第1章　関数の凹凸性と凸集合

関数の凹凸、集合の凸性は高度なミクロ経済学を学ぶために重要な概念です。

1.1　1変数の関数の凹凸

この節では、1 変数の関数の凹凸について学びます。多変数の関数の凹凸に関する殆ど全ては、この節で学ぶ 1 変数の関数の凹凸性の議論から導かれます。

I を $\mathbb{R}$ の区間とします。関数

$$f:\ I \longrightarrow \mathbb{R}$$

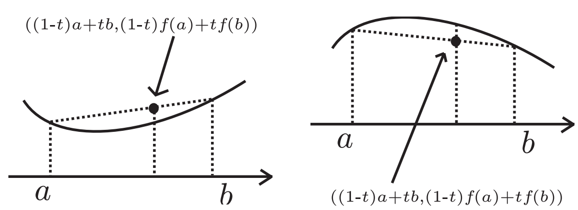
が、任意の $a,\ b \in I$ に対して

$$f((1-t)a+tb) \leq (1-t)f(a)+tf(b) \quad (0 \leq t \leq 1)$$

を満たすとき、f は**凸関数**であるといいます。また $a \neq b$ である任意の $a,\ b \in I$ に対して

$$f((1-t)a+tb) \leq (1-t)f(a)+tf(b) \quad (0 < t < 1)$$

が成立するとき、f は**狭義の凸関数**であるといいます。



次に凹関数を定義します。任意の $a,\ b \in I$ に対して

$$f((1-t)a+tb) \geq (1-t)f(a)+tf(b) \quad (0 \leq t \leq 1)$$

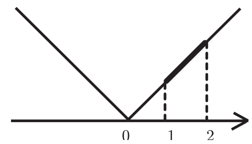
を満たすとき、f は**凹関数**であるといいます。また $a \neq b$ である任意の $a,\ b \in I$ に対して

$$f((1-t)a+tb) \geq (1-t)f(a)+tf(b) \quad (0 < t < 1)$$

が成立するとき、f は**狭義の凹関数**であるといいます。

狭義の凸関数と一般の凸関数の違いを例で考えましょう。$\mathbb{R}$ 上の関数

$$f(x) = |x| = \begin{cases} x & (x \geq 0) \\ -x & (x \leq 0) \end{cases}$$



は凸関数ですが、狭義の凸関数ではありません。実際、$a = 1,\ b = 2$ の場合、

$$f((1-t) \cdot 1 + t \cdot 2) = t$$

$$(1-t) \cdot f(1) + t \cdot f(2) = (1-t) \cdot 1 + t \cdot 2 = t$$

より厳格な不等号'$<$' は $0 < t < 1$ に対して成立しません。

関数 f が開区間 I 上で C^2 級である場合、すなわち 2 階までの微分が I 上の各点で存在して f、f'、f'' が I の各点で連続であるときは、f'' の符号で関数の凹凸を判定することができます。以下、$-\infty \leq a_0 < b_0 \leq b_0$ として、I は開区間 $I = (a_0, b_0)$ とします。そして

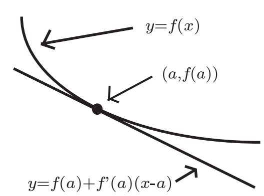
$$f:\ I = (a_0, b_0) \longrightarrow \mathbb{R}$$

は C^2 級と仮定します。

まず次の命題 1.1 を準備とします。

命題 1.1 (1) $f''(t) \geq 0$ が I の各点 t で成立していると仮定します。このとき任意の $a \in I$ に対して

$$f(t) \geq f'(a)(t-a) + f(a) \quad (t \in I)$$



が成立します。

(2) $f''(t) > 0$ が I の各点 t で成立していると仮定します。このとき任意の $a \in I$ に対して

$$f(t) > f'(a)(t-a) + f(a) \quad (t \in I,\ t \neq a)$$

が成立します。

（証明） (1) $f''(t) \geq 0$ が常に成立すると仮定します。Taylor の定理を $t = a$ において用いると

$$f(t) = f(a) + f'(a)(t-a) + \frac{f''(c)}{2}(t-a)^2$$

を満たす c が t と a の間に存在することが分ります。$f''(c) \geq 0$ ですから

$$\frac{f''(c)}{2}(t-a)^2 \geq 0$$

が成立します。従って

$$f(t) = f(a) + f'(a)(t-a) + \frac{f''(c)}{2}(t-a)^2 \geq f(a) + f'(a)(t-a)$$

が導かれます。
(2) (1) の証明をよくみれば、同様に証明できます。**（証明おわり）**

定理 1.1 $f''(t) \geq 0$ が I の各点 t で成立していると仮定します。このとき f は I 上の凸関数となります。

（証明） $x^* = ta + (1-t)b$ $(0 < t < 1)$ と定めます。定理 1.1 を用いると

$$f(a) \geq f(x^*) + f'(x^*)(a - x^*) \tag{1.1}$$

$$f(b) \geq f(x^*) + f'(x^*)(b - x^*) \tag{1.2}$$

を得ます。ここで $t \times (1.1) + (1-t)(1.2)$ から

$$\begin{aligned} & tf(a) + (1-t)f(b) \\ \geq\ & \{tf(x^*) + (1-t)f(x^*)\} + f'(x^*)\{t(a-x^*) + (1-t)(x^*-t)\} \\ =\ & f(x^*) \end{aligned}$$

