経済数学（法政用）: 第 9 章

細矢祐誉

テーマ : 最適化理論（コンパクト集合）

「○○という条件の下で $f(x)$ の値が最大になるような x を求めなさい」という形式の問

題を考える。これを f の（条件付）最大化問題、と呼ぶ。以下、この問題を効率的に解くための方法を考えてみよう。しばらくの間、f は一変数関数とする。

例題 : $f(x) = -x^2 + 5x$ を $0 \leq x \leq 10$ の範囲内で最も大きくするような x を求めなさい。

下のいくつかの記法はすべて、上の問題を考える、ということを宣言するために用いる。（この他にも書き方があると思うが、だいたい下の 3 パターンの書き方が多い。また、$0 \leq x \leq 10$ の代わりに $x \in [0, 10]$ と書かれることもある。）

$$
\begin{aligned}
&1) \quad \max_{0 \leq x \leq 10} \quad -x^2 + 5x \\
&2) \quad \max_{x} \quad -x^2 + 5x \ (0 \leq x \leq 10) \\
&3) \quad \max_{x} \quad -x^2 + 5x \\
&\qquad \text{subject to. } 0 \leq x \leq 10
\end{aligned}
$$

なお、$\max_x$ という書き方は max とだけ略記される場合も多い。

基本的な用語法として、まず関数 $f(x)$ のことを「目的関数」と呼ぶ。また、「x が○○の中にある」という条件を「制約条件」と呼ぶ。$f(x)$ が最大になる x のことを「解」もしくは「最大点」と呼び、その x を選んだときの f の値のことを「最大値」と呼ぶ。

上の例題の場合、目的関数は $-x^2 + 5x$。制約条件は $0 \leq x \leq 10$。最大点は 2.5 ただひとつ。最大値は 6.25 になる。これをどうやって解けばいいか？　というのが、今回学ぶ内容である。

最適化するに当たって現在使える中で最も有効な手段は「グラフを書く」ということである。きちんとしたグラフを手で書けさえすれば、最大化問題はグラフを見るだけで解ける。しかし、前回の講義ノートでやったように、細密なグラフを書くのはとてもたいへんである。できればもっと楽に解ける方法が欲しい。そのためのテクニックを見つけるのが最適化理論の大きな役割のひとつである。

・解の存在

そもそも、最大点は存在するのか？

存在しない典型的な問題をいくつか上げる。

1) $\max_x x^2+1 (0 \le x < 1)$

この問題の場合、x が 1 に近ければ近いほど高い値になるが、$x=1$ 自体は選べないため、解がない。

2) $\max_x x^2+1 (0 \le x)$

いくらでも高い x が選べるせいで、最大になる点というのがなくなってしまっている。

3) $\max_x f(x) (0 \le x \le 1)$、ただし、

$$f(x) = \begin{cases} x & (x < \frac{1}{2} \text{ のとき}) \\ 0 & (\text{それ以外}) \end{cases}$$

とする。

この場合、f は $x=\frac{1}{2}$ の近くで不連続になってしまっている。そして、$x<\frac{1}{2}$ の範囲では高ければ高いほど値が高くなるが、最も高くなるはずだった $x=\frac{1}{2}$ のところで 0 に落ちてしまうため、やはり最大点は存在しない。

このように、最大点が存在しない例はいくつもあるが、幸いにして条件が「$a \le x \le b$」という形であって、さらに f が連続である場合には、前に紹介した「最大最小原理」という名の定理（Rolle の定理を証明するときに使った）によって最大点は存在する。したがって、このタイプの制約のときには最大点は存在するかどうか疑う必要はない。

・解を見つける方法

典型的な最大点の居所は、実は 3 パターンしかない。以下に、それぞれについてひとつずつ例を挙げる。

例 1 ：端。

$\max_x -x^2+4x (0 \le x \le 1)$ という問題を考える。$-x^2+4x$ は区間 $0 \le x \le 1$ 内全体で増加的なので、$x=1$ が解になる。

