経済数学（法政用）：第１０章

細矢祐誉

テーマ：最適化理論（凹関数、準凹関数）

・凹関数の最適化問題

この章でも、引き続き

$$\begin{aligned} &\max && f(x) \\ &\text{subject to.} && x \in A \end{aligned}$$

といった問題を考えていく。ただし、ここでは A はコンパクトであることを仮定しない。ということは、第 9 章で紹介したような「解が確実に存在する」ことを前提とした議論はうまく行かない。以下、例をひとつ挙げる。

例：$\max_x x^2(-1 \le x)$ という問題を考える。x^2 の導関数は $2x$ なので、微分が 0 になるところは 0 のみ。端は -1 のみである。$f(x) = x^2$ とすれば $f(-1) = 1, f(0) = 0$ なので、第 9 章の解法である「微分が 0 になるところか端が解」という論法が使えるのならば、解答は -1 となる。

しかし、$f(2) = 4$ であり、明らかに -1 はこの問題の解ではない。よって、この問題に解答はないことがわかる。

上の例から、解がないケースでは前の解き方は失敗するということはわかった。従ってなにか別の手法が必要である。そこで次にヒントになるのが Fermat の定理である。この定理は「最大点では微分が 0 になる」という定理である。これをひっくり返して、「微分が 0 であれば最大点である」ということは言えないだろうか？

無論、無条件ではこの定理は無理である。なぜなら、f の最大点は $-f$ の最小点でもあり、そして $f'(x) = 0$ ならば $-f'(x) = 0$ が成り立つため、最小点でも微分すると 0 になるからである。無論これは一例であり、実際には微分が 0 であるにもかかわらず極大でも極小でもない点すら存在する。（x^3 の 0 の近くでのグラフを書いてみるといい。）

そこで次に行き着くのが、条件の追加である。といっても制約は $a \le x$ しか形を考えていないので、条件を追加できるのは f のみである。そこで、f になにかの条件を追加すれば、Fermat の定理の逆である「微分が 0 であれば最大点である」が言えるようになるのではないか？　という期待が生まれる。以下で、それが正しいことを示していこう。

まず、凹関数 (concave function) という言葉を定義しておく。実数値関数 f が凹関数

であるとは、どんな x, y と $0 \leq t \leq 1$ を満たすどんな t を持ってきても、次の不等式

$$f((1-t)x+ty) \geq (1-t)f(x)+tf(y)$$

が成り立っていることを言う。

主要な定理は次で与えられる。なお、上の例はすべて f が一変数関数であるときの例であったが、証明を見ればわかるように多変数関数で議論してもまったく問題ない。

定理１：f は凹関数で、$Df(x)=0$ が成り立っていたとすると、x は f の最大点である。

証明：まず y を適当に取り、$g(t)=f((1-t)x+ty)=f(x+t(y-x))$ と置けば、合成関数の微分の公式から

$$g'(0)=Df(x)(y-x)=0$$

である。よって、

$$\begin{aligned}
0 &= g'(0)\\
&= \lim_{t\downarrow 0}\frac{g(t)-g(0)}{t}\\
&= \lim_{t\downarrow 0}\frac{f((1-t)x+ty)-f(x)}{t}\\
&\geq \lim_{t\downarrow 0}\frac{(1-t)f(x)+tf(y)-f(x)}{t}\\
&= \lim_{t\downarrow 0}\frac{-tf(x)+tf(y)}{t}\\
&= \lim_{t\downarrow 0}(f(y)-f(x)) = f(y)-f(x)
\end{aligned}$$

となるので、$f(x) \geq f(y)$ がわかる。y はなんでもよかったので、どんな y に対しても $f(x) \geq f(y)$ が成り立っていることになり、よって x は f の最大点である。 ■

この定理を使えば、f が凹関数でさえあれば、$f'(x)=0$ という式が成り立っている点 x を見つけるだけで、自動的に x が解であるということがわかってしまう。これは前の方法と比べて明らかに計算の手間が減っている上、解が存在するかどうかわからないときでも通用するやり方である。

