ラグランジュの未定乗数法について

京都大学大学院工学研究科航空宇宙工学専攻

大和田　拓

要約

拘束条件下での関数の極値点を探したいときに，その候補となる停留点を探す方法としてLagrange の未定乗数法(以下Ｌ法)が良く知られている．ここでは境界を伴わない場合に対し，その原理を説明する．さらにその応用として拘束条件がある場合のオイラー・ラグランジュ方程式を導出する．

１．問題と処方箋

問題

M個の拘束条件

$$g_k(x_1,\cdots,x_N)=0 \qquad (k=1,\cdots,M)\,, \tag{1}$$

の下で関数$f(x_1,\cdots,x_N)$の停留点を求めよ[1]．ただし$f(x_1,\cdots,x_N)$および$g_k(x_1,\cdots,x_N)$　$(k=1,\cdots,M)$は滑らかで，拘束条件(1)を満たす点で$\nabla g_k=\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_N}\right)\neq \mathbf{0}$　$(k=1,\cdots,M)$とする．

具体的な解法の処方箋は以下に示すように非常に単純である．

処方箋

未定の乗数λ_i $(i=1,\cdots,M)$を導入し，関数$F(x_1,\cdots,x_N,\lambda_1,\cdots,\lambda_M)$を

$$F(x_1,\cdots,x_N,\lambda_1,\cdots,\lambda_M)=\ f(x_1,\cdots,x_N)-\sum_{k=1}^{M}\lambda_k\, g_k(x_1,\cdots,x_N), \tag{2}$$

によって定義する．そして$N+M$個の条件

$$\frac{\partial F}{\partial \alpha}=0 \qquad (\alpha=x_1,\cdots,x_N,\lambda_1,\cdots,\lambda_M), \tag{3}$$

[1] 力学では式(1)のような位置座標の関数(時間を含む場合もある)による制限を holonomic な制限と呼ぶ．

を満たす点$(x_1,\cdots,x_N)$を停留点として求める．

各拘束条件$g_k(x_1,\cdots,x_N)=0$は$\mathbb{R}^N$における(超)曲面を表す(付録２の陰関数定理１を参照). 式(1)で定義される集合(以下では非線形計画法の分野の用語に従ってこれを許容集合，feasible set，と呼ぶことにする)はM個の(超)曲面の共通集合である．許容集合の定義式(1)は式(3)の$\alpha=\lambda_1,\cdots,\lambda_M$の場合から導かれる．そして$\alpha=x_1,\cdots,x_N$の場合からは,

$$\nabla f=\sum_{k=1}^{M}\lambda_k\nabla g_k\ , \tag{4}$$

が導かれる．L 法は式(1)で定義される許容集合上で式(4)を満たす点を停留点とする．従って L 法の原理を説明することは，式(4)の意味を明らかにすることに他ならない.

２．幾何学的な説明

関数fの停留点は微小変位$d\boldsymbol{x}=(dx_1,\cdots,dx_N)$に対する$f$の変動量の主要部(すなわち全微分)がゼロになる点である．すなわち

$$\frac{\partial f}{\partial x_1}dx_1+\cdots+\frac{\partial f}{\partial x_N}dx_N=0. \tag{5}$$

$dx_i=|d\boldsymbol{x}|n_i\ \ (i=1,\cdots,N)$と表すと，式(5)より

$$\frac{\partial f}{\partial x_1}n_1+\cdots+\frac{\partial f}{\partial x_N}n_N=0,$$

が得られ，これはfの$\boldsymbol{n}=(n_1,\cdots,n_N)$方向の微分がゼロであることを意味する.
拘束条件が課されておらず$\boldsymbol{n}$の方向が任意であれば，$\frac{\partial f}{\partial x_i}=0\ (i=1,\cdots,N)$，つまり$\nabla f=\boldsymbol{0}$が従う．しかし拘束条件が課されると，以下で示すように$\boldsymbol{n}$の範囲は任意ではなくなる.

注意

式(5)は$\boldsymbol{dx}$の大きさが微小な場合のfの変動量の一次近似として導かれたが，上の議論においても，さらに以下の式(5)を出発点とする議論においても，$\boldsymbol{dx}$の方向である$\boldsymbol{n}$の範囲だけが問題になり，$\boldsymbol{dx}$の大きさは微小でなくてもかまわない．しかし変位ベクトルとして$\boldsymbol{dx}$を用いる限り，「任意の変位ベクトル$\boldsymbol{dx}$」と言えずに，「微小変位ベクトル$\boldsymbol{dx}$の任意の方向」といちいち言い換えなければならない．これでは面倒なので，以下では微小変位ベクトル$\boldsymbol{dx}$を任意の大きさを持つ変位ベクトル$\boldsymbol{d} = (d_1, \cdots, d_N)^{\mathrm{T}}$に書きかえる．これに従い式(5)は

$$\frac{\partial f}{\partial x_1} d_1 + \cdots + \frac{\partial f}{\partial x_N} d_N = 0, \tag{5$'$}$$

と書きかえられる．「$\boldsymbol{dx}$の方向が任意なら$\nabla f = \boldsymbol{0}$が従う」と言う代わりに，「$\boldsymbol{d}$が任意ならば$\nabla f = \boldsymbol{0}$が従う」とより簡潔に言えるようになる．

