第4章　最適化問題

4.1 最適化と経済学

経済理論では，多くの理論が最適問題を基礎につくられています．

例 4. ある企業は y 単位生産物を作るのに，$c(y)$ の費用がかかるとします．生産物の完全競争市場における価格が p のとき，この企業は利潤 $py - c(y)$ を最大とするように，生産量 y を定めます．この生産量は p に依存します．この依存関係を供給関数とよびます．

例 5. 資金を 1 単位資産もつ投資主体が，その資金から確率 *1* で収益 $1 + r$ 単位得られる安全資産と，確率 1/2 で収益 $1 + r + \varepsilon$ 単位, 確率 1/2 で収益 $1 + r - \varepsilon$ 単位を得る危険資産に $(1 - x) : x$ の比率で分散投資することを考えているとします．この投資主体の危険に対する態度は期待効用であらわされるとして，効用関数を $U(y)$ がわかっているとします．この問題では，x は

$$\frac{1}{2}U((1-x)(1+r)+x(1+r+\varepsilon))+\frac{1}{2}U((1-x)(1+r)+x(1+r-\varepsilon))$$

を最大化することによりえられます．この x は U の形状と関係をもちます．

例 6. n 種の各資源総量が定まっている場合，m 人の経済主体がいる場合のパレート最適資源配分は，適当に定めた社会的厚生関数を，各資源総量を越えないという制約の下で，m 人に配分する資源配分ベクトルを見出すことで求められます．

4.2 最適化問題とは

この本で扱う最適化問題を説明しておきます．X を $\mathbf{R}^n$ の部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ を考えます．このとき

$$(\forall \boldsymbol{x} \in X) f(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$$

をみたす $\boldsymbol{x}^*$ を最大化問題の (大域的) 最適解,

$$(\forall \boldsymbol{x} \in X) f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$$

をみたす $\boldsymbol{x}^*$ を最小化問題の (大域的) 最適解とよびます．

次に (局所的) 最適解を定義します．いま

$$\{\boldsymbol{y} \in \mathbf{R}^n ; ||\boldsymbol{x} - \boldsymbol{y}|| < \varepsilon\}$$

を $B_\varepsilon(\boldsymbol{x})$ と記すことにして，$\boldsymbol{y}$ の ε 近傍をあらわします．この集合は開集合であることに注意してください．($n = 1$ の場合，$B_\varepsilon(x)$ は開区間 $(x - \varepsilon, x + \varepsilon)$ にほかなりません．

$$(\exists \varepsilon > 0)(\forall \boldsymbol{x} \in X \cap B_\varepsilon(\boldsymbol{x}^*)) f(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$$

をみたす $\boldsymbol{x}^*$ を最大化問題の (局所的) 最適解，$f(\boldsymbol{x}^*)$ を極大値,

$$(\exists \varepsilon > 0)(\forall \boldsymbol{x} \in X \cap B_\varepsilon(\boldsymbol{x}^*)) f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$$

をみたす $\boldsymbol{x}^*$ を最小化問題の (局所的) 最適解，$f(x^)$ を極小値とよびます．

定義から大域的最適解は局所的最適解であるけれども，逆は成立しないことがわかります．

最後に，最適値をもたらす vex がたった一つしかないときに，強い意味の最適解とよぶことにします．強い意味の最適解では，上の定義式の等号は $\boldsymbol{x} = \boldsymbol{x}^*$ のときのみに成立します．

例 7. 関数 $f(x) = \frac{x^4}{4} - \frac{7x^3}{6} + \frac{7x^2}{4} - x$ を考えます．定義域を実数直線全体とするとき，$x = 1/2$ と $x = 2$ は最小化問題の局所的最適解，$x = 2$ は最小化問題の大域的最適解をもたらします．

なお，最大化問題の大域的最適解を最大解，最小化問題の大域的最適解を最小解，最大化問題の局所的最適解を極大解，最小化問題の局所的最適解を極小解，とよぶことがあります．

制約条件を考える場合の最適化問題では，最適化の対象となる関数の定義域 X にくわえて，集合 $W \subset \mathbf{R}^n$ を定めて，$W \cap X$ から最適解をさがすことになります．この本では，W を定めるのに，写像 $\boldsymbol{g} : \mathbf{R}^n \longrightarrow \mathbf{R}^\ell$ によって

$$W = \{\boldsymbol{x} \in \mathbf{R}^n ; \boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}\}$$

とあらわさられる場合に限ります．

制約つき最適問題の場合，上記の最適解の定義にある X を $X \cap W$ に変更するだけです．

最適化問題を扱うポイントをここで整理しておくと次のようになります．

存在 そもそも最適解が存在するか．

必要条件 最適解が存在するならば，最適解はどのような条件をみたすのか．

十分条件 最適解が存在するならば，どのような条件をみたすものが最適解になるのか．

計算アルゴリズム 最適解を具体的に計算する手続きはどのようなものか．

最初の存在問題は，ここではきちんと扱いませんが．最適化問題の与え方によっては解が存在しないことは珍しくありません．

例 8. 実数値関数 $f : (0,1) \to \mathbf{R}$ を，$f(x) = x$ とすると，f は定義域である開区間 $(0,1)$ において最大値も最小値もとりません．

例 9. 実数値関数 $f : \mathbf{R} \to \mathbf{R}$ を，$f(x) = \frac{1}{1+e^{-x}}$ とすると，$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} > 0$ ですから単調増加関数で，$\lim_{x\to\infty} f(x) = 1$，$\lim_{x\to-\infty} f(x) = 0$ となり定義域である実数直線上において最大値も最小値もとりません．また極値もありません．

例 10. 実数値関数 $f : \mathbf{R}\backslash\{0\} \to \mathbf{R}$ を，$f(x) = \frac{1}{x}$ とすると，定義域上において最大値も最小値もとりません．また極値もありません．

例 11. 実数値関数 $f : [-1,1] \to \mathbf{R}$ を，

$$f(x) = \begin{cases} 0 & x = 0 \text{ のとき} \\ -x^2 + 1 & x \neq 0 \text{ のとき} \end{cases}$$