しかし、f が凹関数であることをチェックするためにどうすればいいか、この定義だけからではわかりにくい。そこで、簡単にチェックできる方法を与えておきたい。最初に用語をひとつ追加する。A が凸集合であるとは、$x, y \in A$ かつ $0 \leq t \leq 1$ ならば常に $(1-t)x+ty \in A$ である、という性質が成り立っていることを指す。これは x, y がベク

トルでも同様に定義できる概念だが、実は実数の部分集合については、A が凸集合であることは A が「区間」であることと同値であることが知られている（第３章を見よ)。まず、f が凹関数であるためには、f の定義域が凸集合でないと話が始まらないことに注意しよう。次に、f が一変数関数の場合の定理を述べておく。

定理２：f が実数の凸集合 A 上で連続で、A の端点以外で微分可能であるとする。このとき、$f'(x)$ が非増加である（つまり、$x < y$ ならば $f'(x) \geq f'(y)$）ことと、f が A 内で凹関数であることは同値である。特に f が A 上で二階微分可能であれば、f が凹関数であることと、$x \in A$ ならば $f''(x) \leq 0$ となることは同値である。

証明：だいたいどのケースでも同じなので、$A = [a, b]$ という閉区間の場合のみを問題にする。一般の場合は各自で証明を試みること。

まず、f が凹関数であるとしよう。まず、$a \leq x < z < y \leq b$ ならば

$$\frac{f(y)-f(z)}{y-z} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x}$$

であることを示してみよう。最初に、$z = (1-t)x + ty$ となる t がなんであるかを計算してみよう。このとき、

$$z = x - tx + ty = x + t(y-x)$$

であるから、両辺から x を引いて、

$$z - x = t(y-x)$$

であり、よって

$$t = \frac{z-x}{y-x}$$

である。故に $0 < z - x < y - x$ より、$0 < t < 1$ である。次に、

$$f(z) \geq (1-t)f(x) + tf(y) = f(x) + t(f(y) - f(x))$$

であるから、両辺から $f(x)$ を引いて、

$$f(z) - f(x) \geq t(f(y) - f(x)) = \frac{(z-x)(f(y)-f(x))}{y-x}$$

であり、この両辺を $z - x$ で割ると

$$\frac{f(z)-f(x)}{z-x} \geq \frac{f(y)-f(x)}{y-x}$$

が出る。これで不等号の片方が出た。もう片方の不等号

$$\frac{f(y)-f(x)}{y-x} \geq \frac{f(y)-f(z)}{y-z}$$

もまったく同じようにして計算できる（$z=(1-s)y+sx$ となる s を求め、上の t の代わりに使うとよい。）。

さて、そこで $a<x<y<b$ のときに、上の不等式の前者の z を y に向けて極限を取り、後者の z を x に向けて極限を取ることで、

$$f'(y) \leq \frac{f(y)-f(x)}{y-x} \leq f'(x)$$

がわかる。したがって f' は減少関数である。以上で、f が凹関数であるならば f' は減少的であることが示せた。

逆に一階の導関数 f' が減少関数であるとしよう。仮に f が凹関数でなく、$a \leq x < y \leq b$ を満たす x, y と $0<t<1$ を満たす t に対して

$$f((1-t)x+ty) < (1-t)f(x)+tf(y)$$

が成り立っていたとしてみよう。このとき $(1-t)x+ty=z$ とすれば、上の計算と同様に

$$t=\frac{z-x}{y-x}$$

であり、かつ $f(z)-f(x)<t(f(y)-f(x))$ なので、

$$\frac{f(z)-f(x)}{z-x} < \frac{f(y)-f(x)}{y-x}$$

である。同様にして、

$$\frac{f(y)-f(z)}{y-z} > \frac{f(y)-f(x)}{y-x}$$

が示せる。平均値の定理から、z と x の間に $f'(w)=\frac{f(z)-f(x)}{z-x}$ となる w が、また y と z の間に $f'(v)=\frac{f(y)-f(z)}{y-z}$ となる v が存在することになるが、$w<z<v$ で、かつ