例 2 ：微分できないところ。

$\max_x -|x-1| (0 \le x \le 2)$ という問題を考える。$-|x-1|$ はどんな x に対しても 0 以

下の値を返し、$x = 1$ のときだけ 0 と一致するので、$x = 1$ が解になる。1 は $-|x-1|$ のグラフを書くと山の頂点にあることはわかるが、その山は尖っているため、1 のところでこの目的関数は微分できない。

例３：微分した値が 0 になるところ。（一階の条件）

　$\max_x -x^2(-1 \le x \le 1)$ という問題を考える。$-x^2$ はやはりどんな x に対しても 0 以下の値を返し、$x = 0$ のときだけ 0 と一致するので、$x = 0$ が解になる。$-x^2$ はどこでも微分可能だが、その導関数は $-2x$ であり、よって $x = 0$ のときに微分した値は 0 になる。

一般論として、次の定理が成り立つ。

定理１：$a < b$ とし、問題 $\max_x f(x)$ $(a \le x \le b)$ を考える。このとき解は存在する。そして、もし x^* が問題の解であるとすれば、次の３つのうちのいずれか少なくともひとつが成り立つ。（注：ふたつ以上が同時に成り立つ場合もある。）

1) x^* は端点である a か b のどちらか。
2) f は x^* では微分できない。
3) f は x^* で微分可能で、$f'(x^*) = 0$ になる。

証明：解の存在については最大最小原理が保証してくれる。また、Fermat の定理により、「x^* が上の問題の解で、$a < x^* < b$ で、さらに f が x^* で微分可能なときには $f'(x^*) = 0$ が成り立つ」ということがわかる。上の定理はこの結果を書き換えただけである。 ■

・応用：前述の定理ふたつをどう使うか？

　定理 1 を使うことで、多くの $a \le x \le b$ 型の制約を持つ最大化問題で、最大点を発見することができる。いま、この問題には解があることがはっきりしているため、その解は端点か、f が微分できない点か、f を微分して 0 になる点のどこかにある。ということは、上の条件を満たす点（端点、微分できない点、微分して 0 になる点）をすべてリストアップして、その中でいちばん大きな値が出てくる点を取ってくれれば、間違いなくそれが最大点である。

　実際は微分できない点をチェックするのは案外難しいが、微分できない点がない目的関数はとても多い。その場合、「微分して 0 になる点か、端点」をチェックして、その中で一番 f が大きくなるところを持ってくれれば、それが解になる。

例：$\max_x x^3 - 6x^2 + 9x (0 \leq x \leq 2)$

制約の形が $0 \leq x \leq 2$ なので、定理 1 が使える。目的関数 $x^3 - 6x^2 + 9x$ はこの区間内ではどこでも微分可能なので、定理 1 のうち条件 2) については無視していい。

$x^3 - 6x^2 + 9x$ を微分すると、$3x^2 - 12x + 9 = 3(x-1)(x-3)$ となる。なので、微分して 0 になるのは $x = 1$ と $x = 3$ の 2 点のみ。$0 \leq x \leq 2$ の範囲に入っているのはこの中では $x = 1$ だけ、ということになる。なので、定理 1 から最大点の候補は 3) に対応する $x = 1$ か、1) に対応する $x = 0$ と $x = 2$ の 3 通りしかあり得ない。

$f(x) = x^3 - 6x^2 + 9x$ としよう。すると地道な計算により、$f(0) = 0$、$f(1) = 4$、$f(2) = 2$ がわかる。$f(1)$ が最も大きいので、前述のように $x = 1$ が解となる。

・多変数関数の最大化問題

今度は f が多変数関数、つまり $f : \mathbb{R}^n \to \mathbb{R}$ などの場合を考察する。もちろんこの問題も、

$$\begin{array}{rl} \max & f(x) \\ \text{subject to.} & ○○ \end{array}$$

という形式で書かれる。このあたりの記法は一変数の場合と大差ない。

まず解の存在については、一次元と有限次元で大差はない。「○○」という条件を満たす点の集合を A と書くと、これが第 3 章で紹介した「コンパクト」という条件を満たすときには、f が連続であるという仮定の下で、解の存在を示すことができる。

