ラグランジュ乗数法と鞍点

向山敏彦　（ジョージタウン大学）

2016 年 2 月、2026 年 4 月改訂、日本語版 2026 年 4 月

1　一般のケース

次の制約付き最適化問題を考える：

$$\max_{x_1,x_2} \quad f(x_1,x_2) \qquad \text{subject to } g(x_1,x_2)=0.$$

$g(x_1,x_2)=0$ を満たす (x_1,x_2) の組の集合を Ω とする。すると、この問題は次のようにも書ける：

$$\max_{x_1,x_2} \quad f(x_1,x_2) \qquad \text{subject to } (x_1,x_2)\in\Omega.$$

この制約付き問題の解（つまり、これから求めようとしているもの）を (x_1^*,x_2^*) と表す。

次に、この問題の制約なしバージョンを考えよう：

$$\max_{x_1,x_2} \quad f(x_1,x_2).$$

そして、制約なしの最適解 $(\tilde{x}_1,\tilde{x}_2)$ において制約が満たされていない、すなわち $g(\tilde{x}_1,\tilde{x}_2)\neq 0$ であるとする。一般性を失うことなく、

$$g(\tilde{x}_1,\tilde{x}_2)<0 \tag{1}$$

が成り立つと仮定しよう。（そうでない場合には、制約として $-g(x_1,x_2)=0$ を用いればよい。）

ラグランジアンを

$$L(x_1,x_2,\lambda)\equiv f(x_1,x_2)+\lambda g(x_1,x_2) \tag{2}$$

と定義する。$(x_1,x_2)\in\Omega$ のときには $g(x_1,x_2)=0$ が成り立つので、$(x_1,x_2)\in\Omega$ であれば $L(x_1,x_2,\lambda)=f(x_1,x_2)$ となることに注意しよう。

さて、与えられた λ に対して、次の制約なしラグランジュ問題を考える：

$$\max_{x_1,x_2} \quad L(x_1,x_2,\lambda). \tag{3}$$

$L(x_1,x_2,\lambda)$ の最大値を $\mathcal{L}(\lambda)$ と呼び、各 λ に対する問題 (3) の解を $(\hat{x}_1(\lambda),\hat{x}_2(\lambda))$ と呼ぶ。次に、同じ問題の制約付きバージョンを考える：

$$\max_{x_1,x_2} \quad L(x_1,x_2,\lambda) \qquad \text{subject to } (x_1,x_2)\in\Omega. \tag{4}$$

もちろん、$(x_1,x_2)\in\Omega$ のとき $L(x_1,x_2,\lambda)=f(x_1,x_2)$ であるから、制約付きラグランジュ問題 (4) の解は、λ の値に関わらず、元の制約付き問題の解 (x_1^*,x_2^*) と一致する。さらに、

$$L(\hat{x}_1(\lambda),\hat{x}_2(\lambda),\lambda) \geq L(x_1^*,x_2^*,\lambda) \text{ for all } \lambda$$

が成り立つことにも注意しよう。なぜなら、制約を加えることで状況が改善することはないからである（つまり、問題 (3) では問題 (4) よりも広い範囲の (x_1,x_2) の中から選ぶことができる)。左辺は $\mathcal{L}(\lambda)$ に等しく、右辺は $f(x_1^*,x_2^*)$ に等しいので、

$$\mathcal{L}(\lambda) \geq f(x_1^*,x_2^*) \text{ for all } \lambda \tag{5}$$

が成り立つ。ここで重要な点は、**(x_1^*,x_2^*) は任意の λ に対して常に同じ $L(x_1,x_2,\lambda)$ の値を与えるということであり、したがって元の問題の最適値 $f(x_1^*,x_2^*)$ は、λ の値が何であれ、ラグランジアンの値として常に達成可能である**、ということである。

仮定 (1) より、$g(\hat{x}_1(0),\hat{x}_2(0))<0$ が成り立つことを思い出そう。ここで λ を徐々に増加させると何が起こるかを考えよう。λ が増加するにつれて、(2) における $g(x_1,x_2)$ 関数の重みが増加し続けるので、問題 (3) の解は、ある時点で $g(x_1,x_2)$ を負にするような (x_1,x_2) の組よりも、正にするような組を選ぼうとする。（λ が無限大に近づくとき、問題 (3) では $f(x_1,x_2)$ を無視し、$g(x_1,x_2)$ のみに注目することになる。）したがって、十分に大きな λ（これを $\bar{\lambda}$ と呼ぼう）が存在し、すべての $\lambda\geq\bar{\lambda}$ に対して $g(\hat{x}_1(\lambda),\hat{x}_2(\lambda))>0$ となる。

関数 f と g が、$(\hat{x}_1(\lambda), \hat{x}_2(\lambda))$ が λ とともに滑らかに動くようなものであるとしよう(例えば、the Theorem of Maximum の仮定を満たすとしよう)。そのとき、λ を 0 から $\bar{\lambda}$ まで動かしていくと、$g(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) = 0$ となる λ の値 λ^* を見つけることができる。