が従います。**（証明おわり）**

実はこの定理の逆が成立します。

定理 1.2 f が I 上の凸関数ならば $f''(t) \geq 0$ が I の各点 t で成立します。

（証明） f が開区間 $I = (a_0, b_0)$ 上の凸関数であると仮定します。

$$a_0 < a_1 < a_2 < a_3 < b_0$$

であるとき

$$\frac{f(a_2) - f(a_1)}{a_2 - a_1} \overset{(*)}{\leq} \frac{f(a_3) - f(a_1)}{a_3 - a_1} \overset{(**)}{\leq} \frac{f(a_3) - f(a_2)}{a_3 - a_2} \tag{1.3}$$

が成立します[1]。不等式 (*) は次の様に示すことができます。a_2 が区間 $[a_1, a_3]$ を $(1-t) : t$ に内分しているとすると

$$a_2 = tf(a_1) + (1-t)f(a_2)$$

[1] 微分可能性を仮定しなくても、この条件 (1.3) と f が I 上の凸関数であることは同値であることが分かります。注意 1.1 を参照してください。

と表示できます。このとき f は凸関数ですから

$$f(a_2) \leq tf(a_1) + (1-t)f(a_3)$$

が成立します。これを用いると

$$\begin{aligned}\frac{f(a_2)-f(a_1)}{a_2-a_1} &\leq \frac{tf(a_1)+(1-t)f(a_3)-f(a_1)}{\{ta_1+(1-t)a_2\}-a_1} \\ &= \frac{(1-t)(f(a_3)-f(a_1))}{(1-t)(a_3-a_1)} = \frac{f(a_3)-f(a_1)}{a_3-a_1}\end{aligned}$$

から (*) が証明されます。不等式 (**) も同様に示されます。

さて不等式 (*) において $a_2 \to a_1$ とすると

$$f'(a_1) \leq \frac{f(a_3)-f(a_1)}{a_3-a_1} \tag{1.4}$$

が導かれます。そして (**) において $a_2 \to a_3$ とすると

$$\frac{f(a_3)-f(a_1)}{a_3-a_1} \leq f'(a_3) \tag{1.5}$$

が従います。以上から $a_1 < a_3$ ならば $f'(a_1) \leq f'(a_3)$ が成立しますから f' は I 上単調増加です。従って $f'' \geq 0$ が成立します[2]。**（証明おわり）**

命題 1.2 関数

$$f:\ I = (a_0, b_0) \longrightarrow \mathbb{R}$$

が C^1 級の凸関数ならば、$a, b \in I$ に対して

$$f(a) - f(b) \geq f'(b)(a-b) \tag{1.6}$$

が成立します。

証明(1.4) と(1.5) で示したことから $a_0 < a_1 < a_3 < a_0$ ならば

$$f'(a_1) \leq \frac{f(a_3)-f(a_1)}{a_3-a_1} \leq f'(a_3)$$

が分かります。$a_3 - a_1 > 0$ を各辺に掛けて

$$f'(a_1)(a_3-a_1) \leq f(a_3)-f(a_1) \leq f'(a_3)(a_3-a_1)$$

を得ますが、2 番目の不等式は

$$f'(a_3)(a_1-a_3) \leq f(a_1)-f(a_3)$$

[2]ここで、微分可能な関数 $g:\ (a_0, b_0) \longrightarrow \mathbb{R}$ が（広義の）単調増加である、すなわち

$$s < t \implies g(s) \leq g(t)$$

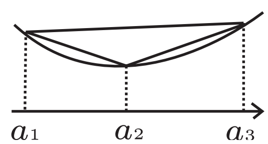
が成立するならば $g'(t) \geq 0$ $(a_0 < t < b_0)$ が成立することを用いました。

と同値です。以上で a と b の大小によらず(1.6) が成立することが示されました。**（証明おわり）**

注意 1.1 条件 (1.3) すなわち、$a_0 < a_1 < a_2 < a_3 < b_0$ を満たす任意の a_1, a_2, a_3 に対して

$$\frac{f(a_2)-f(a_1)}{a_2-a_1} \leq \frac{f(a_3)-f(a_1)}{a_3-a_1} \leq \frac{f(a_3)-f(a_2)}{a_3-a_2}$$

が成立することは、区間 $I=(a_0,b_0)$ 上関数 f が凸であることの十分条件でもあります。実際、(1.3) から



$$\frac{f(a_2)-f(a_1)}{a_2-a_1} \leq \frac{f(a_3)-f(a_2)}{a_3-a_2}$$

が導かれます。$a_2 = ta_1 + (1-t)a_3$ を両辺の分母に代入すると、この不等式は

$$\frac{f(a_2)-f(a_1)}{(1-t)(a_3-a_1)} \leq \frac{f(a_3)-f(a_2)}{t(a_3-a_1)}$$

と変形されます。さらに

$$t(f(a_2)-f(a_1)) \leq (1-t)(f(a_3)-f(a_2))$$

から

$$f(ta_1+(1-t)a_3) = f(a_2) \leq tf(a_1) + (1-t)f(a_3)$$

が導かれます。これは f が凸関数であることを意味します。

2 階の微分係数の符号と狭義の凸関数であることも関係します。定理 1.1 の証明をみれば次の定理も容易に証明できると思います。

定理 1.3 開区間 $I=(a_0,b_0)$ 上の関数 f が 2 階微分可能とします。このとき $f''(t) > 0$ が I の各点 t において成立するならば f は I 上狭義の凸関数であることが分かります。