許容集合の線形近似

まず$M = 1$の場合を考える．拘束条件$g_1(x_1, \cdots, x_N) = 0$が表す超曲面上の一点$\boldsymbol{x} = \boldsymbol{x}^0$の近傍において，この超曲面は$\boldsymbol{x} = \boldsymbol{x}^0$における接平面で近似できる．ここに接平面はいわゆる超平面で∇g_1をその法線方向とする．すなわち

$$\nabla g_1(\boldsymbol{x}^0) \cdot (\boldsymbol{x} - \boldsymbol{x}^0) = 0.$$

$\boldsymbol{x} = \boldsymbol{x}^0$における変位ベクトル$\boldsymbol{d}$の範囲はこの接平面上のベクトルの空間，すなわち接空間(tangent space または tangent vector space)，に制限される．

次に$M \geq 2$の場合を考える．先に述べたように許容集合はM個の超曲面の共通集合である．許容集合の一点$\boldsymbol{x} = \boldsymbol{x}^0$の近傍で各超曲面は$\boldsymbol{x} = \boldsymbol{x}^0$における接平面で近似できる．ここで

「許容集合上の一点$\boldsymbol{x} = \boldsymbol{x}^0$の近傍で，許容集合はその点における$M$個の接平面の共通集合によって近似できる」

という仮定を導入する．この仮定が L 法の核心である．これより$\boldsymbol{x} = \boldsymbol{x}^0$の近傍における許容集合は連立方程式

$$\nabla g_i(\boldsymbol{x}^0) \cdot (\boldsymbol{x} - \boldsymbol{x}^0) = 0 \qquad (i = 1, \cdots, M), \tag{6}$$

の解として表され，許容集合の$\boldsymbol{x}=\boldsymbol{x}^0$における接空間は$\nabla g_i(\boldsymbol{x}^0)(i=1,\cdots,M)$を行ベクトルとする$M$行$N$列の行列$\boldsymbol{G}(\boldsymbol{x}^0)$の核になる．すなわち$M$個の超曲面の共通集合である許容集合上の点$\boldsymbol{x}=\boldsymbol{x}^0$における接空間を$T(\bigcap_{i=1}^{M} g_i(\boldsymbol{x})=0\,,\boldsymbol{x}=\boldsymbol{x}^0)$と表すと

$$T\left(\bigcap_{i=1}^{M} g_i(\boldsymbol{x})=0\,,\boldsymbol{x}=\boldsymbol{x}^0\right)=\mathrm{Ker}\ \boldsymbol{G}(\boldsymbol{x}^0). \tag{7}$$

線形代数の基礎事項(直交補空間，付録 1 参照)より，式(7)から式(4)は直ちに導かれる．行列の核は行ベクトルの張る空間の直交補空間で，
$\boldsymbol{V}=\mathrm{span}(\nabla g_i(\boldsymbol{x}^0),i=1,\cdots,M)$とすると$\mathrm{Ker}\ \boldsymbol{G}(\boldsymbol{x}^0)=\boldsymbol{V}^{\perp}$に等しい．つまり

$$T\left(\bigcap_{i=1}^{M} g_i(\boldsymbol{x})=0\,,\boldsymbol{x}=\boldsymbol{x}^0\right)=\boldsymbol{V}^{\perp}=\left(\mathrm{span}(\nabla g_i(\boldsymbol{x}^0),i=1,\cdots,M)\right)^{\perp}. \tag{7$'$}$$

$\boldsymbol{x}=\boldsymbol{x}^0$における変位ベクトル$\boldsymbol{d}$の範囲はその点における許容集合の接空間$\boldsymbol{V}^{\perp}$上に制限される. 一方, 式(5)'は$\nabla f$が$\boldsymbol{V}^{\perp}$の直交補空間に属することを意味する. $\boldsymbol{V}$の直交補空間($\boldsymbol{V}^{\perp}$)の直交補空間は$\boldsymbol{V}$自身だから，結局$\nabla f\in\boldsymbol{V}$が結論される．これを等式で表したものが式(4)に他ならない.

<u>制約想定</u>

非線形計画法では式(1)のような等式制約条件の他に不等式制約条件も扱うが, 許容集合の線形近似(6)，あるいはそれより導かれる式(7)，がどのような場合に成り立つかは，非線形計画法において現在盛んに研究されている制約想定(Constraint Qualifications)の問題に含まれる．$M=1$の場合や超曲面が接平面と一致する場合，すなわち$g_k(x_1,\cdots,x_N)(k=1,\cdots,M)$が$x_1,\cdots,x_N$の一次式で表される場合，線形近似(6)は明らかに成り立つ．しかし例えば 3 次元空間で二つの曲面がある曲線上で接するような場合では，線形近似(6)は成り立たない．許容集合は曲線であるが，二つの曲面をその曲線上でそれぞれ接平面で近似すれば, 両者は同じ平面だからその共通集合はその接平面自身になってしまうからである. さらに簡単な例として$\mathbb{R}^2$における 2 曲線$y=x^2$と$y=-x^2$の場合を考えると, これら 2 つの曲線は原点で接し許容集合は原点のみとなるので，許容集合の接空間は定義できない．しかしこれら 2 つの曲線を原点近傍で原点における接線で近似すれば，許容集合は直線$y=0$となり，接空間はその直線上のベクトルの

