

ラグランジュ未定乗数法ならびにKuhn-Tucker最適性条件における制約想定

日大生産工 篠原 正明
日大生産工(院) 茂木 渉

1. はじめに

等号制約付き最適化手法であるラグランジュ未定乗数法、さらには、この方法をも含みより一般化した等号・不等号制約付き最適化における1次の最適性必要条件であるKuhn-Tucker条件においては、制約想定(*constraint qualification*)という風変わった聞きなれない名前の条件を制約式が満足することが期待されている。この制約想定が満足されてないと、Kuhn-Tucker条件が良好に動作しないのである。通常の単純な最適化計算では、あまり気にする必要はないとの意見もあるが、多少なりとも巧妙な仕掛けを施した最適化問題では考慮する必要があると思われる。

本論文では、ラグランジュ未定乗数法とKuhn-Tucker条件をその幾何学的解釈と共に紹介し、簡単な例題を通して「制約想定」がなぜ必要となるかを説明する。

2. ラグランジュ未定乗数法

次の等号制約付き最適化問題ECOPを考える。

—ECOP—

$$\begin{aligned} \text{目的関数: } & f(x) \rightarrow \text{最小化} \\ \text{制約条件: } & g_j(x) = 0 \quad (j = 1, \dots, m) \end{aligned} \quad (1)$$

但し、 $x = (x_1, x_2, \dots, x_n)^T$, 変数の数 = n , 制約数 = m である。

[定理1]

ECOPの最適解が満たすべき一次の必要条件は、(2)と(3)で与えられる。

$$\frac{\partial L}{\partial x} = \mathbf{0} \quad (2)$$

$$\frac{\partial L}{\partial \lambda} = \mathbf{0} \quad (3)$$

$$\text{但し, } L = L(x, \lambda) = f(x) - \lambda^T g(x) \quad (4)$$

又、 $\lambda = (\lambda_1, \dots, \lambda_m)^T$, $g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T$ であり、 $L(x, \lambda)$ はラグランジュ関数(*Lagrangian function*)、 λ はラグランジュ(未定)乗数(*Lagrangian multiplier*)と呼ばれる。

(証明略)

(2)式に(4)式を代入すると、(5)を得る。

$$\frac{\partial L}{\partial \lambda} = \frac{\partial f(x)}{\partial x} - \sum_{j=1}^m \lambda_j \frac{\partial g_j(x)}{\partial x} = \mathbf{0} \quad (5)$$

(5)式は最適点においては、目的関数 $f(x)$ の勾配ベクトル $\frac{\partial f}{\partial x}$ と制約関数 $g_j(x)$ の勾配ベクトルの自由重み付き和 $\sum \lambda_j \frac{\partial g_j}{\partial x}$ が等しい。あるいは、等しくなるような自由変数の乗数ベクトルが存在することを主張する。自由変数とは正/負あるいは零値をとるという意味である。(5)式の幾何学的解釈はKuhn-Tucker条件のそれに含まれるので、5章で後述する。(3)式はECOPの制約条件 $g(x) = 0$ に帰着される。

3. Kuhn-Tucker条件

次の不等号制約付き最適化問題ICOPを考える。

—ICOP—

$$\begin{aligned} \text{目的関数: } & f(x) \rightarrow \text{最小化} \\ \text{制約条件: } & g_j(x) \geq 0 \quad (j = 1, \dots, m) \end{aligned}$$

[定理2]

ICOPの最適解が満たすべき一次の必要条件は、(6)~(9)で与えられる。

$$\frac{\partial L}{\partial x} = \mathbf{0} \quad (6)$$

$$\frac{\partial L}{\partial \lambda} \leq \mathbf{0} \quad (7)$$

$$\lambda \geq \mathbf{0} \quad (8)$$

$$\lambda^T g(x) = \mathbf{0} \quad (9)$$

$$\text{但し, } L = L(x, \lambda) = f(x) - \lambda^T g(x) \quad (10)$$

(証明略)

(6)式は、ECOPの場合と同様に、形式的には(5)式となるが、乗数ベクトル λ が非負である($\lambda \geq 0$)。 (7)式はICOPの制約条件 $g(x) \geq 0$ に帰着される。(9)式は相補性条件と呼ばれており、 $\lambda_j \geq 0, g_j(x) \geq 0$ に注意すると、次式(11)と等価である。