演習 1.1 **(1)** $f(t) = -\log t$ が凸関数であることを示してください。
(2) $x>0$, $y>0$, α, $\beta>0$ で $\alpha+\beta=1$ が成立するとき

$$x^\alpha y^\beta \leq \alpha x + \beta y$$

が成立することを示してください。

演習 1.2 **(1)** $p > 1$ のとき $f(t) = t^p$ が凸関数であることを証明してください。
(2) $x > 0,\ y > 0,\ \alpha,\ \beta > 0$ で $\alpha + \beta = 1$ であるとき

$$(\alpha x + \beta y)^p \leq \alpha x^p + \beta y^p$$

が成立することを示してください。

演習 1.3 **（Jensen の不等式）** $\mathbb{R}$ の区間 I 上の凸関数 $f:\ I \to \mathbb{R}$ を考えます。
(1) $t_i \geq 0\ (i = 1, \cdots, n)$ が

$$t_1 + \cdots + t_n = 1$$

を満たし、$x_1, \cdots, x_n \in I$ ならば

$$t_1 x_1 + \cdots + t_n x_n \in I$$

であることを証明してください。
(2) 不等式

$$f(t_1 x_1 + \cdots + t_n x_n) \leq \sum_{j=1}^{n} t_j f(x_j)$$

を証明してください。
(3) $x_1, \cdots x_n > 0$ であるとき、相加相乗平均の不等式

$$\frac{x_1 + \cdots + x_n}{n} \leq (x_1 \cdots x_n)^{\frac{1}{n}}$$

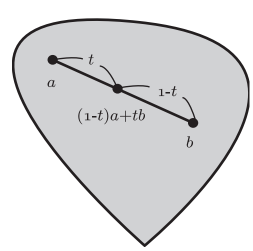
を証明してください。

1.2　凸集合

$\mathbb{R}^n$ の部分集合 C が**凸集合**であるとは、C に任意の 2 点 $a,\ b \in C$ に対して a と b を端点とする線分 $[a, b]$ が C に含まれることを意味します。すなわち

$$[a, b] := \{(1 - t)a + b;\ 0 \leq t \leq 1\} \subset C$$

が成立することを意味します。

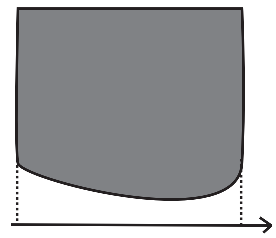


数直線 $\mathbb{R}$ の区間 I 上定義された関数

$$f:\ I \longrightarrow \mathbb{R}$$

があると $\mathbb{R}^2$ 中の集合

$$\mathrm{epi}(f) := \{(y,x);\ y \geq f(x),\ x \in I\}$$



は凸集合です。

実際、$(y_1, x_1),\ (y_2, x_2) \in \mathrm{epi}(f)$ とします。このとき、この 2 点が端点である線分上の任意の点は

$$((1-t)x_1 + tx_2, (1-t)y_1 + ty_2)$$

と $0 \leq t \leq 1$ を満たす $t \in \mathbb{R}$ を用いて表現できます。このとき $y_1 \geq f(x_1)$ と $y_2 \geq f(x_2)$ であることを用いると

$$\begin{aligned} f((1-t)x_1 + tx_2) &\leq (1-t)f(x_1) + tf(x_2) \\ &\leq (1-t)y_1 + ty_2 \end{aligned}$$

より

$$((1-t)x_1 + tx_2, (1-t)y_1 + ty_2) \in \mathrm{epi}(f)$$

が従います。よって $\mathrm{epi}(f)$ は凸集合です。

演習 1.4 (1) C_1 と C_2 が $\mathbb{R}^n$ の凸集合であるとします。このとき C_1 と C_2 の共通部分 $C_1 \cap C_2$ が凸集合であることを証明してください。
(2) C_1 は $\mathbb{R}^n$ の凸集合、C_2 は $\mathbb{R}^m$ の凸集合であるとします。このとき C_1 と C_2 の積集合

$$C_1 \times C_2 := \{(x,y);\ x \in C_1,\ y \in C_2\}$$

が $\mathbb{R}^{n+m}$ 中の凸集合であることを証明してください。

1.3　凸関数・凹関数（多変数の場合）

1.3.1　定義

C を $\mathbb{R}^n$ の凸集合とします。C 上定義された関数

$$f:\ C \longrightarrow \mathbb{R}$$

が**凸関数**であるとは、任意の $a,\ b \in C$ に対して

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b) \quad (0 \leq t \leq 1)$$

が成立することです。C が凸集合ですから、$0 \leq t \leq 1$ のとき $(1-t)a+tb \in C$ となり、$f((1-t)a+tb)$ が定義されていることに注意しましょう。1 変数関数の場合と同様に、f が**狭義の凸関数**であるとは $a \neq b$ を満たす任意の $a,\ b \in C$ に対して

$$f((1-t)a+tb) < (1-t)f(a)+tf(b) \quad (0 < t < 1)$$

が成立することを意味します。

凹関数も同様に定義します。上の f が**凹関数**であるとは任意の $a,\ b \in C$ に対して

$$f((1-t)a+tb) \geq (1-t)f(a)+tf(b) \quad (0 \leq t \leq 1)$$

が成立することです。また、f が**狭義の凹関数**であるとは $a \neq b$ を満たす任意の $a,\ b \in C$ に対して

$$f((1-t)a+tb) > (1-t)f(a)+tf(b) \quad (0 < t < 1)$$

が成立することを意味します。

演習 1.5 C を $\mathbb{R}^n$ の凸集合とします。
(1) このとき

$$f:\ C \longrightarrow \mathbb{R}$$

が凸関数であることと $-f$ が凹関数であることは必要十分です。このことを証明してください。
(2) f と g は C 上の凸関数とします。このとき $f+g$ が凸関数であることを証明してください。