とすると，定義域であるにおいて最小値 *0* をとりますが，最大値はありません．最大値がないのは，f が $x = 0$ において不連続だからです．

例 12. 制約条件 $y \geq \frac{1}{x}, x > 0$ の下で関数 $f(x,y) = -x$ を最大化するという最適化問題に解はない．(なぜか各自，考えよ．)

以上の例からわかるように，関数が連続，あるいは連続微分可能であっても，定義域や制約条件によっては，最適化問題の解は無条件に解の存在が保証されるわけではありません．最適化の解の存在について，高々言える事は次の定理くらいです．

定理 17. $X \subset \mathbf{R}^n$ がコンパクトであるとき，実数値関数 $f : X \to \mathbf{R}$ が連続であるならば，f は最大値と最小値をもつ．

4.3 制約なし最適化問題 (1 次元) の必要条件

一般の次元の問題を扱う前に，1 次元の最適化問題の最適のための必要条件を考えます．次の定理が大切です．

定理 18 (局所最適化の 1 階の必要条件)． 関数 $f:\mathbf{R}\longrightarrow\mathbf{R}$ が微分可能とする．開区間 (a,b) において，f が $x^*\in(a,b)$ で最大値または最小値をとるならば $f'(x^*)=0$

(証明) 最大値の場合のみを証明すれば十分です.（最小値の場合は不等号を適宜逆向きにすることで証明できます.）開区間 (a,b) において，f が $x^*\in(a,b)$ で最大値をとることから，$x^*+\varepsilon\in(a,b)$ となる任意の $\varepsilon>0$ に対して，$f(x^*+\varepsilon)-f(x^*)\le 0$ が成り立ちます．よって

$$\frac{f(x^*+\varepsilon)-f(x^*)}{\varepsilon}\le 0 \tag{4.1}$$

となります．同様に，$x^*+\varepsilon\in(a,b)$ となる任意の $\varepsilon<0$ に対して，$f(x^*+\varepsilon)-f(x^*)\le 0$ が成り立ちます．よって

$$\frac{f(x^*+\varepsilon)-f(x^*)}{\varepsilon}\ge 0 \tag{4.2}$$

となります．

微分の定義より，$\varepsilon>0$ のまま ε をゼロに近づけることにより，(4.1) から $f'(x^*)\le 0$ がえられます．同様に，微分の定義より，$\varepsilon<0$ のまま ε をゼロに近づけることにより，(4.2) から $f'(x^*)\ge 0$ がえられます．両方を考慮することにより，$f'(x^*)=0$ がえられます．(証明おわり）

定理 18 は，最大あるいは最小の近傍に関しての議論なので，最大値を極大値に，最小値を極小値に読み替えても，成立することに注意してください．

定理は，最適問題の解のみたす必要条件をあきらかにしています．この必要条件は，関数の 1 階の微分に関する条件なので，しばしば，1 階の必要条件とよばれます．

演習 21． $f(x)=ex-e^x$ の開区間 $(0,2)$ において最大化のための *1* 階の必要条件をみたす点をもとめなさい．

例 7 の関数をふたたびとりあげましょう．$f(x)=\frac{x^4}{4}-\frac{7x^3}{6}+\frac{7x^2}{4}-x$ は，$f'(1/2)=f'(1)=f'(2)=0$ となっています．また $f(1/2)=-37/192, f(1)=-1/6, f(2)=-1/3$ ですから，$x=1/2$ は最小化問題の解ではありません．また，$x=1$ は極小ですらないことが確認できます．

次の定理が成立します．

定理 19 (局所最適化の 2 階の必要条件)． 関数 $f:\mathbf{R}\longrightarrow\mathbf{R}$ が *2* 階連続微分可能とする．x^* が f の最大化問題の局所最適解をもたらすとき，$f''(x^*)\le 0$，x^* が f の最小化問題の局所的最適解をもたらすとき $f''(x^*)\ge 0$ となる．

(証明) ここでも，最大化の場合のみ証明することにします.（最小化の場合は，適宜符号を換えれば，同様に証明できます.）$f''(x^*)>0$ としてみます．このとき，任意の ε に対してテイラーの定理より

$$f(x^*+\varepsilon)=f(x^*)+f'(x^*)\varepsilon+\frac{1}{2}f''(x^*+\theta\varepsilon)\varepsilon^2$$

となる $\theta\in(0,1)$ が存在する．$f'(x^*)=0$ だから

$$f(x^*+\varepsilon)-f(x^*)=\frac{1}{2}f''(x^*+\theta\varepsilon)\varepsilon^2$$

f'' が連続関数であることから十分小さな絶対値の ε に対して $f''(x^*+\theta\varepsilon)>0$ となるから，すぐ上の左辺は正の値をとります．これは x^* が局所的な最適としたことに矛盾します．(証明おわり)

テイラーの定理を使うことで，強い意味の局所最適化の 2 階の必要条件も得られます．

定理 20 (強い意味の局所最適化の 2 階の必要条件)． 関数 $f:\mathbf{R}\longrightarrow\mathbf{R}$ が *2* 階連続微分可能とする．x^* が f の最大化問題の局所最適解をもたらすとき，$f''(x^*)<0$，x^* が f の最小化問題の局所的最適解をもたらすとき $f''(x^*)>0$ となる．

(証明) x^* が強い意味の局所最大なので，ある $\alpha>0$ が存在して，開区間 $(x^*-\alpha, x^*+\alpha)$ に属する任意の $x\neq x^*$ に対して $f(x)-f(x^*)\leq 0$ となります．よって $0<|\varepsilon|<\alpha$ ならば，$f(x^*+\varepsilon)-f(x^*)<0$ となります．一方，テイラーの定理により，$\theta\in(0,1)$ が存在して

$$f(x^*+\varepsilon)-f(x^*)=f'(x^*)\varepsilon+\frac{1}{2}f''(x^*+\theta\varepsilon)\varepsilon^2$$

定理 18 により $f'(x^*)=0$ ですから

$$f(x)-f(x^*)=\frac{1}{2}f''(x^*+\theta\varepsilon)\varepsilon^2$$

となります．$f''(x^*+\theta\varepsilon)<0$ に注意してください．$\varepsilon\neq 0$ のまま，ε をゼロに近づけても，この式の成立は変わりません．このとき $f''(x^*+\theta(x-x^*))$ は負のまま，$f''(x^*)$ に収束します．よって，2 階の導関数の連続性より，$f''(x^*)\leq 0$ をえます．(証明おわり)