Constraint Qualification in Lagrangian Multiplier Method and Kuhn-Tucker Condition

Masaaki SHINOHARA and Wataru MOGI

$$\lambda_j g_j(\mathbf{x}) = 0 \quad (j=1, \dots, m) \quad (11)$$

相補性条件(11)式は、「もし、 $g_j(\mathbf{x}) > 0$ ならば、対応する乗数 $\lambda_j = 0$ 」あるいは「もし $\lambda_j > 0$ ならば、対応する制約式は等号、すなわち活性(active)となる($g_j(\mathbf{x}) = 0$)」を意味する。

4. Kuhn-Tucker条件の幾何学的解釈

2変数3制約のICOP($n=2, m=3$)の2変数平面上での図的表現を通して、Kuhn-Tucker条件の幾何学的解釈を説明する。

—ICOP—

目的関数： $f(x_1, x_2) \rightarrow$ 最小化

制約条件： $g_1(x_1, x_2) \geq 0$

$$g_2(x_1, x_2) \geq 0$$

$$g_3(x_1, x_2) \geq 0$$

図1に2変数 (x_1, x_2) 平面上での目的関数 $f(x_1, x_2)$ の等高線 L で最小点、3つの制約条件の等号式 $g_1(x_1, x_2) = 0, g_2(x_1, x_2) = 0, g_3(x_1, x_2) = 0$ 、と各不等号領域の共通領域としての実行可能域(灰色表示)を表示する。なお、 $g_j(x_1, x_2) = 0$ で制約条件が等号で満たされる場合の線を示すが、その線の両側に+、- を表示して、 $g_j(x_1, x_2) = 0$ の正負を示した。

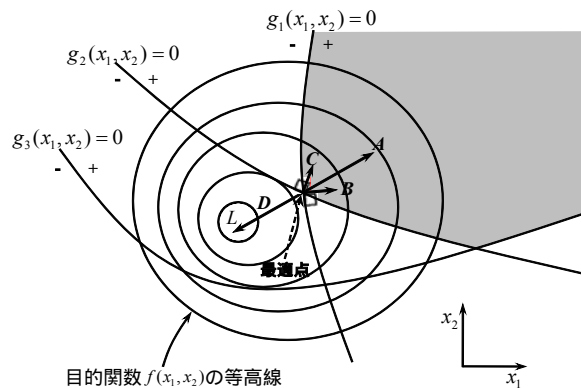


図1. Kuhn-Tucker条件の幾何学的解釈

図1において、 $g_1(x_1, x_2) = 0$ と $g_2(x_1, x_2) = 0$ の交点が最適点となるが、この点に注目する。この点における目的関数 $f(x_1, x_2)$ の勾配ベクトル $\frac{\partial f}{\partial \mathbf{x}}$ をA、第1制約関数 $g_1(x_1, x_2)$ の勾配ベクトル $\frac{\partial g_1}{\partial \mathbf{x}}$ をB、第2制約関数 $g_2(x_1, x_2)$ の勾配ベクトル $\frac{\partial g_2}{\partial \mathbf{x}}$ をCとするならば、 $f(x_1, x_2), g_1(x_1, x_2), g_2(x_1, x_2)$ の増減方向により、A, B, Cいずれも図中に示すようなベクトルとなる。

この場合のKuhn-Tucker条件を記述すると以下の通りである。

$$\frac{\partial f}{\partial \mathbf{x}} - \lambda_1 \frac{\partial g_1}{\partial \mathbf{x}} - \lambda_2 \frac{\partial g_2}{\partial \mathbf{x}} - \lambda_3 \frac{\partial g_3}{\partial \mathbf{x}} = \mathbf{0} \quad (12)$$

$$g_j(x_1, x_2) \geq 0 \quad (j=1, 2, 3) \quad (13)$$

$$\lambda_j \geq 0 \quad (14)$$

$$\lambda_j g_j(x_1, x_2) = 0 \quad (j=1, 2, 3) \quad (15)$$

注目する最適点では、 $g_3(x_1, x_2) > 0$ なので $\lambda_3 = 0$ となり、従って、(12)式の目的関数と制約関数の勾配バランス式は次式となる。

$$\frac{\partial f}{\partial \mathbf{x}} = \lambda_1 \frac{\partial g_1}{\partial \mathbf{x}} + \lambda_2 \frac{\partial g_2}{\partial \mathbf{x}} \quad (16)$$

$$\lambda_1 \geq 0, \lambda_2 \geq 0 \quad (17)$$

すなわち、最適点においては、目的関数の勾配 $\frac{\partial f}{\partial \mathbf{x}}$ と、活性制約条件関数(この場合は g_1 と g_2)の勾配の非負結合($\lambda_1 \frac{\partial g_1}{\partial \mathbf{x}} + \lambda_2 \frac{\partial g_2}{\partial \mathbf{x}}$, $\lambda_1 \geq 0, \lambda_2 \geq 0$)が等しくなる。あるいは、等しくなるような非負の乗数ベクトル $\lambda (\geq 0)$ が存在することを主張している。