空間として定義できてしまう．このような不整合が生じるのは，曲面(あるいは曲線)同士が接するという高階の微係数が主役となる幾何学的状況を接平面(接線)という線形近似では捉えられないからである．L 法ではこのような状況を全く想定していない．例えば$\nabla g_k(\boldsymbol{x}^0)(k=1,\cdots,M)$の中に平行なものが一組もなければこのような不整合は起こらない．次節では，この条件をさらに厳しくした$\nabla g_k(\boldsymbol{x}^0)(k=1,\cdots,M)$が線形独立という条件の下で式(7)が成り立つことを示す．

３．陰関数定理を使った説明

拘束条件下の関数の停留点を実直に探すとすれば，i) 許容集合を求め，ii) その許容集合をあるパラメータの関数として表すことで，問題を拘束条件のない場合に変換し，iii) パラメータ空間における停留点を探すというような手順が考えられるだろう．ここではこの方針に沿ってＬ法の原理を説明をする．その際，fやg_kの具体的な関数形を使わずに抽象的な計算を可能にしてくれる数学の道具が陰関数定理である．ただしこれを使うには以下の条件が必要になる．

条件１．拘束条件の個数MはNより小さい．$1 \le M < N$.

条件２．勾配$\nabla g_k\ (k=1,\cdots,M)$は許容集合上で線形独立である．

$\mathbb{R}^N$における線形独立なベクトルの最大個数はNだから，条件 2 の下では条件 1 は$M=N$の場合を除外することを意味する．$M=N$の場合の許容集合は点集合になるので，わざわざ L 法を使う必要はない．従って実質的な条件は条件 2 のみと言っていい．

Step1　変数の書き換え

付録 2 の連立の陰関数定理(定理 2)の説明で使った記号と整合性を持たせるために，ここでは変数の書き換えを行う．

∇g_kがk行の成分となるようなM行N列の行列$\boldsymbol{G}$を考える．

$$
\boldsymbol{G} = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_N} \\ \vdots & & \vdots \\ \dfrac{\partial g_M}{\partial x_1} & \cdots & \dfrac{\partial g_M}{\partial x_N} \end{pmatrix}.
$$

仮定２より許容集合上の点$(x_1^0, \cdots, x_N^0)$において行列$\boldsymbol{G}$の行ベクトル全体が線形独立である．すなわち$\operatorname{rank} \boldsymbol{G} = M$．一方，$\operatorname{rank} \boldsymbol{G}$は$\boldsymbol{G}$の列ベクトルの中で線形独立なものの最大個数でもある(付録１参照)． $i_p\ (p = 1, \cdots, M)$列のベクトル $\left(\frac{\partial g_1}{\partial x_{i_p}}, \cdots, \frac{\partial g_M}{\partial x_{i_p}}\right)^{\mathrm{T}}$ $(p = 1, \cdots, M)$が線形独立としよう．

付録２の記号に合わせて，x_{i_p}を$z_p\ (p = 1, \cdots, M)$と表し，それ以外の$N - M$個の独立変数$x_j\ (j \notin \{i_p\ (p = 1, \cdots, M)\})$を$y_i\ (i = 1, \cdots, N - M)$と表すことにする．新変数$\boldsymbol{y} = (y_1, \cdots, y_{N-M})$および$\boldsymbol{z} = (z_1, \cdots, z_M)$の関数$\tilde{f}(\boldsymbol{y},\boldsymbol{z})$を$f(x_1, \cdots, x_N)$を用いて$\tilde{f}(\boldsymbol{y},\boldsymbol{z}) = f(x_1, \cdots, x_N)$と定義し，$\mathbb{R}^{N-M} \times \mathbb{R}^M \longrightarrow \mathbb{R}^M$のベクトル値関数

$$
\boldsymbol{H}(\boldsymbol{y}, \boldsymbol{z}) = \begin{pmatrix} H_1(\boldsymbol{y}, \boldsymbol{z}) \\ \vdots \\ H_M(\boldsymbol{y}, \boldsymbol{z}) \end{pmatrix},
$$

を

$$
H_k(\boldsymbol{y}, \boldsymbol{z}) = g_k(x_1, \cdots, x_N) \quad (k = 1, \cdots, M),
$$

によって定義する．

<u>Step２　許容集合の表現</u>

許容集合上の点$(x_1^0, \cdots, x_N^0)$に対応する点$(\boldsymbol{y}^0, \boldsymbol{z}^0)$では$\boldsymbol{H}(\boldsymbol{y}^0, \boldsymbol{z}^0) = \boldsymbol{0}$が成り立つ. 条件２より

$$
\boldsymbol{D} = \begin{pmatrix} \dfrac{\partial H_1}{\partial z_1}(\boldsymbol{y}^0, \boldsymbol{z}^0) & \cdots & \dfrac{\partial H_1}{\partial z_m}(\boldsymbol{y}^0, \boldsymbol{z}^0) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial H_m}{\partial z_1}(\boldsymbol{y}^0, \boldsymbol{z}^0) & \cdots & \dfrac{\partial H_m}{\partial z_m}(\boldsymbol{y}^0, \boldsymbol{z}^0) \end{pmatrix},
$$

