Kobe University Repository : Kernel

PDF issue: 2024-04-30

制約付き最大化の解法と厳密性

宮川，栄一

(Citation)
国民経済雑誌,225(3):31-51

(Issue Date)
2022-03-10

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81013180

(URL)
https://hdl.handle.net/20.500.14094/81013180

ISSN 0387-3129

国民経済雑誌

制約付き最大化の解法と厳密性

宮　川　栄　一

国民経済雑誌　第225巻　第３号　抜刷

2022年３月

神戸大学経済経営学会

『国民経済雑誌』第225巻第３号（2022年３月），31-51

制約付き最大化の解法と厳密性†

宮　川　栄　一[a]

制約付き最大化問題を解くラグランジュ乗数の手法（非線形計画法）は実践的にも理論的にも重要だが，その解説は細かい仮定に無頓着な入門書タイプと，定理の数学的内容に集中した数学書タイプに二極化しているところがある。本稿ではそのギャップを埋めるべく，ラグランジュ乗数の定理に潜む細かい仮定に注意を払いながら，定理をどう使えば最大化の解を厳密に導出できるかの実践方法を細かく説明する。有名なクーン・タッカー（KKT）の定理ではなく，その原型と言えるフリッツ・ジョンの定理を中心に考える。

キーワード　フリッツ・ジョン，クーン・タッカー，KKT，制約想定，非線形計画法

1　は　じ　め　に

経済学の根幹にあるのは「選択肢のなかで一番いいと思うものを人は選ぶ」という考え方であり，それを数学的に表すのが制約付き最大化である。最大にするのは満足度や望ましさを表す指標であり，効用関数や利潤関数など，一般的には目的関数と呼ばれる関数である。目的関数の定義域にあるものすべてが選択可能なわけではなく，予算制約などの制約が付くことが大半で，そうした制約は不等式や等式で表されることが多い。

制約付き最大化問題を解く方法として便利なのが，ラグランジュ乗数を使った手法である。制約が等式の場合のラグランジュ定理，不等式の場合の（カルーシュ・）クーン・タッカー定理（KKT 定理）が有名だが，他にも様々な定理があり，一般的には非線形計画法と呼ばれている。この一連の手法は多くの教科書や専門書で解説されているが，わずらわしい注意事項は省いて使い方を簡単に説明する入門書タイプと，定理の数学的な詳細に集中して使い方には触れない数学書タイプに二極化しているところがある。しかし，定理を見ただけでは使い方は分かりづらいし，使い方は１つではない。しかも定理には仮定があるので，仮定に注意しながら使う必要があり，仮定を無視して使うと誤答に辿り着く危険性がある。

a　神戸大学大学院経済学研究科，miyagawa@econ.kobe-u.ac.jp

そこで本稿では，不等式制約が付いた一般的な最大化問題を考え，非線形計画法の１つの使い方を詳しく解説する。定理に付いている仮定に注意を払いながら，定理に沿う形で厳密に解を求める方法を紹介する。定理の使い方は１通りではないので，紹介する方法はあくまで１つの方法である。最大化問題は一般的なものを扱うが，念頭にあるのは消費者の効用最大化である。説明は初心者にも分かるようゆっくり行う。

本稿の方法の特徴は，制約が不等式なのにクーン・タッカーの定理は使わず，その原型と言えるフリッツ・ジョンの定理を使うことである。フリッツ・ジョンの定理は，入門書にはあまり載っていないが，クーン・タッカー定理の一般化であり，必要な仮定が少ない綺麗な定理なので，クーン・タッカーの落とし穴を回避する方法として有力である。

非線形計画法の定理と使用法の中間を詳しく解説している文献としてはSundaram（1996）がある。入門書のレシピがうまく行く場合や失敗する場合を詳しく解説している。ただフリッツ・ジョン定理への言及はない。フリッツ・ジョンの定理は，非線形計画法に詳しい文献には紹介されていることが多い。例としては西村（1990），小山（2010），永谷（1998），Simon and Blume（1994），マンガサリアン（1972）が挙げられる。

2　フリッツ・ジョン定理

不等式の制約が付いた最大化問題の一般形を考える。まずは記号を説明する。最大化問題において選択するのは $x\equiv(x_1, \ldots, x_n)\in\mathbb{R}^n$ というベクトルである。[1] 消費者問題の場合だと，n は財の種類数であり，x_i は財 i の消費量である。

制約条件は

$$g_j(x)\geqslant 0 \qquad j=1, \ldots, J$$

と書く。ただし J は制約の数で，数に制限はなく，$g_j:\mathbb{R}^n\to\mathbb{R}$ は任意の関数である。例えば消費者問題の場合，予算制約は $\sum_{i=1}^n p_i x_i\leqslant m$ だが，これを $g_j(x)\geqslant 0$ という形で表すには，$g_j(x)\equiv m-\sum_{i=1}^n p_i x_i$ と設定すればいい。ただし m は所得，p_iは財 i の価格であり，これらは外生である。消費者問題の場合，$x_i\geqslant 0$ という制約もあるが，これは $g_i(x)\equiv x_i$ とすればいい（ただし $i=1, \ldots, n$）。

最大化するのは $g_0(x)$ という関数とする。この $g_0:\mathbb{R}^n\to\mathbb{R}$ も任意の関数で，目的関数と呼ばれる。したがって制約付き最大化問題は

$$\begin{aligned}&\max_{x\in\mathbb{R}^n} g_0(x)\\ &\quad \text{s.t. } g_j(x)\geqslant 0 \quad \forall j=1, \ldots, J\end{aligned} \tag{P}$$

である。[2)3)] g_0 を最小にしたい場合は，$-g_0(x)$ を最大にすればいいので，最大化に限定しても問題ない。

ラグランジュ関数とは，すべての $g_j(x)$ の加重和である。「すべての」というのは目的関

数の $g_0(x)$ を含む。$g_j(x)$ に付く重みは λ_j と書き，ラグランジュ乗数と呼ばれる。したがってラグランジュ関数は

$$L(x) \equiv \lambda_0 g_0(x) + \lambda_1 g_1(x) + \cdots + \lambda_J g_J(x) = \sum_{j=0}^{J} \lambda_j g_j(x)$$

である。最初の $g_0(x)$にも乗数 λ_0を付けるのがフリッツ・ジョン定理の特徴である。クーン・タッカーの場合は $\lambda_0=1$ とするが，フリッツ・ジョンではそうしない。

これだけの準備でフリッツ・ジョンの定理（John, 1948）を述べることができる[4)]。証明は付録に掲載している。

定理 1（フリッツ・ジョン定理）　$\hat{x} \in \mathbb{R}^n$ を最大化問題（P）の解とする。この $\hat{x}$ において，すべての $j \geqslant 0$ について g_j が微分可能とする。すると以下を満たす $\lambda = (\lambda_0, \ldots, \lambda_J) \in \mathbb{R}^{J+1}$ が存在する。

(a)　$\dfrac{\partial L(\hat{x})}{\partial x_i} = 0$，つまり $\displaystyle\sum_{j=0}^{J} \lambda_j \frac{\partial g_j(\hat{x})}{\partial x_i} = 0 \quad \forall i = 1, \ldots, n$　（１階条件）

(b)　$\lambda_j \geqslant 0 \quad \forall j \geqslant 0$　（非負条件）

(c)　$\lambda_j > 0$ for some $j \geqslant 0$　（非ゼロ条件）

(d)　$\lambda_j g_j(\hat{x}) = 0 \quad \forall j \geqslant 1$　（相補スラック条件）

この４条件を**フリッツ・ジョン条件**あるいは **FJ 条件**と呼ぶ。最初の(a)がメインの条件で，ラグランジュ関数 $L(x)$を x_iで偏微分するとゼロになるという条件である。次の(b)は乗数 λ_j がすべてゼロ以上という条件である（$j=0$ も含めてである）。この条件だけだと，すべての λ_j がゼロの場合も許容されるが，それは駄目と言っているのが次の(c)である。λ_j のうち最低１つは厳密に正でないといけない。最後の(d)は，ラグランジュ関数の２項目以降がすべてゼロという条件である。λ_j と g_j のうちの片方が正なら，もう片方はゼロと言える（逆は言えない[5)]）。

クーン・タッカー条件との違いは，λ_0 があることと，非ゼロ条件(c)があることである。クーン・タッカーは $\lambda_0=1$ と設定するので，それだけで非ゼロ条件が満たされるから，非ゼロ条件をわざわざ付ける必要がない。