1.3.2　凸関数と凹関数から定まる凸集合

凸関数あるいは凹関数から定まる凸集合について考えます。C を $\mathbb{R}^n$ の凸集合とします。そして

$$f:\ C \longrightarrow \mathbb{R}$$

を凸関数とします。このとき

$$\mathrm{epi}(f) := \{(y,x) \in \mathbb{R} \times C;\ y \geq f(x),\ x \in C\}$$

は凸集合です。これは 1 変数の凸関数の場合と同様に証明できます。また任意の $c_0 \in \mathbb{R}$ に対して

$$\{x \in C;\ f(x) \leq c_0\}$$

は凸集合です。実際、$a,\ b \in C$ が

$$f(a) \leq c_0,\ f(b) \leq c_0$$

を満たしているとすると

$$f((1-t)a+tb) \leq (1-t)f(a)+tf(b) \leq (1-t)c_0+tc_0=c$$

から、線分 $[a,b]$ の点 $(1-t)a+tb$ がこの集合に属することが分かります。

演習 1.6 C が $\mathbb{R}^n$ の凸集合であり、

$$f:\ C \longrightarrow \mathbb{R}$$

が凹関数であるとします。このとき $c_0 \in \mathbb{R}$ に対して

$$\{x \in C;\ f(x) \geq c_0\}$$

が凸集合であることを証明してください。

上で $\mathbb{R}^n$ の凸集合 C の上の凸関数 f に対して

$$\{x \in C;\ f(x) \leq c_0\}$$

が任意の c_0 に対して凸集合であることを示しました。ここでこの性質を満たす関数について定義をします。

定義 1.1 $\mathbb{R}^n$ の凸集合 C の上の関数

$$f:\ C \longrightarrow \mathbb{R}$$

が**準凸関数**であるとは、任意の $c_0 \in \mathbb{R}$ に対して

$$\{x \in C;\ f(x) \leq c_0\}$$

が凸集合であるときです。

この定義から**凸関数は準凸関数である**ことが分かります。ここで準凸関数の定義を少し使いやすい形に書き換えましょう。

命題 1.3 $\mathbb{R}^n$ の凸集合 C の上の関数

$$f:\ C \longrightarrow \mathbb{R}$$

が準凸関数である必要十分条件は、任意の $a,\ b \in C$ に対して

$$f(a) \leq f(b) \Longrightarrow f((1-t)a+tb) \leq f(b) \quad (0 \leq t \leq 1) \tag{1.7}$$

が成立することです。

（証明） $f:\ C \to \mathbb{R}$ が準凸であると仮定します。2 点 $a,\ b \in C$ を任意に取ります。$f(a) \leq f(b)$ とします。このとき $c_0 = f(b)$ として

$$\{x \in C;\ f(x) \leq f(b)\}$$

は凸集合で、両方の点 $a,\ b$ を含みます。このことから $0 \leq t \leq 1$ を満たす $t \in \mathbb{R}$ に対して

$$(1-t)a + tb \in \{x \in C;\ f(x) \leq f(b)\}$$

すなわち

$$f((1-t)a + tb) \leq f(b)$$

が成立します。

逆に(1.7) が任意の 2 点 $a,\ b \in C$ に対して成立すると仮定します。任意の $c_0 \in \mathbb{R}$ をとります。

$$a_0,\ b_0 \in \{x \in C;\ f(x) \leq c_0\}$$

すなわち

$$f(a_0),\ f(b_0) \leq c_0$$

とします。ここで $f(a_0) \leq f(b_0)$ としても一般性を失いません。このとき仮定(1.7) から

$$f((1-t)a_0 + tb_0) \leq f(b_0) \leq c_0$$

が成立します。これは

$$(1-t)a + tb \in \{x \in C;\ f(x) \leq c_0\}$$

に他なりません。以上から $\{x \in C;\ f(x) \leq c_0\}$ が凸集合であることが導かれました。**（証明おわり）**

演習 1.7 C は $\mathbb{R}^n$ の凸集合であるとします。関数

$$f:\ C \longrightarrow \mathbb{R}$$

が**準凹関数**であるとは、任意の $c_0 \in \mathbb{R}$ に対して

$$\{x \in C;\ f(x) \geq c_0\}$$

が凸集合であるときです。

(1) f が準凹関数であることと $-f$ が準凸関数であることは必要十分です。このことを証明してください。

(2) f が準凹関数である必要十分条件は、任意の 2 点 $a,\ b \in C$ に対して

$$f(a) \leq f(b) \Longrightarrow f((1-t)a + tb) \geq f(b) \quad (0 \leq t \leq 1) \tag{1.8}$$

