講義ノート
経済学のための最適化理論

増山　幸一
明治学院大学経済学部

2006 年 9 月

1　始めに

経済分析、とりわけミクロ経済学などのテキストで頻繁に登場する効用最大化問題や費用最小化問題の解法を理解するためには、最適化理論の数学的な構造を理解することが必須である．この講義ノートでは、静学的な最適化理論の基本的な手法を説明する．動学的最適化問題については触れていない．読者に要求されるバックグラウンドに関して、数学的な能力に関しては、簡単な微分法と行列演算の知識を前提としている．経済学の知識に関しては、入門レベルのミクロ経済学の知識を想定している．

「第２節　集合と関数の性質」で、集合と関数の基礎的な性質を解説している．開集合、閉集合、集積点そして集合のコンパクト性、凸性などについて説明してあります．関数に関しては、連続性、微分可能性、そして凸関数、凹関数などの性質を説明しています。集合論と関数論に関して初歩的な知識、例えば、集合の凸性や関数の連続性などの知識を有している諸君は、第２節をスキップして結構です．第 2 節が純粋数学的で抽象的に過ぎるという印象を抱いた諸君も、第 2 節を深く理解せずに、眺める程度にして読み飛ばしても結構です。第 3 節以降を学習するに当たっては、第 2 節の内容に習熟していなくても理解できる構成にはなっています．

第 3 節で、制約条件がないときの関数の最大化問題を取り上げ、最大化のための必要条件と十分条件について説明している．応用例として、企業の利潤最大化問題を考察している．生産関数が凹関数であるという条件がいかに重要になっているかについて理解が進むと思われる．また、ミクロ経済学でおなじみのホテリングの補題も導出されている．第 4 節では、第 3 節での結果を受けて、等式制約条件がある場合の関数の最大化問題を取り上げている．いわゆるラグランジュ乗数法を説明している．第 2 階の必要条件が縁付ヘッセ行列を用いて説明されている．そこでは、行列が負定値になるための必要十分条件が重要な役割を果たしている．直感的には、制約条件を満たす領域に接する平面上でヘッセ行列が負定値になる条件を求めることに等しい．第 5 節で、数理計画法で必須の条件であるキューン・タッカー条件を説明している．ラグランジュ乗数法を用いると、簡単にキューン・タッカーの条件が導出できることを説明している．最大化すべき関数が的確な性質を有していれば、キューン・タッカー条件が最大化のための必要条件になるだけでなく、十分条件にもなることが説明されている．この応用として、第 6 節で、費用最小化問題を取り上げ、第 7 節で効用最大化問題を取り上げ、その解法について説明してあります．

すべての定理に対する証明を記述すると、余りに分量が多くなるため、不必要な証明は省略した．証明に興味のある諸君は掲示してある参考文献を参照してください．表記法は解析学の教科書などでスタンダードな表記を用いています．数学のテキストでは、二つのベクトル x と y の内積を $\langle x, y\rangle$ と表記するケース

が一般的であるが、以下の記述では内積の表記として単純に $x \cdot y$ を用いた．また、定義であることを強調する場合、定義の表記として、

$$R^n := \{(x_1, x_2, ..., x_n) : x_i \in R, i = 1, 2, .., n\}$$

という形式の表現を用いた．R^n は n 次元ユークリッド空間であるが、

$$R^n_+ := \{(x_1, x_2, ..., x_n) : x_i \in R,\ x_i \geq 0,\ i = 1, 2, .., n\}$$

は各座標が非負の n 次元ユークリッド空間を表す．「付録：解析学の基礎」で解析学の基本的な諸概念に関する知識をリフレッシュしているので、最適化理論をより深く理解したい諸君はそこから復習してください．かなり高度かもしれませんが、実った果実は大きいと信じます．

2　集合と関数の性質

2.1　集合のコンパクト性

定義 2.1

m 次元ユークリッド空間 R^m の部分集合 U を考える．集合 U が U の各点の近傍であるとき、開集合であるという．

開集合であることの条件は以下のようにも表現できる．U に属する各点 x に対して、開球 $B(x; r) = \{y \in R^m; \|x - y\| < r\}$ が U の部分集合となるような $r > 0$ が存在するとき、開集合であるという[1]．明らかに、開球 $B(x; r) = \{y \in R^m; \|x - y\| < r\}$ は開集合である．有限個の開球の和集合が開集合になることはほぼ自明である．だから、$Z = \{1, 2, 3 \cdots, n\}$ とするとき、$\bigcup_{\alpha \in Z} B(x_\alpha; r_\alpha)$ は開集合になる．開区間 (a, b) は R の開集合である．区間 $(a, b]$ は開集合ではない．なぜなら、開球 $B(b; \epsilon) = \{x \in R; \|x - b\| < \epsilon\}$ は区間 $(a, b]$ に含まれない．

定理 2.1

開集合の加算個の和は開集合である．有限個の開集合の共通部分は開集合である．

証明：解析学のテキスト, 例えば、Hoffman の Analysis in Euclidean Space, p.56 を参照してください．

定義 2.2

$S \subset R^m, x \in R^m$ とする．点 x のあらゆる近傍が x と異なる S の点を含むとき、点 x は S の蝕点 *(cluster point)* という．

例 2.1

$\{x_n\} = \{0, 1, 0, 1, \cdots\}$ である．$S = \{0, 1\}$ とする．0 と 1 は点列 $\{x_n\}$ の集積点であるが、集合 S の蝕点ではない[2]．

極限、集積点および蝕点は同一になるのであろうか？それらの関係は以下の定理群で明らかになる．証明は、Hoffman の Analysis in Euclidean Space, を参照してください．

---

[1] $\|x - y\| < r$ はユークリッドノルムを表現する．ユークリッド空間の基本的な性質に関しては、付録：解析学の基礎を参照してください．

[2] x のあらゆる近傍が無限の個数の $\{x_n\}$ を含むとき、点 x は点列 $\{x_n\}$ の集積点であるという．

補題 2.1
$S \subset R^m, x \in R^m$ とする．以下の各命題は等価である．

(1). x は集合 S の触点である．

(2). x のあらゆる近傍は S 内の無限個の点を含む．

(3). $x_n \neq x, x = \lim_{n\to\infty} x_n$ となるような点列 $\{x_n\}(x_n \in S)$ が存在する．

補題 2.2
点列 $\{x_n\} \in R^m$ のある部分点列が x に収束するならば、点 x は点列 $\{x_n\}$ の集積点である．

定理 2.2 (Bolzano-Weierstrass)
R^m の有界な点列は必ず集積点を持つ．言い換えると、R^m の有界な点列は必ず収束する部分点列を持つ．

R^m の有界な点列は集積点を持つが、集積点が一つとは限らない．有界な点列が唯一つの集積点を持つならば、その点列は集積点に収束する．反対に、収束する点列はただ一つの集積点を持つ．$\{x_n\}$ を R^m の点列とし、この点列からなる集合を

$$S := \{x_n; n \in Z_+\}$$

とする[3]．点 x が集合 S の触点であるとき、x のすべての近傍は集合 S の無限個の点を含む．明らかに、点 x は点列 $\{x_n\}$ の集積点である．しかし、点 y が点列 $\{x_n\}$ の集積点であるとき、必ずしも集合 S の触点にはならない．この反例は上の例で説明したとおりである．この反例は、同一の値の点（0 と 1) が無限に繰り返されることから生まれる．点列が互いに異なる値の点からなるとき、集積点は触点と一致する．上記の例外を除けば、集積点と触点は一致する．したがって、解析学のテキストでは通常、集積点だけが採用されるケースが多い．

定義 2.3
集合 S のすべての触点が S に含まれるとき、S は閉集合であるという．

注意：R^m と空集合 $\emptyset$ は開集合かつ閉集合である．

補題 2.3
R^m の部分集合 S を考える．S のすべての収束点列の極限が S に含まれるとき、つまり、$x = \lim x_n, x_n \in S$ のとき $x \in S$ ならば、S は閉集合である．

この補題から、閉集合を以下のように定義してもよいことが分かる．S に関する集積点は必ずしも S に属するとは限らないが、もしすべての集積点が S に属するならば、S は閉集合である．

定理 2.3
集合 S は、S の補集合が開集合であるとき、そしてそのときにのみ、閉集合である．

証明：解析学のテキスト, 例えば、Hoffman の Analysis in Euclidean Space, pp.58-59 を参照してください．

[3] Z_+ は正の整数（自然数）を表す．

定義 2.4
S を R^m の部分集合とする．集合 S を含むあらゆる閉集合の共通集合を S の閉包 *(closure)* という．S に含まれるあらゆる開集合の和集合を S の開核 *(interior)* という．S の閉包と S の補集合の閉包の共通集合を境界 *(boundary)* という．

S の閉包を記号 $\bar{S}$ で表現し、S の開核を S^o で表記する．

定義 2.5
S を T の部分集合とする．T に属するあらゆる点が S の閉包に含まれるとき、S は T で稠密 *(dense)* であるという．

S が T で稠密であることは、T のあらゆる点が S の点によって近似できることを意味する、すなわち、T のあらゆる点は S に属する点列の極限になっている．例えば、有理数の集合は R^1 で稠密になっている．一般的に、あらゆる集合はその閉包で稠密である．もし K が閉集合であるならば、部分集合 S が K で稠密になるのは $\bar{S}=K$ のときだけである．

定義 2.6
$S \subset R^m$ とする．$S \subset \bigcup_\alpha U_\alpha$ となるような集合の族 $\{U_\alpha : \alpha \in A\}$ を S の被覆 *(cover)* という．

例えば、$S \subset R^m, U_x = B(x;\epsilon), \epsilon > 0$ とすると、$\{U_x; x \in S\}$ は S の被覆である．U_x は開集合なので、開被覆という．

定義 2.7
集合 K のあらゆる開被覆が有限個の部分被覆をもつとき、K はコンパクト *(compact)* であるという．

例 2.2
$\{x_n\}$ を収束する点列とし、x をその極限とする．集合 S を

$$S := \{x, x_n; n = 1, 2, ..\}$$

とする．いま $\{U_\alpha\}$ を S のある開被覆であるとする．このとき、開集合 U_α の一つは点 x を含まなければならない．$x \in U_{\alpha_1}$ としよう．x_n は x に収束するから、集合 U_{α_1} は有限個数の x_n を除いた点列 $\{x_n\}$ を含んでいる．U_{α_1} に属さない x_n の個数（有限個）を $k-1$ とする．このとき、U_{α_1} に属さない各 x_n を含む $k-1$ 個の開集合 $U_{\alpha_2}, U_{\alpha_3}, .., U_{\alpha_k}$ を選ぶことができる．こうして作成された $U_{\alpha_1}, U_{\alpha_2}, .., U_{\alpha_k}$ は S の有限個の部分被覆となっている．S はコンパクトである．

定理 2.4 (Heine-Borel)
R^m において、すべての有界閉集合はコンパクトである．

証明：これはまりにも有名な定理であるので、証明はすべての解析学のテキストに載っている．コンパクト集合の必要十分条件に関わる定理群の証明として、ここでは、Hoffman, Analysis in Euclidean Space, pp.68-70 を推薦する．

集合 $K = [0, 1]$ は有界で、閉集合である．よって、集合 K はコンパクトである．消費の実行可能領域集合 K

$$K = \{(x_1, x_2) \in R^2 : p_1x_1 + p_2x_2 \leq M, x_1 \geq 0, x_2 \geq 0\}$$

は有界で、閉集合である．だから、コンパクトである．

問題 2.1
(1). 実数の区間 $[1,3)$ は開集合でないことを示しなさい．

(2). 実数の集合 $S=\{0,1\}$ を考える．0 と 1 は S の蝕点ではないことを示しなさい．

(3). 実数の集合 $S=\{0,1\}$ は閉集合であることを示しなさい．

(4). カントール集合は閉集合であることを示しなさい．

2.2 集合の凸性

定義 2.8
M を R^n の部分集合とする．任意の $x,y\in M, x\neq y$ と任意の $\lambda, 0\leq\lambda\leq 1$ に対して、$\lambda x+(1-\lambda)y\in M$ ならば、M は凸 *(convex)* であるという．
また、任意の実数 $\alpha>0$ に対して、$x\in M$ ならば $\alpha x\in M$ が成立するとき、M は原点を頂点とする錘 *(cone)* という．

$w\in R^n, w\neq 0, \alpha\in R$ とするとき、集合 $\{x; x\in R^n, w\cdot x<\alpha\}$ は R^n の半空間で開集合であり、凸集合でもある．また、集合 $\{x; x\in R^n, w\cdot x\geq\alpha\}$ は半空間で、閉集合、凸集合である[4]．

定理 2.5
$\{M_i\}_{i\in I}$ を R^n の凸集合の族とする．積 $\cap_{i\in I}M_i$ は凸集合である．

証明は Mangasarian, Nonlinear Programming, p.41 を参照のこと．

定義 2.9
$p\in R^n, p\neq 0, \alpha\in R$ とする．集合 $H(p,\alpha):=\{x; x\in R^n, p\cdot x=\alpha\}$ は R^n の超平面という．ベクトル p を超平面 $H(p,\alpha)$ の法線ベクトルという．

n 変数の線型方程式の解集合は R^n における超平面になる．

定理 2.6 (分離定理)
K_1 と K_2 を R^n における凸集合とする．ただし、$K_1^o\cap K_2^o=\emptyset$ かつ $K_1^o\neq\emptyset$ とする．このとき、K_1 と K_2 を分離する超平面、つまり

$$p\cdot y\leq H(p,\alpha)\leq p\cdot z, \forall y\in K_1, \forall z\in K_2$$

を満たす超平面が存在する．

証明は Mangasarian, Nonlinear Programming, p.49 を参照のこと．

[4] $w\cdot x$ は二つのベクトル w と x の内積を表現する．数学の教科書では、内積は $<w,x>$ で表現することが一般的である．前者の表現方法は経済数学や数理計画法などのテキストで用いられている．

2.3 関数の連続性

変数 x が限りなく一つの値 a に近づくとき、x の関数 $f(x)$ もまた限りなく $f(a)$ に近づくならば、$f(x)$ は $x = a$ で連続であるという．これを厳密に定義すると、以下の通りである．

定義 2.10
f を R^n の部分集合 D で定義された、D から R^m への写像であるとする．$x_o \in D$ に対して、$f(x_o)$ の近傍 V の逆像 $f^{-1}(V)$ が x_o の近傍であるとき、f は x_o で連続であるという．

$x_o \in D$ とするとき、以下の各命題は等価であることが知られている．

(1). f は x_o で連続である．

(2). 任意の $\epsilon > 0$ に対して、

$$\|f(x) - f(x_o)\| < \epsilon, \forall x \in B(x_0, \delta)$$

となる $\delta > 0$ が存在する．

(3). D における点列 $\{x_n\}$ が x_o に収束するならば、点列 $\{f(x_n)\}$ は $f(x_o)$ に収束する．

証明は Hoffman, Analysis in Euclidean Space, pp.78-79 を参照してください．この関係を用いると、f と g が連続であれば、$f + g$ も連続であること、および、$f \cdot g$ も連続であることが簡単に証明できる．さらに、f が R^n から R^m への連続写像で、g が R^m から R^k への連続写像であるとき、合成写像 $g \circ f$ が連続になることも容易に証明できる．

定理 2.7
$f : D \subset R^n \to R^m$ は連続である．K が定義域 D のコンパクトな部分集合であるならば、$f(K)$ はコンパクトである．