目的関数の勾配 $\frac{\partial f}{\partial \mathbf{x}}$ (=A)の逆方向のベクトルをDとする。

$$\mathbf{D} = -\mathbf{A} = -\frac{\partial f}{\partial \mathbf{x}} \quad (18)$$

するとDは、目的関数を最小化する方向の勾配を表している。すなわち、図1の実行可能域内で運動することが許容されているボールを自由に転がすと、目的関数の減少する方向へと移動し、2つの障壁 $g_1(x_1, x_2) \geq 0$ と $g_2(x_1, x_2) \geq 0$ の境目でもある注目する最適点で静止する。目的関数 $f(x_1, x_2)$ を重力ポテンシャル関数と見れば、降下方向に働く重力D(作用)と、2つの障壁での反発力 $\lambda_1 B$ と $\lambda_2 C$ ($\lambda_1 \geq 0, \lambda_2 \geq 0$) (反作用)が図2に示す「力の平行四辺形」として釣り合っている。

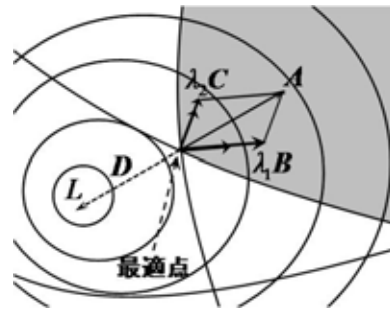


図2. 「力の平行四辺形」としてのKuhn-Tucker条件

5. ラグランジュ未定乗数法の幾何学的解釈

等号制約最適化のラグランジュ未定乗数法 [定理1] は、不等号制約最適化のKuhn-Tucker条件 [定理2] に含まれる。具体的には、ラグランジュ乗数の非負制約が外れて、自由変数ベクトルとなる。従って、(5)式の意味するところは、...最適点においては、目的関数の勾配 $\frac{\partial f}{\partial \mathbf{x}}$ と、制約関数の勾配 $\frac{\partial g_j}{\partial \mathbf{x}}$ ($j=1,$

..., m)の自由結合 $\sum \lambda_j \frac{\partial g_j}{\partial x}$ (λ_j は自由)が等しい、あるいは、等しくなるような自由乗数ベクトル λ が存在することを主張している。

重力ポテンシャルの「力の平行四辺形」による解釈に従えば、 $g_j(x) \geq 0$ の領域ではなく、等号制約 $g_j(x) = 0$ の曲線上をボールが転がるわけで、 $g_j(x) = 0$ は「 $g_j(x) \geq 0$ と $g_j(x) \leq 0$ 」と等価なので、2つの不等号領域のどちらかが真の活性制約式となるので、その真の活性制約式に注目して「反作用」と「力の平行四辺形」を考えればよい。但し、ここで、 $g_j(x) = 0$ は $g_j(x) \geq 0$ と $g_j(x) \leq 0$ は共に活性である。しかし、対応するラグランジュ乗数は一方は必ず0となるので、Kuhn-Tucker条件の意味で正となるのが真の活性制約式である。このように、等号制約では、 $g_j(x) \geq 0$ と $g_j(x) \leq 0$ は共に活性制約であるが、真に活性な制約は1つであり、他は対応する乗数 = 0となるため、以下に示すように幾何学的解釈は、力の平行四辺形は潰れて「作用力 = 反作用力」として考えるのが素直である。