この定理でえられた最適解のための必要条件は，関数の 2 階の微分に関する性質なので，2 階の必要条件とよばれることがあります．

4.4 制約なし最適化問題 (1 次元) の十分条件

前の節の定理 18 と定理 19 は，局所的な最適解の必要条件をあきらかにしています．しかし，十分条件ではありません．例えば $f(x)=x^3$ とおくと，$f'(0)=f''(0)=0$ ですから $x=0$ は定理の必要条件を満たしていますが，f の最大も最小，さらには極値すらもたらしません．

1 変数関数についての最適問題の局所的最適解の十分条件は次の定理によって与えられます．

定理 21 (局所最適化の 2 階の十分条件)． 関数 $f:\mathbf{R}\longrightarrow\mathbf{R}$ が *2* 階連続微分可能とする．$f'(x^*)=0$, かつ $f''(x^*)<0$ であるとき x^* は f の最大化問題の局所最適解となる．また $f'(x^*)=0$, かつ $f''(x^*)>0$ であるとき x^* は f の最小化問題の局所最適解となる．

(証明) ここでも最大化の場合のみを証明します．任意の $\varepsilon>0$ に対してテイラーの定理より

$$f(x^*+\varepsilon)=f(x^*)+\varepsilon f'(x^*)+\frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon)$$

となる $\theta\in(0,1)$ が存在します．ここで $f'(x^*)=0$ と 2 階の導関数の連続性より，十分小さな $\varepsilon>0$ に対して $f''(x^*+\theta\varepsilon)<0$ となりますから，この $\varepsilon>0$ に対して

$$f(x^*+\varepsilon)-f(x^*)=\frac{1}{2}\varepsilon^2 f''(x^*+\theta\varepsilon)<0$$

となります．これは，x^* が最適解であることを示します．(証明おわり)

さて，関数 f に対してある大域的条件を課すと，最適のための十分条件が得られます．

関数 $f: \mathbf{R} \to \mathbf{R}$ について

$$(\forall a \in \mathbf{R})(\forall b \in \mathbf{R})(\forall t \in [0,1])tf(a) + (1-t)f(b) \leq f(ta + (1-t)b)$$

が成立するとき凹関数とよびます．また,

$$(\forall a \in \mathbf{R})(\forall b \in \mathbf{R})(\forall t \in [0,1])tf(a) + (1-t)f(b) \geq f(ta + (1-t)b)$$

が成立するとき凸関数とよびます.

関数 $f: \mathbf{R} \to \mathbf{R}$ について

$$(\forall a \in \mathbf{R})(\forall b \in \mathbf{R})(\forall t \in (0,1))tf(a) + (1-t)f(b) < f(ta + (1-t)b)$$

が成立するとき狭義の凹関数とよびます．また,

$$(\forall a \in \mathbf{R})(\forall b \in \mathbf{R})(\forall t \in [0,1])tf(a) + (1-t)f(b) > f(ta + (1-t)b)$$

が成立するとき狭義の凸関数とよびます.

例 13. アフィン関数 $f(x) = \alpha + \beta x$ は，凹関数でもあり，凸関数でもあります．しかし，狭義とつの凸関数あるいは，狭義の凹関数ではありません.

例 14. 関数 $f(x) = \log x$ は凹関数です.

例 15. 関数 $f(x) = x(x-1)(x+1)$ は凹関数でも，凸関数でもありません.

大局的最適化のための十分条件は以下の定理を用いて得ることができます.

定理 22. $a \in \mathbf{R}$ とする．a において微分可能な関数 $f: \mathbf{R} \to \mathbf{R}$ が凹関数ならば

$$(\forall x \in \mathbf{R})f(x) - f(a) \leq f'(a)(x-a)$$

凸関数ならば,

$$(\forall x \in \mathbf{R})f(x) - f(a) \geq f'(a)(x-a)$$

(証明) 凹関数，凸関数の定義において，$a = y$ とおいたものを整理して,

$$f(x) - f(a) = \frac{f(a + t(x-a)) - f(a)}{t}$$

を得たのち，t をゼロに近づければよい.（証明おわり）

定理 23 (大局的な最適化の十分条件). $f: \mathbf{R} \to \mathbf{R}$ が微分可能な凹関数であるとき，$f'(x^*) = 0$ である $x^* \in \mathbf{R}$ は f の大局的な最大化問題の解となる．同様に，凸関数のとき x^* は大局的な最小化問題の解となる.

(証明)　定理 22 において $a = x^*$ とおくと，凹関数の場合

$$(\forall x \in \mathbf{R})f(x) - f(x^*) \leq 0$$

凸関数の場合

$$(\forall x \in \mathbf{R})f(x) - f(x^*) \geq 0$$

(証明おわり)

実は，定理 22 と関連して，次の定理が成立します.

定理 24. $f:\mathbf{R}\to\mathbf{R}$ が任意の点で微分可能とする．

$$(\forall x\in\mathbf{R})(\forall y\in\mathbf{R})f(x)-f(y)\le f'(y)(x-y)$$

ならば，f は凹関数.

$$(\forall x\in\mathbf{R})(\forall y\in\mathbf{R})f(x)-f(y)\ge f'(y)(x-y)$$

ならば f は凸関数.

(証明)

なお，ここまで f の凹性，凸性に関して，f の各点における微分可能性のみを仮定して議論してきましたが，f が 2 階連続微分可能とすると，2 階の導関数を用いて，f の凹性，凸性を特徴づけることができます.

定理 25. $f:\mathbf{R}\to\mathbf{R}$ が *2* 階連続微分可能とする．f が凹関数であることの必要十分条件は,

$$(\forall x\in\mathbf{R})f''(x)\le 0$$

凸関数であることの必要十分条件は,

$$(\forall x\in\mathbf{R})f''(x)\ge 0$$

4.5 制約なし最適化問題 (2 次元以上) の必要条件

この節では，1 次元の場合にならって，多変数関数の場合の最適のための必要条件を調べます. この節では

$$\nabla f(\boldsymbol{x})=\left(\frac{\partial f}{\partial x_1}(\boldsymbol{x})\cdots\frac{\partial f}{\partial x_n}(\boldsymbol{x})\right)$$

とします.

