\sqrt{5}}{4}$$

ともとまる．したがって，極値をあたえる点は

$$(x, y, z) = \left(\pm\frac{2}{\sqrt{5}}, \pm\frac{1}{\sqrt{5}}, \pm\frac{2}{\sqrt{5}}\right) \quad \text{（復号同順）}$$

であり，極値（最大値）は $\sqrt{5}$，極値（最小値）は $-\sqrt{5}$ と求まったのである[*16]．

5　未定乗数 λ の意味

未定乗数法でもとめた未定乗数にはどのような意味があるのか，を考えてみる．この考察には，3 節で見た「教科書編」での定式化が便利であるので，それを使う[*17]．

拘束条件 $g(\boldsymbol{x}) = 0$ のもとでの $f(\boldsymbol{x})$ の極値は

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

と関数 L を定義して，$dL = 0$ から

$$\frac{\partial L}{\partial x_i} = 0 \quad (i = 1, 2, \ldots n)\,, \quad \frac{\partial L}{\partial \lambda} = 0$$

という偏導関数の条件を出し，それらを計算することによって求まるのであった．ここで拘束条件にひとつの任意性を付与すること，つまり

$$g(\boldsymbol{x}) = h(\boldsymbol{x}) - q = 0$$

というように，q という変量を含んだ形での拘束条件とする場合を考える．$h(\boldsymbol{x})$ という関数が，拘束条件として，様々な値 q を取ると考えれば良い．ただし，この変量 q は少し特殊で，未定乗数法によって極値を求める行為を行うときには定数である，とみなすことにする．以上から，L をあらためて書き下すと

$$L(\boldsymbol{x}, \lambda, q) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda(h(\boldsymbol{x}) - q)$$

であり，ここで λ を変数，q も変数とみなして全微分をとると

$$dL(\boldsymbol{x}, \lambda, q) = \sum_{i=1}^{n} \frac{\partial L}{\partial x_i} dx_i + \frac{\partial L}{\partial \lambda} d\lambda + \frac{\partial L}{\partial q} dq$$

[*15] どこから，こんな計算結果が生じてきたのか？特異点の条件にもなっていないし…

[*16] さて，最大値（最小値）であることは如何様に示されるだろうか？

[*17] 読み進めていけばわかることだが，今まで定数であるとしていたものを変数として扱うことにするので，「ベクトル編」の定式化では厳しい．

となる．極値をとる場合には $\partial L/\partial x_i = 0$, $\partial L/\partial\lambda = 0$ となるのであった（極値をもとめるときには q は定数とみなす）．そしてこの偏導関数の満たすべき条件から求めた未定乗数を μ ，極値をあたえる点を $\boldsymbol{u}$ とすれば

$$L(\boldsymbol{u}, \mu, q) = f(\boldsymbol{u}) - \mu(h(\boldsymbol{u}) - q)$$

となり，極値でも当然拘束条件はなりたつから $h(\boldsymbol{u}) - q = 0$．したがって

$$L(\boldsymbol{u}, \mu, q) = f(\boldsymbol{u})\ .$$

いっぽうで，

$$dL(\boldsymbol{x}, \lambda, q)\Big|_{\boldsymbol{x}=\boldsymbol{u}, \lambda=\mu} = \frac{\partial L}{\partial q}\Big|_{\boldsymbol{x}=\boldsymbol{u}, \lambda=\mu} dq \iff dL(\boldsymbol{u}, \mu, q) = \mu dq \iff df(\boldsymbol{u}) = \mu dq$$

ということがいえる．これは，拘束条件を $h(\boldsymbol{x}) = q$ としたときに $f(\boldsymbol{u})$ を q で微分すると，それは $f(\boldsymbol{x})$ の極値を与えるときの未定乗数 λ の値になる，ということを示しているつまり，f の極値は q の関数でもあって，それの q による変化率が未定乗数 μ なのだ，ということを示しているのである．

* *

例題を解いて確認してみよう[*18]．まず次のように，拘束条件に変量 q が入らない素直な場合を考える．

素直な場合１

$h(x, y) = xy = 36$ という拘束条件があるとき $f(x, y) = 2x + 3y$ の極値を与える点 (x, y) はどこか．またそのときの f の値はいくつか．

$g(x, y) = h(x, y) - 36 = 0$ と拘束条件を書き換えた上で，未定乗数法を用いる関数を

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y) = f(x, y) - \lambda(h(x, y) - 36) = 2x + 3y - \lambda(xy - 36)$$

とすれば，極値を与えるときの λ と点 (x, y) は

$$\begin{cases} \dfrac{\partial L}{\partial x} = 2 - \lambda y = 0 \\ \dfrac{\partial L}{\partial y} = 3 - \lambda x = 0 \\ \dfrac{\partial L}{\partial \lambda} = xy - 36 = 0 \end{cases}$$