図3に2変数1等号制約でのラグランジュ乗数法の幾何学的解釈を示すが、等号制約式 $g(x_1, x_2) = 0$ は2つの特別な不等号制約式で表現でき、等号制約式上をボールが転がることは、この2つの不等号制約式の障壁に常に挟まれている状況である。図3には便宜的に制約関数 $g(x_1, x_2)$ の正負領域を示したが、これと逆の場合でも問題ない(但し、 $\frac{\partial g}{\partial x}$ の方向は逆に向く)。このように最適点では $\frac{\partial f}{\partial x}$ と $\frac{\partial g}{\partial x}$ は正負も含めて方向のみが一致することが要求される。あるいは、最適点では勾配方向と直交している接線方向(接面方向)が一致している、すなわち、目的関数の等高線と等号制約関数が互いに接しているとも解釈できる。

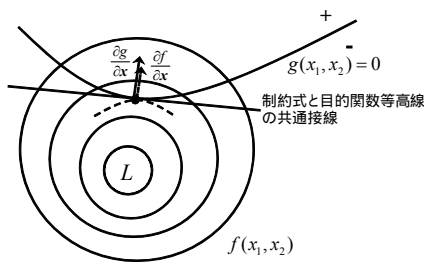


図3. ラグランジュ未定乗数法の幾何学的解釈

6. 線形代数理論としての制約想定

$b = \frac{\partial f}{\partial x}$, $A = \left(\frac{\partial g_1}{\partial x}, \frac{\partial g_2}{\partial x}, \dots, \frac{\partial g_m}{\partial x} \right)$, $\lambda = (\lambda_1, \dots, \lambda_m)^T$ とすると、(5)式は以下で表現できる。

$$A\lambda = b \quad (19)$$

A は $n \times m$ 行列、 λ は m 次元列ベクトル ($m \times 1$ 行列)、 b は n 次元列ベ

クトル ($n \times 1$ 行列)である。この連立線形方程式に解が存在するために、 A と b に課せられた想定条件が「制約想定」とも解釈できる。等号制約付き最適化のラグランジュ未定乗数法では、変数は自由変数であるが、不等号制約付き最適化のKuhn-Tucker条件では変数は非負である($\lambda \geq 0$)点に注意すれば、制約想定は連立線形方程式の可解条件あるいは非負解存在条件となる。

7. 制約想定幾何学的解釈

制約想定を4章5章での幾何学的解釈の立場から再考する。結論から言うならば、実行可能領域が図4に示すように尖った領域から成り、最適点はその先端で実現される状況が制約想定違反である。図4において、制約条件関数1,2の勾配 $B = \frac{\partial g_1}{\partial x}$, $C = \frac{\partial g_2}{\partial x}$ と目的関数の勾配 $D = -\frac{\partial f}{\partial x}$ は直交しており、 D を B と C の適当な非負結合で表すことは不可能である。非負結合($\lambda_1 \geq 0, \lambda_2 \geq 0$)どころか、線形結合で表すことも不可能である。

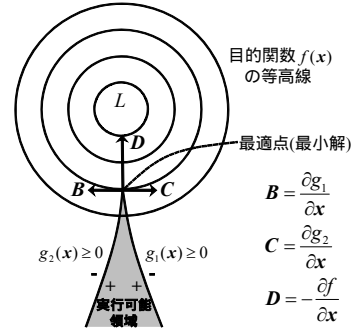


図4. 制約想定が満たされない場合

従って、制約想定は幾何学的に解釈するならば、最適点において図4に示すような、尖頭状の制約領域でない、と言える。

8. 制約想定違反の簡単な例

なるべく制約数 m も変数の数 n も少なく、最適点において制約想定が満たされない例を以下に2つ示す。

[例1: 不等式制約]

—ICOP (1)—

$$\text{目的関数: } f = (x+1)^2 + y^2 \rightarrow \text{最小化} \quad (20)$$

$$\text{制約条件: } y \geq 0 \quad (21)$$

$$x^3 - y \geq 0 \quad (22)$$

ICOP (1)のKuhn-Tucker条件は(21),(22)に加えて以下の通りである。

$$\frac{\partial L}{\partial x} = 2(x+1) - 3\mu x^2 = 0 \quad (23)$$

$$\frac{\partial L}{\partial y} = 2y - \lambda + \mu = 0 \quad (24)$$

$$\lambda \cdot y = 0 \quad (25)$$

$$\mu(x^3 - y) = 0 \quad (26)$$

$$\lambda \geq 0, \mu \geq 0 \quad (27)$$

但し、ラグランジュ関数 $L=L(x,y,\lambda,\mu)$ は(28)式で与えられる。

$$L(x,y,\lambda,\mu) = (x+1)^2 + y^2 - \lambda y - \mu(x^3 - y) \quad (28)$$

ICOP (1)の実行可能領域を図5に示すが、最適点は(0,0)であるが、この点は(23)~(27)のKuhn-Tucker条件を満たしていないことを以下に示す。