が成立することです。このことを示してください。

1.3.3　凸関数と極大・極小

一般に $\mathbb{R}^n$ の部分集合 X 上で定義された関数

$$f:\ X \longrightarrow \mathbb{R}$$

を考えます。f が $x = x^* \in X$ で**最大値**をとるとは

$$f(x) \leq f(x^*) \quad (\forall x \in X)$$

が成立することです。f が $x = x^* \in X$ で**最小値**をとるとは

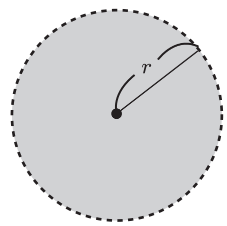
$$f(x) \geq f(x^*) \quad (\forall x \in X)$$

が成立することです。

局所的な (local な) 最大値、最小値も考えることがあります。そのために $r > 0$ に対して中心が x^*、半径が δ である開球

$$B_r(x^*) := \{x \in \mathbb{R}^n;\ ||x - x^*|| < \delta\}$$

を定義します。



関数

$$f:\ X \longrightarrow \mathbb{R}$$

が $x = x^* \in X$ で**極大値**をとるとは、ある $\delta > 0$ が存在して

$$f(x) \leq f(x^*) \quad (\forall x \in X \cap B_\delta(x^*))$$

が成立することです。f が $x = x^* \in X$ で**極小値**をとるとは

$$f(x) \geq f(x^*) \quad (\forall x \in X \cap B_\delta(x^*))$$

が成立することです。

一般には極大と最大、極小と最小は異なるのですが、凸集合上の凸関数と凹関数には著しい性質があります。

定理 1.4 C を $\mathbb{R}^n$ の凸集合として、関数

$$f: \ C \longrightarrow \mathbb{R}$$

を考えます。
(1) f が凸関数であるとします。f が $x = x^* \in X$ で極小値をとるならば、f は $x = x^* \in X$ で最小値をとります。さらに

$$\mathrm{Min}(f) := \{x \in X; \ f \text{ は } x \text{ で最小値をとる }\}$$

は凸集合になります。
(2) f が凹関数であるとします。f が $x = x^* \in X$ で極大値をとるならば、f は $x = x^* \in X$ で最大値をとります。さらに

$$\mathrm{Max}(f) := \{x \in X; \ f \text{ は } x \text{ で最大値をとる }\}$$

は凸集合になります。

（証明） (1) f が $x = x^* \in X$ で極小値をとると仮定します。このとき

$$f(x) \geq f(x^*) \quad (\forall x \in X \cap B_\delta(x^*)) \tag{1.9}$$

を満たす $\delta > 0$ が存在します。

証明は背理法を用い、$f(x^*)$ が最小値でないと仮定して矛盾を導きます。このとき $\tilde{x}$ で

$$f(\tilde{x}) < f(x^*)$$

を満たすものが存在します。C が凸集合で f がその上の凸関数ですから

$$(1-t)x^* + t\tilde{x} \in X \tag{1.10}$$

が $0 < t < 1$ に対して成立して、

$$f((1-t)x^* + t\tilde{x}) \leq (1-t)f(x^*) + tf(\tilde{x})$$

が従います。さらに(1.10) を用いると

$$\begin{aligned} f((1-t)x^* + t\tilde{x}) &\leq (1-t)f(x^*) + tf(\tilde{x}) \\ &< (1-t)f(x^*) + tf(x^*) = f(x^*) \end{aligned}$$

が導かれます。t を十分小さくすると

$$(1-t)x^* + t\tilde{x} \in X \cap B_\delta(x^*)$$

となり(1.9) から

$$f(x^*) \leq f((1-t)x^* + t\tilde{x})$$

が従います。以上から

$$f(x^*) < f(x^*)$$

となり矛盾が導かれました。

次に Min(f) が凸集合であることを証明します。$x^*,\ \tilde{x} \in \mathrm{Min}(f)$ と仮定します。このとき f が凸関数であることから

$$\begin{aligned} f((1-t)x^* + t\tilde{x}) &\leq (1-t)f(x^*) + tf(\tilde{x}) \\ &= (1-t)f(x^*) + tf(x^*) = f(x^*) \end{aligned}$$

が成立します。ここで $f(x^*) = f(\tilde{x})$ であることを用いました。さらに $f(x^*)$ が最小値ですから

$$f(x^*) \leq f((1-t)x^* + t\tilde{x})$$

も成立します。以上から

$$f((1-t)x^* + t\tilde{x}) = f(x^*) = f(\tilde{x}) \tag{1.11}$$

が示され、

$$(1-t)x^* + t\tilde{x} \in \mathrm{Min}(f)$$

が示されました。

(2) 凹関数の場合は、(1) の凸関数と全く同様に証明できます。または f が凹関数の場合、$-f$ が凸関数であることを用いて (1) の結果を適用することによっても証明できます。**（証明おわり）**

つぎに f が狭義の凸関数（resp. 凹関数）である場合に、最小点 (resp. 最大点) が存在すれば一意的であることを証明します。

定理 1.5 $\mathbb{R}^n$ の凸集合 C 上で定義された狭義の凸関数（resp. 狭義の凹関数）

$$f:\ C \longrightarrow \mathbb{R}$$

を考えます。このとき $x^* \in C$ と $\tilde{x} \in C$ で f が最小値（resp. 最大値）をとるならば $x^* = \tilde{x}$ が従います。