は正則である．従って連立の陰関数定理(付録 2)により，$\boldsymbol{y}=\boldsymbol{y}^0$の近傍に

$$\boldsymbol{z}^0=\boldsymbol{\varphi}(\boldsymbol{y}^0), \tag{8}$$

$$\boldsymbol{H}(\boldsymbol{y},\boldsymbol{\varphi}(\boldsymbol{y}))=\boldsymbol{0}. \tag{9}$$

を満たす$\mathbb{R}^{N-M}\longrightarrow\mathbb{R}^M$の滑らかなベクトル値関数$\boldsymbol{\varphi}(\boldsymbol{y})$が唯一つ存在する．つまり$(\boldsymbol{y}^0,\boldsymbol{z}^0)$の近傍で許容集合が$\boldsymbol{z}=\boldsymbol{\varphi}(\boldsymbol{y})$と表される．

式(9)の両辺をy_j $(j=1,\cdots,N-M)$で微分すると

$$\frac{\partial H_i}{\partial y_j}+\frac{\partial H_i}{\partial z_k}\frac{\partial \varphi_k}{\partial y_j}=0. \tag{10}$$

$\boldsymbol{A}$および$\boldsymbol{\Phi}$をそれぞれ$\frac{\partial H_i}{\partial y_j}(\boldsymbol{y}^0,\boldsymbol{z}^0)$および$\frac{\partial \varphi_i}{\partial y_j}(\boldsymbol{y}^0)$を$i$行$j$列の成分とする$M$行$N-M$列の行列とすると式(10)より

$$\boldsymbol{A}+\boldsymbol{D\Phi}=\boldsymbol{0}\,, \tag{11}$$

が成り立つ．

<u>Step３　許容集合上の停留点</u>

$(x_1^0,\cdots,x_N^0)$の近傍の許容集合上の$f(x_1,\cdots,x_N)$は$\boldsymbol{y}\in\boldsymbol{y}^0$の近傍$\subset\mathbb{R}^{N-M}$での関数$F(\boldsymbol{y})=\tilde{f}(\boldsymbol{y},\boldsymbol{\varphi}(\boldsymbol{y}))$として表される．許容集合上の$f(x_1,\cdots,x_N)$の停留点が$(x_1^0,\cdots,x_N^0)$であることは，$F(\boldsymbol{y})$の停留点が$\boldsymbol{y}=\boldsymbol{y}^0$であることと同じで，次式を得る．

$$\frac{\partial F}{\partial y_j}(\boldsymbol{y}=\boldsymbol{y}^0)=\left(\frac{\partial \tilde{f}}{\partial y_j}+\frac{\partial \tilde{f}}{\partial z_k}\frac{\partial \varphi_k}{\partial y_j}\right)(\boldsymbol{y}=\boldsymbol{y}^0)=0\quad(j=1,\cdots,N-M). \tag{12}$$

式(4)は式(12)の単なる書き換えに過ぎない．以下にこれを示そう．

$\frac{\partial \tilde{f}}{\partial y_j}$を成分とする横ベクトルを$\boldsymbol{a}$，$\frac{\partial \tilde{f}}{\partial z_k}$を成分とする横ベクトルを$\boldsymbol{b}$とすると，式(12)は

$$\boldsymbol{a}+\boldsymbol{b\Phi}=0, \tag{13}$$

と表わせる．式(11)より$\boldsymbol{\Phi} = -\boldsymbol{D}^{-1}\boldsymbol{A}$だから，

$$\boldsymbol{a} = \boldsymbol{b}\boldsymbol{D}^{-1}\boldsymbol{A}. \tag{14}$$

従って$\boldsymbol{b}\boldsymbol{D}^{-1} = \tilde{\boldsymbol{\lambda}}$，ただし$\tilde{\boldsymbol{\lambda}}$は$M$成分の横ベクトル，とすると

$$\boldsymbol{a} = \tilde{\boldsymbol{\lambda}}\boldsymbol{A}, \tag{15}$$

$$\boldsymbol{b} = \tilde{\boldsymbol{\lambda}}\boldsymbol{D}. \tag{16}$$

式(15)および(16)を具体的に書くと

$$\begin{aligned}&\left(\frac{\partial \tilde{f}}{\partial y_1}, \cdots, \frac{\partial \tilde{f}}{\partial y_{N-M}}, \frac{\partial \tilde{f}}{\partial z_1}, \cdots, \frac{\partial \tilde{f}}{\partial z_M}\right) \\ &= \left(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_M\right)\begin{pmatrix} \dfrac{\partial H_1}{\partial y_1} & \cdots & \dfrac{\partial H_1}{\partial y_{N-M}} & \dfrac{\partial H_1}{\partial z_1} & \cdots & \dfrac{\partial H_1}{\partial z_M} \\ \vdots & & \vdots & \vdots & & \vdots \\ \dfrac{\partial H_M}{\partial y_1} & \cdots & \dfrac{\partial H_M}{\partial y_{N-M}} & \dfrac{\partial H_M}{\partial z_1} & \cdots & \dfrac{\partial H_M}{\partial z_M} \end{pmatrix}.\end{aligned} \tag{17}$$