次の定理が，最適の必要条件を導く出発点になります.

定理 26. X を $\mathbf{R}^n$ の開部分集合として，関数 $f:X\longrightarrow\mathbf{R}$ が *1* 階連続微分可能とする．$\boldsymbol{x}\in X$, $\boldsymbol{h}\in\mathbf{R}^n$ とする.

$$\nabla f(\boldsymbol{x})\cdot\boldsymbol{h}>0\Longrightarrow(\exists\delta)(\forall t\in(0,\delta))f(\boldsymbol{x}+t\boldsymbol{h})-f(\boldsymbol{x})>0$$

(証明) $\tau>0$ を勝手にとります．このとき

$$g(t)=f(\boldsymbol{x}+t\boldsymbol{h}),(0\le t\le\tau)$$

として，$t=0,t=\tau$ に平均値の定理を適用すると $c\in(0,\tau)$ が存在して

$$g(\tau)-g(0)=\tau g'(c)$$

となります．$\theta = c/\tau < 1$ とおくと $g'(c) = \nabla f(\boldsymbol{x} + \theta\tau\boldsymbol{h}) \cdot \boldsymbol{h}$ に注意すると

$$f(\boldsymbol{x} + \tau\boldsymbol{h}) - f(\boldsymbol{x}) = \tau\nabla f(\boldsymbol{x} + \theta\tau\boldsymbol{h}) \cdot \boldsymbol{h}$$

をえます．∇f が連続であることから $\delta > 0$ が存在して任意の $t \in (0, \delta)$ に対して

$$\nabla f(\boldsymbol{x} + \theta t\boldsymbol{h}) \cdot \boldsymbol{h} > 0$$

が成立します．これにより，$f(\boldsymbol{x} + t\boldsymbol{h}) - f(\boldsymbol{x}) > 0$ がいえます.（証明おわり）

この定理の幾何学的な解釈をあたえておきましょう．$\nabla f(\boldsymbol{x})$ は点 $\boldsymbol{x} = (x_1 x_2 \cdots x_n)$ において最も大きく f の値を増加する方向ベクトルです．$\nabla f(\boldsymbol{x}) \cdot \boldsymbol{h} > 0$ は，ベクトル $\boldsymbol{h}$ がこの方向ベクトルと鋭角をなすことをあらわします．よって，定理は,

この定理を使うと，多変数の場合の局所的最適解の 1 階の必要条件が得られます.

定理 27 (局所最適化の 1 階の必要条件). X を $\mathbf{R}^n$ の開部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ が *1* 階連続微分可能とする．$\boldsymbol{x}^* \in X$ が f の局所的最適解 *(*最大もしくは最小*)* ならば $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$

（証明）最大化の場合を証明する.（最小化の場合も同様.）$\nabla f(\boldsymbol{x}^*) \neq \mathbf{0}$ とします．$\nabla f(\boldsymbol{x}^*) \cdot \nabla f(\boldsymbol{x}^*)$ ですから，定理 26 を用いると,

$$(\exists\delta > 0)(\forall t \in (0, \delta))f(\boldsymbol{x}^* + t\nabla f(\boldsymbol{x}^*)) > f(\boldsymbol{x}^*)$$

これは，$\boldsymbol{x}^*$ が局所最適解であることに矛盾する.（証明おわり）

一般の次元の場合も，定理 19 と同様の 2 階の必要条件に関する定理が成立します.

定理 28 (局所最適化の 2 階の必要条件). X を $\mathbf{R}^n$ の開部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ が *2* 階連続微分可能とする．$\boldsymbol{x}^* \in X$ が

1. f の局所的最大解ならば，$\nabla^2 f(\boldsymbol{x}^*)$ は負の半定符号

2. f の局所的最小解ならば，$\nabla^2 f(\boldsymbol{x}^*)$ は正の半定符号

ここで

$$\nabla^2 f(\boldsymbol{x}^*) = \left(\frac{\partial^2}{\partial x_i \partial x_h} f(\boldsymbol{x}^*)\right)_{ij}$$

(証明) ある $\boldsymbol{h} \in \mathbf{R}^n$ が存在して，${}^t\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^*)\boldsymbol{h} > 0$ とする．$||\boldsymbol{h}|| = 1$ とかんがえてよい.（なんとなれば $\boldsymbol{h}/||\boldsymbol{h}||$ をかんがえればよい.）

多次元のテイラーの定理より，$\theta \in (0, 1)$ が存在して

$$f(\boldsymbol{x}^* + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*) + \varepsilon\nabla f(\boldsymbol{x}^*)\boldsymbol{h} + \frac{1}{2}\varepsilon^{2t}\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^* + \theta\varepsilon\boldsymbol{h})\boldsymbol{h}$$

となるが，定理 27 より $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ なので

$$f(\boldsymbol{x}^* + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}^*) = \varepsilon\nabla f(\boldsymbol{x}^*)\boldsymbol{h} + \frac{1}{2}\varepsilon^{2t}\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^* + \theta\varepsilon\boldsymbol{h})\boldsymbol{h}$$

2 階導関数の連続性より十分小さな ε に対して ${}^t\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^* + \theta\varepsilon\boldsymbol{h})\boldsymbol{h} > 0$ となるしたがって，$f(\boldsymbol{x}^* + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}^*) > 0$ となり $\boldsymbol{x}^*$ は局所最適解にはならない.（証明おわり）

1 次元の場合の定理 20 に対応する定理も成立します.

定理 29 (局所最適化の 2 階の必要条件). X を $\mathbf{R}^n$ の開部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ が *2* 階連続微分可能とする．$\boldsymbol{x}^* \in X$ が

1. f の強い意味の局所的最大解ならば，$\nabla^2 f(\boldsymbol{x}^*)$ は負の定符号

2. f の強い意味の局所的最小解ならば，$\nabla^2 f(\boldsymbol{x}^*)$ は正の定符号

ここで

$$\nabla^2 f(\boldsymbol{x}^*) = \left(\frac{\partial^2}{\partial x_i \partial x_h} f(\boldsymbol{x}^*)\right)_{ij}$$

（証明は演習とします.）

4.6 制約なし最適化問題 (2 次元以上) の十分条件

一般の次元における局所的最適解の十分条件はヘッセ行列 $\nabla^2 f(\boldsymbol{x}^*)$ の定符号性を用いて以下のように表現されます.