一方で、x を中心とする半径 $r > 0$ の開球 $B_r(x) = \{y \in \mathbb{R}^n | \|y - x\| < r\}$ を考えよう。$x \in A$ は、ある半径 r について $B_r(x)$ がすっぽり A に含まれるとき、A の**内部** (interior) にあると呼ばれる。すると定理 1 から、次の定理が成り立つ。

定理２：問題

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

を考え、この解 x^* が A の内部にあるとする。もし f が x_i で偏微分可能ならば、$\frac{\partial f}{\partial x_i}(x^*) = 0$ である。

証明：x^* は A の内部にあるため、ある $r>0$ について $B_r(x^*)\subset A$ である。よって、

$$\begin{aligned}\max_{x_i} \quad & g(x_i)\equiv f(x_1^*,...,x_{i-1}^*,x_i,x_{i+1}^*,...,x_n^*)\\ \text{subject to.} \quad & x_i^*-r\le x_i\le x_i^*+r\end{aligned}$$

を考えれば、この問題の解は x_i^* である。故に定理 1 から

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

となって証明が終わる。 ■

この定理はたしかに正しいのだが、実を言うとあまり有用ではない。理由はふたつある。

一つ目に、たとえば一変数関数の場合を考えて $A=[a,b]$ であれば、この内部に存在しない点は a と b の二点だけだった。したがって、「内部にある」点の解の候補さえ絞ってしまえば、それらと a,b という端点のふたつだけが解の候補である。しかし、これは多変数関数の場合に使えない。なぜ使えないかというと、たとえば A が正方形 $[a,b]\times[a,b]$ だった[*1]とすれば、「内部にない点」が「正方形の縁」という「無限集合」であるため、まだ解の候補が無限に存在する。よって前の例でのやり方は破綻する。

この問題も深刻なのだが、より深刻なのは二つ目の問題である。たとえば典型的なこのタイプの問題は、経済学だと次のような形で出てくる。

$$\begin{aligned}\max \quad & f(x,y)\\ \text{subject to.} \quad & px+qy\le m,\\ & x\ge 0, y\ge 0.\end{aligned}$$

これは消費者の需要を求める問題である。f は効用関数と呼ばれ、p は x で表される商品の価格、q は y で量を表される商品の価格、m は予算である。

この問題の条件を満たす点の集合 $\{(x,y)|x\ge 0,y\ge 0,px+qy\le m\}$ は、$p,q,m>0$ ならば $(0,0),(\frac{m}{p},0),(0,\frac{m}{q})$ を頂点とする直角三角形になることが容易にわかるが、この種の問題が経済学で出たとき、f はたいてい増加関数である。したがって、解は $px+qy=m$ を満たす点にだけ現れる——つまり、解は内部には存在しないのである！

[*1] これはつまり、$\{(x,y)|a\le x\le b, a\le y\le b\}$ という意味である。一般に集合 A と B について、$A\times B$ は集合

$$\{(a,b)|a\in A, b\in B\}$$

を表す。

　というわけで、この種の問題には上の定理は使えない。対処するためにはどうすればよいだろうか？　賢い読者ならばこう考えるかもしれない。どうせ解が $px+qy=m$ を満たすことがわかっているならば、それに焦点を絞ればよい。つまり、$y=\frac{1}{q}[m-px]$ と置いた次の問題

$$\begin{aligned} \max \quad & f(x, \frac{1}{q}[m-px]) \\ \text{subject to. } & 0 \le x \le \frac{m}{p}, \end{aligned}$$

に問題を置き換えてしまえば、これは一変数の問題だから先ほどのやり方で解ける。たしかにこれは有効である。しかし、一変数増やした次の問題

$$\begin{aligned} \max \quad & f(x,y,z) \\ \text{subject to. } & px+qy+rz \le m, \\ & x,y,z \ge 0, \end{aligned}$$