この式を元の記号で表せば，それは式(4)と一致する．

説明が長くなったので, 議論の筋道をまとめておく. 許容集合上の一点$(\boldsymbol{y}^0, \boldsymbol{z}^0)$の近傍で許容集合が$\boldsymbol{y}$のベクトル値関数$\boldsymbol{z} = \boldsymbol{\varphi}(\boldsymbol{y})$として表されることを，陰関数定理は条件 1 および 2 の下で保証する．拘束条件下での$f(x_1, \cdots, x_N)$の停留値を求める問題が拘束条件のない$F(\boldsymbol{y}) = \tilde{f}(\boldsymbol{y}, \boldsymbol{\varphi}(\boldsymbol{y}))$の停留値を求める問題に変換され，$\frac{\partial F}{\partial y_i} = 0\ (i = 1, \cdots, N - M)$として式(4)が導かれたのである．

<u>式(7)の検証</u>

陰関数定理はＬ法の核心となる仮定(6)から導かれる式(7)が成り立つことを保証する．以下にこれを示そう．

$\boldsymbol{z} = \boldsymbol{\varphi}(\boldsymbol{y})$で表される許容集合上の微小変位ベクトル

$$d\boldsymbol{l} = \begin{pmatrix} d\boldsymbol{y} \\ d\boldsymbol{z} \end{pmatrix} = (dy_1, \cdots, dy_{N-M}, dz_1, \cdots, dz_M)^{\mathrm{T}},$$

を考える．$\boldsymbol{I}_{N-M}$ を$N-M$次の単位行列，$\boldsymbol{C}$を$\frac{\partial \varphi_i}{\partial y_j}$を$i$行$j$列成分とする$M$行$N-M$列の行列とし，これらを縦に並べてできる$N$行$N-M$列の行列を$\boldsymbol{B}$として定義する．

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I}_{N-M} \\ \boldsymbol{C} \end{pmatrix}.$$

$d\boldsymbol{l}$は

$$d\boldsymbol{l} = \boldsymbol{B}d\boldsymbol{y}, \tag{18}$$

のように$\boldsymbol{B}$と$d\boldsymbol{y}$との積の形で表される．

式(18)は$\mathbb{R}^{N-M}$から$\boldsymbol{B}$の列ベクトルの張る空間への線形写像を示している．また式(10)は$\boldsymbol{B}$の列ベクトルが

$$\nabla H_i = \left(\frac{\partial H_i}{\partial y_1}, \cdots, \frac{\partial H_i}{\partial y_{N-M}}, \frac{\partial H_i}{\partial z_1}, \cdots, \frac{\partial H_i}{\partial z_M} \right) \qquad (i = 1, \cdots, M),$$

と直交することを示している．すなわち$\boldsymbol{B}$の列ベクトルの張る$\mathbb{R}^N$の部分空間を$\boldsymbol{X}$，$\nabla H_i\ (i = 1, \cdots, M)$の張る$\mathbb{R}^N$の部分空間を$\boldsymbol{Y}$とすると，$\boldsymbol{X} \perp \boldsymbol{Y}$．$\boldsymbol{Y}$と直交する$\mathbb{R}^N$のベクトルの集合が$\boldsymbol{Y}^\perp$であるから，$\boldsymbol{X}$は$\boldsymbol{Y}^\perp$の部分空間である．すなわち$\boldsymbol{X} \subseteq \boldsymbol{Y}^\perp$．一方，$\boldsymbol{B}$の列ベクトルは線形独立だから$\boldsymbol{B}$の列ベクトルの張る空間$\boldsymbol{X}$の次元は$N-M$で，これは$\boldsymbol{Y}^\perp$の次元に等しい(条件２より$\nabla H_i\ (i = 1, \cdots, M)$は線形独立で$\dim \boldsymbol{Y} = M$)．従って$\boldsymbol{X} = \boldsymbol{Y}^\perp$．もとの記号に直せば$\boldsymbol{X}$が許容集合の接空間であり，$\boldsymbol{Y}$が$\boldsymbol{V} = \mathrm{span}(\nabla g_k, k = 1, \cdots, M)$である．すなわち式(7)が条件１および２の下で成り立つ．

４．例題

ここでは３つの簡単な例題を用意した．いずれの問題もわざわざＬ法を使わなくても停留点が求まる簡単な問題である．Ｌ法で停留点が求まるものもあれば，計算が途中で行き詰ってＬ法ではお手上げの問題もある．さらにＬ法がうまく

行く場合でも，未定乗数は一意に定まらない．その理由を考えるのがここでの狙いである．

例題 1 拘束条件$g_1(x,y,z)=x-y=0$，$g_2(x,y,z)=x=0$，$g_3(x,y,z)=y=0$，$g_4(x,y,z)=2x+y=0$の下で関数$f(x,y,z)=(x-1)^2+y^2+z^2$の停留点を探せ．

解答例：まず L 法を使わずに解く．許容集合がz軸であることはすぐに分かり，z軸上で関数fは$1+z^2$と表されるから停留点は原点にある．次に L 法で解く．

$$F(x,y,z,\lambda_1,\lambda_2,\lambda_3)=f(x,y,z)-\lambda_1 g_1(x,y,z)-\lambda_2 g_2(x,y,z)-\lambda_3 g_3(x,y,z)-\lambda_4 g_4(x,y,z),$$

として

$$\frac{\partial F}{\partial \alpha}=0\ (\alpha=x,y,z,\lambda_1,\lambda_2,\lambda_3,\lambda_4).$$