制約集合 Ω の定義より、$g(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) = 0$ ということは、$(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) \in \Omega$ が満たされる。**この事実は、$(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*))$ が制約付きラグランジュ問題 (4) において実行可能であり、したがって問題 (4) の解であることを意味する。**(そうでなければ、$(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*))$ は $\lambda = \lambda^*$ のときの制約なし問題 (3) の解となりえない。もし制約下で f をさらに大きくするような組 (x_1, x_2) が見つかったとしたら、そのような (x_1, x_2) は $L(x_1, x_2, \lambda^*)$ をさらに大きくすることができるためである。) つまり、私たちは元の制約付き問題の解を見つけたことになる。すなわち $(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) = (x_1^*, x_2^*)$ と言うことができる。

以上より、(x_1^*, x_2^*) を求めたいときにはまず λ^* を与えてやり、その上で制約なし問題 (3) を解いて $(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*))$ を計算すればよいことがわかった。ではどうやって λ^* を求めればよいだろうか。まず

$$\mathcal{L}(\lambda^*) = L(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*), \lambda^*) = f(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) = f(x_1^*, x_2^*)$$

に注目しよう。最初の等号は $\mathcal{L}$ 関数の定義によるものであり 2 番目は $g(\hat{x}_1(\lambda^*), \hat{x}_2(\lambda^*)) = 0$ によるもの、3 番目は前段落の結果によるものである。関係式 (5) よりこれは

$$\mathcal{L}(\lambda) \geq \mathcal{L}(\lambda^*) \text{ for all } \lambda$$

を意味する。言い換えれば λ^* は $\mathcal{L}(\lambda)$ 関数を λ について最小化することで求めることができる。$\mathcal{L}(\lambda)$ の定義より λ^* は

$$\min_{\lambda} \max_{x_1, x_2} \quad L(x_1, x_2, \lambda)$$

を解くことで求められる。**つまり L 関数の鞍点を探すのである。**

f と g が微分可能であるときこの問題の 1 階条件は

$$\frac{\partial}{\partial x_1}L(x_1, x_2, \lambda) = 0,$$

$$\frac{\partial}{\partial x_2}L(x_1, x_2, \lambda) = 0,$$

および

$$\frac{\partial}{\partial \lambda}L(x_1, x_2, \lambda) = 0$$

である。最後の条件は $g(x_1, x_2) = 0$ を課すことと同じである。これが教科書的なラグランジュ法の標準的な手順である。

注：図による解説については Fryer and Greenman (1987) を参照されたい。

2　例

ある消費者がスーパーで財 1 と財 2 の 2 つの財を買うとしよう。これらの財の価格を p_1 と p_2 購入量を x_1 と x_2 と表す。消費者の所得は I である。したがって消費者の予算制約は

$$p_1x_1 + p_2x_2 = I \tag{6}$$

となる。ここで消費者はスーパーでいくらでも好きなだけ買い物ができるものとしよう。予算を超えても下回ってもよい。スーパーのレジ係は怠け者で消費者が実際に何個の財を買ったかを正確に確認しない。消費者は、どんなに財を買っても単に I を支払うだけである。消費者は通常の性質を持つ効用関数 $u(x_1, x_2)$ を持っており、もし罰則がなければ予算を超えて買いたいと思うだろう。

レジで I を支払ったあと、スーパーの出口で消費者は門番と対面しなければならないとしよう。門番は消費者に各財を何個買ったかを尋ねる（消費者は正直に答えなければならない)。つまり、門番は I と $p_1x_1 + p_2x_2$ を比較するための情報を持っている。門番

は消費者に予算制約 (6) を守らせたいと考えている。しかし門番にはそのための手段が 1 つしかない。彼は消費者から

$$F(x_1, x_2, \lambda) \equiv -\lambda(I - p_1x_1 - p_2x_2)$$

util の手数料を徴収できる。(つまり手数料は消費者の効用単位で測られる。) ここで $\lambda \geq 0$ とする。門番は λ の値を決められるとしよう。消費者が予算を超過した場合 $(I - p_1x_1 - p_2x_2 < 0)$、手数料 $F(x_1, x_2, \lambda)$ は正の値となる。したがって消費者が予算を超えた場合には門番は消費者から罰金を得ることとなる。$I - p_1x_1 - p_2x_2 > 0$ の場合には $F(x_1, x_2, \lambda)$ が負になるので門番は消費者に手数料を支払わなければならない。この設定の下で、出口における消費者の効用は

$$L(x_1, x_2, \lambda) \equiv u(x_1, x_2) - F(x_1, x_2, \lambda) = u(x_1, x_2) + \lambda(I - p_1x_1 - p_2x_2) \quad (7)$$