にはどう対処すればよいかは依然としてわからない。もっと一般に $x \in \mathbb{R}^n$ として、$p \in \mathbb{R}^n, p \gg 0$ と $m \in \mathbb{R}, m>0$ に対して[*2]

$$\begin{aligned} \max \quad & f(x) \\ \text{subject to. } & p \cdot x \le m, \\ & x \ge 0, \end{aligned}$$

としたらどうかもわからない。さらにこれの双対 (dual) 問題と呼ばれる次の問題

$$\begin{aligned} \min \quad & p \cdot y \\ \text{subject to. } & f(y) \ge f(x), \\ & x \ge 0 \end{aligned}$$

への対処などは、二変数ですらわからない。なぜかといえば、$f(y)=f(x)$ が成り立つ y がどう書けるかがわからないからである。

　というわけで、山積みの問題があることがわかった。しかしこれらは統一的に、ある原理によって解決することができる。

・ラグランジュの未定乗数法（線形制約）

[*2] この $p \gg 0$ という記号は $p_i > 0$ がすべての i に対して成り立つことを意味する。

一般に問題

$$\begin{aligned} \max \quad & f(x) \\ \text{subject to.} \ & g(x) = 0, \\ & x \in A \end{aligned}$$

を考える。ここで、次のような形で新しい関数 L（ラグランジュ関数と呼ぶ）を作る。

$$L(x, \lambda) = f(x) + \lambda g(x).$$

ラグランジュの原理とは、上の問題の解は、L を微分して 0 になる点であるという原理である。より詳しく述べると、上の問題の解は、$DL(x^*, \lambda^*) = 0$ となる λ^* が存在するような x^* として特徴付けできるという原理である。

この原理は、実を言うと正しいことも間違っていることもある。しかし、大筋で正しいことが多い。一般の場合は秋学期に考えるとして、まずはこの $g(x)$ が、我々が考えている問題に合う場合、つまり

$$g(x) = m - p \cdot x$$

となる場合を考える。この場合、

$$Dg(x) = -p^T$$

であるため（p^T は、本来縦のベクトルである p を横に並べたベクトルである）、

$$DL(x, \lambda) = (Df(x) + \lambda Dg(x), g(x)) = (Df(x) - \lambda p^T, g(x))$$

となることに注意する。

いま、x^* が上の問題の解で、さらに x^* は A の内部に属していたとする。そして f は x^* で全微分可能だとしよう。ここで $p \cdot v = 0$ となるようなベクトル v を取ってきて、$x(t) = x^* + tv$ としよう。すると、x^* が A の内部に属していることから、十分小さな $a > 0$ に対して

$$-a \le t \le a \Rightarrow x(t) \in A$$

となる。そこで、問題

$$\begin{aligned} \max \quad & g(t) \equiv f(x(t)) \\ \text{subject to.} \ & -a \le t \le a, \end{aligned}$$

を考えると、0 がこの問題の解である。故に定理 1 から、

$$g'(0) = 0$$

であるが、一方で合成微分の公式から

$$g'(0) = Df(x^*)x'(0) = Df(x^*)v$$

である。故に $Df(x^*)v = 0$ である。これで、$p \cdot v = 0 \Rightarrow Df(x^*)v = 0$ が示された。

一般に、$p \cdot v = 0 \Rightarrow Df(x^*)v = 0$ であるならば、$Df(x^*)$ は p の定数倍である。これは次元の公式と呼ばれる線形代数の有名な定理から示すことができるのだが、今回はそういうことをあえてせずに、直接示してみよう。まず、$v = (p_2, -p_1, 0, ..., 0)$ とする。このとき、$p \cdot v = 0$ なので、$Df(x^*)v = 0$ である。これは

$$f_{x_1}(x^*)p_2 - f_{x_2}(x^*)p_1 = 0$$

という意味になるから、これを整理すると

$$\frac{f_{x_1}(x^*)}{p_1} = \frac{f_{x_2}(x^*)}{p_2}$$

という結果を得る。

まったく同じようにして、すべての i と j について

$$\frac{f_{x_i}(x^*)}{p_i} = \frac{f_{x_j}(x^*)}{p_j}$$

という結果を得ることができるので、この値を λ^* とすれば、

$$f_{x_i}(x^*) = \lambda^* p_i$$

がすべての i に同時に成り立つことになる。これは

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

であるという意味になる。一方で x^* は解なので当然ながら条件 $g(x^*) = 0$ を満たし、よって

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

となる。結論として以下の定理を得る。

定理３：問題

$$\begin{array}{rl} \max & f(x) \\ \text{subject to.} & g(x) = 0, \\ & x \in A \end{array}$$