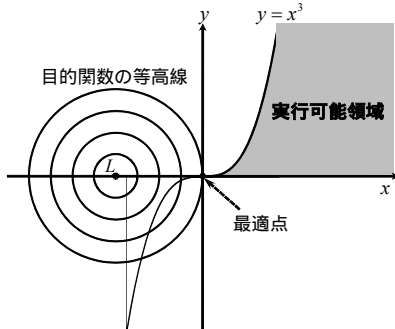


図5 . ICOP (1)の実行可能領域と最適点

未定乗数 (λ, μ) を正/零値の組合せにより場合分けする。

[場合 : $\lambda=0, \mu=0$]

(23),(24)より、 $x=-1, y=0$ となるが、(22)を満足せず。よって、この場合 は解無し。

[場合 : $\lambda>0, \mu=0$]

(25)より、 $y=0$ 、(24)より $\lambda=0$ となるが、これは仮定に矛盾する。よってこの場合 は解無し。

[場合 : $\lambda=0, \mu>0$]

(25)より、 $y=x^3$ 、(24)より $2y+\mu=0$ 、従って、

$$\mu = -2x^3 \quad (29)$$

(29)を(23)に代入すると、次の5次方程式(30)を得る。

$$6x^5 + 2x + 2 = 0 \quad (30)$$

方程式(30)は $x=-0.65\dots$ の実根を持つが、 $(x=-0.65, y=-0.27)$ は、(21)を満足しない。よって、この場合 は解無し。

[場合 : $\lambda>0, \mu>0$]

(25),(26)より、 $y=0, y=x^3$ 、すなわち、 $x=0, y=0$ を得る。(23)、(24)に代入すると、 (λ, μ) に関する次の連立方程式を得るが、これは解けない。

$$\lambda = \mu \quad (31)$$

$$3\mu \times 0 = 2 \quad (32)$$

場合 ~ のすべての場合において、解無しとなり、Kuhn-Tucker条件を満足する解は存在しない。

[例2 : 等式制約]

—ECOP (1)—

$$\text{目的関数: } f = (x+1)^2 + y^2 \rightarrow \text{最小化} \quad (33)$$

$$\text{制約条件: } y = h(x) \quad (34)$$

$$y = -h(x) \quad (35)$$

$$\text{但し、} \quad h(x) = \begin{cases} (x-1)^2 & 1 < x \\ 0 & 0 \leq x \leq 1 \\ x^2 & x < 0 \end{cases} \quad (36)$$

ラグランジュの未定乗数法を適用すると、(34),(35)に加えて以下の条件を得る。

$$\frac{\partial L}{\partial x} = 2(x+1) + \lambda h'(x) - \mu h'(x) = 0 \quad (37)$$

$$\frac{\partial L}{\partial y} = 2y - \lambda - \mu = 0 \quad (38)$$

(34),(35)を満たす解は(39)である。

$$0 \leq x \leq 1, \quad y = 0 \quad (39)$$

$y=0$ を(38)に代入し、 $0 \leq x \leq 1$ で $h'(x)=0$ に注意すると、 (λ, μ) に関する次の連立方程式を得るがこれは解けない(但し、 $0 \leq x \leq 1$)。

$$\lambda + \mu = 0 \quad (40)$$

$$0 \times (\mu - \lambda) = 2x + 2 \quad (41)$$

9. おわりに

制約想定違反のケースは等号の制約条件式が重畳しており、退化の特殊ケースとも考えられる。Kuhn-Tucker条件のみならずラグランジュ未定乗数法についても簡単な制約想定違反の例題を作成し、制約想定の一必要性を理解するための教材を与えた。ラグランジュ未定乗数法の例については、もっと単純な例が望まれる。

制約想定違反のケースは、落下する(極小)ボールに例えれば図6に示すように、左右から壁に挟まれて、静止している状態に対応している(ボールの「真剣白刃止め」)ことを、幾何学的解釈として与えた。

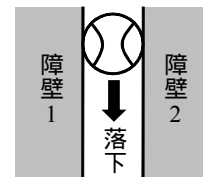


図6 . 制約想定違反は落下ボールの真剣白刃止め