証明は Hoffman, Analysis in Euclidean Space, pp.86-85 を参照してください．この定理から、以下の重要な性質が導出される．
関数 $f : R^n \to R$ はコンパクトな集合 $D \subset R^n$ 上で連続であるとする．このとき、

$$f(x_1) = \sup\{f(x) : x \in D\},\ f(x_2) = \inf\{f(x) : x \in D\}$$

となる x_1, x_2 が存在し、$x_1 \in D, x_2 \in D$ である[5]．つまり、最大値と最小値が存在する．式で表現すれば、$f(x_2) \leq f(x) \leq f(x_1), \forall x \in D$ を満たす x_1, x_2 が存在する．
連続関数という条件は経済学では非常にきつい条件とみなされているので、連続性よりもゆるい条件、例えば、上半連続というゆるい制約で考えるケースが多い．こうしたケースでの最大最小問題は上級レベルの議論となるのでここでは省略する．

定理 2.8 (ブラウワーの不動点定理)
C を R^n の空でない、コンパクトな凸集合、f を C から C への連続写像とする．このとき、$\hat{x} = f(\hat{x})$ となる f の不動点 $\hat{x}$ が存在する．

[5] sup と inf の定義は第 8 節付録：解析学の基礎を参照してください．簡単に説明すると、集合の最小な上界を sup で、最大な下界を inf で表現する．

証明は Hoffman, Analysis in Euclidean Space, pp.265-266 を参照してください．経済学への応用はほとんどすべての数理経済学の教科書で取り上げられている．例えば、Takayama Akira, Mathematical Economics などを参照ください．

定義 2.11
f を $D \subset R$ から R^m への写像 (関数) とする．極限

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x},\ t,\ x \in D$$

が存在するとき、f は点 x で微分可能であるという．この極限を f の導関数といい、$f'(x), Df(x)$ で表現する．関数 f が定義域のすべての点で微分可能であるとき、単に微分可能であるという．

f が点 x で微分可能であるとき、$\lim_{t \to x} \|f(t) - f(x)\| = 0$ なので、点 x で連続である．k 回微分可能で、k 階導関数が連続であるような関数の集合を C^k で表現する．

しかし、関数 f が点 x で連続であるとしても、必ずしも微分可能であるとは限らない．このとき、点 x の右側から極限をとるとき得られる微係数を右側微係数 (right-hand derivative) といい、

$$D^+ f(x) = \lim_{t \to x^+} \frac{f(t) - f(x)}{t - x}$$

で定義する．同様に、点 x の左側から極限をとるとき得られる微係数を左側微係数 (left-hand derivative) といい、

$$D^- f(x) = \lim_{t \to x^-} \frac{f(t) - f(x)}{t - x}$$

で定義する．例えば、$f(x) = |x|$ のケースを考えよう．$D^+ f(0) = 1$ であるが、$D^- f(0) = -1$ となる．つまり、この関数 f は原点で下に尖っている．この場合、極限を右側から取ったときと左側からとったときでは極限が異なる．このとき、微分可能ではないという．よって、$D^+ f(x)$ と $D^- f(x)$ とが存在し、かつ、$D^+ f(x) = D^- f(x)$ であるときにのみ微分可能であるといえる．

定理 2.9
$f : (a, b) \to R$ とする．もし
(1) f は x で極大 *(*極小*)* である、
(2) f は x で微分可能である、
が成立すれば、$f'(x) = 0$ である．

証明：f が x で極大をとるので、

$$f(t) \leq f(x),\ |x - t| < \delta$$

をみたす $\delta > 0$ が存在する．よって、

$$D^+ f(x) = \lim_{t \to x^+} \frac{f(t) - f(x)}{t - x} \leq 0,$$
$$D^- f(x) = \lim_{t \to x^-} \frac{f(t) - f(x)}{t - x} \geq 0.$$

微分可能だから、$D^+ f(x) = D^- f(x)$ である．それ故、$f'(x) = 0$ となる．(QED)

定理 2.10 (平均値の定理)
$f : R \to R$ が $[a, b]$ で連続で、(a, b) で微分可能ならば、

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

となるような $x \in (a, b)$ が存在する．

証明：

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

とおく．g は (a, b) で微分可能で、$g(a) = g(b) = 0$ である．また、g は $[a, b]$ で連続であり、$[a, b]$ がコンパクトであることより、g は $[a, b]$ 上で最大値、最小値をとる．最大値または最小値のいずれかは開区間 (a, b) 内部の点で実現するので、

$$g'(x) = 0, x \in (a, b)$$

を満たす x が存在する．したがって、

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

(QED)

この定理は関数の第 1 階導関数を前提にしているが、第 n 階導関数が存在するとき、平均値の定理をより一般化することができる．関数 f の第 n 階導関数を $f^{(n)}$ で表記する．

定理 2.11 (Taylor の定理)
$f : R \to R$ が $[a, b]$ で第 n 階まで微分可能であるとする．点 x を区間 (a, b) 内のある点、y を任意の点とするとき、

$$f(y) = f(x) + \frac{f'(x)}{1!}(y - x) + \frac{f''(x)}{2!}(y - x)^2 + \cdots + \frac{f^{(n-1)}(x)}{(n - 1)!}(y - x)^{n-1} + \frac{f^{(n)}(z)}{n!}(y - x)^n$$

が成立する．ただし、$z = x + \theta(y - x), 0 < \theta < 1$ ．

証明は高木貞治『解析概論』62 ページを参照のこと．Taylor の定理で、$n = 1$ とすると、

$$f(y) = f(x) + f'(z)(y - x)$$

が得られる．これは平均値の定理そのものである．$n = 2$ とおくと、

$$f(y) = f(x) + f'(x)(y - x) + \frac{f''(z)}{2!}(y - x)^2$$

となる．これは、点 y での関数値 $f(y)$ を y の 2 次関数で近似するときの近似式を表現する．$n \to \infty$ とするとき、関数値 $f(y)$ は y の無限級数で表現されることになる．これをテイラー級数展開 (テイラーの公式) という．関数が多変数関数である場合にも、テイラーの定理を拡張することができる．この拡張を議論するためには、多変数関数の微分を定義する必要に迫られる．

定義 2.12
$f : R^n \to R$ とする．ある $x \in R^n$ に対して、

$$\lim_{\|h\| \to 0} \frac{|f(x + h) - f(x) - \langle a, h \rangle|}{\|h\|} = 0, \ h \in R^n$$

となるようなベクトル $a=(a_1,a_2,\cdots,a_n)$ が存在すとき、f は x で微分可能であるという．このとき、各 a_j を偏導関数 *(partial derivatives)* といい、

$$\frac{\partial f(x)}{\partial x_j}, \text{あるいは}, D_j f(x)$$

と表現する．

各偏導関数 (偏微分) は通常

$$\frac{\partial f(x)}{\partial x_j} = \lim_{t\to 0} \frac{f(x_1,x_2,\cdots,x_j+t,\cdots,x_n)-f(x_1,\cdots,x_j,\cdots,x_n)}{t}$$

にしたがって計算される．

関数 f のベクトル x による偏導関数は

$$\frac{\partial f(x)}{\partial x} := (\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n})$$

と表現され、関数 f のグラディエント・ベクトルともいわれる．関数 f のグラディエント・ベクトルは $\mathrm{grad} f(x)$ あるいは $Df(x)$ と表記される．C^k を k 階微分可能で、k 階偏導関数が連続であるような関数の集合を表現する．$f \in C^2$ であるとき、

$$D_i D_j f = D_j D_i f$$

が成立することが知られている．以下では、$D_i D_j f = D_{ij} f$ と表現することにする．関数 f の 2 階偏導関数

$$\frac{\partial^2 f(x)}{\partial x^2} = \begin{pmatrix} D_{11}f(x) & D_{12}f(x) & \cdots & D_{1n}f(x) \\ D_{21}f(x) & D_{22}f(x) & \cdots & D_{2n}f(x) \\ \vdots & \vdots & \vdots & \vdots \\ D_{n1}f(x) & D_{n2}f(x) & \cdots & D_{nn}f(x) \end{pmatrix}$$

をヘッセ行列と呼ぶ．

定理 2.12
$f: U \subset R^n \to R,\ f \in C^1$ であるとする．関数 f が開集合 U の上の点 x で極大値を持つならば、

$$Df(x) = 0,\ \text{すなわち}\ D_j f(x) = 0, j = 1, 2, \cdots, n$$

である．

証明は簡単なので、読者の練習問題とする．この定理から以下の重要な性質が導出できる．

定理 2.13
$f: U \subset R^n \to R,\ f \in C^2$、$U$ は凸開集合であるとする．x、$y \in U$ に対して、

$$f(y) = f(x) + Df(z)(y-x),$$
$$f(y) = f(x) + Df(x)(y-x) + \frac{1}{2}(y-x)^t \frac{\partial^2 f(w)}{\partial x^2}(y-x)$$

となるような $z = \alpha x + (1-\alpha)y,\ 1 \geq \alpha \geq 0$ および $w = \beta x + (1-\beta)y,\ 1 \geq \beta \geq 0$ が存在する[6]．

[6] x^t はベクトル x の転置ベクトルを表現する．

証明は 1 変数関数のケースと同様な手続きで行うことができる．$g(y) = f[x + \alpha(y - x)],\ 1 \geq \alpha \geq 0$ とおいて証明してください．この定理は 1 変数関数での「テイラーの定理」を拡張したもので、多変数関数におけるテイラー級数展開を保障する役割を果たしている．

問題 2.2

(1). $\sin x$ および $\cos x$ は $(-\infty, +\infty)$ で連続であることを証明しなさい．

(2). $f(x) = x \sin \frac{1}{x},\ f(0) = 0$ とするとき、$f(x)$ は $x = 0$ で連続であるが、微分可能ではないことを示しなさい．

2.4　凹関数と凸関数

定義 2.13
任意の $x, y \in D \subset R^n$ と任意の $\lambda (1 > \lambda > 0)$ に対して、

$$f(\lambda x + (1 - \lambda) y) \geq \lambda f(x) + (1 - \lambda) f(y)$$

であるならば、関数 $f : D \to R$ は D 上で凹 *(concave)* であるという．ただし、D は凸集合であるとする．等号が成立しないとき、f は厳密に凹であるという．

定義 2.14
任意の $x, y \in D \subset R^n$ と任意の $\lambda (1 > \lambda > 0)$ に対して、

$$\lambda f(x) + (1 - \lambda) f(y) \geq f(\lambda x + (1 - \lambda) y)$$

であるならば、関数 $f : D \to R$ は D 上で凸 *(convex)* であるという．ただし、D は凸集合であるとする．等号が成立しないとき、f は厳密に凸であるという．

補題 2.4
$f : D \subset R^n \to R$ が D 上 *(D は凸)* で凹関数であることは、集合

$$H_f := \{(x, \alpha); x \in D, \alpha \in R, f(x) \geq \alpha\}$$

が R^{n+1} で凸集合であることと等価である．

証明は Mangasarian, Nonlinear programming, pp.58-59 を参照のこと．

定義 2.15
任意の $x, y \in D \subset R^n$ と任意の $\lambda (1 \geq \lambda \geq 0)$ に対して、

$$f(x) \leq f(y) \text{ ならば } f(x) \leq f(\lambda x + (1 - \lambda) y)$$

であるとき、$f : D \subset R^n \to R$ は D 上で準凹 *(quasi concave)* であるという．

補題 2.5
$f : D \subset R^n \to R$ が D 上 *(D は凸)* で準凹関数であることは、各 $\alpha \in R$ に対して、集合

$$H_\alpha := \{x; x \in D, f(x) \geq \alpha\}$$

が R^n で凸集合であることと等価である．

証明は Mangasarian, Nonlinear programming, p.133 を参照のこと．

これらの関係から以下のような命題が導出できる．つまり、$f : D \subset R^n \to R$ が凸な定義域 D 上で凹関数になることは、

$$f(x) - f(y) \leq Df(y) \cdot (x - y),\ \forall x, y \in D$$

が成立することと等価である．さらに、$f : D \subset R^n \to R$ が凸な定義域 D 上で準凹関数になることは、

$$f(x) \geq f(y),\ x, y \in D \text{ のとき } Df(y) \cdot (x - y) \geq 0,$$

が成立することと等価である．証明は Mangasarian, Nonlinear programming, pp.83-84,134 を参照のこと．

定理 2.14
以下の各命題は等価である．D は凸集合であるとする．
(1) $f : D \subset R^n \to R$ は D 上で凹関数である．
(2) f のヘッセ行列 H は D 上で半負値 (*negative semidefinite*)、つまり、$h^t H h \leq 0, \forall h \in R^n$ である．

定理 2.15
以下の各命題は等価である．D は凸集合であるとする．
(1) $f : D \subset R^n \to R$ は D 上で準凹関数である．
(2) $Df \cdot h = 0$ となるすべてのベクトル $h \in R^n$ に対して、$h^t H h \leq 0$ である．

凹関数および準凹関数に関わるこれらの定理群の証明は、Mangasarian, Nonlinear programming、第 6 章、第 9 章などで与えられている．証明に興味のある諸君は上記の箇所を参照してください．なお、凹関数と準凹関数との間の関係は以下のようになっている．関数 f が厳密に凹関数であるならば、凹関数でもある．凹関数であれば、厳密に準凹関数であり、厳密に準凹関数であれば準凹関数である．簡単に要約すれば、凹関数の集合は準凹関数の集合に含まれる．従って、準凹関数に対して成立する定理はすべて凹関数に対しても成立する．

3 関数の最大化問題

3.1 内点解

関数の最大化問題は以下のように定式化される．

$$\max_x \quad f(x) \quad \text{subject to} \quad x \in X$$

ただし、X は R^n のコンパクト部分集合[7]で、 $f : R^n \to R, \quad f \in C^2$ である[8]．

定義 3.1
(1) $x^* \in X$ かつ $f(x^*) \geq f(x),\ \forall x \in X$ のとき、f は x^* で最大値 $f(x^*)$ をとり、x^* は最大解であるという．

[7]集合 X のあらゆる開カバーが有限個の部分カバーをもつとき、集合 X はコンパクトであるという．あらゆるコンパクト集合は閉集合で、有界である．また、X がユークリッド空間の部分集合であるとき、すべての有界閉集合はコンパクトである．
[8]2 階導関数が連続であるような実数関数の集合を C^2 で表記する．

(2) $x^* \in X$ かつ $f(x^*) \geq f(x),\ \forall x \in X \cap N_\epsilon(x^*)$ であるならば、f は x^* で極大解 $f(x^*)$ をとり、x^* を極大解という．ただし、$N_\epsilon(x^*)$ は x^* の近傍 $N_\epsilon(x^*) = \{x \in R^n;\ \|x - x^*\| < \epsilon\}$ である．

(A)　$n = 1$ のケース：$X \subset R$

定理 3.1
$f(x)$ が X の内点 x^* で最大値をとるならば、$f'(x^*) = 0,\ f''(x^*)\underline{<}0$ が成立する．

証明：
$f(x^*) \geq f(x),\ \forall x \in X$ だから、微小な $\delta x (\neq 0)$ に対して

$$f(x^*) \geq f(x^* + \delta x).$$

平均値の定理より、$0 < \theta < 1$ を満たす θ が存在し、

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

$f'(x^*) = 0$ であるから、

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

よって、$\delta x f''(x^* + \theta\delta x)\delta x\underline{<}0.$　$\delta x^2 > 0$ だから、

$$f''(x^* + \theta\delta x)\underline{<}0.$$

f'' は連続で、δx は任意なので、$\delta x \to 0$ とすれば、

$$f''(x^*)\underline{<}0.$$

(QED)

定理 3.2
$f'(x^*) = 0,\ f''(x^*) < 0$ ならば、f は x^* で極大値をとる．

証明：
平均値の定理から、$x^* + \Delta x \in N_\epsilon(x^*)$ となる微小な Δx に対して、

$$f(x^* + \Delta x) = f(x^*) + f'(x^* + \theta\Delta x)\Delta x \tag{1}$$

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

$$\Delta x > 0 \text{ に対して},\ f'(x^* + \theta\Delta x) < 0;\ \Delta x < 0 \text{ に対して},\ f'(x^* + \theta\Delta x) > 0.$$

よって、$f'(x^* + \theta\Delta x)\Delta x < 0.$
式 (1) から、$f(x^* + \Delta x) < f(x^*)$
(QED)

(B)　$n > 1$ のとき： $X \subset X$.