（証明） f が狭義の凸関数である場合だけを考えます。(1.11) にあるように $0 \leq t \leq 1$ に対して

$$f((1-t)x^* + t\tilde{x}) = f(x^*) = f(\tilde{x})$$

が従います。もし $x^* \neq \tilde{x}$ ならば、f が狭義の凸関数ですから $0 < 1 < t$ を満たす t に対して

$$f((1-t)x^* + t\tilde{x}) < (1-t)f(x^*) + tf(\tilde{x}) = f(x^*)$$

となり矛盾が生じます。従って $x^* = \tilde{x}$ となります。**（証明おわり）**

1.3.4　微分可能な関数の凹凸

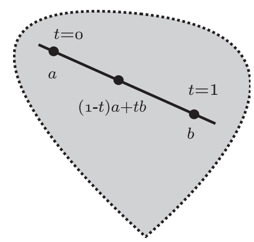
1 変数の C^2 級関数の凹凸が 2 階の微分係数の符号で判定できたように、多変数の関数が C^2 級であるときはその 2 階の偏微分係数を用いて凹凸を判定できます。このことを解説するために U が $\mathbb{R}^n$ の開集合で凸集合であると仮定します。そして

$$f:\ U \longrightarrow \mathbb{R}$$

が C^2 級であると仮定します。まず U の相異なる 2 点 $a,\ b \in U$ を取ります。そして

$$F(t) := f((1-t)a+tb) = f((1-t)a_1+tb_1, \cdots, (1-t)a_n+tb_n)$$

と 1 変数関数 $F(t)$ を定義します。$t=0$ のときが $x=a$ に、$t=1$ がに対応していることにまず注意しましょう。このことから閉区間 $[0,1]$ で $F(t)$ が定義されていることが分かります。さらに、U が開集合ですから a の周りと b の周りが U に含まれていることから、$\delta>0$ を小さくとれば F が $[0,1]$ を含む開区間 $(-\delta, 1+\delta)$ で



$$F:\ (-\delta, 1+\delta) \longrightarrow \mathbb{R} \tag{1.12}$$

と定義されていることも導かれます。

このとき

$$\begin{aligned} F'(t) &= \sum_{i=1}^{n} f_{x_i}((1-t)a+tb)\cdot\frac{d}{dt}\{(1-t)a+tb\} \\ &= \sum_{i=1}^{n} f_{x_i}((1-t)a+tb)\cdot(b_i-a_i) \\ F''(t) &= \sum_{1\le i,j\le n} f_{x_ix_j}((1-t)a+tb)\cdot(b_i-a_i)(b_j-a_j) \end{aligned}$$

と計算されます。ここで n 次元ベクトル

$$\vec{\xi} = {}^t(b_1-a_1, \cdots, b_n-a_n)$$

と $x \in U$ における f の Hesse 行列

$$H(f)(x) = \big(f_{x_ix_j}(x)\big)_{i,j}$$

を定めると、$F''(t)$ は

$$F''(t) = \Big(H(f)((1-t)a+tb)\vec{\xi}, \vec{\xi}\Big)$$

と 2 次形式の形に書くことができます。従って

$$F''(t) \geq 0 \quad (t \in (-\delta, 1+\delta))$$

の必要十分条件は任意の $t \in (-\delta, 1+\delta)$ に対して

$$H(f)((1-t)a+tb) \text{ が非負定値である}$$

ことが分かります。以上の準備から次の定理 1.6 を証明することができます。

定理 1.6 f が凸関数である必要十分条件は任意の $x \in U$ に対して x における f の Hesse 行列が $H(f)(x)$ が非負定値であることです。

（証明） 任意の $x \in U$ に対して $H(f)(x)$ が非負定値であると仮定します。このとき $a \neq b$ である $a, b \in U$ を任意にとり(1.12) の $F(t)$ を定めます。このとき $F''(t) \geq 0$ が $t \in (-\delta, 1+\delta)$ に対して成立します。定理 1.1 より $F(t)$ は凸関数となり

$$F((1-t)\cdot 0 + t\cdot 1) \leq (1-t)F(0) + tF(1) \quad (0 \leq t \leq 1)$$

より

$$f((1-t)a+tb) \leq (1-t)f(a) + tf(b) \quad (0 \leq t \leq 1)$$

が従います。このことから f は U 上の凸関数であることが従います。

逆に f が U 上の凸関数とします。任意の $a \in U$ と n 次元ベクトル $\vec{\eta} \neq \vec{0}$ を任意にとります。このとき

$$b = a + s\vec{\eta} \in U$$

が成立するように $s > 0$ をとります。この a と b に対して(1.12) の $F(t)$ を定めます（このとき $\vec{xi} = s\vec{\eta}$ となります)。すると F が凸関数となりますから

$$F''(t) \geq 0 \quad (t \in (-\delta, 1+\delta))$$

が成立します。特に

$$F''(0) = (H(f)(a)\vec{\eta}, \vec{\eta}) = s^2 \left(H(f)(a)\vec{\xi}, \vec{\xi}\right) \geq 0$$