定理 30 (局所最適化の 2 階の十分条件). X を $\mathbf{R}^n$ の開部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ が *2* 階連続微分可能とする．$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ をみたす $\boldsymbol{x}^* \in X$ は,

1. $\nabla^2 f(\boldsymbol{x}^*)$ は負の定符号ならば f の局所的最大解,

2. $\nabla^2 f(\boldsymbol{x}^*)$ は正の定符号ならば f の局所的最小解,

(証明)　ここでも最大の場合のみを証明する．任意の $\varepsilon > 0$ と $||\boldsymbol{h}|| = 1$ となる任意の $\boldsymbol{h} \in \mathbf{R}^n$ に対して多次元のテイラーの定理より，$\theta \in (0, 1)$ が存在して

$$f(\boldsymbol{x}^* + \varepsilon \boldsymbol{h}) - f(\boldsymbol{x}^*) = \varepsilon \nabla f(\boldsymbol{x}^*)\boldsymbol{h} + \frac{1}{2}\varepsilon^{2t}\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^* + \theta\varepsilon\boldsymbol{h})\boldsymbol{h}$$

となります．ここで，$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ と f の 2 階の導関数の連続性から十分小さな $varepsilon > 0$ に対して ${}^t\boldsymbol{h}\nabla^2 f(\boldsymbol{x}^* + \theta\varepsilon\boldsymbol{h})\boldsymbol{h} < 0$ となるので，この十分小さな $\varepsilon > 0$ に対して

$$(\forall \varepsilon' \in (0, \varepsilon)) f(\boldsymbol{x}^* + \varepsilon'\boldsymbol{h}) - f(\boldsymbol{x}^*) < 0$$

がいえるので，$\boldsymbol{x}^*$ は強い意味の局所的最大解となります.（証明おわり）

大域的性質である凹性・凸性を関数に課すことで，最大の十分条件に関する定理が得られます.

ここで，多次元の場合の凹関数と凸関数を定義して，凹性と凸性の条件を導いておきましょう. 多次元の場合も，基本的に 1 次元の場合と同じです.

関数 $f : \mathbf{R}^n \to \mathbf{R}$ について

$$(\forall \boldsymbol{a} \in \mathbf{R}^n)(\forall \boldsymbol{b} \in \mathbf{R}^n)(\forall t \in [0, 1]) t f(\boldsymbol{a}) + (1-t) f(\boldsymbol{b}) \leq f(t\boldsymbol{a} + (1-t)\boldsymbol{b})$$

が成立するとき凹関数とよびます．また,

$$(\forall \boldsymbol{a} \in \mathbf{R}^n)(\forall \boldsymbol{b} \in \mathbf{R}^n)(\forall t \in [0, 1]) t f(\boldsymbol{a}) + (1-t) f(\boldsymbol{b}) \geq f(t\boldsymbol{a} + (1-t)\boldsymbol{b})$$

が成立するとき凸関数とよびます.

関数 $f : \mathbf{R}^n \to \mathbf{R}$ について

$$(\forall \boldsymbol{a} \in \mathbf{R}^n)(\forall \boldsymbol{b} \in \mathbf{R}^n)(\forall t \in (0,1))\boldsymbol{a} \neq \boldsymbol{b} \Longrightarrow tf(\boldsymbol{a}) + (1-t)f(\boldsymbol{b}) < f(t\boldsymbol{a} + (1-t)\boldsymbol{b})$$

が成立するとき狭義の凹関数とよびます．また，

$$(\forall \boldsymbol{a} \in \mathbf{R}^n)(\forall \boldsymbol{b} \in \mathbf{R}^n)(\forall t \in (0,1))\boldsymbol{a} \neq \boldsymbol{b} \Longrightarrow tf(\boldsymbol{a}) + (1-t)f(\boldsymbol{b}) > f(t\boldsymbol{a} + (1-t)\boldsymbol{b})$$

が成立するとき狭義の凸関数とよびます．

例 16． 線形関数 $f(\boldsymbol{x}) = {}^t\beta\boldsymbol{x}$ は，凹関数でもあり，凸関数でもあります．しかし，狭義の凸関数あるいは，狭義の凹関数ではありません．

例 17． 関数 $f(x_1, x_2) = x_1 + \log x_2$ は凹関数です．

例 18． 関数 $f(x_1, x_2) = x_1^{\frac{1}{3}} x_2^{\frac{2}{3}}$ は凹関数です．

例 19． 関数 $f(x_1, x_2) = \log x_1 + \log x_2$ は凹関数です．

演習 22． 上記の各例の関数が実際に凹関数であることを確認してください．また，どれが狭義の凹関数となるかを調べましょう．

大局的最適化のための十分条件は以下の定理を用いて得ることができます．

定理 31 (凹関数と凸関数の一階条件による特徴づけ)． $\boldsymbol{a} \in \mathbf{R}^n$ とする．$\boldsymbol{a}$ について微分可能な関数 $f : \mathbf{R}^n \to \mathbf{R}$ が凹関数であるとき，またそのときに限って

$$(\forall \boldsymbol{x} \in \mathbf{R}^n) f(\boldsymbol{x}) - f(\boldsymbol{a}) \leq \nabla f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$$

凸関数であるとき，またそのときに限って，

$$(\forall \boldsymbol{x} \in \mathbf{R}^n) f(\boldsymbol{x}) - f(\boldsymbol{a}) \geq \nabla f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})$$

(証明) 必要性を示します.（十分性の証明は演習とします.）凹関数，凸関数の定義において，$\boldsymbol{a} = \boldsymbol{y}$ とおいたものを整理して，

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \leq \frac{f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})) - f(\boldsymbol{a})}{t}$$

を得たのち，t をゼロに近づければよい.（証明おわり）

演習 23． 定理 *31* の証明のうち十分性の部分を完成させてください．

定理 32 (大局的な最適化の十分条件)． $f : \mathbf{R}^n \to \mathbf{R}$ が微分可能な凹関数であるとき，$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ である $\boldsymbol{x}^* \in \mathbf{R}^n$ は f の大局的な最大化問題の解となる．同様に，凸関数のとき $\boldsymbol{x}^*$ は大局的な最小化問題の解となる．