定義 3.2
任意の n 次元ベクトル $x,\ x \neq 0$ に対して、$x^t Ax < 0$ ならば、$n \times n$ 行列 A は負値行列 *(negative definite)* という．任意の n 次元ベクトル $x,\ x \neq 0$ に対して、$x^t Ax \leq 0$ ならば、行列 A は半負値行列 *(negative semidefinite)* という．ここで、x^t は x の転置行列を表す．

定理 3.3 (最大解の必要条件)
$f(x)$ が X の内点 x^* で最大値をとるならば、$\partial f(x^*)/\partial x = 0$ であり、 $\partial^2 f(x^*)/\partial x^2$ は半負値行列である．

証明：
$f(x^*) \geq f(x),\ \forall x \in X$ が成立している．$x^* + h\Delta x \in X$ となる任意のベクトル Δx とスカラー h に対して

$$f(x^*) \geq f(x^* + h\Delta x).$$

平均値の定理から、

$$f(x^* + h\Delta x) = f(x^*) + \frac{\partial f(x^*)}{\partial x} h\Delta x + \frac{1}{2}(h\Delta x)^t \frac{\partial^2 f(x^* + \theta h\Delta x)}{\partial x^2}(h\Delta x).$$

となる θ が存在する[9]．

$$f(x^* + h\Delta x) - f(x^*) = h\frac{\partial f(x^*)}{\partial x}\Delta x + \frac{1}{2}h^2(\Delta x)^t \frac{\partial^2 f(x^* + \theta\Delta x)}{\partial x^2}(\Delta x) \leq 0$$

なので、

$$h\frac{\partial f(x^*)}{\partial x}\Delta x + \frac{1}{2}h^2(\Delta x)^t \frac{\partial^2 f(x^* + \theta\Delta x)}{\partial^2 x}(\Delta x) \leq 0.$$

また、x^* で最大になるので、$\partial f(x^*)/\partial x = 0$ である．したがって、

$$(\Delta x)^t \frac{\partial^2 f(x^* + \theta\Delta x)}{\partial x^2}(\Delta x) \leq 0.$$

(QED)

定理 3.4 (極大解の十分条件)
$\partial f(x^*)/\partial x = 0$ かつ、$\partial^2 f(x^*)/\partial x^2$ が負値行列ならば、f は x^* で極大値をとる．*(*$\partial^2 f(x^*)/\partial x^2$ をヘッセ行列という．*)*

証明：
$x^* + h\Delta x \in X$ となる任意のベクトル Δx とスカラー h に対して,

$$f(x^* + h\Delta x) = f(x^*) + \frac{\partial f(x^*)}{\partial x} h\Delta x + \frac{1}{2}(h\Delta x)^t \frac{\partial^2 f(x^* + \theta h\Delta x)}{\partial x^2}(h\Delta x).$$

$\partial f(x^*)/\partial x = 0$ だから、$(\partial f(x^*)/\partial x)\Delta x = 0$. ヘッセ行列は負値であるから、

$$(\Delta x)^t \frac{\partial^2 f(x^* + \theta\Delta x)}{\partial x^2}(\Delta x) < 0.$$

よって、

$$f(x^* + h\Delta x) < f(x^*).$$

[9] 多変数関数の平均値の定理については、第２節を参照してください．

(QED)
注意：極大解の十分条件で、ヘッセ行列が負値であるという条件を、半負値である場合に拡大すると、定理は成り立たない．例えば、$f(x) = x^4$ の例では、$f'(0) = 0,\ f''(0) = 0$ であるが、$x^* = 0$ は極大点ではなく、極小点となっている．

定理 3.5
$n \times n$ 行列 H が負値行列になるための必要十分条件は、行列 H の k 次主要小行列式 (首座行列式) の符号が $(-1)^k,\ k = 1, 2, .., n$ となることである．

(証明は線形代数のテキストを参照ください．)

定理 3.6 (最大解の十分条件)
集合 X は空でない、コンパクト、凸であるとする[10]．目的関数 f は 2 階連続微分可能で、X 上で凹であるとする[11]．このとき、極大解は最大解である．最大解の集合は凸である．さらに、f が厳密に凹であるならば、解はユニークである．

証明：
極大解 x^* が最大解でないとする．このとき、$f(x^0) > f(x^*)$ となる $x^0 \in X,\ x^0 \neq x^*$ が存在する．f は凹だから、任意の $0 < \lambda < 1$ に対して、

$$f(\lambda x^* + (1-\lambda)x^0) \geq \lambda f(x^*) + (1-\lambda)f(x^0) > f(x^*).$$

X は凸だから、

$$\lambda x^* + (1-\lambda)x^0 \in X.$$

λ は任意だから、

$$x' = \lambda x^* + (1-\lambda)x^0 \in N_\epsilon(x^*)$$

とできる．よって、

$$f(x') > f(x^*),\ x' \in N_\epsilon(x^*) \cap X$$

が成立する．これは矛盾である．
最大解の集合を M とする．$x^*, x^{**} \in M$ とする．つまり

$$f(x^*) = f(x^{**}) = \max_{x \in X} f(x).$$

f は凹だから、任意の $0 < \lambda < 1$ に対して、

$$f(\lambda x^* + (1-\lambda)x^{**}) \geq \lambda f(x^*) + (1-\lambda)f(x^{**}) = \max_{x \in X} f(x).$$

よって、

$$f(\lambda x^* + (1-\lambda)x^{**}) = \max_{x \in X} f(x).$$

したがって、

$$\lambda x^* + (1-\lambda)x^{**} \in M.$$

[10]任意の $x, y \in X, x \neq y$ と任意の $\alpha (0 \leq \alpha \leq 1)$ に対して、$\alpha x + (1-\alpha)y \in X$ ならば、X は凸であるという．
[11]任意の $x, y \in X$ と任意の $\alpha, 0 < \alpha < 1$ に対して、$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$ であるならば、関数 f は X 上で凹であるという．等号が成立しないとき、f は厳密に凹であるという．

(解のユニーク性の証明は省略)
(QED)

問題 3.1
以下の関数の極値を、もし存在すれば、求めなさい.

(1). $f(x) = x^3 + 3x^2 + 3x$

(2). $f(x) = \frac{1}{3}x^3 + \frac{1}{2}x^2 + x + 10$

(3). $f(x) = 4x^3 - 6x^2 - 24x + 120$

(4). $f(x) = (4 + x^2)/x$

(5). $f(x) = \frac{1}{5}x^3 - \frac{3}{10}x^2 - 2x + 10$

(6). $f(x) = \sqrt{x^2 + 5}$

問題 3.2
関数 $F(x_1, x_2) = (x_1 - a_1^2 x_1^2)(x_1 - a_2^2 x_2^2)$ の最大値を求めなさい. ただし、$a_1^2 \neq a_2^2$ である.

3.2 利潤最大化問題

企業の生産関数を $y = f(x)$ とする. x は n 次元投入ベクトルで、$x \in R^n$ であるとする. 財の市場価格を p とし、投入要素 i の市場価格を $w_i,\ i = 1, 2, .., n$ とする. 要素価格ベクトルを $w = (w_1, w_2, ..., w_n),\ w \in R_+^n$ と表記する. このとき企業の利潤 θ は

$$\theta(x) = pf(x) - w \cdot x$$

である. 利潤 $\theta(x)$ の最大化を求めよう. 第 1 階の必要条件は

$$\frac{\partial \theta(x)}{\partial x_i} = 0,\ i = 1, 2, ..., n$$

であり、第 2 階の必要条件は

$$\frac{\partial^2 \theta(x)}{\partial x^2} \text{が半負値行列である}$$

となる.
第 1 階の条件は

$$\frac{\partial \theta(x)}{\partial x_i} = p\frac{\partial f(x)}{\partial x_i} - w_i = 0,\ i = 1, 2, ..., n$$

となる. この式を変形すると

$$p\frac{\partial f(x)}{\partial x_i} = w_i,\ i = 1, 2, ..., n.$$

左辺は i 要素の限界生産物価値であり、右辺は i 要素の価格である.

$$\frac{\partial^2 \theta(x)}{\partial x_i \partial x_j} = p\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

なので、

$$\frac{\partial^2\theta(x)}{\partial x^2} = p\frac{\partial^2 f(x)}{\partial x^2}$$

となる．定理 3.3 より、第 2 階の条件は、ヘッセ行列 $\partial^2 f(x)/\partial x^2$ が半負値行列になることである．前節で説明したとおり、関数の凹性とヘッセ行列の間には、以下のような簡潔な関係が成立する．

f が凹関数であるとき、そしてそのときにのみ、関数 $f(x)$ のヘッセ行列 $\partial^2 f/\partial x^2$ は半負値である．ただし、関数 f が厳密に凹関数であるとき、関数 $f(x)$ のヘッセ行列 $\partial^2 f/\partial x^2$ は半負値になるが、必ずしも負値になるとは限らない[12]．他方、そのヘッセ行列が負値であれば、関数 f は厳密に凹である．

したがって、生産関数が凹関数であることが最大解の第 2 階の必要条件になる．

最大解は市場価格 $p,\ w$ に依存するので、最大解を $x^* = x(p,w)$ と表現する．
$n = 1$ のケース：
第 1 階の条件は

$$pf'(x(p,w)) - w = 0.$$

この式を w で微分すると、

$$pf''(x(p,w))\frac{\partial x(p,w))}{\partial w} - 1 = 0$$

が得られる．f が厳密に凹であれば、$f''(x) < 0$ である．したがって、

$$\frac{\partial x(p,w))}{\partial w} = \frac{1}{pf''(x(p,w))} < 0$$

が得られる．生産関数が厳密に凹関数であるならば、要素需要曲線は右下がりとなる．

ここで、一般的なケースを考えよう．この場合、第 1 階の条件は

$$p\frac{\partial f(x(p,w))}{\partial x} - w = 0,\ w \in R^n_+, x \in R^n$$

となる．この式を要素価格ベクトルで微分すると、

$$p\frac{\partial^2 f(x(p,w))}{\partial x^2} \cdot \frac{\partial x(p,w)}{\partial w} - I = 0,$$

ここで、I は単位行列である．これを変形して、

$$\frac{\partial x(p,w)}{\partial w} = p^{-1} \cdot [\frac{\partial^2 f(x(p,w))}{\partial x^2}]^{-1}$$

を得る．生産関数が凹関数であるとき、そのヘッセ行列は半負値である．ヘッセ行列が逆行列を持つと仮定すれば、ヘッセ行列は非負でなければならない．つまり、生産関数は厳密に凹関数でなければならない．ヘッセ行列 $H := \partial^2 f/\partial x^2$ が負値であれば、その逆行列 H^{-1} も負値である．したがって、要素需要の導関数 $\partial x(p,w)/\partial w$ は負値行列である．ヘッセ行列は明らかに対称行列である．対称行列の逆行列も対称行列となるので、$\partial x(p,w)/\partial w$ も対称行列である．

$$dx = \frac{\partial x(p,w)}{\partial w} \cdot dw$$

[12]詳しくは、Mangasarian, Nonlinear Programming の 90 ページを参照のこと．

という関係式を用いると、

$$dw \cdot dx = dw \cdot \frac{\partial x(p,w)}{\partial w} \cdot dw < 0$$

となることが分かる．$dw = (0, 0, .., dw_i, 0, .., 0)$ とするとき、

$$dw_i dx_i < 0$$

が成立することは自明である．この不等式から、i 要素の市場価格の上昇が、i 要素に対する需要量を減少させるという結論を得ることができる．

次に、生産関数が凹ならば、供給曲線は右上がりになることを証明しよう．
利潤が最大になっているとき、企業の生産量は

$$y = y(p,w) = f(x(p,w))$$

利潤は

$$\pi = \pi(p,w) = pf(x(p,w)) - w \cdot x(p,w)$$

と表現される．利潤最大化の条件 $p\partial f(x(p,w))/\partial x - w = 0$ を p で微分すると、

$$\frac{\partial f(x^*)}{\partial x} + p\frac{\partial^2 f(x^*)}{\partial x^2} \cdot \frac{\partial x}{\partial p} = 0$$

となる．生産量に関する恒等式 $f(x(p,w)) = y(p,w)$ を p で微分すると、

$$\frac{\partial f(x^*)}{\partial x}\frac{\partial x}{\partial p} = \frac{\partial y}{\partial p}$$

が得られる．これらをまとめて行列で表現すると、

$$\begin{pmatrix} p\partial^2 f(x^*)/\partial x^2 & 0 \\ \partial f(x^*)/\partial x & -I \end{pmatrix} \begin{pmatrix} \partial x/\partial p \\ \partial y/\partial p \end{pmatrix} = \begin{pmatrix} -\partial f(x^*)/\partial x \\ 0 \end{pmatrix}$$

となる．この方程式の一般解は

$$\frac{\partial x}{\partial p} = -p^{-1}H^{-1}\frac{\partial f(x^*)}{\partial x}$$

および

$$\frac{\partial y}{\partial p} = -p^{-1}(\frac{\partial f}{\partial x})^t \cdot H^{-1} \cdot (\frac{\partial f}{\partial x})$$

となることは容易に分かる．ここで、H は f のヘッセ行列である．生産関数が凹であるならば、ヘッセ行列は半負値行列となるので、その逆行列の存在を仮定すれば、負値となる．したがって、右辺は正となる．つまり、生産関数が厳密に凹であれば、供給曲線は右上がりとなる．

定理 3.7 (ホテリングの補題)
要素価格 w および財価格 p は正であるとする．このとき、利潤関数は以下の性質を満たす．

$$y(p,w) = \frac{\partial \pi(p,w)}{\partial p}, \tag{2}$$

$$x_i(p,w) = -\frac{\partial \pi(p,w)}{\partial w_i}. \tag{3}$$

証明：利潤関数は

$$\pi(p,w) = pf(x(p,w)) - w \cdot x(p,w) \tag{4}$$

であるから、これを w_i で微分すると

$$\begin{aligned}\frac{\partial \pi(p,w)}{\partial w_i} &= p\frac{\partial f(x^*)}{\partial x}\frac{\partial x(p,w)}{\partial w_i} - x_i(p,w) - w\frac{\partial x(p,w)}{\partial w_i} \\ &= \{p\frac{\partial f(x^*)}{\partial x} - w\}\frac{\partial x(p,w)}{\partial w_i} - x_i(p,w)\end{aligned}$$

となる．第 1 階の条件より、p$\partial f(x^*)/\partial x = w$　だから、

$$\frac{\partial \pi(p,w)}{\partial w_i} = -x_i(p,w).$$

(4) を p で微分すると、

$$\begin{aligned}\frac{\partial \pi(p,w)}{\partial p} &= f(x^*) + p\frac{\partial f(x^*)}{\partial x}\frac{\partial x(p,w)}{\partial p} - w\frac{\partial x(p,w)}{\partial p} \\ &= f(x^*) + \{p\frac{\partial f(x^*)}{\partial x} - w\}\frac{\partial x(p,w)}{\partial p}.\end{aligned}$$

よって、

$$\frac{\partial \pi(p,w)}{\partial p} = f(x^*) = y(p,w).$$

(QED)