より

$$\left(H(f)(a)\vec{\xi}, \vec{\xi}\right) \geq 0$$

が従います。以上で $H(f)(a)$ が非負定値であることが証明できました。

（証明おわり）

上の定理 1.6 の証明をみれば、次の定理 1.7 を証明できます。

定理 1.7 $\mathbb{R}^n$ の開凸集合 U 上に定義された C^2 級関数

$$f\colon\ U \longrightarrow \mathbb{R}$$

が狭義の凸関数である十分条件は任意の $a \in U$ において f の Hesse 行列 $H(f)(a)$ が正定値であることです。

1.4　凸集合の分離定理

第2章　不等式制約の最適値問題

2.1　問題の設定

この章では不等式で定まる制約条件の下での最適値問題を考えます。すなわち $\mathbb{R}^n$ の開集合 U の上で定義された関数 f、g_j $(j = 1, 2, \cdots, \ell)$ に対して

$$y = f(x) = f(x_1, \cdots, x_n) \tag{2.1}$$

の最大値を制約条件

$$g_j(x) \geq 0 \ (j = 1, \cdots, \ell) \tag{2.2}$$

の下で求める非線形計画問題 (NLP) を考えます。

2.2　鞍点条件

前節で設定した非線形計画問題 (NLP) に対して Lagrange 関数

$$L(x, \lambda) = f(x) + \sum_{j=1}^{\ell} \lambda_j g_j(x) \tag{2.3}$$

を考えます（ここで $\lambda = (\lambda_1, \cdots, \lambda_\ell)$ としています)。点 $(x^*, \lambda^*) \in U \times (\mathbb{R}_+)^\ell$ が Lagrange 関数 $L(x, \lambda)$ の**鞍点**(*saddle point*) であるとは

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \quad (\forall x \in U) \tag{2.4}$$

$$L(x^*, \lambda^*) \leq L(x^*, \lambda) \quad (\forall \lambda \in (\mathbb{R}_+)^\ell) \tag{2.5}$$

これから Lagrange 関数の鞍点と非線形計画問題 (NLP) の解の関係について調べていきます。次の命題 2.1 はそのための準備にあたります。

命題 2.1 点 $(x^*, \lambda^*) \in U \times (\mathbb{R}_+)^\ell$ が Lagrange 関数 $L(x, \lambda)$ の鞍点であることの必要十分条件は

$$g_j(x^*) \geq 0 \ (1 \leq j \leq \ell), \ \sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) = 0 \tag{2.6}$$

$$「x \text{ の関数 } L(x, \lambda^*) \text{ は } x^* \text{ で最大値をとる}」 \tag{2.7}$$

です。

（証明） (x^*, λ^*) が鞍点であると仮定します。このとき(2.5) から

$$f(x^*) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \le f(x^*) + \sum_{j=1}^{\ell} \lambda_j g_j(x^*)$$

すなわち

$$\sum_{j=1}^{\ell} (\lambda_j - \lambda_j^*) g_j(x^*) \ge 0 \tag{2.8}$$

が従います。もし $g_1(x^*) < 0$ ならば $\lambda_1 > \lambda_1^*$ である λ_1 を用いて $\lambda = (\lambda_1, \lambda_2^*, \cdots, \lambda_\ell^*)$ と(2.8) の左辺に代入すれば

$$\sum_{j=1}^{\ell} (\lambda_j - \lambda_j^*) g_j(x^*) = (\lambda_1 - \lambda_1^*) g_1(x^*) < 0$$

となり(2.8) に矛盾します。したがって $g_1(x^*) \ge 0$ が従います。このようにして

$$g_j(x^*) \ge 0 \quad (1 \le j \le \ell)$$

も成立します。ここで $\lambda_j^* \ge 0$ $(1 \le j \le \ell)$ であることも使うと

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \ge 0$$

が得られます。他方(2.8) において $\lambda = 0$ と代入すると

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \le 0$$

が従います。以上で

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) = 0$$

も示されて、(2.6) を導きました。(2.7) は(2.4) と同値です。

逆に点 $(x^*, \lambda^*) \in U \times (\mathbb{R}_+)^\ell$ が(2.6) と(2.7) を満たすと仮定します。(2.4) は(2.7) と同値ですから、(2.6) を示します。$\lambda = (\lambda_1, \cdots, \lambda_\ell) \in (\mathbb{R}_+)^\ell$ に対して

$$\sum_{j=1}^{\ell} \lambda_j g(x^*) \ge 0$$

が(2.6) の $g_j(x^*) \ge 0$ から従います。これに(2.6) にある

$$\sum_{j=1}^{\ell} \lambda_j^* g(x^*) = 0$$

をも加味して

$$\sum_{j=1}^{\ell} \lambda_j g(x^*) \geq \sum_{j=1}^{\ell} \lambda_j^* g(x^*)$$

が分かります。この両辺に $f(x^*)$ を加えると

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

すなわち(2.6) が導かれます。**（証明おわり）**

この命題 2.1 から、Lagrange 関数の鞍点である条件は非線形最適問題 (NLP) の解の十分条件であることが導かれます。

定理 2.1 $(x^*, \lambda^*) \in U \times (\mathbb{R}_+)^\ell$ が Lagrange 関数の鞍点であると仮定します。このとき x^* は (NLP) の解となります。すなわち x^* は制約条件

$$g_j(x^*) \geq 0 \quad (1 \leq j \leq \ell) \tag{2.9}$$

を満たし、

$$x \in U,\ g_j(x) \geq 0\ (1 \leq j \leq \ell) \Longrightarrow f(x) \leq f(x^*)$$

が成立します。

（証明） 命題 2.1 の(2.6) から(2.9) が成立することが導かれます。他方命題 2.1 の(2.7) から

$$L(x, \lambda^*) \leq L(x^*, \lambda^*)$$

を得ます。これと(2.6) から

$$f(x) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x) \leq f(x^*)$$

が従います。また $\lambda^* \in (\mathbb{R}_+)^\ell$ であることと $x \in U$ が制約条件を満たすことから

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x) \geq 0$$

も成立しますから

$$f(x) \leq f(x) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x) \leq f(x^*)$$