$$f'(w) < \frac{f(y)-f(x)}{y-x} < f'(v)$$

が成り立つことになり、f' が減少関数であるという事実と矛盾する。したがってこれはあり得ず、f は凹関数である。 ■

これによって、凹関数であることを確かめるためには、単に二階微分が常に 0 以下であることを確かめればよいことがわかる。

なお、$f(x)=\sqrt{x}$ などを考えれば、これは開半直線 $(0,+\infty)$ 上で微分

$$f'(x)=(\sqrt{x})'=(x^{\frac{1}{2}})'=\frac{1}{2}x^{-\frac{1}{2}}=\frac{1}{2\sqrt{x}}$$

を持つ。この右辺は $x=0$ ならば分母が 0 になるので、$f'(0)$ は存在しないことがわかる。上の定理で、「端点を除いて微分可能」という条件にこだわったのは、この場合をカバーするためであることを注意しておく。

一般の多変数の凹関数については、その判定法は簡単ではない。ヘッセ行列

$$D^2f(x)=\begin{pmatrix} f_{x_1x_1}(x) & \dots & f_{x_1x_n}(x)\\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(x) & \dots & f_{x_nx_n}(x)\end{pmatrix}$$

が、線形代数で言われる**半負値定符号** (negative semi-definite) と呼ばれる条件を満たす、というのがその条件になるのだが、難しいので省略する。代わりに、次の定理が成り立つことを指摘しておく（証明は省略するが、案外面倒であることを指摘しておく）。

定理３：$A\subset\mathbb{R}^n$ が非空な内部を持つ凸集合であり、f が A 上で定義された連続な実数値関数であるとする。このとき、f が凹関数であることは、A の内部の任意の x と、任意の $v\in\mathbb{R}^n$ に対して、関数

$$g(t)=f(x+tv)$$

が凹関数であることである。

これを使うと、二変数の場合の判定定理が容易に導ける。

定理４：$A\subset\mathbb{R}^2$ が非空な内部を持つ凸集合であり、f が A 上で定義された連続な実数値関数で、A の内部で二階連続微分可能であるとする。このとき、f が凹関数であることは、A の内部で次の 3 つの不等式が成り立つことと同値である。

$$f_{xx}\le 0,$$

$$f_{yy}\le 0,$$

$$f_{xx}f_{yy}-f_{xy}f_{yx}\ge 0.$$

証明：定理３の仮定はすべて満たされているので、A の内部にある (x,y) と、任意の (v,w) を取ってきて、

$$g(t)=f(x+tv,y+tw)$$

としよう。f が凹関数であることは、すべての v と w について g が凹関数であることと同値である。そこで g を二階微分すると、

$$g'(t)=f_x(x+tv,y+tw)v+f_y(x+tv,y+tw)w,$$
$$g''(t)=f_{xx}v^2+2f_{xy}vw+f_{yy}w^2$$

となる（$f_{xy}=f_{yx}$ に注意。最後、変数省略）。よって、f が凹関数であることは、すべての v と w について

$$f_{xx}v^2+2f_{xy}vw+f_{yy}w^2\le 0$$

であることと同値である。

以下、最初に f が凹関数だとしよう。すると上の不等式がすべての v と w について成り立つ。$v=0$ とすれば $f_{yy}\le 0$ が、$w=0$ とすれば $f_{xx}\le 0$ が出てくるので、定理４の不等式のうち２つはすぐ出てくる。もし $f_{xx}=0$ であるとすれば、$v=-\frac{f_{yy}}{2f_{xy}}+f_{xy}, w=1$ と代入すれば、$f_{xy}^2\le 0$ を得て、ここから $f_{xy}=0$ を得る。したがって $f_{xx}f_{yy}-f_{xy}f_{yx}=0$ であり、定理４の３番目の不等式が得られる。一方、$f_{xx}=0$ ならば、上の不等式の両辺を w^2 で割って $s=\frac{v}{w}$ とすれば