また、利潤関数は以下の性質を満たすことも知られている[13]．

(1). $p' \geq p,\ w' \leq w$ ならば、$\pi(p',w') \geq \pi(p,w)$ である．

(2). 任意の $t > 0$ に対して、$\pi(tp,tw) = t\pi(p,w)$ である．

(3). $(p'',w'') = (tp + (1-t)p', tw + (1-t)w'),\ 0 \leq t \leq 1$ ならば、

$$\pi(p'',w'') \leq t\pi(p,w) + (1-t)\pi(p',w')$$

が成立する．

(4). $p > 0, w > 0$ であるとき、$\pi(p,w)$ は (p,w) に関して連続である．

(証明は省略する)

問題 3.3

企業の生産関数が $f(x_1,x_2) = x_1^\alpha x_2^\beta$ であるとする．

(1). 利潤を最大にする生産要素の組がユニークに定まるための条件を $\alpha,\ \beta$ を用いて表現しなさい．

(2). 利潤最大化のための第 *1* 階の条件を具体的に表現しなさい．

[13] Varian, Microeconomic Analysis、第 1 章を参照のこと．

(3). 利潤最大化のための第 *1* 階の条件を要素価格で微分して、要素価格の上昇が生産要素需要量および生産量にどのような影響を与えるか考察しなさい．

(4). 利潤最大化のための第 *1* 階の条件を生産物価格で微分して、生産物価格の上昇が要素需要量および生産量にどのような影響を与えるか考察しなさい．

問題 3.4
企業の生産関数が $f(x_1, x_2) = (x_1^\rho + x_2^\rho)^{1/\rho}$ であるとする．

(1). 利潤最大化のための第 *1* 階の条件を具体的に表現しなさい．

(2). 利潤最大化のための第 *1* 階の条件を要素価格で微分して、要素価格の上昇が生産要素需要量および生産量にどのような影響を与えるか考察しなさい．

(3). 利潤最大化のための第 *1* 階の条件を生産物価格で微分して、生産物価格の上昇が要素需要量および生産量にどのような影響を与えるか考察しなさい．

4　等式制約の下での最大化問題:ラグランジュ乗数法

以下のような等号制約条件の下での関数の最大化問題を考える．

$$\max\ f(x) \text{ subject to } g_i(x) = b_i,\ i = 1, 2, .., m,$$

ただし、$f : R^n \longrightarrow R,\ f \in C^2,\quad g_i : R^n \longrightarrow R,\ g_i \in C^2,$　で、b_i は実数である．制約条件を満たす x の集合 X は

$$X := \{x \in R^n; g_i(x) = b_i,\ i = 1, 2.., m\}$$

と与えられる．この集合を機会集合または実行可能集合という．以下での表記法を簡単化するために、以下のような表記を用いる．ベクトルは縦ベクトルとする．変数 x は n 次元ベクトルだから、$x^t := (x_1, x_2, .., x_n)$ である．また、$g^t := (g_1, g_2, .., g_m)$ と表記する．微分記号について、

$$D_i f := \frac{\partial f}{\partial x_i},\ D_{ij} f := \frac{\partial^2 f}{\partial x_i \partial x_j},\ (Df)^t := (D_1 f, D_2 f, .., D_n f)$$

を採用する．

定理 4.1 (陰関数の定理)
関数 g は $W \subset R^{n+m}$ から R^m への写像で、$g \in C^1$ なる関数である．$g(x, y) = z,\ x \in R^n, y \in R^m,\ z \in R^m$ という表現を持っている．$g(a, b) = 0$ であるとする．$m \times m$ 行列

$$M = \begin{pmatrix} D_{n+1}g_1(a,b) & D_{n+2}g_1(a,b) & \dots & D_{n+m}g_1(a,b) \\ D_{n+1}g_2(a,b) & D_{n+2}g_2(a,b) & \dots & D_{n+m}g_2(a,b) \\ \vdots & \vdots & \dots & \vdots \\ D_{n+1}g_m(a,b) & D_{n+2}g_m(a,b) & \dots & D_{n+m}g_m(a,b) \end{pmatrix}$$

の行列式がゼロでないとする．このとき、a の開近傍 U の点 x に対して $(x \in U \subset R^n)$、

$$g(x, h(x)) = 0,\ y = h(x) \in V$$

となるようなユニークな関数 $h(x)$ と b の近傍 V $(V \in R^m)$ が存在する．

証明は Hoffman, Analysis in Euclidean Space, p.402. を参照ください．

例 4.1
最大化問題で $n=2, m=1$ のケースを考えよう．問題は

$$\max_{x_1,x_2} f(x_1,x_2) \ s.t.\ g(x_1,x_2)=b.$$

制約条件を全微分すると、

$$dg = D_1 g \cdot dx_1 + D_2 g \cdot dx_2 = 0.$$

$D_2 g \neq 0$ とすると、

$$\frac{dx_2}{dx_1} = -\frac{D_1 g}{D_2 g}.$$

陰関数の定理から、$g(x_1,x_2)=b$ を x_2 について解いて、

$$x_2 = h(x_1)$$

とおける．よって、

$$\frac{dx_2}{dx_1} = \frac{dh(x_1)}{dx_1} = -\frac{D_1 g}{D_2 g}.$$

また、$x_2 = h(x_1)$ を目的関数に代入すると、

$$f(x_1,x_2) = f(x_1,h(x_1)) =: H(x_1)$$

なので、最大化問題から制約条件が消去されたことになる．$H(x_1)$ の最大解を x_1^* とすると、最大解の第 *1* 階の条件は $H'(x_1^*)=0$ だから、

$$H'(x_1^*) = D_1 f(x_1^*, h(x_1^*)) + D_2 f(x_1^*, h(x_1^*)) h'(x_1^*) = 0$$

が成立する．したがって、$x_2^* = h(x_1^*)$ を代入し、$dx_2/dx_1 = h'(x_1^*) = -D_1 g/D_2 g$ を用いれば、

$$D_1 f(x_1^*, h(x_1^*)) - D_2 f(x_1^*, h(x_1^*)) \cdot D_1 g(x_1^*, x_2^*)/D_2 g(x_1^*, x_2^*) = 0.$$

これを変形すると、

$$D_1 f(x_1^*, x_2^*) - \frac{D_2 f(x_1^*, x_2^*)}{D_2 g(x_1^*, x_2^*)} D_1 g(x_1^*, x_2^*) = 0$$

が得られる．ここで、

$$\lambda = \frac{D_2 f(x_1^*, x_2^*)}{D_2 g(x_1^*, x_2^*)}$$

とおくと、

$$D_1 f(x_1^*, x_2^*) - \lambda D_1 g(x_1^*, x_2^*) = 0 \tag{5}$$

が成立する．また、

$$D_2 f(x_1^*, x_2^*) - D_2 f(x_1^*, x_2^*) \frac{D_2 g(x_1^*, x_2^*)}{D_2 g(x_1^*, x_2^*)} = 0$$

が常に成り立つことは自明である．この式に λ を代入すると、

$$D_2 f(x_1^*, x_2^*) - \lambda D_2 g(x_1^*, x_2^*) = 0 \tag{6}$$

が得られる．

次の関数 L

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda(b - g(x_1, x_2)) \tag{7}$$

を新たに定義すると、*(5)* と *(6)* はそれぞれ以下のようにコンパクトに表現される．

$$D_1 L(x_1^*, x_2^*, \lambda) = 0,$$
$$D_2 L(x_1^*, x_2^*.\lambda) = 0.$$

ここで新しく定義された関数 $L(x_1, x_2, \lambda)$ をラグランジュ関数、変数 λ をラグランジュ乗数という．

定理 4.2 (第 1 階の必要条件)

$n > m$ とする．$g(x^*)$ のヤコビ行列の階数が m であるとする[14]．このとき、x^* が最大化問題の解であるならば、m 次元ベクトル λ が存在して、

$$\begin{aligned} L(x_1, x_2, \lambda) &= f(x_1, x_2) + \lambda(b - g(x_1, x_2)), \\ \frac{\partial L(x^*, \lambda^*)}{\partial x} &= 0, \\ \frac{\partial L(x^*, \lambda^*)}{\partial \lambda} &= 0 \end{aligned}$$

が成立する．

(証明)

ヤコビ行列は m ヶの独立な列ベクトルをもつので、陰関数の定理から、制約式 $g(x) = b$ において、

$$x = (x^1, x^2),\ x^1 \in R^{n-m}, x^2 \in R^m$$

とおくと、

$$x^2 = h(x^1),\ dx^2 = \frac{\partial h(x^1)}{\partial x^1} \cdot dx_1$$

と書ける．ここで、

$$f(x) = f(x^1, x^2) = f(x^1, h(x^1)) =: H(x^1)$$

であるから、

$$\frac{\partial H(x^1)}{\partial x^1} = \frac{\partial f(x^1, h(x^1))}{\partial x^1} + \frac{\partial f(x^1, h(x^1))}{\partial x^2} \cdot \frac{\partial h(x^1)}{\partial x^1} = 0$$

が得られる．また、$g(x^1, x^2) = b$ より、

$$\frac{\partial g}{\partial x^1} dx^1 + \frac{\partial g}{\partial x^2} dx^2 = 0.$$

よって、

$$dx^2 = -(\frac{\partial g}{\partial x^2})^{-1} \frac{\partial g}{\partial x^1} dx^1,$$

つまり

$$\frac{\partial h(x^1)}{\partial x^1} = -(\frac{\partial g}{\partial x^2})^{-1} \frac{\partial g}{\partial x^1}$$

[14] 関数 g をベクトル x で微分して得られる $m \times n$ 行列 $\partial g(x)/\partial x$ をヤコビ行列 (Jacobian matrix) という．ヤコビ行列を A とするとき、その特性方程式 $|A - \mu I| = 0$ の解 (つまり特性根 μ) のうちゼロでない特性根の個数を行列 A の階数という．

が成立している．従って、最大解の第 1 階の条件は

$$\frac{\partial f(x^1, h(x^1))}{\partial x^1} - \frac{\partial f(x^1, h(x^1))}{\partial x^2} \cdot (\frac{\partial g}{\partial x^2})^{-1} \frac{\partial g}{\partial x^1} = 0$$

となる．ここで、

$$\lambda = \frac{\partial f(x^1, x^2)}{\partial x^2} (\frac{\partial g}{\partial x^2})^{-1}$$

とおくと、

$$\frac{\partial f(x^1, x^2))}{\partial x^1} - \lambda \frac{\partial g(x^1, x^2)}{\partial x^1} = 0.$$

(QED)

定理 4.3 (第 2 階の必要条件)
x^* が最大化問題の解であるならば、ラグランジュ関数のヘッセ行列は

$$\frac{\partial g}{\partial x} \cdot h = 0$$

を満たすようなベクトル h に対して半負値、つまり、

$$h^t \frac{\partial^2 L}{\partial x^2} h \leq 0$$

である．さらに、$(\partial g/\partial x) \cdot h = 0$ である h に対して、$h^t(\partial^2 L/\partial x^2)h < 0$ ならば、第 *1* 階の必要条件は十分条件となる．そのとき解はユニークである．

(証明は省略)

補題 4.1
A を $n \times n$ ヘッセ行列、B を $n \times m$ 行列とする．$Bh = 0$ のもとで、$h^t A h < 0$ となるための必要十分条件は、縁付ヘッセ行列

$$H = \begin{pmatrix} 0 & B \\ B^t & A \end{pmatrix}$$

の最後の $n - m$ 個の首座小行列式の符号が交代することである．ただし、最初の符号は $(-1)^{m+1}$ でなければならない．つまり、

$$d_k = \begin{pmatrix} 0 & B_k \\ (B_k)^t & A_k \end{pmatrix}, \quad (-1)^k d_k > 0,\ k = m+1, m+2, .., n$$

となることである．ここで、B_k は行列 B の最初の k 行からなる行列、A_k は行列 A の最初の k 行と k 列からなる行列である[15]．

(証明は省略)

この補題をラグランジュ関数に適用すると、以下の定理が得られる．

[15] 縁付ヘッセ行列を

$$H = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix}$$

と定式化するテキストも多いが、条件は同じになる．

定理 4.4

mヶのベクトル $\{Dg_i(x^*), i = 1, 2, .., m\}$ が線形独立である．$Dg(x^*) \cdot h = 0$ をみたすすべてのベクトル $h, h \in R^n$ に対して、$h^t(\partial^2 L/\partial x^2)h < 0$ となるための必要十分条件は L の縁付ヘッセ行列

$$H = \begin{pmatrix} 0 & (\partial g/\partial x)^t \\ \partial g/\partial x & \partial^2 L/\partial x^2 \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & D_1 g_1 & \cdots & D_n g_1 \\ 0 & \ldots & 0 & \ldots & \ldots & \ldots \\ 0 & \cdots & 0 & D_1 g_m & \cdots & D_n g_m \\ D_1 g_1 & \cdots & D_1 g_m & D_{11} L & \cdots & D_{1n} L \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ D_n g_1 & \cdots & D_n g_m & D_{n1} L & \cdots & D_{nn} L \end{pmatrix}$$

の最後の $n - m$ 個の首座小行列式の符号が交代することである．ただし、最初の符号が $(-1)^{m+1}$ となる．つまり、最後の $n - m$ 個の首座小行列式 d_k の符号が $(-1)^k d_k > 0,\ k = m + 1, m + 2, \cdots, n$ をみたすことである．

例 4.2

以下の問題を考える．

$$\max\ \ f(x) = x_1 + x_2\ \ s.\ t.\ \ \ x_1^2 + x_2^2 = 1.$$

この場合、$n = 2, m = 1$ である．ラグランジュ関数は

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

であるから、第 *1* 階の条件は

$$\partial L/\partial x_1 = 1 - 2\lambda x_1 = 0$$
$$\partial L/\partial x_2 = 1 - 2\lambda x_2 = 0$$
$$\partial L/\partial \lambda = 1 - x_1^2 - x_2^2 = 0$$

となる．上の第 *1* 式と第 *2* 式から、

$$\lambda = \frac{1}{2x_1} = \frac{1}{2x_2}$$

なので、$x_1 = x_2$ である．これを第 *3* の式に代入すると、$x_1 = \pm\frac{1}{\sqrt{2}}$ が得られる．よって第 *1* 階の条件をみたす解は

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \text{ または } \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

である．しかし、第 *2* 階の条件をみたす解はどれでしょうか．

$$\frac{\partial^2 L}{\partial x^2} = \begin{pmatrix} -2\lambda & 0 \\ 0 & -2\lambda \end{pmatrix}, \quad \frac{\partial g}{\partial x} = (2x_1, 2x_2)$$

だから、

$$\langle \partial g/\partial x, h \rangle = 2x_1 h_1 + 2x_2 h_2 = 0,$$
$$h^t(\partial^2 L/\partial x^2)h = -2\lambda(h_1^2 + h_2^2) \leq 0$$

を満たさなければならない．この条件をみたすのは、$\lambda > 0$ のときである．よって、$x_1 > 0, x_2 > 0$ でなければならない．この最大化問題の解はしたがって、$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ だけである．

同じ結論を縁付ヘッセ行列を用いて導出しよう．ラグランジュ関数の縁付ヘッセ行列は

$$H = \begin{pmatrix} 0 & 2x_1 & 2x_2 \\ 2x_1 & -2\lambda & 0 \\ 2x_2 & 0 & -2\lambda \end{pmatrix}$$

である．$n - m = 2 - 1 = 1$ だから、考慮すべき首座行列式は一つしかない．

$$d_{m+1} = d_2 = \begin{vmatrix} 0 & 2x_1 & 2x_2 \\ 2x_1 & -2\lambda & 0 \\ 2x_2 & 0 & -2\lambda \end{vmatrix}$$

だから、

$$d_2 = -2x_1 \begin{vmatrix} 2x_1 & 0 \\ 2x_2 & -2\lambda \end{vmatrix} + 2x_2 \begin{vmatrix} 2x_1 & -2\lambda \\ -2x_2 & 0 \end{vmatrix} = 8\lambda(x_1^2 + x_2^2).$$