フリッツ・ジョン定理によると，４つの FJ 条件は最大解の必要条件である。微分可能性の仮定が満たされていれば，最大化問題の解 $\hat{x}$ は必ず FJ 条件を満たす。必要条件であって十分条件ではないので，FJ 条件を満たすものが解とは言えない。しかし悲観する必要もない。解が必ず満たす条件というのは強力な情報である。

3　必要条件の使い方

必要条件の使い道は消去法である。必要条件を満たさないものは，解でないと言い切れるので，すべて忘れていい。必要条件を満たすものに集中すればいい。FJ 条件は，微分したものがゼロというシャープな条件なので，通常はほとんどの候補が消去され，残るのはごく少数である。典型的な効用最大化だと，消去法で残るのは1つだけということが多い。残りが1つだけの場合，それ以外は解でないので，もし解があるなら，その1つが解だと言える。だから，あとは解が存在することが分かればいい。

解の存在については，非線形計画法のような微分を使う手法よりも一般的な手法で証明できることが多い。典型的なケースだと，ワイエルシュトラスの定理（最大値・最小値の定理）が使える。ワイエルシュトラスの定理とは，$f: X \to \mathbb{R}$ が連続関数で，$X \subseteq \mathbb{R}^n$ が有界かつ閉な集合（一般的にはコンパクト集合）で空でないなら，X 内で f を最大にするものが存在する，という定理である。標準的な消費者問題の場合，予算集合は（価格がすべて正なら）有界かつ閉であり，効用関数が微分可能なら自動的に連続関数なので，効用最大化の解は存在すると言える。

ワイエルシュトラス定理等を使って解の存在が証明できて，FJ 条件によって x が1つに絞れたら，その x は解であると言い切れる。それ以外は消去済みだからである。しかも同じ理由で，x が唯一の解であることも言えてしまう。必要条件であって十分条件でないのだが，解の存在証明と組み合わせることでシャープな答えを出せる。解の存在はどのみちチェックしておくのが望ましいので，これは1つの有効な方法である。

FJ 条件が解の候補を1つに絞り切れない場合もある。しかし残ったものが少数なら，それらを目的関数 $g_0(x)$ の値で直接比較すればいい。残ったものの中で $g_0(x)$ が一番大きいものが解の候補であり，解の存在が証明済みなら，それが解だと言い切れる。

4　クーン・タッカーとの関係

フリッツ・ジョンの定理では，乗数のベクトル $\lambda = (\lambda_0, \ldots, \lambda_J)$ は1つに決まらない。というのは，もし $\lambda = (\lambda_0, \ldots, \lambda_J)$ が FJ 条件を満たすなら，すべての λ_j を2倍したものも FJ 条件を満たすからである。すべての λ_j を2倍すると，1階条件の左辺はすべて2倍になるが，ゼロの2倍はゼロのままである。非負条件についても同様で，非負なものは2倍しても非負である。残り2つの条件も同様である（正の数を2倍しても正だし，ゼロを2倍してもゼロである）。2倍ではなく，任意の $k>0$ 倍でも同様である。

もし $\lambda_0 > 0$ であれば，すべての λ_j を λ_0 で割ってもいい。そうして出来た新たな乗数を λ'_j とすると

$$\lambda_0'=\frac{\lambda_0}{\lambda_0}=1,\ \lambda_1'=\frac{\lambda_1}{\lambda_0},\ \ldots,\ \lambda_J'=\frac{\lambda_J}{\lambda_0}$$

となるので，$\lambda_0'=1$ となる。つまり，もし $\lambda_0>0$ なら，$\lambda_0=1$ と設定してもいい。λ_j の間の相対的な比率だけが重要なので，$\lambda_0>0$ であれば，λ_0 を基準にして，λ_j は λ_0 の何倍かで表示してもいいということである。[6)]

もし $\lambda_0>0$ なら，$\lambda_0=1$ と設定できるし，非ゼロ条件も満たされるので，残る条件は

(a) $\dfrac{\partial g_0(\hat{x})}{\partial x_i}+\sum_{j=1}^{J}\lambda_j\dfrac{\partial g_j(\hat{x})}{\partial x_i}=0 \quad \forall i=1,\ldots,n$

(b) $\lambda_j\geqslant 0 \quad \forall j\geqslant 1$

(d) $\lambda_j g_j(\hat{x})=0 \quad \forall j\geqslant 1$

となる。これは**クーン・タッカー条件**と呼ばれるものに他ならない。[7)] つまりクーン・タッカー条件とは，FJ 条件に $\lambda_0>0$ を追加したものである。$\lambda_0>0$ は解の必要条件ではないので，クーン・タッカー条件も解の必要条件ではない。

クーン・タッカー条件が解の必要条件と言えないのは，彼らの定理（Kuhn and Tucker, 1951）を見ても分かる。

定理 2（クーン・タッカー定理） $\hat{x}\in\mathbb{R}^n$ を最大化問題（P）の解とする。この $\hat{x}$ において，制約想定（後述）が満たされるとする。また $\hat{x}$ において，すべての $j\geqslant 0$ について，g_j が微分可能とする。するとクーン・タッカー条件を満たす $(\lambda_1,\ldots,\lambda_J)\in\mathbb{R}^J$ が存在する。

フリッツ・ジョンとの違いは，制約想定と呼ばれる仮定が付いていることである。この仮定には様々なバリエーションがあるが，代表的なものは以下の通りである。偏微分 $\partial g_j(x)/\partial x_i$ をすべての i について列挙したベクトルを $\nabla g_j(x)$ と書く。つまり

$$\nabla g_j(x)\equiv\left(\frac{\partial g_j(x)}{\partial x_1},\ldots,\frac{\partial g_j(x)}{\partial x_n}\right)\in\mathbb{R}^n$$

である。また，J 個ある不等式制約のうち，x において等号成立しているものの集合を $E(x)$ と書く。つまり $E(x)\equiv\{j\geqslant 1: g_j(x)=0\}$ である。すると，

《定義》ベクトルの集合 $\{\nabla g_j(\hat{x}): j\in E(\hat{x})\}$ が 1 次独立のとき，**制約想定**が $\hat{x}$ で満たされると言う。つまり任意の $a\in\mathbb{R}^J$ について，

$$\left[\sum_{j\in E(\hat{x})}a_j\nabla g_j(\hat{x})=(0,\ldots,0)\right] \text{ ならば } [a_j=0\ \forall j\in E(\hat{x})]$$

が成立するということである。

制約想定の難点は，それを満たすべきなのが最大化の解 $\hat{x}$ だということである。最大化

の解 $\hat{x}$ は普通は未知なので，未知の解が制約想定を満たすことを事前に確認しなければならない。まさにこれから求めようとしている未知の解が制約想定を満たしてなければ，クーン・タッカー条件は解の必要条件ではない。しかも解の必要条件かどうかはクーン・タッカー条件を計算しても分からない。

未知の解が制約想定を満たさない場合でも，クーン・タッカー条件を計算すれば，それらしい答えは出るかもしれない。しかもその答えは制約想定を満たしているかもしれない。その場合，本人は安心して分析を進めてしまうかもしれないが，安心してはいけない。制約想定を満たすべきは最大化問題の解であって，クーン・タッカー条件の解ではない。とりあえずクーン・タッカーを計算して，その答えが制約想定を満たすかどうか確認すれば大丈夫，というわけではないのである。

クーン・タッカー条件がフリッツ・ジョン条件と違うのは $\lambda_0>0$ だけなので，制約想定も $\lambda_0>0$ を得るために仮定されている。それは以下の証明を見ても分かる。この証明は，フリッツ・ジョン定理を使ってクーン・タッカー定理を証明したものである。

《クーン・タッカー定理の証明》フリッツ・ジョン定理より，$\hat{x}$ において FJ 条件を満たす $(\lambda_0, \lambda_1, \ldots, \lambda_J)\in\mathbb{R}^{J+1}$ がある。$\lambda_0>0$ を証明すればいいので，背理法を使って $\lambda_0=0$ と仮定する。すると FJ の1階条件は

$$\sum_{j=1}^{J}\lambda_j\frac{\partial g_j(\hat{x})}{\partial x_i}=0 \quad \forall i=1, \ldots, n$$