$$f_{xx}s^2+2f_{xy}s+f_{yy}\le 0$$

となるので、判別式の条件から

$$f_{xy}^2-f_{xx}f_{yy}\le 0$$

を得るが、これはやはり３番目の不等式と同値である。よって、定理４の３つの不等式は、いずれの場合もたしかに成り立つ。

逆に、f が定理４の３つの不等式を全部満たしているとすれば、

$$f_{xx}v^2+2f_{xy}vw+f_{yy}w^2\le 0$$

が成り立つことも、同様にして示すことができる。が、これは読者への宿題とする。■

・凹関数の問題の解き方

例１：$f(x)=\sqrt{x}-x$ とし、問題 $\max_x f(x)(0\le x)$ を考える。
　計算すれば、

$$f''(x)=-\frac{1}{4}x^{-\frac{3}{2}}<0$$

ということがわかるので、系からこの関数は凹関数である。一方、

$$f'(x)=\frac{1}{2}x^{-\frac{1}{2}}-1$$

がわかる。したがって、

$$f'(x)=0\Leftrightarrow x^{-\frac{1}{2}}=2\Leftrightarrow x=\frac{1}{4}$$

がわかる。
　定理１と２から、$x=\frac{1}{4}$ は上の問題の最大点であることがわかる。

例２：問題

$$\begin{array}{rl}\max & f(x,y)=(xy)^{\frac{1}{3}}-x-y\\ \text{subject to.} & x\ge 0,y\ge 0.\end{array}$$

を考える。このとき、$x>0,y>0$ ならば

$$f_{xx}=-\frac{2}{9}x^{-\frac{5}{3}}y^{\frac{1}{3}}<0,$$

$$f_{xy}=f_{yx}=\frac{1}{9}x^{-\frac{2}{3}}y^{-\frac{2}{3}},$$

$$f_{yy}=-\frac{2}{9}x^{-\frac{5}{3}}y^{\frac{1}{3}}<0,$$

$$f_{xx}f_{yy}-f_{xy}f_{yy}=\frac{1}{27}x^{-\frac{4}{3}}y^{-\frac{4}{3}}>0$$

となるので、定理４から f は凹関数である。そして

$$f_x=\frac{1}{3}x^{-\frac{2}{3}}y^{\frac{1}{3}}-1=0,$$

$$f_y=\frac{1}{3}x^{\frac{1}{3}}y^{-\frac{2}{3}}-1=0,$$

を解けば、最初の式から

$$y^{\frac{1}{3}}=3x^{\frac{2}{3}},$$

つまり

$$y=27x^2$$

を得る。同様に二番目の式から

$$x = 27y^2$$

を得る。代入して、

$$y = (27)^3 y^4$$

を得るので、ここから

$$y = \frac{1}{27}, x = \frac{1}{27}$$

を得る。定理 1 から、これが解である。

・限定付き最大化と準凹関数

ここでは、第 9 章と同様に

$$\begin{aligned} \max \quad & f(x) \\ \text{subject to.} \quad & p \cdot x \leq m, \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

を考える。すでに述べたように、ラグランジュの原理では解けない問題があるのだが、その解法は一般に簡単ではない。そこで、ラグランジュの原理だけで解ける f はどんなものか、という疑問が生まれる。凹関数だとどうだろうか？　実は凹関数ならば、ラグランジュの原理で出てきた解は必ずこの問題の解である、という定理がある（Sundaram の本の第 7 章を参照）。しかしこれは、実はこの種の問題には使えない。というのは、たとえば前の章で述べた関数

$$f(x) = x_1 x_2 ... x_n$$

などは凹関数ではないからである。第 9 章で示したようにこの関数に対する上の問題の解はラグランジュの原理で計算できるのだが、これを処理できないようでは力不足である。というわけで、ここでは凹関数を拡張した「準凹関数 (quasi-concave function)」という概念を定義しておきたい。