$(-1)^{m+1} d_{m+1} = (-1)^2 d_2 > 0$ より、$d_2 > 0$ でなければならない．よって、$\lambda > 0$ が成立しなければならない．

定理 4.5 (包絡面の定理)
最適化問題

$$\max \ f(x, a) \ s.\ t.\ g(x, a) = 0, f : R^n \times R \to R,\ g : R^n \times R \to R^m$$

の解を $x^* = x(a)$ とする．さらに、$M(a) = f(x(a), a)$ とする．このとき、ラグランジュ関数を

$$L(x, \lambda, a) = f(x, a) - \lambda \cdot g(x, a)$$

とおけば、

$$\frac{dM(a)}{da} = \frac{\partial f(x(a), a)}{\partial a} - \lambda \cdot \frac{\partial g(x(a), a)}{\partial a} = \frac{\partial L(x(a), \lambda, a)}{\partial a}$$

が成立する．

証明：
第 1 階の必要条件は、

$$\frac{\partial L(x, \lambda, a)}{\partial x} = \frac{\partial f(x, a)}{\partial x} - \lambda \cdot \frac{\partial g(x, a)}{\partial x} = 0$$
$$\frac{\partial L(x, \lambda, a)}{\partial \lambda} = -g(x, a) = 0$$

である．最大値 $M(a) = f(x(a), a)$ を a で微分すると、

$$\begin{aligned} \frac{dM(a)}{da} &= \frac{\partial f(x(a), a)}{\partial x} \cdot \frac{\partial x(a)}{\partial a} + \frac{\partial f(x(a), a)}{\partial a} \\ &= \sum_{i=1}^{n} \frac{\partial f(x(a), a)}{\partial x_i} \cdot \frac{\partial x_i(a)}{\partial a} + \frac{\partial f(x(a), a)}{\partial a} \end{aligned} \tag{8}$$

が得られる．第 1 階の条件から、

$$\frac{\partial f(x, a)}{\partial x_i} = \lambda \cdot \frac{\partial g(x, a)}{\partial x_i}, i = 1, 2, .., n. \tag{9}$$

制約条件式を a で微分すると、

$$\frac{\partial g(x,a)}{\partial x} \cdot \frac{\partial x(a)}{\partial a} + \frac{\partial g(x,a)}{\partial a} = 0. \tag{10}$$

(9) 式を (8) 式に代入すると、

$$\frac{dM(a)}{da} = \lambda \sum_{i=1}^{n} \frac{\partial g(x(a),a)}{\partial x_i} \cdot \frac{\partial x_i(a)}{\partial a} + \frac{\partial f(x(a),a)}{\partial a}$$

が得られる．さらに、この式に (10) を代入すると、

$$\frac{dM(a)}{da} = -\lambda \frac{\partial g(x(a),a)}{\partial a} + \frac{\partial f(x(a),a)}{\partial a}$$

となる．(QED)

問題 4.1

以下の最大化問題に解が存在するならば、解を導出しなさい．

(1). $2x_1 + 3x_2 = 4$ という制約条件のもとで $\exp\{-(x_1^2 + 2x_2^2)\}$ を最大にする．

(2). $x_1 - x_2 = 0$ という制約条件のもとで $\sin x_1 \cdot \cos x_2$ を最大にする．

(3). $(x_1 - 1)^3 - x_2^2 = 0$ という制約条件のもとで $x_1^2 + x_2^2$ を最大にする．

(4). $x_1 + x_2 + x_3 = 6$ という制約条件のもとで $x_1 x_2^2 x_3^3$ を最大にする．

5　不等式制約の下での最適化:キューン・タッカーの条件

非負条件の下での最大化問題

$$\max\ f(x) \text{ subject to } x \geq 0, x \in R^n, f \in C^1$$

を最初に考える．この問題の解を x^* と表記する．$x^* > 0$ のとき、つまり内点解であるとき、第 1 階の必要条件は

$$\frac{\partial f(x^*)}{\partial x} = 0$$

となる．端点解であるとき、例えば、$x_i^* = 0$ $(x_j > 0, j \neq i)$ であるとき、必要条件は

$$\frac{\partial f(x^*)}{\partial x_i} < 0 \text{ かつ } \frac{\partial f(x^*)}{\partial x_i} \cdot x_i^* = 0, \text{ または } \frac{\partial f(x^*)}{\partial x_i} = 0$$

である．この結果をまとめてコンパクトに表現すると、一般的な第 1 階の必要条件は

$$\frac{\partial f(x^*)}{\partial x} \leq 0, \quad \frac{\partial f(x^*)}{\partial x} \cdot x^* = 0$$

と表現できる．

不等式制約下の最大化問題は以下のように一般的に定式化できる．

$$[\text{P1}] \quad \max_x\ f(x) \text{ subject to } g(x) \leq b, x \geq 0$$

ただし、$x \in R^n,\ b \in R^m,\ f \in C^1,\ g \in C^1$ であり、$f : R^n \longrightarrow R,\ g : R^n \longrightarrow R^m$ である．制約条件を満たす機会集合 X は

$$X := \{x \in R^n : g(x) \leq b\}$$

と定義される．

スラック変数 s を

$$s := b - g(x)$$

で定義する．このスラック変数を用いると問題 [P1] は以下の最大化問題と等価になる．

$$[\mathrm{P2}] \quad \max_{(x,s)} \ f(x) \ \mathrm{s.\ t.}\ g(x) + s = b, x \geq 0, s \geq 0.$$

これは等号制約下での最大化問題なので、ラグランジュ関数を用いて解くことができる．λ をラグランジュ乗数とすると、ラグランジュ関数は

$$L(x, s, \lambda) = f(x) + \lambda(b - s - g(x))$$

となるので、第 1 階の必要条件は

$$\begin{aligned}
\frac{\partial L(x^*, s^*, \lambda^*)}{\partial x} &= \frac{\partial f(x^*)}{\partial x} - \lambda^* \frac{\partial g(x^*)}{\partial x} \leq 0 \\
\frac{\partial L(x^*, s^*, \lambda^*)}{\partial x} \cdot x^* &= (\frac{\partial f(x^*)}{\partial x} - \lambda^* \frac{\partial g(x^*)}{\partial x}) \cdot x^* = 0 \\
\frac{\partial L(x^*, s^*, \lambda^*)}{\partial s} &= -\lambda \leq 0 \\
\frac{\partial L(x^*, s^*, \lambda^*)}{\partial s} \cdot s^* &= -\lambda^* s^* = 0 \\
\frac{\partial L(x^*, s^*, \lambda^*)}{\partial \lambda} &= b - s^* - g(x^*) = 0
\end{aligned}$$

$s^* = b - g(x^*)$ だから、

$$\lambda^* s^* = \lambda^*(b - g(x^*)) = 0.$$

これらを整理すると、

$$\begin{aligned}
\frac{\partial f(x^*)}{\partial x} - \lambda^* \frac{\partial g(x^*)}{\partial x} &\leq 0 \\
(\frac{\partial f(x^*)}{\partial x} - \lambda^* \frac{\partial g(x^*)}{\partial x}) \cdot x^* &= 0 \\
\lambda^* &\geq 0 \\
\lambda^*(b - g(x^*)) &= 0 \\
x^* &\geq 0 \\
g(x^*) &\leq b
\end{aligned}$$

となる．ラグランジュ関数を改めて、

$$L(x, \lambda) = f(x) + \lambda(b - g(x)) \tag{11}$$

とおくと、上の条件式は

$$\frac{\partial L(x^*,\lambda^*)}{\partial x} \leq 0 \tag{12}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial x} \cdot x^* = 0 \tag{13}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial \lambda} \geq 0 \tag{14}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial \lambda} \cdot \lambda^* = 0 \tag{15}$$
$$x^* \geq 0 \tag{16}$$
$$\lambda \geq 0 \tag{17}$$

と整理できる．これらをキューン・タッカーの条件という．条件式 (12) と (13) は、ラグランジュ関数 L が変数 x によって最大化されているための第 1 階の必要条件であり、条件式 (14) と (15) はラグランジュ関数 L が変数 λ によって最小化されているための第 1 階の必要条件になっている．(x^*,λ^*) は関数 L の鞍点 (saddle point) になっている．

キューン・タッカーの条件は最大解の必要条件であること、またある条件が満たされれば必要十分条件となることを主張する定理群を以下で説明する．以下の議論では一般性を失うことなく、$b=0$ とおくことにする．

定義 5.1 (スレーターの条件)
$g(x^o)<0, x^o \geq 0$ を満たす x^o が存在するとき、関数 g はスレーターの条件を満たすという．

スレーターの条件が満たされればキューン・タッカーの条件が最大化問題の必要条件となることを保障する定理が以下の定理である．

定理 5.1 (必要条件としてのキューン・タッカー条件)
(1) $f(x)$ は R^n 上の実数値関数、
(2) $g(x)$ は R^n から R^m への写像、
(3) f と g は R^n 上で微分可能であるとする．
また、$x^* \in R^n$ を以下の *2* 条件をみたす有界点であるとする．
(i) $g(x^*) \leq 0$ かつ、ある $h(h \in R^n)$ に対して、$g(x^*) + (\partial g(x^*)/\partial x) \cdot h < 0$ である．
(ii) $g(x) \leq 0$ をみたす任意の $x \in R^n$ に対して、$f(x^*) \geq f(x)$ である．
このとき、ラグランジュ関数を

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

とするとき、

$$\frac{\partial L(x^*,\lambda^*)}{\partial x} = \frac{\partial f(x^*)}{\partial x} - \lambda^* \cdot \frac{\partial g(x^*)}{\partial x} = 0$$

をみたす非負のベクトル $\lambda^* \in R^n$ が存在する．

証明：簡単に証明の要約を説明する．二つの集合

$$K_1 = \{k=(s,z) \in R \times R^n : s \geq o, z \leq 0\}$$

と

$$K_2 = \{k = (r, \lambda) \in R \times R^n : \text{ある } h \text{ に対して}, r \leq \frac{\partial f(x^*)}{\partial x} \cdot h,\ \lambda \leq g(x^*) + \frac{\partial g(x^*)}{\partial x} \cdot h$$

を考える．K_1 は $(0,0)$ を頂点とする凸錘であり、K_2 は $(0, g(x^*))$ を頂点とする凸錘である．K_2 は K_1 の内点を含まない．この事実は証明が必要だが、ここでは省略する．
よって、二つの凸集合を分離する超平面の存在を保証する分離定理が成り立つ．つまり、

$$s\sigma + \mu \cdot y \leq \delta,\ \forall (s, \mu) \in K_1$$

と

$$r\sigma + \lambda \cdot y \geq \delta,\ \forall (r, \lambda) \in K_2$$

を満たすゼロでない $(\sigma, y) \in R \times R^n$ が存在する．$(0,0)$ は集合 K_1 と K_2 の共有点だから、$\delta = 0$．よって、

$$s\sigma + \mu \cdot y \leq 0,\ \forall (s, \mu) \in K_1.$$

ここで、$s = 0, z = 0$ とおくと、

$$\sigma \leq 0, \quad y \geq 0$$

が成立することは容易に理解できる．さらに　$\sigma \neq 0$ を証明できる．ここで、

$$\lambda^* = -\frac{y}{\sigma} \geq 0$$

とおく．一方、

$$r\sigma + \lambda \cdot y \geq 0, \forall (r, \lambda) \in K_2.$$

任意の $h, h \in R^n$ に対して、

$$\frac{\partial f(x^*)}{\partial x} \cdot h \geq r,\ \lambda \geq g(x^*) + \frac{\partial g(x^*)}{\partial x} \cdot h.$$

この関係式を利用して、

$$[\frac{\partial f(x^*)}{\partial x} - \lambda^* \cdot \frac{\partial g(x^*)}{\partial x}]h \leq 0$$

を導出すると証明はほぼ完成する．詳しい証明を省略する．詳しい証明については非線形計画法のテキスト、例えば、O.L. Mangasarian, Nonlinear Programming, McGraw-Hill, 1969 などを参照ください．この定理から以下の系が証明できる．

定理 5.2 (系)
問題 *[P1]* を考える *(*ただし、$b = 0$ とおく*)*．x^* が有界最大解であり、ある $h (h \in R^n), h > -x^*$ に対して、$g(x^*) + (\partial g(x^*)/\partial x) \cdot h < 0$ であるとする．このとき、$L(x, \lambda) = f(x) - \lambda g(x)$ とすれば、

$$\begin{aligned} \frac{\partial L(x^*, \lambda^*)}{\partial x} &\leq 0 \\ \frac{\partial L(x^*, \lambda^*)}{\partial x} \cdot x^* &= 0 \\ \frac{\partial L(x^*, \lambda^*)}{\partial \lambda} \cdot \lambda^* &= 0 \end{aligned}$$

をみたす非負のベクトル $\lambda^* \in R^n$ が存在する．

証明：上記の定理を用いれば、証明は容易に行われる．
以下の定理はキューン・タッカーの定理とも呼ばれる重要な定理である．

定理 5.3 (必要十分条件としてのキューン・タッカー条件)
問題 *[P1]* を考える（ただし、$b = 0$ とおく）．目的関数 f は凹関数であり、g は凸写像であるとする[16]．このとき、

(1). スレーターの条件がみたされるとする．x^0 が *(*有界*)* 最大解であるならば、キューン・タッカー条件が成立するような $\lambda^0 \in R^m_+$ が存在する．

(2). キューン・タッカー条件をみたす $x^0 \in R^n_+$ と $\lambda^0 \in R^m_+$ が存在することは、x^0 が *(*有界*)* 最大解であるための十分条件である．

証明：$g(x^1) < 0$ となる x^1 が存在する．g が凸であることから、

$$g(x^0) + g'(x^0)(x^1 - x^0) \leq g(x^1).$$

$h = x^1 - x^0$ とおくと、

$$g(x^0) + g'(x^0)h \leq g(x^1) < 0.$$

$h > -x^0$ だから、上記の系が成立して、(1)．が成り立つ．
f は凹であるから、任意の $x \in R^n_+$ に対して

$$f'(x^0)(x - x^0) \geq f(x) - f(x^0).$$

g が凸であることから

$$g'(x^0)(x - x^0) \leq g(x) - g(x^0).$$

以上の 2 式から、ある非負の $\lambda^0 \in R^m$ に対して

$$\{f'(x^0) - \lambda^0 g'(x^0)\}(x - x^0) \geq f(x) - f(x^0) - \lambda^0\{g(x) - g(x^0)\}.$$

キューン・タッカーの条件より、

$$\lambda^0 g(x^0) = 0, \quad \{f'(x^0) - \lambda^0 g'(x^0)\}x^0 = 0$$

が成立している．よって、

$$\{f'(x^0) - \lambda^0 g'(x^0)\}x \geq f(x) - f(x^0) - \lambda^0 g(x).$$

また、

$$f'(x^0) - \lambda^0 g'(x^0) \leq 0$$

だから、$x \in R^n_+$ に対して、

$$f(x^0) + \{f'(x^0) - \lambda^0 g'(x^0)\}x \leq f(x^0).$$

ゆえに、

$$f(x^0) \geq f(x) - \lambda^0 g(x).$$

仮定より、

$$g(x) \leq 0, \quad \lambda^0 \geq 0.$$

[16] $g : D \subset R^n \to R^m$ であるとき、(定義域 D は凸集合とする) 任意の $x, y \in D$ と $\alpha, 0 < \alpha < 1$ に対して、$f(\alpha x + (1 - \alpha)y)) \leq \alpha f(x) + (1 - \alpha)f(y)$ であるならば、g は凸写像であるという．

従って、

$$f(x^0) \geq f(x)$$

が成立する．(QED)

この定理から、目的関数 $f(x)$ が凹であり、関数 $g(x)$ が凸である場合、$g(x)$ がスレーターの条件を満たすならば、キューン・タッカーの条件は最大解のための必要十分条件になることが分かった．

定理 5.4 (鞍点条件)
問題 *[P1]* を考える（ただし、$b=0$ とおく）．目的関数 f は凹関数であり、g は凸写像であるとする．$g(x^1)<0$ となる $x^1>0$ が存在するとする．このとき、制約条件 $g(x) \leq 0$ のもとで関数 $f(x)$ を R^n_+ 上で最大化する解 x^0 が満たすべき必要十分条件は、

$$L(x^0,\lambda) \geq L(x^0,\lambda^0) \geq L(x,\lambda^0)$$

が成立するような $\lambda^0 \in R^m_+$ が存在することである．ただし、$L(x,\lambda)=f(x)-\lambda \cdot g(x)$.