となる。$\nabla g_j(x)$ を使ってベクトル表記すると

$$\sum_{j=1}^{J}\lambda_j\nabla g_j(\hat{x})=(0, \ldots, 0) \tag{1}$$

となる。$\sum$ にはすべての制約 j が含まれているが，厳密な不等号で成立している制約 j については，相補スラック条件より $\lambda_j=0$ なので，$\sum$ に入る項もゼロである。よって(1)は

$$\sum_{j\in E(\hat{x})}\lambda_j\nabla g_j(\hat{x})=(0, \ldots, 0) \tag{2}$$

と書ける。制約想定が $\hat{x}$ で成立しているという仮定より

$$\lambda_j=0 \quad \forall j\in E(\hat{x})$$

を得る。つまり等号成立している制約についても $\lambda_j=0$ となる。残る λ_0 は背理法でゼロと仮定しているので，結局 λ_j はすべてゼロとなり，非ゼロ条件に矛盾する。背理法より $\lambda_0>0$ を得るので，$\lambda_0=1$ としていい。残りの $(\lambda_1, \ldots, \lambda_J)$ がクーン・タッカー条件を満たすのは前述の通りである。　（証明終）

5　フリッツ・ジョン vs クーン・タッカー

フリッツ・ジョンとクーン・タッカーのどちらを使ってもいいのはもちろんだが，本稿ではフリッツ・ジョンの使用を勧めたい。一番の理由はフリッツ・ジョンの単純さである。微分可能な関数であれば必ず解の必要条件を出してくれる。制約想定の確認を行ってからラグランジュを計算するという2段階でないのも分かりやすい。もちろん，その代わりに後で λ_0 の処理が必要になるので，違いは表面的なことと言える。しかしフリッツ・ジョンの場合はラグランジュを処理するだけでよく，すぐに微分に取りかかれるのが単純である。加えて，計算結果に λ_0 が出てくるので，その処理を忘れることがないというメリットもある。

フリッツ・ジョンを使って $\lambda_0>0$ を処理すると言っても話が抽象的なので，標準的な消費者問題の場合で具体的に実践してみる。効用関数は一般形にして $u(x)$ と書く。ただし $x=(x_1, \ldots, x_n)$ は消費ベクトルである。制約は

$$\sum_{i=1}^{n} p_i x_i \leqslant m$$

という予算制約と

$$x_i \geqslant 0 \quad \forall i=1, \ldots, n$$

という非負制約である。$x_i \geqslant 0$ という非負条件を $g_i(x)$ にして，$g_i(x)=x_i(\forall i=1, \ldots, n)$ とすると添字が揃って分かりやすい。予算制約は最後の $n+1$ 番目の制約式にして，$g_{n+1}(x)=m-\sum_{i=1}^{n} p_i x_i$ とする。

普通は価格はすべて正と仮定するが，ここの証明は価格がゼロでも問題ないので，$p_i \geqslant 0$ とする。$p_i>0$ を満たす財の集合を $S \equiv \{i : p_i>0\}$ と書く。所得は $m>0$ と仮定する。

フリッツ・ジョンのラグランジュ関数は

$$L(x)=\lambda_0 u(x)+\sum_{i=1}^{n} \lambda_i x_i+\lambda_{n+1}\left\{m-\sum_{i=1}^{n} p_i x_i\right\}$$

である。1階条件は

$$\lambda_0 \frac{\partial u(x)}{\partial x_i}+\lambda_i-\lambda_{n+1} p_i=0 \quad \forall i=1, \ldots, n \tag{3}$$

となる。非負条件より $\lambda_j \geqslant 0$ がすべての j で成り立ち，非ゼロ条件より λ_j のうちの最低1つは $\lambda_j>0$ を満たす。相補スラック条件より，ラグランジュの2項目以降の $\lambda_j g_j(x)$ がすべてゼロなので，以下を得る。

$$\lambda_i x_i=0 \quad \forall i=1, \ldots, n \tag{4}$$

$$\lambda_{n+1}\left\{m-\sum_{i=1}^{n} p_i x_i\right\}=0 \tag{5}$$

問題の $\lambda_0>0$ だが，背理法で $\lambda_0=0$ と仮定する。(3)の初項がゼロになるので，

$$\lambda_i=\lambda_{n+1}p_i \quad \forall i=1, \ldots, n \tag{6}$$

を得る。ここから矛盾を引き出せばいい。2つのケースに分けると簡単である。

《ケース1》 $\lambda_{n+1}=0$ の場合。(6)より $\lambda_i=0\ \forall i=1, \ldots, n$ を得る。すると λ_0 から λ_{n+1} まで全部ゼロになるので，非ゼロ条件に矛盾する。

《ケース2》 $\lambda_{n+1}>0$ の場合。価格が $p_i>0$ となる財 i（つまり $i\in S$）については，(6)の右辺は $\lambda_{n+1}p_i>0$ である。よって左辺も正となり，$\lambda_i>0\ \forall i\in S$ を得る。したがって相補スラック条件(4)より，$x_i=0\ \forall i\in S$ を得る。つまり価格が正の財は全く消費しない。一方，$\lambda_{n+1}>0$ に相補スラック条件(5)を適用すると

$$m=\sum_{i=1}^{n}p_ix_i \tag{7}$$

を得る。つまり予算はすべて使い切る。しかし価格が正の財をまったく消費しないので，これは矛盾である。式で確認すると，価格が正の財 $i\in S$ については $x_i=0$ だし，価格がゼロの財 $i\notin S$ については $p_i=0$ なので，どちらの財についても $p_ix_i=0$ となるので，(7)の右辺はゼロになるが，左辺は仮定より $m>0$ なので矛盾である。

結局どちらのケースも矛盾になるので，背理法より $\lambda_0>0$ を得る。　　（証明終）

この証明で分かる通り，標準的な消費者問題では $\lambda_0>0$ が成立する。なので安心してクーン・タッカーを使っていい。しかし実際の研究で直面する最大化問題で同じことが言える保証はない。$\lambda_0>0$ と分かっていないなら，$\lambda_0>0$ かどうかの確認は必要である。

ただ，研究で直面する問題のほとんどでは $\lambda_0>0$ が成立する。というのは，上の背理法でも分かる通り，もし $\lambda_0=0$ なら，効用関数などの目的関数が FJ 条件から消えてしまう。そんな条件で解が決まるとすれば，最大化問題に何か問題があると考えるのが自然である。しかし，だからと言って確認が省略できるわけでないのは当然である。

ただ $\lambda_0>0$ が簡単に分かる場合もある。1番簡単なのは制約条件がすべて x について線形（アフィン関数）の場合である。

命題1　すべての $j\geqslant 1$ で $g_j(x)$ がアフィン関数ならば，つまり $g_j(x)=\alpha_{j0}+\sum_{i=1}^{n}\alpha_{ji}x_i$ ならば（$\alpha_{j0}, \ldots, \alpha_{jn}$ は定数），フリッツ・ジョン定理の結論に $\lambda_0>0$ を追加できる。[8)]

標準的な消費者問題では予算制約も非負条件もアフィン関数なので，この命題が使える（上記で証明したので使う必要はないが）。この命題の証明は，フリッツ・ジョン定理を少し拡張することで得られる（付録の定理6参照）。

制約式がアフィン関数でない場合も経済学ではよくある。その場合も制約式がすべて凹関数なら次の命題が使える。

命題 2　$g_j(x)$ がすべての $j \geqslant 1$ で凹関数で，かつ

$$g_j(x^*) > 0 \quad \forall j \geqslant 1 \qquad \text{（スレーター条件）}$$

を満たす $x^* \in \mathbb{R}^n$ が存在する場合，FJ 条件が成立すれば自動的に $\lambda_0 > 0$ が成立する。

《証明》FJ 条件を満たす $\hat{x}$ と λ を考え，背理法を使って $\lambda_0 = 0$ と仮定する。クーン・タッカー定理の証明の（2）までは同じなので，（2）を再掲する。

$$\sum_{j \in E(\hat{x})} \lambda_j \nabla g_j(\hat{x}) = (0, \ldots, 0) \qquad (8)$$