$\alpha=x,y,z$の場合より

$$\begin{pmatrix}2(x-1)\\2y\\2z\end{pmatrix}-\lambda_1\begin{pmatrix}1\\-1\\0\end{pmatrix}-\lambda_2\begin{pmatrix}1\\0\\0\end{pmatrix}-\lambda_3\begin{pmatrix}0\\1\\0\end{pmatrix}-\lambda_4\begin{pmatrix}2\\1\\0\end{pmatrix}=0, \tag{a}$$

が得られ，$\alpha=\lambda_1,\lambda_2,\lambda_3,\lambda_4$の場合から

$$x=y=0, \tag{b}$$

が得られる．式(a)から直ちに

$$z=0. \tag{c}$$

つまり停留点は原点という正解が導かれた．未定乗数は式(a)，(b)，(c)より，sおよびtをパラメータとして

$$\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\\\lambda_4\end{pmatrix}=s\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}+t\begin{pmatrix}0\\-1\\-1\\1\end{pmatrix}+\begin{pmatrix}0\\2\\0\\0\end{pmatrix},$$

という関係で表されるが，相変わらずこれらの未定乗数の値は未定のままである．

例題 2 拘束条件$g_1(x,y,z)=x^2+y^2-1=0$，$g_2(x,y,z)=x^2+y^2+z^2-1=0$の下で関数$f(x,y,z)=(x-2)^2+y^2+(z-1)^2$の停留点を探せ．

解答例：まず L 法を使わずに解く．許容集合が$\{z=0,\ x^2+y^2=1\}$であることはすぐ分かる．従って問題はこの円周上で$f(x,y,z)=(x-2)^2+y^2+(z-1)^2$の停留点を探すことに帰着し，円周上の点を$(\cos\theta,\sin\theta)$とパラメータ$\theta\ (0\le\theta<2\pi)$を用いて表すと，円周上で$f=-2\cos\theta+6$だから，$\frac{df}{d\theta}=0$を満たす停留点として$(x,y,z)=(\pm1,0,0)$を得る．次に L 法で解く．

$$F(x,y,z,\lambda_1,\lambda_2)=f(x,y,z)-\lambda_1 g_1(x,y,z)-\lambda_2 g_2(x,y,z),$$

$$F_\alpha(x,y,z,\lambda_1,\lambda_2)=0\ (\alpha=x,y,z,\lambda_1,\lambda_2),$$

とすると，$\alpha=\lambda_1,\lambda_2$の場合からは$z=0$かつ$x^2+y^2=1$が導かれ，$\alpha=z$の場合からは$2(z-1)-2\lambda_2 z=0$が導かれる．しかしこれらは両立しない．つまりこの問題はL法ではお手上げである．

最後の例題として例題2に拘束条件$g_3(x,y,z)=z=0$を追加した例題3を用意した．追加した拘束条件によって許容集合は変わらないから，正解は例題2と同じである．

<u>例題3</u> 拘束条件$g_1(x,y,z)=x^2+y^2-1=0$，$g_2(x,y,z)=x^2+y^2+z^2-1=0$，$g_3(x,y,z)=z=0$の下で関数$f(x,y,z)=(x-2)^2+y^2+(z-1)^2$の停留点を探せ．

<u>解答例</u>

$$F(x,y,z,\lambda_1,\lambda_2,\lambda_2)=f(x,y,z)-\lambda_1 g_1(x,y,z)-\lambda_2 g_2(x,y,z)-\lambda_3 g_3(x,y,z),$$
$$F_\alpha(x,y,z,\lambda_1,\lambda_2)=0\ (\alpha=x,y,z,\lambda_1,\lambda_2,\lambda_2),$$

において，$\alpha=\lambda_1,\lambda_2,\lambda_3$の場合からは$z=0$かつ$x^2+y^2=1$が導かれる．つまり拘束条件を追加しても許容集合は変わらない．$\alpha=z$の場合からは$2(z-1)-2\lambda_2 z-\lambda_3=0$が導かれる．これに$z=0$を代入すれば$\lambda_3=-2$が得られる．$\alpha=x$の場合からは$(1-\lambda_1-\lambda_2)x=2$が，$\alpha=y$の場合からは$(1-\lambda_1-\lambda_2)y=0$が導かれるが，$1-\lambda_1-\lambda_2=0$は前者を満たさないので$y=0$が結論される．これを$x^2+y^2=1$に代入すれば$x=\pm1$．つまり正解$(x,y,z)=(\pm1,0,0)$が得られる．しかし$\lambda_1$と$\lambda_2$に関しては，$\lambda_1+\lambda_2=-1\pm2$という関係が導かれるが，これらは未定乗数のままである．

L法は許容集合が式(6)で近似できることを仮定した方法である．式(6)から導かれる許容集合の接空間の公式(7)は例題1および3では満たされるが，例題2では満たされない．その結果，例題1および3ではL法は停留点を求めることができ，例題2では計算が途中で破綻し正解にたどり着けなかった．一方，$\nabla g_k(\boldsymbol{x}^0)$が線形独立でないことは3つの例題に共通する．それでも例題1および3でL法が機能したのは，$\nabla g_k(\boldsymbol{x}^0)$の線形独立性は式(7)を満たすための十分条件であって必要条件ではないからである．例題1および3における拘束条件には冗長性がある．例題1では4つの拘束条件が課せられたが，実質的な拘束条件は2つである．残りの2つは許容集合や$\boldsymbol{V}$に影響を及ぼさない冗長な拘束条件である．冗長性より$\nabla g_k(\boldsymbol{x}^0)$は線形従属となり，その結果，未定乗数の値は一意に定まらなかった．例題2では2つの拘束条件には冗長性はないが，この2つの拘束条件が表す曲面は接触するので式(7)を満たさない．ところが例題3では，追加された拘束条件$g_3(x,y,z)=z=0$によって許容集合は変わらないが，$\boldsymbol{V}$が変わって