証明：g の凸性から、

$$g(x^0)+g'(x^0)(x^1-x^0) \leq g(x^1) < 0.$$

定理 5.1 から、

$$f'(x^0)-\lambda^0 g'(x^0)=0, \quad \lambda^0 g(x^0)=0$$

となる $\lambda^0 \in R^n_+$ が存在する．$\lambda^0 \geq 0$ だから、$L(x,\lambda^0)$ は凹関数である．

$$L(x,\lambda^0)-L(x^0,\lambda^0) \leq \frac{\partial L(x^0,\lambda^0)}{\partial x}(x-x^0)=0.$$

よって、

$$L(x,\lambda^0) \leq L(x^0,\lambda^0).$$

ところで、

$$L(x^0,\lambda)-L(x^0,\lambda^0)=f(x^0)-\lambda g(x^0)-\{f(x^0)-\lambda^0 g(x^0)\}=-\lambda g(x^0)+\lambda^0 g(x^0).$$

キューン・タッカーの条件より、$\lambda^0 g(x^0)=0$ なので、

$$L(x^0,\lambda)-L(x^0,\lambda^0)=-\lambda g(x^0) \geq 0,\ \forall \lambda \in R^m_+$$

が成り立つ．従って、

$$L(x^0,\lambda) \geq L(x^0,\lambda^0) \geq L(x,\lambda^0).$$

十分条件も同様に証明できる．(QED)

キューン・タッカー条件が最大化問題の解の必要 (十分) 条件にならない例が存在する．この例外的な現象は、制約条件式 $g_i(x)=b_i, i=1,2,...,m$ が互いに接しているような場合に起こる．スレーターの条件はこうした例外を排除するために必要な条件である．正規性条件という別の条件を課す議論も行われている．ここではその詳細には立ち入らないことにする[17]．

[17] キューン・タッカー条件が最適解を性格づけるとは必ずしもいえない例外的なケースの詳細については、西村『経済学のための最適化理論入門』71 − 76 頁を参照してください．

例 5.1
次のような具体例を取り上げる[18].

$$\max_{x_1,x_2} f(x_1,x_2) = -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2, \quad \textit{subject to}$$
$$g_1(x_1,x_2) = x_1 + x_2 \leq 1,$$
$$g_2(x_1,x_2) = 8x_1^2 + x_2^2 \leq 2,$$
$$x_1 \geq 0,\ x_2 \geq 0.$$

この問題で、f は厳密に凹で、g_1, g_2 が凸関数であることは容易に分かる．スレーターの条件も成立する．したがって、キューン・タッカーの条件が最適解のための必要十分条件となる．

例 5.2
効用最大化問題：

$$\max\ U(x)\ \textit{subject to} \quad p \cdot x \leq m, x \geq 0$$

を考える．ここで、p は価格ベクトルで、m は所得である．効用関数 $U(x)$ は微分可能で、凹であるとする．キューン・タッカーの定理における条件は満たされている．ラグランジュ関数は

$$L(x,\lambda) = U(x) + \lambda \cdot (m - g(x)).$$

キューン・タッカーの条件は

$$\frac{\partial L(x^*,\lambda^*)}{\partial x} = \frac{\partial U(x^*)}{\partial x} - \lambda^* \cdot p \leq 0 \tag{18}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial x} \cdot x^* = (\frac{\partial U(x^*)}{\partial x} - \lambda^* \cdot p)x^* = 0 \tag{19}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial \lambda^*} = m - p \cdot x^* \geq 0 \tag{20}$$
$$\frac{\partial L(x^*,\lambda^*)}{\partial \lambda^*} \cdot \lambda^* = (m - p \cdot x^*) \cdot \lambda^* = 0 \tag{21}$$
$$x^* \geq 0, \lambda^* \geq 0 \tag{22}$$

である．*(18)* 式と *(19)* 式から、$x_i^* > 0$ のとき $\partial U(x^*)/\partial x = \lambda^* \cdot p_i$ であり、$x_i^* = 0$ のとき $\partial U(x^*)/\partial x \leq \lambda^* \cdot p_i$ である．限界効用が正である限り、$\lambda^* > 0$ でなければならない．*(20)* 式と *(21)* 式から、$\lambda^* > 0$ のとき、$m = p \cdot x$、$\lambda^* = 0$ のとき、$m > p \cdot x$ である．$\lambda^* > 0$ でなければならないので、消費の費用が所得以下になることはありえない．また、例えば、$x_1^* = 0, x_2^* > 0$ であるとき、つまりコーナー解が起きるとき、

$$\partial U/\partial x_1 \leq \lambda p_1, \partial U/\partial x_2 = \lambda p_2$$

が成り立つ．したがって、

$$(\partial U/\partial x_1)/(\partial U/\partial x_2) \leq p_1/p_2$$

が成り立ち、第 *1* 財と第 *2* 財の限界代替率が価格比以下となる．
内点解の場合には限界代替率と価格比は均等するが、端点解のケースでは、不等式となる．後者のケースにはラグランジュ乗数法は適用できない．

問題 5.1
以下の最適化問題を解きなさい．*(*解を幾何学的に描きながら考えるとよい．*)*

[18] この例は、M. D. Intriligator, Mathematical Opitimization and Economic Theory, Prentice-Hall, 1971. による．

(1).

$$\max_{x_1,x_2} f(x_1,x_2) = 6x_1 - 2x_1^2 + 2x_1x_2 - 2x_2^2,\ \textit{subject to}$$
$$g_1(x_1,x_2) = 3x_1 + 4x_2 \le 6,$$
$$g_2(x_1,x_2) = -x_1 + 4x_2^2 \le 2,$$
$$x_1 \ge 0,\ x_2 \ge 0$$

(2).

$$\max_{x_1,x_2} f(x_1,x_2) = 3x_1x_2 - x_2^3,\ \textit{subject to}$$
$$g_1(x_1,x_2) = 2x_1 + 5x_2 \ge 20,$$
$$g_2(x_1,x_2) = x_1 - 2x_2 \ = 5,$$
$$x_1 \ge 0,\ x_2 \ge 0$$

(3).

$$\max_{x_1,x_2} f(x_1,x_2) = x_1 + 2x_2,\ \textit{subject to}$$
$$g_1(x_1,x_2) = 3x_1^2 + x_2^2 \le 1,$$
$$g_2(x_1,x_2) = x_1 - 8x_2 \le -1,$$
$$x_1 \ge 0,\ x_2 \ge 0$$

6　費用最小化問題

以下の費用最小化問題：

$$\min_x\ w \cdot x \quad \text{subject to} \quad f(x) = y,\ w \in R_+^n, x \in R_+^n$$

を考える．ここで、y は所与の生産量で $f(x)$ は生産関数、x は投入要素ベクトル、w は要素価格ベクトルである．この問題は以下の最大化問題と等価である．

$$\max_x -w \cdot x \quad \text{subject to} \quad f(x) = y,\ w \in R_+^n, x \in R_+^n.$$

ラグランジュ関数を $L(x,\lambda) = -w \cdot x + \lambda(y - f(x))$ とおく．ここでは、最適解が内点解になることを前提として、ラグランジュ乗数法で解くことにする．このとき、第 1 階の条件は

$$\frac{\partial L(x,\lambda)}{\partial x} = 0;\ \frac{\partial L(x,\lambda)}{\partial \lambda} = 0$$

となる．つまり、

$$w_i = \lambda D_i f(x);\ y - f(x) = 0.$$

この解を (x^*, λ^*) とすると、x^* と λ^* は要素価格ベクトル w と生産量 y に依存するので、

$$x^* = x(w,y), \lambda^* = \lambda(w,y)$$

と表現できる．したがって、生産量 y を生産するための総費用は

$$C(w,y) := w \cdot x^* = w \cdot x(w,y)$$

と表記できる．これを費用関数という．
第1階の条件から、$w_i = -\lambda D_i f(x), i = 1, 2, .., n$ なので、

$$\frac{w_i}{w_j} = \frac{D_i f(x)}{D_j f(x)}, i, j = 1, 2, .., n, i \neq j.$$

この式は技術的限界代替率が要素価格比に等しくなるという条件を意味する．
第2階の条件は

$$\frac{\partial f(x^*)}{\partial x} \cdot h = 0 \text{ となるすべての } h \in R^n \text{ に対して}, h^t \frac{\partial^2 L(x^*, \lambda^*)}{\partial x^2} h \leq 0$$

で与えられる．$D_{ij} L = -\lambda D_{ij} f$ だから、

$$\frac{\partial^2 L(x^*, \lambda^*)}{\partial x^2} = -\lambda^* \frac{\partial^2 f(x^*, \lambda^*)}{\partial x^2}$$

である．よって、第2階の条件は

$$\frac{\partial f(x^*)}{\partial x} \cdot h = 0 \text{ となるすべての } h \in R^n \text{ に対して}, -\lambda^* \cdot h^t \frac{\partial^2 f(x^*, \lambda^*)}{\partial x^2} h \leq 0$$

となる．$w_i = -\lambda D_i f(x), i = 1, 2, .., n$ なので、$\lambda^* < 0$ でなければならない．よって、

$$h^t \frac{\partial^2 f(x^*, \lambda^*)}{\partial x^2} h \leq 0.$$

この条件は、生産関数 $f(x)$ が凹であるとき、満たされる．これは通常ミクロ経済学の教科書で仮定される条件であり、この場合には最適解の第2階の条件は満たされることになる．生産関数が厳密に凹であるときには、解はユニークに定まることになる．

ラグランジュ乗数の経済的な意味を考えるために、ラグランジュ乗数を $\mu = -\lambda$ とする．ラグランジュ関数は

$$L(x, \mu) = -wx - \mu(y - f(x))$$

と変形される．このとき、第1階の条件は

$$-w + \mu^* \cdot \frac{\partial f(x^*)}{\partial x} = 0 \tag{23}$$

$$f(x^*) = y \tag{24}$$

最適解を $x^* = x(w,y), \mu^* = \mu(w,y)$ と表記する．最大化された目的関数は $-w \cdot x(w,y) \equiv -C(w,y)$ と表現できる．ここで、

$$M(w,y) := -C(w,y)$$

とおくと、包絡面の定理から、

$$\frac{\partial M(w,y)}{\partial y} = \frac{\partial L(x^*, \mu^*)}{\partial y} = -\mu^*$$

が得られる．よって、

$$\frac{\partial C(w,y)}{\partial y} = -\frac{\partial M(x^*, \mu^*)}{\partial y} = \mu^*$$

が成立する．この式は、ラグランジュ乗数 μ^* が生産量が1単位拡大するときの費用の増分、つまり限界費用に等しいことを含意する．

補題 6.1 (シェパードの補題)
費用関数の要素価格による微分は要素需要量である．すなわち、

$$\frac{\partial C(w,y)}{\partial w} = x(w,y).$$

証明：
包絡面の定理から、

$$\frac{\partial M(w,y)}{\partial w} = \frac{\partial L(x^*,\mu^*)}{\partial w} = -x^*$$

が得られる．よって、

$$\frac{\partial C(w,y)}{\partial w} = -\frac{\partial M(x^*,\mu^*)}{\partial w} = x^*$$

が成立する．(QED)

第 1 階の条件式 (23,24) を w で微分すると、

$$-I + \frac{\partial \mu}{\partial w}\cdot\frac{\partial f(x^*)}{\partial x} + \mu^*\frac{\partial^2 f(x^*)}{\partial x^2}\cdot\frac{\partial x}{\partial w} = 0,\ \frac{\partial f(x^*)}{\partial x}\cdot\frac{\partial x}{\partial w} = 0$$

となる．ここで、$L(x,\mu)$ の縁付ヘッセ行列は

$$H = \left(\begin{array}{cc} 0 & \partial f(x^*)/\partial x \\ (\partial f(x^*)/\partial x)^t & \mu^*(\partial^2 f(x^*)/\partial x^2) \end{array}\right)$$

となることは自明である．ヘッセ行列を用いて上記の 2 式を整理すると、

$$H\cdot\left(\begin{array}{c} \partial\mu/\partial w \\ \partial x/\partial w \end{array}\right) = \left(\begin{array}{c} 0 \\ I \end{array}\right). \tag{25}$$

最適解であるための第 2 階の条件は (解がユニークであることを前提として)、

$$\frac{\partial f(x^*)}{\partial x}\cdot h = 0 \text{ となるすべての } h\in R^n \text{ に対して}, h^t\cdot\frac{\partial^2 L}{\partial x^2}\cdot h < 0$$

で与えられる．$n = 2$ の場合に、定理 3.4 を適用してみよう．$n = 2, m = 1$ なので、第 2 階の条件が満たされるためには、縁付ヘッセ行列 H の小行列式が $(-1)^{m+1}d_{m+1} = d_2 > 0$ を満たせばよい．よって、$d_2 = \mathrm{det}H > 0$ が成立する．この条件はまた、逆行列の存在も保証する．式 (25) から、

$$\left(\begin{array}{c} \partial\mu/\partial w \\ \partial x/\partial w \end{array}\right) = H^{-1}\cdot\left(\begin{array}{c} 0 \\ I \end{array}\right)$$

が得られる．
第 1 階条件式 (23,24) を y で微分すると、

$$\frac{\partial \mu}{\partial y}\cdot\frac{\partial f(x^*)}{\partial x} + \mu^*\frac{\partial^2 f(x^*)}{\partial x^2}\cdot\frac{\partial x}{\partial y} = 0,\ \frac{\partial f(x^*)}{\partial x}\cdot\frac{\partial x}{\partial y} = 1$$

となる．この式をマトリックス表示すると、

$$H\cdot\left(\begin{array}{c} \partial\mu/\partial y \\ \partial x/\partial y \end{array}\right) = \left(\begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array}\right) \tag{26}$$

とかける．この式を解いて、

$$\begin{pmatrix} \partial\mu/\partial y \\ \partial x/\partial y \end{pmatrix} = H^{-1}\cdot\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

が得られる．

$n=2$ のときを考える．縁付ヘッセ行列は

$$H=\begin{pmatrix} 0 & D_1f & D_2f \\ D_1f & \mu^*D_{11}f & \mu^*D_{12}f \\ D_2f & \mu^*D_{21}f & \mu^*D_{22}f \end{pmatrix}$$

である．(25) 式は、

$$H\cdot\begin{pmatrix} \partial\mu/\partial w_1 & \partial\mu/\partial w_2 \\ \partial x_1/\partial w_1 & \partial x_1/\partial w_2 \\ \partial x_2/\partial w_1 & \partial x_2/\partial w_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

となる．クラメールの公式から、

$$\frac{\partial x_1}{\partial w_1} = \begin{vmatrix} 0 & 0 & D_2f \\ D_1f & 1 & \mu^*D_{12}f \\ D_2f & 0 & \mu^*D_{22}f \end{vmatrix} /\mathrm{det}H = (-1)^4\begin{vmatrix} 0 & D_2f \\ D_2f & \mu^*D_{22}f \end{vmatrix} /\mathrm{det}H = -(D_2f)^2/\mathrm{det}H < 0$$

が得られる．したがって、$\partial x_1/\partial w_1 < 0$ が成立する．つまり、生産量が与えられるとき、生産要素への需要量はその要素価格の上昇に伴って減少する．同じく、

$$\frac{\partial x_2}{\partial w_1} = \begin{vmatrix} 0 & D_1f & 0 \\ D_1f & \mu^*D_{11}f & 1 \\ D_2f & \mu^*D_{21}f & 0 \end{vmatrix} /\mathrm{det}H = (-1)^5\begin{vmatrix} 0 & D_1f \\ D_2f & \mu^*D_{21}f \end{vmatrix} /\mathrm{det}H = D_1fD_2f/\mathrm{det}H > 0$$

が成立する．

(26) も同様にして解くことができる．(26) から

$$H\cdot\begin{pmatrix} \partial\mu/\partial y \\ \partial x_1/\partial y \\ \partial x_2/\partial y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

クラメールの公式から、

$$\frac{\partial x_1}{\partial y} = \begin{vmatrix} 0 & 1 & D_2f \\ D_1f & 0 & \mu^*D_{12}f \\ D_2f & 0 & \mu^*D_{22}f \end{vmatrix} /\mathrm{det}H = \mu^*(D_2fD_{12}f - D_1fD_{22}f)/\mathrm{det}H$$

が得られる．生産関数は通常、$D_if > 0, D_{ij}f > 0, i \neq j$ と想定される．f が凹であるならば、$D_{22}f \leq 0$ なので、

$$\frac{\partial x_1}{\partial y} > 0$$

となる．生産量の増加は生産要素への投入量を拡大させるという結論が得られる．

費用関数は以下のような性質を持つことが知られている．

(1). $w' \geq w$ ならば、$C(w', y) \geq C(w, y)$,

(2). $C(tw, y) = t(C(w, y), t > 0,$

(3). $C(tw + (1-t)w', y) \geq tC(w, y) + (1-t)C(w', y),\ 0 < t < 1$,

(4). $C(w, y)$ は w に関して連続である.