どの $j \in E(\hat{x})$ でも話は同じなので，任意の $j \in E(\hat{x})$ を考える。仮定より $g_j(x)$ は凹関数なので，凹関数の一般的な性質より

$$\nabla g_j(\hat{x}) \cdot (x^* - \hat{x}) \geqslant g_j(x^*) - g_j(\hat{x})$$

が成り立つ。右辺を見ると，x^* の定義より $g_j(x^*) > 0$ だし，$j \in E(\hat{x})$ なので $g_j(\hat{x}) = 0$ である。よって右辺は正なので，以下を得る。

$$\nabla g_j(\hat{x}) \cdot (x^* - \hat{x}) > 0 \quad \forall j \in E(\hat{x}) \qquad (9)$$

最後に，（8）の両辺と $(x^* - \hat{x})$ との内積をとると

$$\sum_{j \in E(\hat{x})} \lambda_j \nabla g_j(\hat{x}) \cdot (x^* - \hat{x}) = 0$$

を得る。しかし先ほどの（9）と非負条件の $\lambda_j \geqslant 0$ より，左辺がゼロになるには $\sum$ の中の λ_j は全部ゼロでなければならない。つまり $\lambda_j = 0 \ \forall j \in E(\hat{x})$ である。しかし $E(\hat{x})$ に含まれない制約については $g_j(\hat{x}) > 0$ なので相補スラック条件より $\lambda_j = 0$ だし，背理法の仮定から $\lambda_0 = 0$ なので，結局 λ_j はすべてゼロになる。これは非ゼロ条件に矛盾する。　（証明終）

こうした命題を使えば，経済学で出てくる問題の多くで $\lambda_0 > 0$ となることが分かる。したがってそういう問題では最初からクーン・タッカー条件を使って問題ない。しかし目の前の問題でも $\lambda_0 > 0$ と言えるかは分からないので，注意が必要である。特に目的関数や制約式が外生ではなく，何か別の問題を解いて導出された式の場合，それらが望ましい性質を満たしているとは限らない。

6　微分可能性の仮定

フリッツ・ジョン定理の唯一の仮定は微分可能性である。解 $\hat{x}$ においてすべての g_j が微分可能でなければならない。関数を微分できるかは自明なことが多いし，微分できる関数しか使わないことがほとんどなので，この仮定は大して問題にならない，と言いたいところだが，そうとは言い切れない。実感的には微分できる関数でも，実は微分可能でない，ということがよくある。

どういうことかと言うと，「微分可能」の定義には微分が有限ということが含まれている

のである。微分の値が無限の場合，微分する行為は実感上はできているかもしれないが，無限というのは実数ではないので，定義上は微分可能でないのである。変数 x がベクトルの場合，すべての偏微分が有限でなければならない。[9)]

フリッツ・ジョン定理で考えると，解 $\hat{x}$ における偏微分

$$\frac{\partial g_j(\hat{x})}{\partial x_i}$$

が（どの $j \geqslant 0$ と i についても）有限でなければならない。しかし偏微分が無限になることはよくある。例えば経済学では $g_j(x)=\sum_{i=1}^{n}\sqrt{x_i}$ というような関数がよく使われるが，$\sqrt{x_i}$ を微分すると $x_i=0$ で無限大になる。念のため計算すると

$$\frac{\partial g_j(x)}{\partial x_i}=\frac{1}{2\sqrt{x_i}}$$

となるので，$x_i=0$ の場合は分母がゼロになる。ゼロで割ることは出来ないので，この分数は実数ではない（意味的には無限大である）。なので，$x_i=0$ となる i が1つでもある x では，この関数 $g_j(x)$ は微分可能ではない。したがって，そういう x が最大化の解である場合，フリッツ・ジョン定理の仮定が満たされないので，その解がFJ条件を満たすとは言えない。FJ条件やクーン・タッカー条件を計算しても意味がないかもしれないのである。意味があるかどうかは，今まさに求めようとしている未知の解の値次第である。

偏微分が無限になるケースが問題だと言うが，そういうケースは1階条件によって排除できるのではないか，という議論がある。偏微分が無限になる点だと，1階条件のどこかが無限になるので，それを使って1階条件の左辺が無限になると示せばいい。そうすれば左辺がゼロにならないことが分かるので，1階条件を満たさないことになり，解でないと言えるのではないかという議論である。直感的には強力な議論であり，内容的にも的外れとは言えないが，厳密な議論とは言えない。というのは，偏微分が無限になる点では，フリッツ・ジョン定理やクーン・タッカー定理の仮定が成立しないので，1階条件は解の必要条件ではない。解の必要条件でないものを満たさないからと言って，解でないとは主張できない。偏微分が無限になる点は，1階条件を満たさなくても，解の可能性はある。

ではどうすればいいか。1つの対処法は，微分可能でない点は最大化の解でない，ということを事前に証明してしまうことである。ほとんどの場合，微分不可能なのは限られた点だけである。経済学でよく問題になるのは，上記のように $x_i=0$ となる i が存在する点（軸上の点）である。そういう点は最大解にならないことを示せばいい。そうすれば最大解はFJ条件を満たすと言える。偏微分が無限大になる点だと，x_i をわずかに増減させるだけで効用が無限の速度で増えたり，コストが無限の速度で減ったりする。内容を直感的に考えれば，そういう点が解にならないことも，どうしたらそれを証明できるかも分かることが多い。

この対処法でうまく行かない場合なら，まず最大解が微分可能性を満たすと想定して，FJを使って解の候補を絞る。その解の候補と，微分可能性を満たさない点の全部について目的関数 $g_0(x)$ の値を比較する，という手がある。解が微分可能性を満たしていれば解の候補に残るからである。ただここまで必要となるケースは少ない。

具体例として，消費者問題の場合で最初の対処法を実践してみる。

《問題》標準的な消費者問題を考える。効用関数は $u(x)=\sum_{i=1}^{n}u_i(x_i)$ である。すべての i について

$$\lim_{x_i\to 0}u_i'(x_i)=+\infty$$

である。また $x_i>0$ では $u_i'(x_i)$ は有限で，しかも連続とする（微分が連続に変化する）。例えば $u_i(x_i)=\alpha_i x_i^{\beta}$（ただし $\alpha_i>0$ かつ $0<\beta<1$）である。また，少なくとも $u_1(x_1)$ は厳密な増加関数とする。制約は通常の予算制約と非負制約である。所得 $m>0$ も価格 $p_i>0$ もすべて正とする。このとき効用最大化の解 x は $x_i>0\ \forall i$ を満たすことを示しなさい。

《例解》x を最大化の解とし，背理法より，$x_i=0$ となる財 i があると仮定する。この場合でも $x=(0,\ldots,0)$ではない。というのは，所得が正なので，$\varepsilon>0$ が十分小さいと $y=(\varepsilon,0,\ldots,0)$は予算制約を満たし，$u_1(x_1)$ が増加関数なので，$u(y)>u(0,\ldots,0)$となるからである。よって $x_j>0$ を満たす財 j がある。当然 $j\neq i$ である。

そこで x_j を少し減らして，その代わりに x_i を正にすることを考える。具体的には x を以下の x' に修正する。

$$x_i'=\varepsilon>0$$

$$x_j'=x_j-\frac{\varepsilon p_i}{p_j}>0$$

$$x_k'=x_k\quad \forall k\neq i,j$$

つまり財 i の消費量をゼロから $\varepsilon>0$ に増やして，そのために必要な金額 εp_i を捻出するために必要なだけ財 j を減らす。そうすれば予算制約は維持できるし，$\varepsilon>0$ が小さいと x_j' は負にならない。念のために支出額の変化を計算すると

$$\sum_{k=1}^{n}p_k x_k'-\sum_{k=1}^{n}p_k x_k=p_i(x_i'-x_i)+p_j(x_j'-x_j)=p_i\varepsilon-p_j\frac{\varepsilon p_i}{p_j}=0$$

である。x は最大化問題の解なので予算制約を満たすから，x' も予算制約を満たす。

あとは効用が増えることを示せばいい。効用の変化を

$$f(\varepsilon)\equiv u(x')-u(x)=u_i(\varepsilon)-u_i(0)+u_j\left(x_j-\frac{\varepsilon p_i}{p_j}\right)-u_j(x_j)$$

と定義して，ε で微分すると

$$f'(\varepsilon)=u_i'(\varepsilon)-u_j'\left(x_j-\frac{\varepsilon p_i}{p_j}\right)\frac{p_i}{p_j}$$