(証明)　定理 31 において，$\boldsymbol{a} = \boldsymbol{x}^*$ とおくと，$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ 凹関数の場合

$$(\forall \boldsymbol{x} \in \mathbf{R}^n) f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \leq 0$$

凸関数の場合

$$(\forall \boldsymbol{x} \in \mathbf{R}^n) f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \geq 0$$

が得られます．これは，それぞれ大局的な最適化になっていることを示しています．(証明おわり)

実は，定理 22 と関連して，次の定理が成立します．

定理 33. $f : \mathbf{R} \to \mathbf{R}$ が任意の点で微分可能とする．

$$(\forall x \in \mathbf{R})(\forall y \in \mathbf{R}) f(x) - f(y) \leq f'(y)(x-y)$$

ならば，f は凹関数.

$$(\forall x \in \mathbf{R})(\forall y \in \mathbf{R}) f(x) - f(y) \geq f'(y)(x-y)$$

ならば f は凸関数.

(証明)

なお，ここまで f の凹性，凸性に関して，f の各点における微分可能性のみを仮定して議論してきましたが，f が 2 階連続微分可能とすると，2 階の導関数を用いて，f の凹性，凸性を特徴づけることができます.

定理 34. $f : \mathbf{R} \to \mathbf{R}$ が *2* 階連続微分可能とする．f が凹関数であることの必要十分条件は,

$$(\forall \boldsymbol{x} \in \mathbf{R}) \nabla^2 f(x) \text{ が負の半定符号},$$

凸関数であることの必要十分条件は,

$$(\forall \boldsymbol{x} \in \mathbf{R}) \nabla^2 f(x) \text{ が正の半定符号}.$$

f が狭義の凹関数であることの必要十分条件は,

$$(\forall \boldsymbol{x} \in \mathbf{R}) \nabla^2 f(x) \text{ が負の定符号},$$

狭義の凸関数であることの必要十分条件は,

$$(\forall \boldsymbol{x} \in \mathbf{R}) \nabla^2 f(x) \text{ が正の定符号}.$$

定理 35 (最適化の 2 階の十分条件). X を $\mathbf{R}^n$ の開部分集合として，関数 $f : X \longrightarrow \mathbf{R}$ が連続微分可能とする．$\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ をみたす $\boldsymbol{x}^* \in X$ は,

1. f が凹関数ならば f の最大解,
2. f が凸関数ならば f の最小解

4.7 制約条件つき最適化問題の必要条件

4.7.1 等式制約の場合

経済学で多く使われるのが，制約条件つき最適化問題です．経済学が，限られた資源のもとでの個人の効用もしくは集団の厚生の最大化の解明を目指すため，この設定の重要性は言うまでもありません.

制約条件を集合が 30 ページにあるようにあるベクトル値写像 $\boldsymbol{g} : \mathbf{R}^n \to \mathbf{R}^\ell$ を用いて

$$W = \{\boldsymbol{x} \in \mathbf{R}^n ; \boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}\}$$

として表される場合を考えているわけですが，まず経済学上のよくある例を用いて制約条件つき最適化問題を考えていきましょう.

例 20. 予算制約

$$2x + y = 1$$

のもとで，対数線形型の効用関数

$$u(x, y) = \log x + 2 \log y$$

を最大化する，消費の組み合わせ (x, y) を求めよ．

この問題は，制約式を

$$y = 1 - 2x$$

と変形し，

$$f(x) = u(x, 1 - 2x) = \log x + 2 \log(1 - 2x)$$

の最大化を考えれば，$x = 1/6$ さらに，制約条件式より $y = 2/3$ が求まります．

この例では，制約条件式をうまく変形できたので，制約条件なしの問題に帰着することができました．しかし次の例のように，制約条件式から制約条件なしのにきちゃくさせることがやっかいな場合もあります．

例 21. *CES* 生産関数の制約

$$f(x, y, z) = \left(3x^{-1/2} + 2y^{-1/2}\right)^2$$

の制約の下で，利潤

$$2x + y + 3z$$

を最大化する，正の値をとる投入の組み合わせ (x, y) を求めよ．この問題は，制約式をたとえば y について解いて制約条件なしの最適化に帰着させることができなくはありませんが，かなり計算がやっかいです．

ラグランジュの未定乗数法は，制約条件式を使って，いくつかの変数を他の変数の関数として表して，制約条件式のない極値問題に帰着させるという，上の例のプロセスを効率的に行なう手法です．以下，ラグランジュ未定乗数法の仕組みを，2 変数の目的関数 $f(x_1, x_2)$ を，1 個の陰関数 $g(x_1, x_2) = 0$ という制約のもとで最適化するという簡単な設定の下でみていきましょう．ともに，連続微分可能としておきます．

2 つの変数 x_1,x_2 が $g(x_1, x_2) = 0$ という関係にありますから，もし x_2 が x_1 について解けるとすると，

$$x_2 = \varphi(x_1)$$

のように書けて，

$$g(x_1,\ \varphi(x_1)) = 0$$

が恒等的に成立するはずです．実は

$$\frac{\partial g}{\partial x_1} \neq 0$$

という条件の下で，こうした φ の存在を点 (x_1, x_2) の近傍において保証するのが陰関数定理でした．ですから，陰関数定理によって $\varphi(\cdot)$ の存在が保証されるときには,「$f(x_1, x_2)$ を $g(x_1, x_2) = 0$ の下で極大化する」という問題は,「$f(x_1, \varphi(x_1))$ を極大化する」という制約条件なしの問題に帰着されます．制約条件なし最適化問題の局所的な解の必要条件は，すでにみたように，x_1 について微分してゼロとなることです．

実際に微分してゼロとおいて必要条件をもとめると，

$$\frac{\partial f}{\partial x_1}(x_1, \varphi(x_1)) + \frac{\partial f}{\partial x_2}(x_1, \varphi(x_1))\frac{\partial \varphi}{\partial x_1}(x_1) = 0 \tag{4.3}$$