証明は省略する. 興味のある諸君は、中級レベルのミクロ経済学の教科書, 例えば、H. Varian の Microeconomic Analysis などを参照してください.

問題 6.1
生産関数が $f(x) = x_1^\alpha x_2^{1-\alpha}$ であるとき、費用関数を求めなさい. さらに、シェパードの補題が成立することを示しなさい.

7 効用最大化問題

効用関数を $u(x), x \in R^n, u \in C^2$ とする. i 財の市場価格を p_i で表記すると、市場価格ベクトルは $p := (p_1, p_2, .., p_n) \in R_+^n$ である. 所得を m で表現する. 効用最大化問題は

$$\max_x \ u(x) \text{ subject to } p \cdot x = m$$

と定式化される. 最適解が内点解になることを前提とする. ラグランジュ関数を

$$L(x, \lambda) = u(x) + \lambda(m - p \cdot x)$$

とおくと、最適解の第 1 階の条件は

$$\frac{\partial L(x^*, \lambda^*)}{\partial x} = 0;\ \frac{\partial L(x^*, \lambda^*)}{\partial \lambda} = 0$$

となる. この条件の第 2 番目の関係式は予算制約式そのもである. 第 1 番目の式は

$$D_i u(x^*) = \lambda^* p_i, i = 1, 2, .., n$$

となる. これより、

$$\frac{D_i u(x^*)}{D_j(x^*)} = \frac{p_i}{p_j}, i \neq j$$

が得られる. この式は、i 財に対する j 財の限界代替率が価格比に等しいという条件を意味する. 最適解 x^*, λ^* はパラメータ p, m に依存するので、

$$x^* = x(p, m);\ \lambda = \lambda(p, m)$$

と表現する. $x(p, m)$ はいわゆる需要関数といわれるものである. 最適解を効用関数に代入すると、

$$u(x^*) = u(x(p, m)) = v(p, m)$$

と表現できる. これを間接効用関数という. 間接効用関数は価格ベクトルと所得の関数である.

第 2 階の条件は、

$$p \cdot h = 0 \text{ となるすべての } h \in R^n \text{に対して}, h^t \cdot \frac{\partial^2 L}{\partial x^2} \cdot h \leq 0$$

である．

$$\frac{\partial^2 L(x^*,\lambda^*)}{\partial x^2}=\frac{\partial^2 u(x^*)}{\partial x^2}$$

だから、上の不等式は

$$h^t\cdot\frac{\partial^2 u(x^*)}{\partial x^2}\cdot h\le 0$$

となる．また、第 1 階の条件から、$p=(\lambda^*)^{-1}\partial u(x^*)/\partial x$ だから、第 2 階の条件は

$$\frac{\partial u(x^*)}{\partial x}\cdot h=0\ \text{となるすべての}\ h\in R^n\ \text{に対して},h^t\cdot\frac{\partial^2 u(^*)}{\partial x^2}\cdot h\le 0$$

と整理される．

すでに何度も説明している通り、f が準凹関数であるならば、

$$\frac{\partial f}{\partial x}\cdot h=0\ \text{となるすべての}\ h\in R^n\ \text{に対して},h^t\cdot\frac{\partial^2 f}{\partial x^2}\cdot h\le 0$$

であることが知られている．したがって、消費可能領域は凸集合になることは既に知られているので、$u(x)$ が準凹であれば、最大解は存在する．ミクロ経済学の教科書では通常、効用関数は準凹であると仮定されている．

包絡面の定理から、

$$\frac{\partial v(p,m)}{\partial m}=\frac{\partial L(x^*,\lambda^*)}{\partial m}=\lambda^*$$

が成立する．この式は、所得が 1 単位増加するときの効用の増加分が λ に等しいことを含意する．つまり、λ とは所得の限界効用の大きさである．同じようにして

$$\frac{\partial v(p,m)}{\partial p}=\frac{\partial L(x^*,\lambda^*)}{\partial p}=-\lambda^* x^*$$

が成立することが分かる．よって、

$$\frac{\partial v(p,m)}{\partial p}/\frac{\partial v(p,m)}{\partial m}=-x^*=-x(p,m)$$

を得ることができる．これをロイの公式という．

価格 p と所得 m が変化したとき消費量がどのように変化するかを考察する．第 1 階の条件式をここで再掲すると、

$$\frac{\partial u(x(p,m))}{\partial x}-\lambda(p,m)p=0;\ p\cdot x(p,m)=m$$

である．これらを p で微分すると、

$$\begin{aligned}\frac{\partial^2 u(x^*)}{\partial x^2}\cdot\frac{\partial x(p,m)}{\partial p}-\frac{\partial\lambda(p,m)}{\partial p}\cdot p-\lambda^* I&=0\\ x^*+p\cdot\frac{\partial x(p,m)}{\partial p}&=0\end{aligned}$$

$L(x^*,\lambda^*)$ の縁付ヘッセ行列

$$H=\begin{pmatrix}0 & -p^t\\ -p & \frac{\partial^2 u(x^*)}{\partial x^2}\end{pmatrix}$$

を用いると、

$$H\begin{pmatrix} \frac{\partial\lambda}{\partial p} \\ \frac{\partial x}{\partial p} \end{pmatrix} = \begin{pmatrix} x^* \\ \lambda^* I \end{pmatrix}$$

と整理できる．第２階の条件は、

$$d_2 = \det H > 0$$

である．したがって、行列 H は逆行列を持つ．よって、

$$\begin{pmatrix} \frac{\partial\lambda}{\partial p} \\ \frac{\partial x}{\partial p} \end{pmatrix} = H^{-1} \cdot \begin{pmatrix} x^* \\ \lambda^* I \end{pmatrix}.$$

第１階の条件を m で微分すると、

$$\begin{aligned} \frac{\partial^2 u(x^*)}{\partial x^2} \cdot \frac{\partial x(p,m)}{\partial m} - \frac{\partial\lambda(p,m)}{\partial m} \cdot p = 0 \\ p \cdot \frac{\partial x(p,m)}{\partial m} - 1 = 0 \end{aligned}$$

である．だから、

$$H\begin{pmatrix} \dfrac{\partial\lambda}{\partial m} \\ \\ \dfrac{\partial x}{\partial m} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

得られる．これを解いて、

$$\begin{pmatrix} \dfrac{\partial\lambda}{\partial m} \\ \\ \dfrac{\partial x}{\partial m} \end{pmatrix} = H^{-1} \cdot \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

となる．

次に、以下のような最小化問題を考える．

$$\min_x \ p \cdot x \text{ subject to } u(x) = \bar{u}.$$

これは、効用値を $\bar{u}$ を達成するときの支出を最小化するという問題である．ラグランジュ関数を

$$L(x,\mu) = -p \cdot x - \mu(\bar{u} - u(x))$$

とおくと、最適解の第１階の条件は

$$\frac{\partial L(x^*,\mu^*)}{\partial x} = -p + \mu^* \frac{\partial u(x^*)}{\partial x} = 0 \tag{27}$$

$$\frac{\partial L(x^*,\mu^*)}{\partial \mu} = -(\bar{u} - u(x^*)) \tag{28}$$

であり、$\frac{\partial^2 L(x^*,\mu^*)}{\partial x^2} = \mu^* \frac{\partial^2 u(x^*)}{\partial x^2}$ なので、第２階の条件は,

$$\frac{\partial u(x^*)}{\partial x} \cdot h = 0 \text{ となるすべての } h \in R^n \text{に対して}, h^t \cdot \frac{\partial^2 u(x^*)}{\partial x^2} \cdot h \leq 0$$

である．($\mu^* >$ でなければならないことは第 1 階条件から明らかである．) この条件は効用最大化の条件と同一である．(27,28) から、

$$p_i = \mu^* \frac{\partial u(x^*)}{\partial x_i}, i = 1, 2, .., n;\ u(x^*) = \bar{u}$$

が得られる．最適解 x^*, μ^* は明らかにパラメータ $p, \bar{u}$ に依存しているので、

$$x^* = \phi(p, \bar{u}), \mu^* = \mu(p, \bar{u})$$

と表現できる．関数 ϕ を補償需要関数 (ヒックスの需要関数) という．このとき、最小化された支出は

$$p \cdot x^* = p \cdot \phi(p, \bar{u}) =: e(p, \bar{u})$$

と表現できる．e を最小支出関数という．新しい変数 s_{ij} を

$$s_{ij} = \frac{\partial \phi_i(p, \bar{u})}{\partial p_j}$$

と定義すると、これは効用値一定のもとで、価格 p_j の変化に伴う i 財の消費量の変化を表現している．つまり、代替効果の大きさを表している．

包絡面の定理から、

$$\frac{\partial e(p, \bar{u})}{\partial p_j} = -\frac{\partial L(x^*, \mu^*)}{\partial p_j} = x_j^* = \phi_j(p, \bar{u})$$

が得られる．効用最大化問題の解を $x(p, m)$(通常の需要関数、マーシャルの需要関数ともいう) と表現するとき、

$$\bar{u} = u(x(p, m))$$

とすれば、最小化された支出はちょうど所得に等しい．

$$p \cdot \phi(p, \bar{u}) = m.$$

言い換えると、

$$x(p, m) = \phi(p, u(x(p, m))$$

である．支出関数 $e(p, \bar{u})$ と間接効用関数 $v(p, m)$ との間に、

$$e(p, v(p, m)) = m;\ v(p, e(p, \bar{u})) = \bar{u}$$

という関係が成立することは明らかである．また、以下の関係が成立することも自明である．

$$x_i(p, m) = \phi_i(p, v(p, m)),\ \phi_i(p, \bar{u}) = x_i(p, e(p, \bar{u})), i = 1, 2, \cdots, n.$$

ここで、$\phi_i(p, \bar{u}) = x_i(p, e(p, \bar{u}))$ を p_j で微分すると、

$$\frac{\partial \phi_i(p, \bar{u})}{\partial p_j} = \frac{\partial x_i(p, m)}{\partial p_j} + \frac{\partial x_i(p, m)}{\partial m} \frac{\partial e(p, \bar{u})}{\partial p_j}$$

が得られる．上で見たとおり、$\frac{\partial e(p, \bar{u})}{\partial p_j} = \phi_j(p, \bar{u})$ だから、

$$\frac{\partial \phi_i(p, \bar{u})}{\partial p_j} = \frac{\partial x_i(p, m)}{\partial p_j} + \phi_j(p, \bar{u}) \frac{\partial x_i(p, m)}{\partial m}.$$

よって、

$$\frac{\partial x_i(p,m)}{\partial p_j} = \frac{\partial \phi_i(p,\bar{u})}{\partial p_j} - \phi_j(p,\bar{u})\frac{\partial x_i(p,m)}{\partial m}.$$

この式に、$\phi_j(p,\bar{u}) = x_j(p,m)$ を代入すれば

$$\frac{\partial x_i(p,m)}{\partial p_j} = \frac{\partial \phi_i(p,\bar{u})}{\partial p_j} - x_j(p,m)\frac{\partial x_i(p,m)}{\partial m}$$

が成立することが分かる．この等式はスルツキー方程式と呼ばれている関係式に他ならない．第 1 項が代替効果を表現し、第 2 項が所得効果を表わしている．

問題 7.1

効用最大化問題を $n = 2$ のケースについて、解いてみよう．このとき、縁付ヘッセ行列は

$$H = \begin{pmatrix} 0 & -p_1 & -p_2 \\ -p_1 & D_{11}u & D_{12}u \\ -p_2 & D_{21}u & D_{22}u \end{pmatrix}$$

となる．また、第 2 階の条件から $detH > 0$ とならなければならない．比較静学をクラメールの公式を用いて解いてください．

問題 7.2

桃子はパンとクッキーを消費している．パンの消費量を x、クッキーの消費量を y とする．彼女の効用関数は

$$u(x,y) = 3\log(x) + 2\log(y+2)$$

である．彼女は以下の制約の下で効用を最大化する．

(1) 消費量は非負である、

(2) パンの価格とクッキーの価格はそれぞれ *100* 円で予算額は *1000* 円である、

(3) *1* 単位のパンは *150* カロリー、*1* 単位のクッキーは *150* カロリーを含んでおり、*1550* カロリー以上を消費してはいけない．

この問題を制約条件付最大化問題として定式化し、それを解きなさい．

参考文献

この講義ノートを理解するために必要となるレベルの数学のテキスト・ブック（解析学や関数解析のテキスト類) は数え切れないほど多種類のものが出版されている．最適化理論 (数理計画法) のテキスト類も結構多数の種類の教科書が出ていますが、ここで紹介するものは、著者が学生時代 (日本およびアメリカの大学等で) 教科書として使用した、親しみのある懐かしいテキストです．

(1). 高木　貞治、『解析概論』、改定第 3 版、岩波書店、1961.

(2). 西村　清彦、『経済学のための最適化理論入門』、東京大学出版会、1990.

(3). Kenneth Hoffman, *Analysis in Euclidean Space*, Prentice-Hall, 1975.

(4). Michael D. Intriligator, *Mathematical Optimization and Economic Theory*, Prentice-Hall, 1971.

(5). Olvi L. Mangasarian, *Nonlinear Programming*, Robert E. Krieger Publishing, 1979.

(6). H.L. Royden, *Real Analysis*, Macmillan, 1963.

(7). Akira Takayama, *Mathematical Economics*, Second Edition, Cambridge Univ. Press, 1985.

(8). Hal R. Varian, *Microeconomic Analysis*, W.W. Norton, 1984.