式(7)が満たされるようになったのである．言い換えれば，追加された拘束条件が拘束条件$g_1(x,y,z)=x^2+y^2-1=0$および$g_2(x,y,z)=x^2+y^2+z^2-1=0$の一方を冗長な拘束条件として追いやったのである．拘束条件の冗長性より$\nabla g_k(\boldsymbol{x}^*)$は線形従属で未定乗数の値が一意に定まらないのは例題１の場合と同じである．例題１および３で冗長な拘束条件を取り除けば，$\nabla g_k(\boldsymbol{x}^0)$の線形独立性が満たされ，未定乗数の値が一意に定まるのは言うまでもない．

よくＬ法は便利な解法と言われる．しかしＬ法は式(6)（或いは式(7)）を仮定した方法なので，これが満たされることを確認しなければならない．未定乗数の値が一意に定まっても，それは停留点において$\nabla g_k(\boldsymbol{x}^0)$の線形独立性が満たされることを意味するのであって，許容集合全体における確認が不要であることを意味しない．確認の作業まで含めればＬ法は必ずしも便利な方法とは言えないのではないか，つまり確認の作業を省略するから便利な方法と思えるのではないかと筆者は思うのである．

５．オイラー・ラグランジュ方程式への応用

最後に Ernst Mach が思考の経済と絶賛した解析力学への橋渡しとして，拘束条件下の汎関数の停留曲線(最大値あるいは最小値を与える曲線の候補)を求める問題へのＬ法の応用を説明する．汎関数とは，例えば$\sin x$ に対して5の値を返し，$\cos x$に対しては2の値を返すというように，入力が関数で出力が実数(複素数)値である写像である．ここでは力学への応用を考慮して次の形の汎関数を考える．

$$J[x(t)]=\int_a^b L(x(t),\dot{x}(t),t)dt, \tag{19}$$

ここに$\dot{x}(t)=\frac{dx(t)}{dt}$であり，$L(x(t),\dot{x}(t),t)$は関数形が既知の滑らかな関数で，さらに$x(t)$には境界条件

$$x(a)=\alpha,\qquad x(b)=\beta, \tag{20}$$

が課せられているとする．この汎関数の停留曲線を滑らかな曲線の範囲から探すのが問題である．Ｌ法は拘束条件下での汎関数の停留曲線を求めるのに応用されるが，ここではまず拘束条件がない場合を考えよう．

停留曲線$x^*(t)$が存在すると仮定し，検索する曲線の範囲をεをパラメータとする曲線群

$$y(t,\varepsilon) = x^*(t) + \varepsilon\eta(t), \tag{21}$$

に絞る．ただし$y(t,\varepsilon)$は滑らかでなければならないので$\eta(t)$も滑らかな関数で，さらに$x(t) = y(t,\varepsilon)$が境界条件(20)を満たすように境界条件

$$\eta(a) = \eta(b) = 0, \tag{22}$$

が課されているとする．この曲線群を引数とする汎関数$J[y(t,\varepsilon)]$はεの関数になり，これを$D(\varepsilon)$と表すと，$x_*(t)$が停留曲線であるための必要条件として

$$\frac{dD(\varepsilon)}{d\varepsilon}|_{\varepsilon=0} = 0, \tag{23}$$

を得る. 定積分における微分と積分の順序交換と部分積分を行って式(23)の右辺を計算すると

$$\begin{aligned}\frac{\mathrm{d}D(\varepsilon)}{\mathrm{d}\varepsilon}|_{\varepsilon=0} &= \int_a^b \left(\frac{\partial L}{\partial x}\eta + \frac{\partial L}{\partial \dot{x}}\dot{\eta}\right)(x^*(t),\dot{x}^*(t),t)dt \\ &= \int_a^b \left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right)(x^*(t),\dot{x}^*(t),t)\eta(t)dt\ . \end{aligned} \tag{24}$$

$\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right)(x^*(t),\dot{x}^*(t),t)$が$t$の連続関数で，式(23)を満たす任意の滑らかな関数$\eta(t)$に対して式(23)が成り立つならば，以下に示す変分学の基本補題と呼ばれるロジックが使えて

$$\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right)(x^*(t),\dot{x}^*(t),t) = 0, \tag{25}$$

を得る.