となる。

門番は消費者が予算制約 (6) を満たすような (x_1, x_2) を自発的に選ぶようにさせたいと考えているとしよう。門番はそれを手数料の水準 λ を適切に選ぶことで達成したい。上で述べたように λ が 0 に非常に近ければ消費者は予算を超えた上で門番に手数料を払うことを好む：$I - p_1x_1 - p_2x_2 < 0$。λ が非常に大きければ消費者は逆に予算を下回って門番から手数料を受け取ることを選ぶ：$I - p_1x_1 - p_2x_2 > 0$。したがってその中間のどこかに予算制約 (6) を等号で成立させる適切な λ の水準があるはずである。この λ を λ^* と呼ぶ。**ところで、λ^* のもとでの最適な選択 (x_1^*, x_2^*) は $\max_{x_1, x_2} u(x_1, x_2)$ subject to $p_1x_1 + p_2x_2 = I$ の解である**。なぜなら、$p_1x_1 + p_2x_2 = I$ のもとで $u(x_1, x_2)$ をさらに大きくするような (x_1, x_2) があるとするならば、そのような (x_1, x_2) は $L(x_1, x_2, \lambda^*)$ も $L(x_1^*, x_2^*, \lambda^*)$ よりも大きな値にするはずだからである。

門番はどうすればそのような λ^* を見つけられるだろうか。つまり、消費者を予算超過にも予算未満にもしないための適切な罰則の水準はどのくらいだろうか。まず注意したいのは、消費者の立場から見ると、**λ が何であれ**、予算制約 (6) を等号で満たしつつ $u(x_1, x_2)$ を最大化するような x_1 と x_2 を選べば (そのような x_1 と x_2 の組を (x_1^*, x_2^*)

と呼んだ)、$L(x_1^*, x_2^*, \lambda) = u(x_1^*, x_2^*)$ **という値は常に達成できる**、ということである。したがって、どんな λ でも消費者は $L(x_1^*, x_2^*, \lambda) = u(x_1^*, x_2^*)$ を達成することはできるわけだから、もし消費者がある λ^* のもとで実際に自発的に (x_1^*, x_2^*) を選ぶ（あらゆる (x_1, x_2) のもとでそれが最善の選択である）状況に置かれているならば、**そのような λ^* はあらゆる λ の中で消費者にとって最悪のもの**、ということである。他の λ の値であれば消費者にとってより良い $L(x_1, x_2, \lambda)$ の値を得られる可能性があるのに対して λ^* のもとでは消費者は出口における L の値として高々（常に達成可能な）$L(x_1^*, x_2^*, \lambda) = u(x_1^*, x_2^*)$ しか得られないからである。

したがって門番は $\max_{x_1, x_2} L(x_1, x_2, \lambda)$ を最小化するような λ を選ぶことで消費者に自発的に予算制約を満たさせることができる。つまり L 関数を (7) で定義したとき λ と (x_1, x_2) が

$$\min_{\lambda} \max_{x_1, x_2} \quad L(x_1, x_2, \lambda)$$

の解であれば予算制約は満たされかつ $u(x_1, x_2)$ は予算制約の下で最大化される。これは上の一般のケースで導いた結果に他ならない。そして、λ^* の値は消費者がこのときに予算を 1 ドル超過することに対して受け入れる罰則の大きさ（を効用単位で測ったもの）ということがわかる。言い換えれば λ^* **は消費者にとっての 1 ドルを効用単位で測った価値**、ということである。

3　ラグランジュ乗数の意味

上記の例を使ってラグランジュ乗数の意味をフォーマルに示そう。x_1 と x_2 の最適な選択を $(x_1^*(I), x_2^*(I))$ と書くとき（x_1^* と x_2^* は p_1 と p_2 にも依存するがここではそれらを動かさないので表記を省略する）1 ドルの効用単位で測った価値は

$$\frac{d}{dI} u(x_1^*(I), x_2^*(I)) = x_1^{*\prime}(I) u_1(x_1^*(I), x_2^*(I)) + x_2^{*\prime}(I) u_2(x_1^*(I), x_2^*(I)) \qquad (8)$$

となる。ここで $u_i(x_1^*(I), x_2^*(I))$ は $u(x_1, x_2)$ を i 番目の項について偏微分し $(x_1^*(I), x_2^*(I))$ で評価したものであり、プライムは微分を表す。

次に 1 階条件には

$$\frac{\partial}{\partial x_i}L(x_1^*, x_2^*, \lambda^*) = u_i(x_1^*(I), x_2^*(I)) - \lambda^* p_i = 0 \tag{9}$$

が含まれていることを思い出そう。すべての I について、最適な選択のもとでは予算制約

$$x_1^*(I)p_1 + x_2^*(I)p_2 = I$$

が満たさなければならないので

$$x_1^{*\prime}(I)p_1 + x_2^{*\prime}(I)p_2 = 1$$

が成り立つ。(9) よりこれは

$$x_1^{*\prime}(I)\frac{u_1(x_1^*(I), x_2^*(I))}{\lambda^*} + x_2^{*\prime}(I)\frac{u_2(x_1^*(I), x_2^*(I))}{\lambda^*} = 1$$

と書き換えられる。したがって (8) より

$$\frac{d}{dI}u(x_1^*(I), x_2^*(I)) = \lambda^*$$

となる。つまりラグランジュ乗数 λ^* は追加的な所得の限界効用である。一般に**ラグランジュ乗数は制約が 1 単位緩められたときの目的関数の限界的な増分を表す。**

参考文献

[1] Fryer, M. J. and J. V. Greenman (1987). *Optimisation Theory*, Edward Arnold.