となります．

一方で，陰関数定理が存在を保証する $\varphi(\cdot)$ の下で，恒等的に成立しているはずの，$g(x_1, \varphi(x_1)) = 0$ の両辺を x_1 について微分すると右辺が定数だから，

$$\frac{\partial g}{\partial x_1}(x_1, \varphi(x_1)) + \frac{\partial g}{\partial x_2}(x_1, \varphi(x_1))\frac{\partial \varphi}{\partial x_1}(x_1) = 0$$

がえられます．つまり，

$$\frac{\partial \varphi}{\partial x_1} = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \tag{4.4}$$

が，制約条件式 $g(x_1, x_2) = 0$ をみたす点 (x_1, x_2) の近傍で成立します．

注意 15. 式 *(4.4)* を，*(4.3)* に代入して，

$$\frac{\partial f}{\partial x_2} \neq 0$$

を仮定してまとめると，いわゆる接点条件

$$\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \tag{4.5}$$

がえられます．これは，経済学において，予算制約のもとで効用を最大化する消費者の例で言うと，限界代替率が価格に等しいという条件にと解釈することができます．

今度は，式 (4.4) を

$$\frac{\partial g}{\partial x_1} \neq 0$$

の仮定の下で

$$\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial g}{\partial x_1}} = \frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}} \tag{4.6}$$

と書き直してみます．書き直した (4.6) は，制約条件の下での極大化のための必要条件となっています．もともとの制約式 $g(x_1, x_2) = 0$ も，必要条件になっていることに注意しましょう．

さて，この両辺の値を $-\lambda$ と定義することにすると，

$$\begin{aligned}
\frac{\partial f}{\partial x_1} + \lambda\frac{\partial g}{\partial x_1} &= 0 \\
\frac{\partial f}{\partial x_2} + \lambda\frac{\partial g}{\partial x_2} &= 0
\end{aligned}$$

という式が得られます．この式は，

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \tag{4.7}$$

という 3 変数関数を定義したとき，L を x_i，$(i = 1, 2)$ について偏微分してゼロとおいた式に他ならなりません．

また，$L(x_1, x_2, \lambda)$ を λ について偏微分してゼロとおくと

$$\frac{\partial L}{\partial \lambda} = g(x_1, x_2) = 0$$

という元々の制約式がえられます．

以上の議論は，設定を極大から極小に置き換えても，成立することが明らかです．そこで，以下のように結論をまとめることができます．

命題 2． $f(x_1, x_2)$，$g(x_1, x_2)$ を連続微分可能な関数とする．

$$\frac{\partial g}{\partial x_1}(x_1^*, x_2^*) \neq 0, \quad \frac{\partial g}{\partial x_2}(x_1^*, x_2^*) \neq 0$$

が (x_1^*, x_2^*) において成立していると仮定する．このとき，(x_1^*, x_2^*) において，$f(x_1, x_2)$ が $g(x_1, x_2) = 0$ という条件の下で，極値をとるための必要条件は，

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

と定義するとき，

$$\begin{cases} \frac{\partial L}{\partial x_1}(x_1^*, x_2^*) = 0 \\ \frac{\partial L}{\partial x_2}(x_1^*, x_2^*) = 0 \\ \frac{\partial L}{\partial \lambda}(x_1^*, x_2^*) = 0 \end{cases} \tag{4.8}$$

と与えられる．

式 (4.7) で定義した x_1, x_2 と λ の関数である L を Lagrange 関数とよびます．また，λ を Lagrange の未定乗数といいます．命題 4.8 のように，条件つき極値問題の必要条件を求める方法を，Lagrange の未定乗数法とよびます．

例 22． *2 変数関数として* $f(x_1, x_2) = x_1 x_2$，*制約条件式として* $x_1^2 + 2x_2^2 - 1 = 0$ *を考える場合の最大化問題の必要条件を Lagrange の未定乗数法でもとめてみましょう．*

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda(x_1^2 + 2x_2^2 - 1)$$

とおいて，*(4.8)* の各式を計算すると，

$$x_2 + 2\lambda x_1 = 0, x_1 + 4\lambda x_2 = 0, x_1^2 + 2x_2^2 - 1 = 0$$

これを x_1, x_2, λ に関する連立方程式として解くと，極大値の候補が得られるというわけです．実際には

$$(x_1, x_2) = \left(\frac{\sqrt{2}}{2}, \frac{1}{2}\right), \quad (x_1, x_2) = \left(-\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$$

が極大値になります．

上の例では，$\varphi(\cdot)$ として $x_1^2+2x_2^2-1=0$ を $x_1=\sqrt{1-2x_2^2}$ あるいは，$x_1=-\sqrt{1-2x_2^2}$ を計算して，制約条件なしの極値問題を具体的に求めなくてもよいために，計算がずっと，簡便になっていることに注目してください．しかし，背後にある考え方に変化はないのです．

以上，制約条件のある極大化問題は，陰関数定理によって制約なしの極大化（極小化）問題に帰着され，そこから得られる最適のための必要条件が Lagrange 未定乗数法から得られる必要条件に等しいことを，2 変数関数・1 制約式の場合に確認しました．さらに，以下のようにすれば，極大あるいは極小のための二階の必要条件がえられます．命題 2 の前提が成立しているとすると，陰関数定理から $g(a,b)=0$ を満たす点 (a,b) の近くで $g(x_1,x_2)=0$ を満たす任意の点は，x_1 をパラメーターとして

$$(x_1,\varphi(x_1))$$

と表現されます．つまり，$\frac{\partial g}{\partial x_1}(a,b)\neq 0$ がみたされると陰関数定理によってパラメーター表示されます．

さて，

$$g(x_1,\varphi(x_1))=0$$

が $x_1=a$ の近くで恒等的に成立するのでこの両辺を x_1 で微分して

$$\frac{\partial g}{\partial x_1}+\frac{\partial g}{\partial x_2}\cdot\varphi'(x_1)=0 \tag{4.9}$$

が $x_1=a$ の近くで恒等的に成立します．従って

$$\varphi'(x_1)=-\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}$$

が成立します．式 (4.9) をもう一度 x_1 で微分してみましょう．

$$\frac{\partial^2 g}{\partial x_1^2}+2\frac{\partial^2 g}{\partial x_1\partial x_2}\cdot\varphi'(x_1)+\frac{\partial^2 g}{\partial x_2^2}\cdot(\varphi'(x_1))^2+\frac{\partial g}{\partial x_2}\cdot\varphi''(x_1)=0$$