を解くことによって得られる．実際に計算をすれば，$x = 3/\lambda$, $y = 2/\lambda$ であるから

$$xy = \frac{6}{\lambda^2} = 36 \quad \therefore\ \lambda = \pm\sqrt{\frac{1}{6}}$$

ともとまる．λ の解をそれぞれ $\mu_+ = \sqrt{1/6}$, $\mu_- = -\sqrt{1/6}$ と書くことにしよう．μ_+, μ_- に応じて極値を与える点は２つあって，それらを (u_+, v_+), (u_-, v_-) とすれば

$$\begin{aligned} (u_+, v_+) &= (3\sqrt{6},\ 2\sqrt{6}) \quad &&\text{そのとき} \quad f(u_+, v_+) = 12\sqrt{6}\ , \\ (u_-, v_-) &= (-3\sqrt{6},\ -2\sqrt{6}) \quad &&\text{そのとき} \quad f(u_-, v_-) = -12\sqrt{6}\ . \end{aligned}$$

つぎに，拘束条件の定数部分が少し異なる場合を考えてみる．

素直な場合２

$h(x, y) = xy = 42$ という拘束条件があるとき $f(x, y) = 2x + 3y$ の極値を与える点 (x, y) はどこか．またそのときの f の値はいくつか．

[*18] 小島著『ゼロから学ぶ微分積分』 [1] の 150 ページから，経済学の流儀での λ の意味が説明されている．それを学習して書き起こしてみたものである．経済学での「シャドウプライス」というものらしいが，わたくしは「シャドウプライス」についてはよくわかっていない．

これは，上の解法で $\partial L/\partial\lambda = xy - 42 = 0$ が異なるだけだから

$$xy = \frac{6}{\lambda^2} = 42 \quad \therefore\ \lambda = \pm\sqrt{\frac{1}{7}}$$

つまり，λ の解は $\mu_+ = \sqrt{1/7}$, $\mu_- = -\sqrt{1/7}$ であり，極値を与える2つの点は

$$\begin{aligned}(u_+, v_+) &= (3\sqrt{7},\ 2\sqrt{7}) &\quad \text{そのとき}\quad & f(u_+, v_+) = 12\sqrt{7}\ ,\\ (u_-, v_-) &= (-3\sqrt{7},\ -2\sqrt{7}) &\quad \text{そのとき}\quad & f(u_-, v_-) = -12\sqrt{7}\ .\end{aligned}$$

拘束条件の定数部分を変量 q とした場合を考えてみよう．

変量 q を用いる場合

$h(x, y) = xy = q$ という拘束条件があるとき $f(x, y) = 2x + 3y$ の極値を与える点 (x, y) はどこか．またそのときの f の値はいくつか．

今までとの違いは，やはり $\partial L/\partial\lambda = xy - q = 0$ のところだけであるから

$$xy = \frac{6}{\lambda^2} = q \quad \therefore\ \lambda = \pm\sqrt{\frac{6}{q}}$$

で，λ の解は $\mu_+ = \sqrt{6/q}$, $\mu_- = -\sqrt{6/q}$ であり，極値を与える2つの点は

$$\begin{aligned}(u_+, v_+) &= \left(3\sqrt{\frac{q}{6}},\ 2\sqrt{\frac{q}{6}}\right) &\quad \text{そのとき}\quad & f(u_+, v_+) = 12\sqrt{\frac{q}{6}}\ ,\\ (u_-, v_-) &= \left(-3\sqrt{\frac{q}{6}},\ -2\sqrt{\frac{q}{6}}\right) &\quad \text{そのとき}\quad & f(u_-, v_-) = -12\sqrt{\frac{q}{6}}\end{aligned}$$

ともとまる．もちろん $q = 36$ のときは「素直な場合１」の結果をあたえるし，$q = 42$ のときは「素直な場合２」の結果を導いている．

この結果をみると，f の極値は q 次第で変わることがわかる．つまり，q の関数であると考えられるのだ．それゆえ，

$$\frac{d}{dq}f(u_+, v_+) = \frac{d}{dq}\left(12\sqrt{\frac{q}{6}}\right) = \sqrt{\frac{6}{q}} = \mu_+$$

となっている．同様にして

$$\frac{d}{dq}f(u_-, v_-) = \frac{d}{dq}\left(-12\sqrt{\frac{q}{6}}\right) = -\sqrt{\frac{6}{q}} = \mu_-$$

が得られる．つまり，拘束条件 $h(x, y) = q$ の q が変化する時の極値 $f(u_\pm, v_\pm)$ の変化率 df/dq は，拘束条件のもとで極値をもとめた未定乗数 $\mu_\pm$ に他ならないのである．