凸集合 $A \subset \mathbb{R}^n$ 上で定義された関数 f が準凹であるとは、任意の $x, y \in A$ と $t \in [0, 1]$ に対して

$$f((1-t)x + ty) \geq \min\{f(x), f(y)\}$$

が成り立つことである。

図で書けばわかるが、f が準凹であるとは、f の等高線の上側集合が凸集合であるというのと同値である。したがって一変数ならば、単調な関数はすべて準凹である。一方で、

$f(x,y) = x^2 + y^2$ は準凹ではない——というより、まさに第9章で述べたように、「予算線に下側から接する」等高線を持つ関数は準凹ではないのである。これが、次の定理が成り立つ本質的な理由になっている。

定理５：A は内部が非空な凸集合で、$f : A \to \mathbb{R}$ は連続かつ準凹であるとし、x^* で全微分可能、$Df(x^*) \gg 0$ であり、

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

としたときにある λ^* が存在して

$$DL(x^*,\lambda^*) = (0,0)$$

が成り立つとする。さらに、ある $x^+ \in A$ において $p\cdot x^+ < m$ になるとする。このとき、x^* は次の問題：

$$\begin{array}{rl} \max & f(x) \\ \text{subject to.} & p\cdot x \le m \\ & x \in A \end{array}$$

の解である。

証明：仮に x^* が解でなかったとしよう。このとき、うまく x を取れば、$x \in A, p\cdot x \le m$ で、さらに $f(x) > f(x^*)$ であるようなものを取ってくることができる。f は連続なので、必要ならば十分小さな $s > 0$ に対して $(1-s)x + sx^+$ を x と取り替えてやっても、相変わらず $f((1-s)x + sx^+) > f(x^*)$ が成り立ち、しかも

$$p\cdot[(1-s)x + sx^+)] = (1-s)p\cdot x + sp\cdot x^+ < (1-s)m + sm = m$$

である。よって、$(1-s)x + sx^+$ をあらかじめ x と交換しておくことで、$p\cdot x < m$ と最初から仮定してよい。

　さて、

$$DL(x^*,\lambda^*) = (Df(x^*) - \lambda^* p^T, m - p\cdot x^*) = (0,0)$$

である。特に最後の座標から、$p\cdot x^* = m$ であり、また残りから

$$Df(x^*) = \lambda^* p^T$$

である。$Df(x^*) \gg 0$ なので $\lambda^* > 0$ であることに注意。そこで

$$g(t) = f((1-t)x^* + tx)$$

と定義しよう。f は x^* で全微分可能なので g は $t=0$ で微分可能で、

$$\begin{aligned} g'(0) &= Df(x^*)(x-x^*) \\ &= \lambda^* p\cdot(x-x^*) \\ &= \lambda^*[p\cdot x - p\cdot x^*] \\ &= \lambda^*[p\cdot x - m] \\ &< 0 \end{aligned}$$

がわかる。

一方、f は準凹なので、

$$g(t) \ge \min\{f(x^*), f(x)\} = f(x^*) = g(0)$$

であり、したがって微分の定義から、

$$g'(0) = \lim_{t\downarrow 0}\frac{g(t)-g(0)}{t} \ge 0$$

でなければならない。この 2 つの不等号は互いに反対向きであるため、矛盾である。よってこのようなことはありえず、x^* は解でなければならない。 ■

なお、証明はしないが、二変数の準凹関数には以下のような判定定理がある。証明は Debreu (1952) と Otani (1983) を参照（おそらく、この定理の証明がきちんと載っている教科書は、市販レベルでは存在しない)。

定理６：f が開凸集合 A 上で二階連続微分可能であり、また $Df(x)\ne 0$ が常に成り立つとき、

$$2f_{xy}f_xf_y - f_{xx}f_y^2 - f_{yy}f_x^2 \ge 0$$

であれば、f は準凹である。