付録：解析学の基礎

A　実数の集合

実数とは有理数と無理数のすべてからなる．有理数は、0 および $\pm a/b$(ただし、a, b は自然数) で表現できる数である．無理数は有理数以外の実数、例えば $\sqrt{2} = 1.4142135\cdots, e = 2.718281828\cdots, \pi = 3.1415926535\cdots$ などである．実数全体からなる集合を通常 R で表記する．

すべての数を A, B の二組に分け、A に属する各数を B に属する各数よりも小さくできたとするとき、このような組分け (A, B) を Dedekind の切断という．A を下組、B を上組という．以下の定理 (Dedekind の定理) が成立することが知られている．

定理 A.1
実数の切断は、下組と上組の境界として、一つの数を確定する．

この定理によって実数の連続性が保証される．つまり、切断 (A, B) が与えられるとき、一つの数 s が存在して、s は A の最大数または B の最小数になる．また、一つの数 s をとるとき、s よりも小さい数を下組に入れ、s よりも大きい数を上組に入れるとき、切断を完成するためには、s 自身も下組または上組に入らなければならない．

定義 A.1
A を R の部分集合とする．

(1). A の任意の $x \in A$ に対して、$x \leq b$ となる実数値 b が存在するとき、A は上に有界であるといい、この b を A の上界 *(upper bound)* という．

(2). A の任意の $x \in A$ に対して、$c \leq x$ となる実数値 c が存在するとき、A は下に有界であるといい、この c を A の下界 *(lower bound)* という．

(3). A が上に有界、下に有界であるとき、A は有界 *(bounded)* という．

例 A.1
正の実数値の集合 $A := \{x \in R; x > 0\}$ を考える．
(1) 上に有界でない．*(2)* 下に有界である．$c \leq 0$ となる実数 c は A の下界である．

補題 A.1 (Weierstrass)
A を R の空でない部分集合であるとする．

(1). A は上に有界であるとする．このとき、A は最小な上界 *(smallest upper bound)* を持つ．

(2). A は下に有界であるとする．このとき、A は最大な下界 *(greatest lower bound)* を持つ．

証明は解析学の教科書、例えば、高木貞治『解析概論』岩波書店発行、5 ページを参照のこと．

例 A.2
正の実数値の集合 $A := \{x \in R; x > 0\}$ は下に有界である．$c{\leq}0$ となる実数 c は A の下界であり、その最大値は 0 である．これは集合 A の最小値ではない．他方、$A := \{x \in R; x \leq 1\}$ は上に有界で、上界は $b > 1$ となるような b である．したがって、最小の上界は 1 である．

A の最小な上界 (上限ともいう) を通常、$\sup A$ と表記し、最大な下界 (下限ともいう) を $\inf A$ と表記する．$A := \{x \in R; x > 0\}$ のとき、$\inf A = 0$、$A := \{x \in R; x \leq 1\}$ のとき、$\sup A = 1$ である．

定義 A.2
実数の数列 $\{x_n\}$ を考える．自然数 n が限りなく増大するとき、$\{x_n\}$ が一定の数 x に限りなく近づくならば、数列 $\{x_n\}$ は x に収束するといい、x を $\{x_n\}$ の極限という．記号では

$$\lim_{n\to\infty} x_n = x$$

と書く．別の表現をすれば、任意の正数 ϵ が与えられるとき、それに対応して一つの番号 N_ϵ が

$$n > N_\epsilon \text{ ならば } |x_n - x| < \epsilon$$

であるように定められるとき、数列 $\{x_n\}$ は x に収束するという．

定理 A.2
$\{x_n\}, \{y_n\}$ が収束するとき、以下の等式が成立する．

(1). $\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n,$

(2). $\lim_{n\to\infty}(x_n \cdot y_n) = (\lim_{n\to\infty} x_n)(\lim_{n\to\infty} y_n),$

(3). $\lim_{n\to\infty}\dfrac{x_n}{y_n} = \dfrac{\lim_{n\to\infty} x_n}{\lim_{n\to\infty} y_n}$, ただし、$y_n \neq 0,\ \lim_{n\to\infty} y_n \neq 0$.

証明は解析学の教科書、例えば、高木貞治『解析概論』7 ページを参照のこと．

定理 A.3
実数の数列 $\{x_n\}$ を単調な数列 (単調増加、もしくは単調減少) であるとする．このとき、数列 $\{x_n\}$ が有界であることは、この点列が収束するための必要十分条件である．

証明は解析学の教科書、例えば、高木貞治『解析概論』7-8 ページを参照のこと．

例 A.3
$a > 0$ ならば、$\lim_{n\to\infty} \sqrt[n]{a} = 1$ である．$a > 1$ であるとき、$\sqrt[n]{a} > 1$ でありかつ、$\sqrt[n]{a} > \sqrt[n+1]{a}$ である．よって、$\lim_{n\to\infty} \sqrt[n]{a}$ は単調減少数列で、1 が下界である．それ故、1 以上の極限を持つ．極限を $\alpha > 1$ とすると、$\alpha > 1 + h, h > 0$ とおける．$\lim_{n\to\infty} \sqrt[n]{a} > 1 + h$ だから、$a > (1 + h)^n$ である．n が限りなく大きくなると、右辺は無限になる．これは矛盾である．よって、極限値は 1 となる．$a < 1$ のとき、$a' = 1/a$ とおくと、同様にして、証明できる．

例 A.4
$x_n = (1+\frac{1}{n})^n$ とすれば、数列 $\{x_n\}$ は単調に増大し、かつ有界である．よって、$\{x_n\}$ は収束する．この極限値を指数定数 e と定義する．

定義 A.3
実数の点列 $\{x_n\}$ を考える．x のあらゆる近傍が無限の個数の $\{x_n\}$ を含むとき、点 x は点列 $\{x_n\}$ の集積点 *(accumulation point)* であるという．

この定義は実数の集合だけでなく、2 次元以上の集合にも適用される．集積点の別の表現を用いれば、任意の $\epsilon > 0$ と任意の正整数 n に対して、$|x - x_k| < \epsilon$ となる $k > n$ が存在するならば、点 x は点列 $\{x_n\}$ の集積点であるという．点列 $\{x_n\}$ が x に収束するとき、x は集積点であるが、その逆は成り立たない．ある集合 A に関して点 x が集積点であるとは、点 x にどれほど近いところにも集合 A に属する点が無数にあることである．ただし、x が集合 A に属するとは限らない．

例 A.5
m, n を任意の自然数とし、点 $(\frac{1}{m}, \frac{1}{n})$ の集合を A とする．$(\frac{1}{m}, 0)$、$(0, \frac{1}{n})$、および $(0, 0)$ は集合 A の集積点である．n が奇数のとき $x_n = 0$、n が偶数のとき $x_n = 1$ とする．0 と 1 は点列 $\{x_n\}$ の集積点である．

問題 A.1
以下の集合の上限あるいは下限を求めなさい．

(1). $A = \{x \in R; x > 0\}$

(2). $A = \{x \in R; x \leq 1\}$

(3). $A = \{x \in R; 1 \leq x < 2\}$

問題 A.2
(1). 数列 $\{1, -1, 1, -1, \cdots\}$ は収束しないことを示しなさい．

(2). 実数の数列 $\{x_n\}(x_n \geq 0)$ が x に収束するならば、$\{\sqrt{x_n}\}$ は $\sqrt{x}$ に収束することを示しなさい．

(3). $\{x_n\}(x_n \in R)$ が x に収束するならば、$s_n = (x_1 + x_2 + x_3 + \cdots + x_n)/n$ は x に収束することを示しなさい．

(4). $x_n = (1+\frac{1}{n})^n$ とすれば、数列 $\{x_n\}$ は単調に増大し、かつ有界であることを示しなさい．

B　ユークリッド空間

n 組の実数 $x_i, i = 1, 2, .., n$ を要素とする n 次元ベクトル $x := (x_1, x_2, ..., x_n)$ からなる集合を R^n と表記する．つまり、

$$R^n := \{(x_1, x_2, ..., x_n) : x_i \in R, i = 1, 2, .., n\}.$$

定義 B.1
$x \in R^n, y \in R^n$ に対して、以下のように加算とスカラー積を定義する．
(1) $x + y = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$
(2) $cx = (cx_1, cx_2, \cdots, cx_n),\ c \in R$.

加算とスカラー積が定義されるとき、R^n はベクトル空間である．つまり、$x, y, z \in R^n$ に対して、

(1). $x + y = y + x$(交換則)

(2). $(x + y) + z = x + (y + z)$(結合則)

(3). $x + 0 = x$ となる 0 がユニークに定まる

(4). $x + (-x) = 0$ となる $-x$ がユニークに定まる

(5). $c(x + y) = cx + cy, c \in R$(ベクトルの分配則)

(6). $(b + c)x = bx + cx, b, c \in R$(スカラーの分配則)

(7). $1 \cdot x = x$

が成立する．

定義 B.2
S を R^n の部分集合とする．
(i) $x, y \in S \Rightarrow x + y \in S,$
(ii) $x \in S, c \in R \Rightarrow cx \in S,$
が成立するとき、S は R^n の部分空間であるという．

この部分空間の定義から、次の性質が成立することが分かる．
S を R^n の部分空間であるとする．$x^1, x^2, \cdots, x^k$ を S に属する k 個のベクトルとするとき、あらゆる線形結合

$$c_1x^1 + c_2x^2 + \cdots + c_kx^k, \forall c_i \in R$$

は S に属する．

R^n 内の任意の k 個のベクトル $x^1, x^2, \cdots, x^k$ をとり、これらの線形結合のすべてを S とする．

$$S := \{x = c_1x^1 + c_2x^2 + \cdots + c_kx^k, \forall c_i \in R\}.$$

S は k 個のベクトル $x^1, x^2, \cdots, x^k$ によって張られた空間であり、R^n の部分空間となる．$x, y \in R^n$ とし、$S = \{tx + (1-t)y; \forall t \in R\}$ とするとき、集合 S は R^n の部分空間でしょうか？S が原点を含まないならば、部分空間の定義を満たさないので、S は部分空間ではない．

定義 B.3
$x, y \in R^n$ であるとき、$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$ を内積という．$x \cdot y$ とも書く．

内積は以下の性質を持つ．$x, y, z \in R^n, c \in R$ に対して、

(1). $\langle x, y\rangle = \langle y, x\rangle,$

(2). $\langle cx + y, z\rangle = c\langle x, y\rangle + \langle y, z\rangle,$

(3). $\langle x, x\rangle \geq 0;\ \langle x, x\rangle = 0$ ならば $x = 0$ である．

証明は簡単だから読者の練習問題とする.

定義 B.4
$x \in R^n$ に対して、$\|x\| = \langle x, x\rangle^{1/2}$ を x のノルムという.

ここで定義されたノルムをユークリッド・ノルムという. ユークリッド・ノルム以外に定義されるノルムも存在する. 例えば、$\|x\| = |x_1| + |x_2| + \cdots + |x_n|$ のようなノルムが定義されている. いずれにしろ、点 x と y の距離は $\|x - y\|$ で測られる. ノルムが定義された空間をノルム空間、距離が導入された空間を距離空間という.

補題 B.1 (コーシー・シュワルツの不等式)
$x, y \in R^n$ のとき、$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ が成立する. 等号は $x = cy, c \in R$ のときに成立する.

証明は簡単だから読者の練習問題とする. $\langle x - cy, x - cy\rangle \geq 0, c \in R$ を展開し、$c = \langle x, y\rangle / \|y\|^2$ とおくと、証明できる.

　ノルムは以下の性質を満たす. $x, y \in R^n, c \in R$ とする.

(1). $\|x\| \geq 0,\ \|x\| = 0$ ならば $x = 0$.

(2). $\|cx\| = |c|\|x\|$.

(3). $\|x + y\| \leq \|x\| + \|y\|$.

性質 (1)、(2) は自明である. (3) は簡単に証明できる.

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$$

だから、コーシー・シュワルツの不等式から、

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

が得られる. よって、(3) が成立する.

問題 B.1
以下の問に答えなさい.

(1). n 次元ユークリッド空間における内積 $\langle x, y\rangle$ は $\langle x, y\rangle = \langle y, x\rangle$ という性質を満たすことを示しなさい.

(2). ユークリッド空間において、コーシー＝シュワルツの不等式 $|\langle x, y\rangle| \leq \|x\|\|y\|$ が成立することを示しなさい.

(3). ユークリッド空間において、３角不等式 $\|x + y\| \leq \|x\|\|y\|$ が成立することを示しなさい.

C　ユークリッド空間の完備性

定義 C.1
$x^o \in R^m,\ \epsilon > 0$ とする．x^o の周りの開球 $B(x^o;\epsilon)$ は

$$B(x^o;\epsilon) = \{x \in R^m; \|x - x^o\| < \epsilon\}$$

と定義される．開球 $B(x^o;\epsilon)$ を含む R^m の部分集合を x^o の近傍という．x^o の近傍を $N_\epsilon(x^o)$ と表現する．

定義 C.2
R^m 内の点列 $\{x^n\}$ を考える．R^m の点 x の任意の近傍が有限個の n を除いて点列 $\{x^n\}$ を含むならば、$\{x^n\}$ は x に収束するという．このとき、$x = \lim_{n\to\infty} x^n$ と表現する．
以下のような代替的な定義も用いられる．任意の正の実数 $\epsilon > 0$ に対して、

$$\|x - x^n\| < \epsilon, \forall n > N_\epsilon$$

となる整数 N_ϵ が存在するとき、点列 $\{x^n\}$ は x に収束するという．

補題 C.1
$\{x^n\} \in R^m, \{y^n\} \in R^m, x = \lim x^n, y = \lim y^n$ とする．このとき、
(1) $cx + y = \lim_{n\to\infty}(cx^n + y^n), c \in R$
(2) $\langle x, y\rangle = \lim_{n\to\infty}\langle x^n, y^n\rangle$
が成立する．

証明：解析学のテキスト, 例えば、Hoffman, Analysis in Euclidean Space などを参照してください．

定理 C.1
R^m の点列 $\{x^n\}$ が収束するための必要十分条件は点列 $\{x^n\}$ の各座標 $x_i^n (i = 1, 2, .., m)$ がそれぞれ収束することである．

証明：解析学のテキスト, 例えば、Hoffman, Analysis in Euclidean Space, p.34 を参照してください．

定義 C.3
R^m の点列 $\{x^n\}$ に対して、$\|x^n\| \le M, n = 1, 2, 3, \cdots$ をみたす実数値 M が存在するとき、点列 $\{x^n\}$ は有界であるという．

　明らかに、すべての収束する点列は有界である．つまり、有界性は収束するための必要条件である．しかし、有界であることは十分条件を構成しない．定理 1.3 によれば、収束するための必要十分条件は点列が単調で、有界であることである．

定義 C.4
任意の $\epsilon > 0$ に対して、

$$\|x^k - x^n\| < \epsilon, \forall k, n > N_\epsilon$$

が成立するような正の整数 N_ϵ が存在するとき、点列 $\{x^n\}$ はコーシー列であるという．

収束する点列は明らかにコーシー列になっている．しかし、コーシー列が収束する点列であるとは限らない．

定理 C.2
R^m に属するすべてのコーシー列は収束する．*(ユークリッド空間の完備性)*

証明：解析学のテキスト, 例えば、Kenneth Hoffman, Analysis in Euclidean Space, Prentice-Hall,1975(学部レベルの教科書) あるいは H.L.Royden, Real Analysis, Macmillan,1963(大学院レベルの教科書) を参照してください．
一般的に、コーシー列が収束するような性質を持つノルム空間を完備空間と呼んでいる．完備なノルム空間をバナッハ空間という．

問題 C.1
(1). 収束する点列はコーシー列になることを証明しなさい．

(2). 点列 $\{x_n\}$ は m 次元ユークリッド空間における点列である．この点列が $\|x_n - x_{n+1}\| < 2^{-n}$ を満たすとき、点列 $\{x_n\}$ は収束することを示しなさい．