このことがらは，次のように解釈することができる．拘束条件 $h(x, y) = q$ で安定状態にある時，$f(x, y)$ の極値（最大値でも最小値でもよいだろう）は，未定乗数法をつかってもとめた $\lambda = \mu$ のときの値である．これはいままで実際に計算してきた事柄である．そこで，その安定状態を若干撹乱し $h(x, y) = q'$ に変化させようと試みた場合，つまり $q \to q'$ と変化させた場合では，極値 f の変化の割合（変化率）は，q のときに求めておいた未定乗数 μ に他ならないということなのである[*19]．

[*19] 往々にして，この事柄は（とくに物理の方面では）

$$q \to q + \Delta q \quad \text{のとき} \quad \Delta f = f(q + \Delta q) - f(q) \approx \mu$$

などと書かれるが，場合によってはかなり思い切って

$$q \to q + dq \quad \text{のとき} \quad df = f(q + dq) - f(q) \approx \mu \quad \text{すなわち} \quad \frac{df}{dq} = \mu$$

と書かれる傾向がある．

付録 A　5 節の例題再考

5 節の例題は $(x, y) = (0, 0)$ という点での扱いが特殊なので，その形態を簡単に見ておく．また，極値と最大値最小値がどのような関係になっているのかも確認しておく．

再び 5 で見た次の問題を考える．

> $h(x, y) = xy = 36$ という拘束条件があるとき $f(x, y) = 2x + 3y$ の極値を与える点 (x, y) はどこか．またそのときの f の値はいくつか．

未定乗数法の根幹は，ベクトルでの定式化（2 節）でみたように

$$\boldsymbol{\nabla}_n f(\boldsymbol{x}) = \lambda \boldsymbol{\nabla}_n g(\boldsymbol{x}) \tag{付録 A.1}$$

という「ベクトルの平行関係」にあった．そしてこの計算には $\boldsymbol{\nabla}_n g(\boldsymbol{x})$ が使われるので

$$\boldsymbol{\nabla}_n g(\boldsymbol{x}) = \mathbf{0}$$

となるようなときには，(付録 A.1) の計算を進めることができなくなる．今の場合，拘束条件は $g(x, y) = xy - 36 = 0$ であるから

$$\boldsymbol{\nabla}_n g(\boldsymbol{x}) = \overrightarrow{(y, x)}$$

であり，$(x, y) = (0, 0)$ のときには $\boldsymbol{\nabla}_n g(\boldsymbol{x}) = \mathbf{0}$ となる．しかしながら，この点では $g(0, 0) = -36$ となって拘束条件が満たされないことがわかる．なので，2 節での定義に則れば，この点は特異点とは言えないことになる．それゆえ，極値かどうかの吟味をするまでもないのかもしれないが，なんとはなしに次のようね論理をひねり出してみた．これは極値でないことの証左になっているだろうか？こういう理路はありなのだろうか？

> $f(0, 0) = 0$ は明らかであるが $df = 2dx + 3dy$ であるから
>
> $$df\Big|_{(0,0)} = (2dx + 3dy)\Big|_{(0,0)} = 2dx + 3dy$$
>
> となるので $df = 0$ とはなりえない．したがって特異点 $(x, y) = (0, 0)$ は極値とはなりえないことがわかる．

＊ ＊

次に最大最小の考察である．特異点でない場合には

$$\mu_+ = \sqrt{1/6}, \quad \mu_- = -\sqrt{1/6}$$

という 2 つの未定乗数がもとまった．そして μ_+, μ_- に応じて極値を与える点が定まって，それらを (u_+, v_+), (u_-, v_-) とすれば

$$\begin{aligned}
(u_+, v_+) &= (3\sqrt{6},\ 2\sqrt{6}) \qquad \text{そのとき} \quad f(u_+, v_+) = 12\sqrt{6}\ , \\
(u_-, v_-) &= (-3\sqrt{6},\ -2\sqrt{6}) \quad \text{そのとき} \quad f(u_-, v_-) = -12\sqrt{6}\ .
\end{aligned}$$

となるのであった．

一方で $xy = 36$ という拘束条件から $y = 36/x$ となるので，$f(x, y)$ を 1 変数関数として

$$k(x) = f(x, 36/x) = 2x + \frac{108}{x}$$

と見立てると，導関数は

$$k'(x) = 2 - \frac{108}{x^2} = \frac{2}{x^2}(x - 3\sqrt{6})(x + 3\sqrt{6})$$

となる（この導関数のグラフイメージを図 1 に示す）．$k'(x)$ の正負の状況を反映して増減表を書いてみると

x	$\cdots$	$-3\sqrt{6}$	$\cdots$	0	$\cdots$	$3\sqrt{6}$	$\cdots$
$k'(x)$	$+$	0	$-$	$(undef)$	$-$	0	$+$
$k(x)$	$\nearrow$		$\searrow$	$(undef)$	$\searrow$		$\nearrow$