したがって，

$$\begin{aligned}
&\varphi''(x_1)\\
&=-\frac{1}{\frac{\partial g}{\partial x_2}}\{\frac{\partial^2 g}{\partial x_1^2}+2\frac{\partial^2 g}{\partial x_1\partial x_2}\cdot\varphi'(x_1)+\frac{\partial^2 g}{\partial x_2^2}\cdot(\varphi'(x_1))^2\}\\
&=-\frac{1}{\frac{\partial g}{\partial x_2}^3}\left\{\frac{\partial^2 g}{\partial x_1^2}\left(\frac{\partial g}{\partial x_2}\right)^2-2\frac{\partial^2 g}{\partial x_1\partial x_2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_2}+\frac{\partial^2 g}{\partial x_2^2}\left(\frac{\partial g}{\partial x_1}\right)^2\right\} \qquad (4.10)
\end{aligned}$$

同様に $\Psi(x_1)=f(x_1,\varphi(x_1))$ とおくと

$$\begin{aligned}
\Psi''(x_1) &= \frac{\partial^2 f}{\partial x_1^2}+2\frac{\partial^2 f}{\partial x_1\partial x_2}\cdot\varphi'(x_1) && (4.11)\\
&+ \frac{\partial^2 f}{\partial x_2^2}\cdot(\varphi'(x_1))^2+\frac{\partial f}{\partial x_2}\cdot\varphi''(x_1) && (4.12)
\end{aligned}$$

となります．さて，(a,b) において Lagrange の未定乗数法の必要条件

$$\frac{\partial f}{\partial x_1}+\lambda\frac{\partial g}{\partial x_1}=0,\ \frac{\partial f}{\partial x_2}+\lambda\frac{\partial g}{\partial x_2}=0 \tag{4.13}$$

が満たされるとします．Lagrangean を作るために $L = f + \lambda g$ とおきます．このとき (4.11) より

$$
\begin{aligned}
&\Psi''(x_1) \\
&= \frac{\partial^2 f}{\partial x_1^2} + 2\frac{\partial^2 f}{\partial x_1 \partial x_2} \cdot \left(-\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}\right) + \frac{\partial^2 f}{\partial x_2^2} \cdot \left(-\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}\right)^2 \\
&+ \frac{\partial f}{\partial x_2} \cdot \left(-\frac{1}{\frac{\partial g}{\partial x_2}^3}\right)\left(\frac{\partial^2 g}{\partial x_1^2}\frac{\partial g}{\partial x_2}\frac{\partial g}{\partial x_2} - 2\frac{\partial^2 g}{\partial x_1 \partial x_2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial^2 g}{\partial x_2^2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_1}\right) \\
&= \frac{1}{\left(\frac{\partial g}{\partial x_2}\right)^2}\left(\frac{\partial^2 f}{\partial x_1^2}\frac{\partial g}{\partial x_2}^2 - 2\frac{\partial^2 f}{\partial x_1 \partial x_2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial^2 f}{\partial x_2^2}\frac{\partial g}{\partial x_1}^2\right) \\
&+ \frac{1}{\left(\frac{\partial g}{\partial x_2}\right)^2} \cdot \left(-\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}\right)\left(\frac{\partial^2 g}{\partial x_1^2}\frac{\partial g}{\partial x_2}^2 - 2\frac{\partial^2 g}{\partial x_1 \partial x_2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial^2 g}{\partial x_2^2}\frac{\partial g}{\partial x_1}^2\right)
\end{aligned}
$$

$$
= \frac{1}{\left(\frac{\partial g}{\partial x_2}\right)^2}\left(\frac{\partial^2 L}{\partial x_1^2}\frac{\partial g}{\partial x_2}^2 - 2\frac{\partial^2 L}{\partial x_1 \partial x_2}\frac{\partial g}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial^2 L}{\partial x_2^2}\frac{\partial g}{\partial x_1}^2\right) \tag{4.14}
$$

$$
= -\frac{1}{\left(\frac{\partial g}{\partial x_2}\right)^2}\begin{vmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_2^2} \end{vmatrix} \tag{4.15}
$$

となります．ここで，

$$
B(x, y, \lambda) = \begin{vmatrix} 0 & \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \\ \frac{\partial g}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial g}{\partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_2^2} \end{vmatrix} \tag{4.16}
$$

とおくと，

$$B(x, y, \lambda) < 0 \Leftrightarrow \Psi''(x_1) > 0$$

$$B(x, y, \lambda) > 0 \Leftrightarrow \Psi''(x_1) < 0$$

が成立しています．以上のことをまとめると次の命題が成立します．

命題 3．

$$g(x_1, x_2) = 0$$

の下で，

$$z = f(x_1, x_2)$$

の極値を求める問題を考える．

ここで

$$
\frac{\partial g}{\partial x_1}(a, b)^2 + \frac{\partial g}{\partial x_2}(a, b)^2 \neq 0,\ g(a, b) = 0 \tag{4.17}
$$

を仮定する．

また

$$L(x_1, x_2, \lambda) := f(x_1, x_2) + \lambda g(x_1, x_2)$$

とおく．さらに， *Lagrange* 未定乗数法の必要条件 *(4.13)*， すなわち

$$\frac{\partial L}{\partial x_1}(a, b, \lambda) = \frac{\partial L}{\partial x_2}(a, b, \lambda) = 0$$

を仮定する．このとき，

1. $B(a,b,\lambda) < 0$ ならば，　$g(x_1,x_2) = 0$ の下で $z = f(x_1,x_2)$ は (a,b) で極小値をとる．

2. $B(a,b,\lambda) > 0$ ならば，　$g(x_1,x_2) = 0$ の下で $z = f(x_1,x_2)$ は (a,b) で極大値をとる．

上の命題は，制約条件つきの極値問題を Lagrange 関数とよばれる関数 $L(x_1,x_2,\lambda) = f(x_1,x_2) + \lambda g(x_1,x_2)$ を用いて必要条件をもとめることは，制約条件を用いて一部の変数を消去して，制約条件のない極値問題に帰着させて解く場合の必要条件をもとめることと等しいことを，2 変数の場合について示したことになります．