(変分学の基本補題：$F(t)$が連続関数で式(21)を満たす任意の滑らかな関数$\eta(t)$に対して

$$\int_a^b F(t)\eta(t)dt = 0,$$

を満たすならば，区間(a,b)上で$F(t) = 0$である．

$\because F(t_0) > 0\ (a < t_0 < b)$ならば，$F(t)$は連続だから$t = t_0$の近傍で$F(t) > 0$であり，この近傍内で正の値をとりその外部はゼロとなる滑らかな関数$\eta(t)$を考えれば，定積分は正になり矛盾する．$F(t_0) < 0$の場合も同様．■）

式(25)はオイラー・ラグランジュの方程式と呼ばれる．この式はしばしば引数を省略して

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0,$$

と表されるが，Lに対する偏微分方程式ではなく，停留曲線$x^*(t)$が満たすべき 2 階の常微分方程式であることに注意しよう．

多自由度の場合への拡張

汎関数

$$J[q_1(t),\cdots,q_N(t)] = \int_a^b L(q_1(t),\cdots,q_N(t),\dot{q}_1(t),\cdots,\dot{q}_N(t),t)dt, \tag{26}$$

を考える．ただし$q_1(t),\cdots,q_N(t)$には境界条件

$$q_i(a) = \alpha_i, \qquad q_i(b) = \beta_i, \quad (i = 1,\cdots,N), \tag{27}$$

が課せられているとする．$q_i^*(t)$を停留曲線とし，εをパラメータとする曲線群

$$y_i(t,\varepsilon) = q_i^*(t) + \varepsilon\eta_i(t), \tag{28}$$

を考える．ここに$\eta_i(t)$は

$$\eta_i(a) = \eta_i(b) = 0, \tag{29}$$

を満たす滑らかな関数である．$q_i(t) = y_i(t,\varepsilon)$は境界条件(25)を満たすことに注

意しよう．汎関数$J[y_i(t,\varepsilon)]$はεの関数で，$q_i^*(t)$が停留曲線ならば，境界条件(29)を満たす任意の滑らかな関数$\eta_i(t)$に対して

$$\frac{d}{dx}J[y_i(t,\varepsilon)]|_{\varepsilon=0}=\int_a^b\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\boldsymbol{q}^*(t),\dot{\boldsymbol{q}}^*(t),t)\eta_i(t)dt=0, \tag{30}$$

を満たす．式(30)の導出は式(24)の導出と同様である．ここで$t=t_0$（$a<t_0<b$）において，式(30)右辺の被積分関数が正(あるいは負)としよう．$s(t_0)=1$および$t=t_0$の近傍でのみゼロでない値を持つ滑らかな関数$s(t)$と$\eta_i(t)$との積$s(t)\eta_i(t)$は境界条件(29)を満たす滑らかな関数であるので，式(30)においてこれを$\eta_i(t)$の代わりに用いることが許される．すると$\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\boldsymbol{q}^*(t),\dot{\boldsymbol{q}}^*(t),t)$は連続であるので，変分学の基本補題が使えて，式(30)の被積分関数は$t=t_0$においてゼロが結論される．t_0は区間(a,b)内に任意にとれるので，結局，$a<t<b$において

$$\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\boldsymbol{q}^*(t),\dot{\boldsymbol{q}}^*(t),t)\eta_i(t)=0. \tag{31}$$

さらに$\eta_i(t)(i=1,\cdots,N)$には任意性があり，$\eta_j(t)=0\ (j\neq i)$の場合を考えれば，$a<t<b$において

$$\left(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\boldsymbol{q}^*(t),\dot{\boldsymbol{q}}^*(t),t)=0\quad (i=1,\cdots,N), \tag{32}$$

が結論される．これは$q_i^*(t)$（$i=1,\cdots,N$）に対する連立の常微分方程式である．式(32)もしばしば引数が省略されるが誤解は生じないだろう．

L 法の応用

汎関数の停留曲線を求める問題(25)，(26)を holonomic な拘束条件

$$g_k(\boldsymbol{q}(t),t)=0\ \ (k=1,\cdots,M), \tag{33}$$

を課して考える．ただし$g_k\ (k=1,\cdots,M)$は滑らかな関数とする．拘束条件(33)の下では$\eta_j(t)\ (j=1,\cdots,N)$ に

$$\frac{\partial g_k}{\partial q_j}(\boldsymbol{q}^*(t), t)\eta_j(t) = 0 \qquad (k = 1, \cdots, M), \tag{34}$$

なる条件が課されるが，式(31)までの導出は同じである．式(31)および式(4)を導いた議論，つまり$(V^{\perp})^{\perp} = V = \mathrm{span}\left(\nabla g_k \ (k = 1, \cdots, M)\right)$（ただし$\nabla$は$q_j$ $(j = 1, \cdots, N)$に関する勾配)，より区間(a, b)の各点において

$$\left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)(\boldsymbol{q}^*(t), \dot{\boldsymbol{q}}^*(t), t) = \sum_{k=1}^{M} \lambda_k(t) \frac{\partial g_k}{\partial q_i}(\boldsymbol{q}^*(t), t), \tag{35}$$

が成り立たなければならない. 式(35)の左辺および$\frac{\partial g_k}{\partial q_i}(\boldsymbol{q}^*(t), t)$は滑らかだから，$\lambda_k(t)$ $(k = 1, \cdots, M)$も滑らかである． これが holonomic な拘束条件下のオイラー・ラグランジュの方程式である．ただしこの式は$\boldsymbol{q}^*(t)$が停留曲線であるための必要条件であり， さらに式(7)が成り立つかどうかの確認(例えば∇g_k $(k = 1, \cdots, M)$が線形独立であるかどうか)が必要ということに注意しよう.