となり，そして

$$k(-3\sqrt{6}) = -12\sqrt{6}\ ,\quad k(3\sqrt{6}) = 12\sqrt{6}$$

であるから，$x > 0$ のときは極値であるところが最小値，$x < 0$ のときは極値が最大値となっていることが諒解できる（$k(x)$ のグラフイメージは図 2 ）．

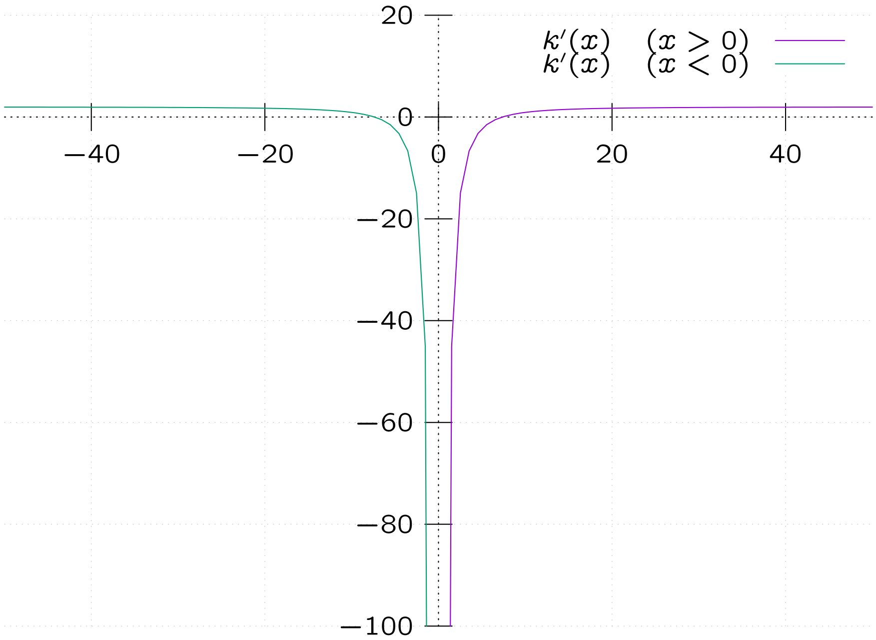
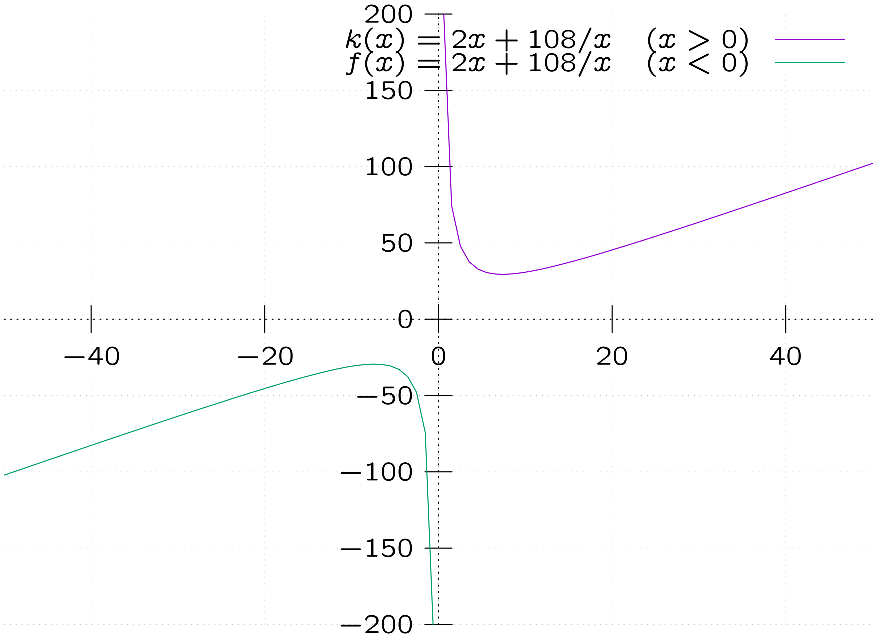


図 1: $k'(x)$ のグラフ



200
150
100
50
0
−50
−100
−150
−200
−40
−20
0
20
40
$k(x) = 2x + 108/x \quad (x > 0)$
$f(x) = 2x + 108/x \quad (x < 0)$

図 2: $k(x)$ のグラフ

参考文献

[1] 小島 寛之.『ゼロから学ぶ微分積分』. 講談社, 2001.（第 16 刷（2014）を参照した）.

目次