を得る。$\varepsilon \to 0$ とすると，第1項の $u_i'(\varepsilon)$ は $+\infty$ になる一方，第2項の極限は $u_j'(x_j)p_i/p_j$ という有限値である[10]。よって $\lim_{\varepsilon \to 0} f'(\varepsilon) = +\infty$ となり，$\varepsilon > 0$ が十分小さいと $f'(\varepsilon) > 0$ となる。厳密に書くと，$\varepsilon_1 > 0$ が存在して，どんな $\varepsilon \in (0, \varepsilon_1)$ でも $f'(\varepsilon) > 0$ となる。

$\varepsilon = 0$ の場合は $x' = x$ なので，$f(0) = 0$ である。$\varepsilon \in (0, \varepsilon_1)$ の範囲では，前の段落より微分は正なので，$f(\varepsilon) > 0$ となりそうだ。厳密に証明すると，平均値の定理より

$$\frac{f(\varepsilon) - f(0)}{\varepsilon - 0} = f'(\varepsilon_2)$$

を満たす $\varepsilon_2 \in (0, \varepsilon)$ が存在する。$0 < \varepsilon_2 < \varepsilon < \varepsilon_1$ なので，前段落より右辺は正である。よって左辺も正なので，$f(\varepsilon) > f(0) = 0$ を得る。したがって，この ε では $u(x') > u(x)$ となる。x' は予算制約を満たすので，x が最大化の解であることに矛盾する。　（例解終）

7　最大解を求める手順

以上で議論してきた最大解の導出手順をまとめると以下のようになる。

(1)　最大解が存在することが分かっていなければ証明する。

(2)　g_j が微分可能性を満たさない点 x があれば，それらの点は最大解にならないことを証明する。証明できない場合は(5)で再検討する。

(3)　FJ 条件と制約条件を満たす (x, λ) の組をすべて求める。$\lambda_0 > 0$ が証明済みでなければ，証明を試みる。最大解が満たすと分かっている追加の必要条件があれば，それを満たす x に限定して構わない。

(4)　FJ 条件と制約条件（および追加の必要条件）を満たす (x, λ) が1つしかなければ，その x が唯一の最大解である。もし (x, λ) が複数あるなら，そのなかで g_0 を最大にするものが解である。

(5)　もし(2)の微分可能性の証明ができていないなら，微分可能でない点 x のすべてと(3)で求れられた x すべてとで g_0 を直接比較し，最大にするものが解である。

8　例　　題

例題としては，大学院レベルのミクロ経済学で重要になる支出最小化問題の典型的な計算問題を考える。支出最小化問題とは，所与の値 w 以上の効用を得るために必要な所得（支出額）の最小値を求める問題である。

効用関数は

$$u(x) = \sum_{i=1}^{n} \alpha_i x_i^{\beta}$$

とする。ただし $\alpha_i > 0\ \forall i$ かつ $0 < \beta < 1$ である。必要な効用水準は $w > 0$ とする。価格は p_i

$>0\ \forall i$ とする。最小化するのは支出額 $\sum_{i=1}^{n} p_i x_i$ である。これにマイナスを掛けた値を最大にすればいいので，最大化問題は

$$\begin{aligned}&\max_{x\in\mathbb{R}^n} -\sum_{i=1}^{n} p_i x_i\\ &\quad \text{s.t. } u(x)\geqslant w\\ &\qquad\quad x_i\geqslant 0 \quad \forall i=1,\ldots,n\end{aligned}$$

である。$x_i\geqslant 0$ に付ける乗数を λ_i とすれば添字が揃うのは以前と同じなので，$u(x)-w\geqslant 0$ に付ける乗数は λ_{n+1} とする。ラグランジュ関数は

$$L=-\lambda_0\sum_{i=1}^{n} p_i x_i+\sum_{i=1}^{n}\lambda_i x_i+\lambda_{n+1}\{u(x)-w\}$$

となる。1階条件は以下の通りである。

$$-\lambda_0 p_i+\lambda_i+\lambda_{n+1}\frac{\partial u(x)}{\partial x_i}=0 \quad \forall i=1,\ldots,n$$

《解の存在》計算を進める前に，まずは解が存在するかを確認する。効用最大化ではワイエルシュトラス定理が使えるが，今の問題では制約を満たす x の集合が有界ではない。x がいくら大きくても制約は満たし得るからである。解決策としては，制約を満たすベクトル $\hat{x}$ を1つ固定する。そういう $\hat{x}$ の存在は今の効用関数では自明である。すると，最低でもその $\hat{x}$ を選べばいいので，$\hat{x}$ より支出額が大きいものは選択肢から外していい。つまり

$$\sum_{i=1}^{n} p_i x_i\leqslant\sum_{i=1}^{n} p_i\hat{x}_i$$

という制約を加えてもよい。するとこれを加えた制約を満たす x の集合は有界性を満たす。閉集合でもあるし，目的関数は x の連続関数なので，ワイエルシュトラスの定理が問題なく使えて，解の存在が言える。

《微分可能性》$x_i=0$ となる i があると微分可能性が満たされないので，そういう点は解にならないことを示す。証明は第6節の例解とほぼ同じなので繰り返しは避ける。ただ x' を以前と同じように作ると，支出額が変わらないので矛盾が得られない。支出額が厳密に減るよう数量を多めに減らす必要がある。具体的には

$$x'_j=x_j-\frac{2\varepsilon p_i}{p_j}>0$$

などとすればよい。$\varepsilon>0$ が十分小さいと $x'_j>0$ だし，$f'(\varepsilon)\to\infty$ となるのも同じである。

《$\boldsymbol{\lambda_0>0}$》次は $\lambda_0>0$ を証明する。前の段落より $x_i>0\ \forall i$ なので，相補スラック条件より $\lambda_i=0\ \forall i=1,\ldots,n$ である。背理法を使って $\lambda_0=0$ と仮定すると，1階条件は

$$\lambda_{n+1}\frac{\partial u(x)}{\partial x_i}=0 \quad \forall i$$

である。$x_i>0$ より $\partial u(x)/\partial x_i>0$ なので，$\lambda_{n+1}=0$ を得る。したがって λ_j はすべてゼロになるので，非ゼロ条件に矛盾する。よって $\lambda_0>0$ であり，$\lambda_0=1$ と設定して構わない。

《x の計算》あとは普通に計算して x を求める。1階条件は

$$p_i=\lambda_{n+1}\frac{\partial u(x)}{\partial x_i}=\lambda_{n+1}\alpha_i\beta x_i^{\beta-1}\quad\forall i$$

である。x_i について求めると

$$x_i=\left(\frac{\lambda_{n+1}\alpha_i\beta}{p_i}\right)^{\frac{1}{1-\beta}}\quad\forall i \tag{10}$$

を得る。$x_i>0$ なので，$\lambda_{n+1}>0$ となり，相補スラック条件より $u(x)-w=0$ を得る。これに先ほど求めた x_i を代入すると

$$w=u(x)=\sum_{i=1}^{n}\alpha_i\left(\frac{\lambda_{n+1}\alpha_i\beta}{p_i}\right)^{\frac{\beta}{1-\beta}}=(\lambda_{n+1}\beta)^{\frac{\beta}{1-\beta}}\sum_{i=1}^{n}\alpha_i\left(\frac{\alpha_i}{p_i}\right)^{\frac{\beta}{1-\beta}}$$

となるので，書き換えると

$$(\lambda_{n+1}\beta)^{\frac{\beta}{1-\beta}}=\frac{w}{\sum_{i=1}^{n}\alpha_i\left(\frac{\alpha_i}{p_i}\right)^{\frac{\beta}{1-\beta}}}$$

となる。これを(10)に戻すと

$$x_i=(\lambda_{n+1}\beta)^{\frac{1}{1-\beta}}\left(\frac{\alpha_i}{p_i}\right)^{\frac{1}{1-\beta}}=\left[\frac{w}{\sum_{j=1}^{n}\alpha_j\left(\frac{\alpha_j}{p_j}\right)^{\frac{\beta}{1-\beta}}}\right]^{\frac{1}{\beta}}\left(\frac{\alpha_i}{p_i}\right)^{\frac{1}{1-\beta}}\quad\forall i$$

を得る。これで必要条件を満たす唯一の x が求まった。解が存在することは証明済みなので，この x が唯一の最大解である。　（例解終）

9　最大解の十分条件

本稿では解の必要条件を使うアプローチを取り上げたが，非線形計画法には解の十分条件もある。必要とする仮定は強いが，仮定が満たされていれば使いやすいし，経済学では満たされるケースが大多数である。