が従います。**（証明おわり）**

次に、U が凸集合で、$f(x)$、$g_j(x)$ が U 上の凹関数であるという前提の下で、$(x^*, \lambda^*) \in U \times (\mathbb{R}_+)^\ell$ が Lagrange 関数の鞍点であることが x^* が非線形計画問題 (NLP) の解である必要条件であることを示します。

定理 2.2 U を $\mathbb{R}^n$ の凸集合、$f(x)$、$g_j(x)$ $(1 \leq j \leq \ell)$ が U 上の凸関数であると仮定します。さらに **Slater の制約想定**、すなわち

$$g_j(\bar{x}) > 0 \quad (1 \leq j \leq \ell)$$

を満たす $\bar{x} \in U$ が存在することを仮定します。このとき $x^* \in U$ が (NLP) の解であることの必要十分条件は、

$$\text{「ある } \lambda^* \in (\mathbb{R}_+)^\ell \text{ に対して } (x^*, \lambda^*) \text{ が Lagrnage 関数の鞍点である」} \tag{2.10}$$

ことです。

(証明) 定理 2.1 によって(2.10) が十分条件であることが分かっています。(2.10) が必要条件であることを示しましょう。そのために $\mathbb{R}^{1+\ell}$ 中の集合

$$C_1 := \{(y, z);\ y \leq f(x) \text{ かつ } z_j \leq g_j(x)\ (j = 1, \cdots, \ell)\ (\exists x \in U)\}$$

$$C_2 := \{(y, z);\ y \geq f(x^*),\ z_j \geq 0\ (j = 1, \cdots, \ell)\}$$

を定めます。明らかに C_2 は $\mathbb{R}^{1+\ell}$ 中の凸集合です。さらに f と g_j が凹関数であることから C_1 も凸であることが分かります。実際、$(y, z),\ (y', z') \in C_1$ とします。このとき

$$y \leq f(x),\ z_j \leq g_j(x)\ (j = 1, \cdots, \ell)$$

$$y' \leq f(x'),\ z'_j \leq g_j(x')\ (j = 1, \cdots, \ell)$$

を満たす $x,\ x' \in U$ が存在します。ここで $t \in [0, 1]$ をとると f と g_j が凹関数であることから

$$\begin{aligned} ty + (1-t)y' &\leq tf(x) + (1-t)f(x') \leq f(tx + (1-t)x') \\ tz_j + (1-t)z_j &\leq tg_j(x) + (1-t)g_j(x') \leq g_j(tx + (1-t)x') \quad (1 \leq j \leq \ell) \end{aligned}$$

が従います。また U が凸集合であることから

$$tx + (1-t)x' \in U$$

も従います。以上で C_2 が凸集合であることが分かりました。

　ここで仮定

$$f(x) \leq f(x^*) \quad (x \in U)$$

を使います。すなわち $(y, z) \in C_1 \cap C_2$ とすると $y = f(x^*)$ が従います。$(f(x^*), z)$ は C_2 の内点とはなりません。従って C_1 と C_2 は共通の内点を持ちません。このことから凸集合の分離定理を適用することができます。すなわち

$$ay + \alpha \cdot z \leq ay' + \alpha\alpha' \quad ((y, z) \in C_1,\ (y', z') \in C_2) \tag{2.11}$$

が成立します。y' と z'_j はいくらでも大きくなりますから

$$a \geq 0,\ \alpha_j \geq 0$$

が従います。

ここで $a > 0$ が従うことを示します。任意の $x \in U$ に対して

$$(f(x), g_1(x), \cdots, g_\ell(x)) \in C_1,\ (f(x^*), 0) \in C_2$$

が成立します。このことから(2.11) によって

$$af(x) + \sum_{j=1}^{\ell} \alpha_j g_j(x) \leq af(x^*) \tag{2.12}$$

を得ます。もし $a = 0$ が成立すると

$$\sum_{j=1}^{\ell} \alpha_j g_j(x) \leq 0 \quad (x \in U)$$

が従います。ここで $x = \bar{x}$ と Slater 想定の場合を考えると

$$\sum_{j=1}^{\ell} \alpha_j g_j(\bar{x}) \leq 0$$

となります。ところが $\alpha \neq 0$ と $\alpha_j \geq 0$ $(1 \leq j \leq \ell)$ が成立しますから、Slater 想定の条件から導かれる

$$\sum_{j=1}^{\ell} \alpha_j g_j(\bar{x}) > 0$$

と矛盾します。従って $a = 0$ となることは不可能で、$a > 0$ が成立します。(2.11) を $a > 0$ で割って

$$f(x) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x) \leq f(x^*) \quad (x \in U) \tag{2.13}$$

を得ます。ここで $\lambda^* = a^{-1}\alpha$ と定めましたが、$\lambda_j^* \geq 0$ $(1 \leq j \leq \ell)$ が成立します。(2.13) において $x = x^* \in U$ とすると

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \leq 0$$

を得ます。他方 $\lambda_j^* \geq 0$ と $g_j(x^*) \geq 0$ より

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \geq 0$$

も成立します。よって

$$\sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) = 0$$

となります。これを(2.13) の右辺に加えると

$$f(x) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x) \leq f(x^*) + \sum_{j=1}^{\ell} \lambda_j^* g_j(x^*) \quad (x \in U) \tag{2.14}$$

が従います。これは(2.6) すなわち

$$L(x, \lambda^*) \leq L(x^*, \lambda^*)$$

に他なりません。**（証明おわり）**