を考え、ただし $g(x) = m - p \cdot x$ とする。ここで、

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

とする。もし x^* が問題の解であり、A の内部にあって、さらに f が x^* で全微分可能だとすると、ある λ^* が存在して、

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

が成り立つ。

・不等式問題への拡張

一見して、これで問題が解決したように見える。しかしまだ詰めが足りない。我々の目標は

$$\begin{aligned} \max \quad & f(x) \\ \text{subject to.}\ & p \cdot x \leq m, \\ & x \geq 0 \end{aligned}$$

という問題である。この問題において、

$$A = \{x \in \mathbb{R}^n | x \geq 0\},$$

と置き（経済学では普通、この集合を $\mathbb{R}^n_+$ と書く）、

$$g(x) = m - p \cdot x$$

と書くと、

$$\begin{aligned} \max \quad & f(x) \\ \text{subject to.}\ & g(x) \geq 0, \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

となって、前の問題と似た形になる。似た形になるが、同じ形ではない。$g(x) = 0$ ではなく、$g(x) \geq 0$ である。

前に述べたように、$f(x)$ が増加関数であれば、$g(x) > 0$ となるところは解ではない——実際、いま $g(x) > 0$ だとして、x よりすべての座標がほんの少しだけ大きなベクトル y を取れば、$g(y) > 0$ を保ちつつ $f(y) > f(x)$ になるようにできるだろうから、そこは解

ではない。よって、実質的には $g(x)=0$ と同じ話になる。さらに、x が $\mathbb{R}^n_+$ であることと、$x \gg 0$ は同値である。こうして我々は次の定理を得る。

定理４：問題

$$\begin{aligned}\max \quad & f(x)\\ \text{subject to.}\ & g(x) \ge 0,\\ & x \in \mathbb{R}^n_+\end{aligned}$$

を考え、ただし $g(x)=m-p\cdot x$ とする。ここで、

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

とする。もし x^* が問題の解であり、$x^* \gg 0$ で、さらに f は x^* で全微分可能であったとすれば、ある λ^* が存在して、

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

が成り立つ。

例：$f(x)=x_1x_2...x_n$ の場合の問題

$$\begin{aligned}\max \quad & f(x)\\ \text{subject to.}\ & p\cdot x \le m,\\ & x \ge 0\end{aligned}$$

を考えてみよう。f は連続なので、最大最小原理から解は必ず存在する。また、ある i について $x_i=0$ ならば $f(x)=0$ であり、すべての i について $x_i>0$ ならば $f(x)>0$ なので、解は内部にしか存在し得ず、さらに解においては $f(x)>0$ が成り立つ。そこで $DL(x,\lambda)=0$ となる点を計算してみると、

$$L(x,\lambda)=x_1x_2...x_n-\lambda(m-p\cdot x)$$

なので、$DL(x,\lambda)=(0,0)$ を書き下すと

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

$$\frac{\partial L}{\partial \lambda}=m-p\cdot x=0$$

である。上の式を使えば

$$\lambda p_i x_i = f(x)$$

を得るが、解では $f(x) > 0$ が成り立つので $\lambda > 0$ であり、よって

$$p_i x_i = \frac{f(x)}{\lambda}$$

となる。この式はすべての i について成り立つが、右辺が i に依存しないので、

$$p_1 x_1 = p_2 x_2 = ... = p_n x_n$$

がわかったことになる。

　後は、

$$0 = m - p \cdot x = m - [p_1 x_1 + p_2 x_2 + ... + p_n x_n] = m - n p_i x_i$$

から、

$$x_i = \frac{m}{n p_i}$$

を得る。こうして我々は $DL(x, \lambda) = 0$ となるためには

$$x = \left(\frac{m}{np_1}, \frac{m}{np_2}, ..., \frac{m}{np_n}\right)$$

でなければならないという結論を得た。内部に解がなければならない以上、これが解である。

・二変数関数についての注記

　$f(x, y)$ の場合の問題

$$\begin{array}{rl} \max & f(x, y) \\ \text{subject to.} & (p, q) \cdot (x, y) \le m, \\ & x \ge 0,\ y \ge 0 \end{array}$$