標準的な消費者問題の場合，無差別曲線が予算線に接する点ならば，その点は効用最大化の解だと言える。つまり，接するという1階微分条件さえ満たせば，最大解と判定するには十分である。これは予算集合が凸集合で，無差別曲線が原点に向かって凸だから（無差別曲線の望ましい側が凸集合だから）である。無差別曲線が原点に対して凸でなかったり，予算集合が凸集合でないと，1階条件を満たす点が解とは言い切れない。これを一般化した定理の代表例が以下である（Arrow and Enthoven, 1961）。

定理 3（アロー・エントーベン定理） 最大化問題（P）の制約条件を満たす $\hat{x}\in\mathbb{R}^n$ を考える。すべての $j\geqslant 0$ について g_j は準凹関数で，$\hat{x}$ において微分可能とする[11]。$\hat{x}$ において FJ 条件と $\lambda_0>0$ を満たす $\lambda\in\mathbb{R}^{J+1}$ が存在して

$$\nabla g_0(\hat{x})\neq 0 \tag{11}$$

が成り立つなら，$\hat{x}$ は最大化問題の解である。

つまり目的関数も制約関数もすべて準凹関数の場合は，FJ 条件と $\lambda_0>0$（つまりクーン・タッカー条件）および制約条件を満たす $\hat{x}$ であれば，それが $\nabla g_0(\hat{x})\neq 0$ さえ満たしていれば，最大解だと言い切れる。$\nabla g_0(\hat{x})\neq 0$ が成立するには，

$$\frac{\partial g_0(\hat{x})}{\partial x_i}\neq 0$$

を満たす $i=1, \ldots, n$ が 1 つでもあればいい。しかもそれは解の候補である $\hat{x}$ で満たせばよく，すべての x で満たす必要はない。証明も綺麗なので紹介しておく。

《証明》 制約条件と FJ 条件を満たす $(\hat{x}, \lambda)$ を任意にとり，$\lambda_0>0$ と(11)を満たすとする。$\lambda_0>0$ なので，$\lambda_0=1$ としてよい。制約条件を満たす任意の $y\in\mathbb{R}^n$ を考え，$g_0(y)\leqslant g_0(\hat{x})$ となることを証明する。

1 階条件をベクトル表記すると

$$\nabla g_0(\hat{x})+\sum_{j=1}^{J}\lambda_j\nabla g_j(\hat{x})=(0, \ldots, 0)$$

である。両辺に $(y-\hat{x})$ を掛ける（内積をとる）と

$$\nabla g_0(\hat{x})\cdot(y-\hat{x})+\sum_{j=1}^{J}\lambda_j\nabla g_j(\hat{x})\cdot(y-\hat{x})=0$$

となる。以前と同様，$\hat{x}$ において等号成立している制約 j の集合を $E(\hat{x})$ と書く。すると，任意の制約 $j\notin E(\hat{x})$ については，$g_j(\hat{x})>0$ なので，相補スラック条件より $\lambda_j=0$ を得る。したがって，それらの j は上の $\sum$ から削除できて，以下を得る。

$$\nabla g_0(\hat{x})\cdot(y-\hat{x})+\sum_{j\in E(\hat{x})}\lambda_j\nabla g_j(\hat{x})\cdot(y-\hat{x})=0 \tag{12}$$

任意の $j\in E(\hat{x})$ を考える。定義より $g_j(\hat{x})=0$ である。y も制約条件を満たすので，$g_j(y)\geqslant 0$ である。g_j は準凹関数なので，任意の $t\in(0, 1)$ について

$$g_j(\hat{x}+(y-\hat{x})t)\geqslant\min\{g_j(\hat{x}), g_j(y)\}=0$$

である。両辺から $g_j(\hat{x})=0$ を引いて t で割ると

$$\frac{g_j(\hat{x}+(y-\hat{x})t)-g_j(\hat{x})}{t}\geqslant 0$$

を得る。これは任意の $t\in(0, 1)$ で成り立つので，$t\to 0$ の極限でも成立する。よって

$$0 \leqslant \lim_{t \to 0} \frac{g_j(\hat{x}+(y-\hat{x})t)-g_j(\hat{x})}{t} = \nabla g_j(\hat{x}) \cdot (y-\hat{x})$$

を得る。これは任意の $j \in E(\hat{x})$ で成立するので，(12)に代入して以下を得る。

$$\nabla g_0(\hat{x}) \cdot (y-\hat{x}) = -\sum_{j \in E(\hat{x})} \lambda_j \nabla g_j(\hat{x}) \cdot (y-\hat{x}) \leqslant 0 \tag{13}$$

仮定より $\nabla g_0(\hat{x}) \neq 0$ なので，$z \in \mathbb{R}^n$ が存在して

$$\nabla g_0(\hat{x}) \cdot (z-\hat{x}) < 0$$

を満たす。任意の $s \in (0, 1)$ を左辺に掛け，(13)の左辺に $(1-s)$ を掛けて合計すると

$$\nabla g_0(\hat{x}) \cdot (sz+(1-s)y-\hat{x}) < 0$$

を得る。g_0 は準凹関数なので，その特性より

$$\nabla g_0(\hat{x}) \cdot (a-\hat{x}) < 0 \text{ ならば } g_0(a) < g_0(\hat{x})$$

が任意の a で成り立つ。$a = sz+(1-s)y$ だと左側が成り立つので，右側も成立して

$$g_0(sz+(1-s)y) < g_0(\hat{x})$$

を得る。これは任意の $s \in (0, 1)$ で成立するので，$s \to 0$ の極限を取ると，

$$\lim_{s \to 0} g_0(sz+(1-s)y) \leqslant g_0(\hat{x})$$

を得る（等号が付く）。g_0 は連続関数なので，左辺は $g_0(y)$ に等しく，よって $g_0(y) \leqslant g_0(\hat{x})$ を得る。y は制約条件を満たせば任意なので，$\hat{x}$ は最大化の解である。　　（証明終）

こうした十分条件の定理が便利になる場合として，最大化問題に解が存在するかが（ワイエルシュトラス定理などを使っても）分からない場合が挙げられる。解の存在が不明だと，最大解の必要条件を満たす x が1つしかなくても，それが解とは言えない。解があるなら x だが，解がないのかもしれない。しかしアロー・エントーベン定理の条件をすべて満たしていれば，x が解と言い切れる。g_0 が狭義の準凹関数なら，解が1つしかないことも言える。強力な定理だが，その代償は g_j がすべて準凹関数でないといけないことである。

十分条件を使う場合の注意事項として，その条件を使って x が求まった場合でも，実は条件の一部しか使っていない場合がある，ということが挙げられる。その場合，求まった x が条件をすべて満たしているとは限らない。十分条件はすべて満たさないといけないので，x が求まったら，条件をすべて満たすことの確認が必要である。

10　お　わ　り　に

フリッツ・ジョンを使うと，λ_0 という変数が1つ増えてしまう，と思うかもしれない。しかし非ゼロ条件より $\lambda_j > 0$ となる j は存在する。したがってそれがどの j か分かった後であれば $\lambda_j = 1$ と設定できるので，変数が増えるわけではない。

たとえて言えば，選挙における各候補者の得票数を相対比率で表そうとして，得票数を1

に基準化する候補者を事前に決めてしまう，みたいな話である。だれを1にしてもほとんどの場合は問題ないのだが，得票数がゼロの人を1にしたら駄目である。得票数が正の人がだれか分かってから基準となる人を選定すべきである。

付録：フリッツ・ジョン定理の証明

フリッツ・ジョン定理の証明はシンプルで美しいので，完結性のために紹介する。

《証明》最大化問題の解を x とし，この点においてすべての g_j が微分可能とする。本文のときと同様，偏微分 $\partial g_j(x)/\partial x_i$ をすべての i について列挙したベクトルを $\nabla g_j(x)$ と書く。また x において等号成立している制約の集合を $E(x)$ とする。

《ステップ 1》 $z\cdot\nabla g_j(x)>0$ をすべての $j\in E(x)\cup\{0\}$ で満たす $z\in\mathbb{R}^n$ は存在しない。

背理法を使って，そのような z が存在すると仮定する。つまり，すべての $j\in E(x)\cup\{0\}$ について $z\cdot\nabla g_j(x)>0$ が成立する。当然ながら $z\neq(0,\ldots,0)$ である（掛けて正になるので）。任意の $j\in E(x)\cup\{0\}$ を考える。小さい $\varepsilon>0$ を考えて，x を $x+\varepsilon z$ にすることを考える。$g_j(x)$ は微分可能なので，微分可能性の定義より

$$\lim_{\varepsilon\to 0}\frac{|g_j(x+\varepsilon z)-g_j(x)-\varepsilon z\cdot\nabla g_j(x)|}{\varepsilon\|z\|}=0$$

である。ただし $\|z\|$ は原点から z までの距離である。$z\neq 0$ なので $\|z\|>0$ である。$\varepsilon\to 0$ としても $\|z\|$ は変化しないので，上の式の $\|z\|$ は無視してよく，すると上の式は

$$\lim_{\varepsilon\to 0}\left|\frac{g_j(x+\varepsilon z)-g_j(x)}{\varepsilon}-z\cdot\nabla g_j(x)\right|=0$$

となる。つまり

$$\lim_{\varepsilon\to 0}\frac{g_j(x+\varepsilon z)-g_j(x)}{\varepsilon}=z\cdot\nabla g_j(x)$$

である。背理法の仮定より右辺は正なので，左辺も正である。よって $\varepsilon>0$ が十分小さいと

$$g_j(x+\varepsilon z)-g_j(x)>0$$

となる。これは所与の $j\in E(x)\cup\{0\}$ についてであったが，$\varepsilon>0$ が十分に小さければ，すべての $j\in E(x)\cup\{0\}$ で成り立つ，つまり

$$g_j(x+\varepsilon z)>g_j(x)\quad\forall j\in E(x)\cup\{0\}\tag{14}$$

を得る。$j\in E(x)$ だと $g_j(x)=0$ なので

$$g_j(x+\varepsilon z)>0\quad\forall j\in E(x)$$

である。$E(x)$ に入っていない制約については，$g_j(x)>0$ なので，$\varepsilon>0$ を十分に小さく取っておけば，$g_j(x+\varepsilon z)>0$ となる（g_j は連続関数なので）。まとめると，$\varepsilon>0$が十分に小さければ，$x+\varepsilon z$ は制約をすべて満たす。しかも(14)は $j=0$ についても成り立つので，$x+\varepsilon z$ は x よりも目的関数 g_0 の値が大きい。ということは，制約を満たすもので x より目的関数の値が大きいものがあることになる。これは x が最大化の解という仮定に矛盾する。これでステップ1の証明ができた。

《ステップ 2》いま証明したことを集合の形で書き直す。g_j の添字 j の順番に特に意味はないので，$E(x)=\{1,\ldots,m\}$ と仮定しても構わない(ただし $m\leqslant J$)。つまり m 番目までの制約が x で等号成立している。それで

$$A \equiv \{(z \cdot \nabla g_0(x), \ldots, z \cdot \nabla g_m(x)) : z \in \mathbb{R}^n\}$$

と定義する。つまり，

$$a = (z \cdot \nabla g_0(x), \ldots, z \cdot \nabla g_m(x)), \quad z \in \mathbb{R}^n$$

というベクトル a の全体集合が A である。前段落で証明したことは，A 内のベクトル a で

$$a_j > 0 \quad \forall j = 0, \ldots, m$$

を満たすものはないということである。正の実数の集合を $\mathbb{R}_{++}$ と書くと，$\mathbb{R}^{m+1}_{++}$ は A と共通部分を持たないということである。しかも A も $\mathbb{R}^{m+1}_{++}$ も凸集合である。したがって共通部分を持たない2つの凸集合が存在する。すると，分離定理によって，ベクトル $p \neq 0$（ただし $p = (p_0, \ldots, p_m) \in \mathbb{R}^{m+1}$）が存在して，

$$p \cdot a \leqslant p \cdot b \quad \forall a \in A, \ \forall b \in \mathbb{R}^{m+1}_{++} \tag{15}$$

が成立する。予算制約のような平面で2つの凸集合を分離することができるのである。

ベクトル b の成分がどんなに巨大でも不等式は成立するので，p に負の成分はない（もし負の成分があると，対応する b の成分が巨大なら左辺＞右辺となってしまう）。つまり

$$p_j \geqslant 0 \quad \forall j = 0, \ldots, m \tag{16}$$

である。分離定理の結論より $p \neq 0$ なので（$p = 0$ でいいなら(15)は当たり前だ），

$$p_j > 0 \ \text{ for some } \ j = 0, \ldots, m \tag{17}$$

を得る。

$b \in \mathbb{R}^{m+1}_{++}$ であれば b はどんなに小さくてもいいのだから，$p \cdot b$ がどんなに小さい正の数でも $p \cdot a$ はそれ以下ということであり，それはつまり

$$p \cdot a \leqslant 0$$

を意味する。しかも，$a \in A$ であれば $-a \in A$ でもある。z の代わりに $-z$ を掛ければいいからだ。したがって $p \cdot (-a) \leqslant 0$ でもあるわけで，合わせて

$$p \cdot a = 0 \quad \forall a \in A$$

を得る。A の定義に立ち返ってこれを書き換えると

$$z \cdot (p_0 \nabla g_0(x) + \cdots + p_m \nabla g_m(x)) = 0 \quad \forall z \in \mathbb{R}^n$$

となる。どんな z でも内積がゼロということは，後半のベクトルはゼロである。よって

$$p_0 \nabla g_0(x) + \cdots + p_m \nabla g_m(x) = 0$$

である。p_j をラグランジュ乗数とする。といっても $p = (p_0, \ldots, p_m)$ なので，等号成立していない制約，つまり $j > m$ については p_j は未定義である。そこで $p_j = 0 \ \forall j > m$ と定義する。すると

$$p_0 \nabla g_0(x) + \cdots + p_m \nabla g_m(x) + \cdots + p_j \nabla g_j(x) = 0$$

を得る。これは1階条件に他ならない。相補スラック条件については，等号成立している制約，つまり $j \in \{1, \ldots, m\}$ については，$g_j(x) = 0$ なので $p_j g_j(x) = 0$ は当然である。等号成立していない制約，つまり $j > m$ については，$p_j = 0$ と定義したので $p_j g_j(x) = 0$ はこれまた当然だ。非負条件と非ゼロ条件も(16)と(17)より成立している。（証明終）

ステップ2は「ゴルダンの定理」と呼ばれる二者択一定理の証明を焼き直したものである。

定理4（ゴルダンの定理） r 個のベクトル $h^j \in \mathbb{R}^n (j = 1, \ldots, r)$ を任意に考えると，以下のどちらか片方は必ず成立し，両方が成立することはない。

(a)　$z \in \mathbb{R}^n$ が存在して，$z \cdot h^j > 0 \ \forall j$ が成り立つ。

(b)　$p\in\mathbb{R}^r$ が存在して，$\sum_j p_j h^j=0$ かつ

$p_j\geqslant 0$ for all j

$p_j>0$ for some j

が成り立つ。

この定理を使えばステップ 2 は簡単である。$h^j=\nabla g_j(x)$（ただし $j=0, \ldots, m$）としてゴルダンの定理を使うと，(a)が成立しないので(b)が成立し，$\sum_{j=0}^m p_j\nabla g_j(x)=0$ を満たす $p=(p_0, \ldots, p_m)$ で非負条件と非ゼロ条件を満たすものが存在する。あとは等号成立していない制約 $j>m$ について $p_j=0$ と定義すればいい。

ゴルダンの定理を少し強化したものに以下の定理がある。

定理 5（モツキンの定理）　r 個のベクトル $h^j\in\mathbb{R}^n(j=1, \ldots, r)$ を任意に考える。また $1<r_{++}\leqslant r_+\leqslant r$ とする。すると以下のどちらか片方は必ず成立し，両方が成立することはない。

(a)　$z\in\mathbb{R}^n$ が存在して，任意の j について

$z\cdot h^j>0$　if $j\leqslant r_{++}$

$z\cdot h^j\geqslant 0$　if $r_{++}<j\leqslant r_+$

$z\cdot h^j=0$　if $j>r_+$

が成り立つ。

(b)　$p\in\mathbb{R}^r$ が存在して，

$$\sum_{j=1}^r p_j h^j=0$$

$p_j\geqslant 0$　for all $j\leqslant r_+$

$p_j>0$　for some $j\leqslant r_{++}$

が成り立つ。

このモツキンの定理を使えば，フリッツ・ジョンの定理を少し強めた以下の定理が得られる（Abadie, 1967）。

定理 6　制約式 $g_j(x)$ がアフィン関数である $j\geqslant 1$ の集合を E_A とする。このときフリッツ・ジョン定理の非ゼロ条件は

$\lambda_j>0$　for some $j\in\{0, \ldots, J\}\setminus E_A$

に強化できる。[12)]