に目を向けよう。f が増加的であれば、定理 4 から、(x^*, y^*) が解で、$x^* > 0, y^* > 0$ で、さらに f が (x^*, y^*) で微分可能であるならば、

$$Df(x^*, y^*) = \lambda^*(p, q)$$

となる λ^* が存在することになる。この状況を理解するために、ある関数 $h(x)$ を考えよう。いま、$h(x^*) = y^*$ であり、また h は方程式

$$f(x, h(x)) \equiv a$$

を常に満たしていると仮定する。さらに $h'(x^*)$ が存在しているとすれば、合成微分の公式から

$$f_x(x^*, y^*) + f_y(x^*, y^*)h'(x^*) = 0$$

が導ける。もし $f_y(x^*, y^*) \neq 0$ であれば

$$h'(x^*) = -\frac{f_x(x^*, y^*)}{f_y(x^*, y^*)}$$

である。一方でラグランジュの原理 $Df(x^*, y^*) = \lambda^*(p, q)$ を思い出すと、$f_y(x^*, y^*) \neq 0$ なのだから $\lambda^* \neq 0$ で、よって上の式は

$$h'(x^*) = -\frac{p}{q}$$

という意味であることがわかる。

前に、上の問題を考えるときに、$y = \frac{1}{q}[m - px]$ を代入してしまえばいいという考え方について述べた。この式は、制約 $(p, q) \cdot (x, y) \leq m$ を等号で満たす (x, y) の条件である。ところがこの式の右辺を x で微分すると、やはり $-\frac{p}{q}$ が出てくる。したがって、実はラグランジュの原理というのは、f の等高線 h の微分と、制約等高線（予算線と呼ばれる）

$$(p, q) \cdot (x, y) = m$$

が (x^*, y^*) で接している、という図形的な条件である。

これは非常にわかりやすいのだが、問題がここで発生する。つまり、等高線が予算線と接するという条件しか解いていないのだから、その等高線は予算線と下から接するか、上から接するかが、これだけでは判断できない。たとえば $f(x, y) = xy$ であればその等高線は $y = \frac{1}{x}$ であり、グラフを書いてみるとわかるとおりこれは予算線に上から接する。しかし $f(x, y) = x^2 + y^2$ であればその等高線のグラフは円の一部であり、よって予算線に下から接する。下から接した場合、グラフを書けば明らかなように、そこは解ではない。つまり、ラグランジュの方法で解けない問題の典型例としては、このように等高線が下から予算線と接している問題が挙げられる。このような場合では、$x = 0$ と $y = 0$ のどちらかが成り立つ点が、解である。

例：問題

$$\begin{array}{rl} \max & x^2 + y^2 \\ \text{subject to.} & x + 2y \leq 10, \\ & x \geq 0,\ y \geq 0 \end{array}$$

を考える。このとき、ラグランジュ関数

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(10 - x - 2y)$$

を微分して 0 と置くと

$$2x - \lambda = 0,$$

$$2y - 2\lambda = 0,$$

$$10 - x - 2y = 0$$

が出てくる。上のふたつから

$$y = \lambda = 2x$$

がわかり、これを下に代入すれば

$$x = 2, y = 4$$

が答えだと出てくる。

　しかし $x^2 + y^2$ は $(2, 4)$ のときは値が 20 であるのに対して、$(0, 5)$ だと 25 になる。したがってこれは解ではない。本当の解は $(10, 0)$ であり、そのときの値は 100 である。