結論が強くなっているのは，正の乗数がみつかる j の範囲が狭まっているところである。E_A に含まれない $j\geqslant 0$ の中にあるということなので，アフィン関数でない制約と目的関数に付く乗数のなかに少なくとも 1 つは正の乗数があるということである。[13)] 標準的な消費者問題の場合のように，制約式がすべてアフィン関数の場合は，残るは目的関数しかないので，$\lambda_0>0$ と言えることになる。これが前述の命題 1 である。

《証明》最大化問題の解を x とする。ステップ 1 では，

$$z\cdot\nabla g_j(x)\geqslant 0\quad \forall j\in E(x)\cap E_A \tag{18}$$

$$z\cdot\nabla g_j(x)>0\quad \forall j\in\{0\}\cup E(x)\setminus E_A \tag{19}$$

を満たす $z\in\mathbb{R}^n$ が存在しないことを示す。以前と同様，そのような z が存在すると仮定する。(19)

は少なくとも $j=0$ については成り立つので，$z\neq 0$ である。任意の $j\in E(x)\cup\{0\}$ を考える。$g_j(x)$ の微分可能性より，以前と同じ議論より

$$\lim_{\varepsilon\to 0}\frac{g_j(x+\varepsilon z)-g_j(x)}{\varepsilon}=z\cdot\nabla g_j(x) \tag{20}$$

を得る。制約が $j\in E(x)\backslash E_A$ の場合，右辺は正なので，$\varepsilon>0$ が十分小さいと

$$g_j(x+\varepsilon z)>0 \quad \forall j\in E(x)\backslash E_A \tag{21}$$

である。

制約が $j\in E(x)\cap E_A$ の場合，同じ議論は出来ないが，$g_j(x)$ がアフィン関数なので

$$g_j(x+\varepsilon z)=g_j(x)+\varepsilon z\cdot\nabla g_j(x)$$

である。$j\in E(x)$ なので初項 $g_j(x)$ はゼロであり，$j\in E_A$ でもあるので $z\cdot\nabla g_j(x)\geqslant 0$ である（(18)より）。合計すると

$$g_j(x+\varepsilon z)\geqslant 0 \quad \forall j\in E(x)\cap E_A$$

を得る。先ほどの(21)と合わせると，$j\in E(x)$ の制約はすべて $x+\varepsilon z$ で満たされることが分かる。

$E(x)$ に含まれない制約 j については，元々 $g_j(x)>0$ であり，$g_j(x)$ が連続関数なので，$\varepsilon>0$ が十分小さければ，$g_j(x+\varepsilon z)>0$ となる。よって $\varepsilon>0$ が十分小さければ，すべての制約は $x+\varepsilon z$ において満たされる。目的関数については，(19)と(20)より，$\varepsilon>0$ が十分小さければ，$g_0(x+\varepsilon z)>g_0(x)$ である。したがって x が最大化の解であることに矛盾する。これでステップ1が証明できた。

あとはモツキンの定理を適用する。制約の添字を並び変えれば，$E(x)=\{1,\ldots,m\}$ とできる（$m\leqslant J$）。すると(18)と(19)には $m+1$ 本の不等式がある。この連立不等式の解 $z\in\mathbb{R}^n$ が存在しないので，モツキンの定理より，$p\in\mathbb{R}^{m+1}$ が存在して，

$$\sum_{j=0}^{m}p_j\nabla g_j(x)=0 \tag{22}$$

$$p_j\geqslant 0 \quad \text{for all } j\in\{0\}\cup E(x) \tag{23}$$

$$p_j>0 \quad \text{for some } j\in\{0\}\cup E(x)\backslash E_A \tag{24}$$

が成り立つ。$E(x)$ に含まれない制約 $j>m$ については $p_j=0$ と定義すれば，(22)より

$$\sum_{j=0}^{J}p_j\nabla g_j(x)=0$$

という1階条件を得る。相補スラック条件が成立するのは以前と同様で，$j\leqslant m$ については $g_j(x)=0$ だし，$j>m$ については $p_j=0$ だからである。非負条件は(23)より得られる。非ゼロ条件については(24)と $E(x)\subseteq\{1,\ldots,J\}$ より得られる。（証明終）

注

† 清水崇氏と宮原泰之氏には有益なコメントをいただきました。本誌の事務局の方にもお世話になりました。感謝いたします。

1）$\mathbb{R}$ は実数の集合であり，$\mathbb{R}^n$ は実数が n 個並んだベクトルの集合である。

2）$\forall$ は「for all」を表す記号である。

3）x の定義域は $\mathbb{R}^n$ にしているが，開集合 $X\subseteq\mathbb{R}^n$ でもいい。

4）フリッツ・ジョン（Fritz John）というのはドイツ生まれの数学者のフルネームであり，ジョンが姓である。

5）$\partial L(\hat{x})/\partial\lambda_j=0$ という条件は含まれていない。$g_j(\hat{x})=0$ とは限らないので，含めるのは正しくない。また $g_j(\hat{x})\geqslant 0\forall j\geqslant 1$ という条件も含まれていない。$\hat{x}$ は最大化問題の解なので，制約条件を満たすのは当然であり，わざわざ含める必要はない。

6）市場均衡では相対価格しか決まらないので，１つの価格を１に固定してもいいのと同じである。

7）Kuhn and Tucker（1951）以前に同様の定理を証明した修士論文の Karush（1939）を先頭に加えて KKT 条件とも呼ばれる。

8）こういう関数のことを経済学界では「線形関数（linear function）」と呼ぶ風習があるが，定数項があるので $g_j(0)=0$ とは限らないので，数学の定義によると線形関数とは言えない。定数項以外の部分が線形関数なだけである。非線形計画法という名称も本当は非アフィン計画法（non-affine programming）と呼ぶ方が適切かも知れない。

9）すべての偏微分が有限であっても微分可能とは限らない。偏微分がすべて有限で，かつ連続関数になっていれば微分可能と言い切れる。

10）$x_j>0$ の領域では微分が連続という仮定を使用している。

11）g_j の定義域は引き続き $\mathbb{R}^n$ とするが，凸な開集合 $X\subseteq\mathbb{R}^n$ でもいい。

12）$A\backslash B$ とは，A の要素のうちで B に含まれないものの集合 $\{x\in A : x\notin B\}$ を表す。

13）アフィン関数に付く乗数の中にも正のものはあるかもしれないが，ないかもしれない。アフィン関数に付く乗数以外の中には正のものが必ず１つはある。

参　考　文　献

Abadie, J.（1967）. “On the Kuhn-Tucker Theorem,” in “Nonlinear Programming,”（J. Abadie, ed.）, Amsterdam: North-Holland.

Arrow, K. J. and Enthoven, A. C.（1961）. “Quasi-Concave Programming”, *Econometrica*, Vol. 29, No. 4, pp. 779–800.

John, F.（1948）. “Extremum Problems with Inequalities as Subsidiary Conditions,” in “Studies and Essays presented to R. Courant on his 60th Birthday, January 8, 1948,”（K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, eds.）, New York: Interscience.

Karush, W.（1939）. *Minima of Functions of Several Variables with Inequalities as Side Conditions*, Master’s thesis, University of Chicago.

Kuhn, H. W. and Tucker, A. W.（1951）. “Nonlinear Programming,” in “Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,”（J. Neyman, ed.）, Berkeley, California: University of California Press.

Simon, C. P. and Blume, L.（1994）. *Mathematics for Economists*, New York: W. W. Norton.

Sundaram, R. K.（1996）. *A First Course in Optimization Theory*, Cambridge: Cambridge University Press.

小山昭雄（2010）『微分積分の基礎 下（新装版 経済数学教室６）』岩波書店.

永谷裕昭（1998）『経済数学（有斐閣ブックス）』有斐閣.

西村清彦（1990）『経済学のための最適化理論入門』東京大学出版会.

マンガサリアン, O. L.（1972）『非線形計画法』培風館，関根智明（訳）.