ラグランジュの未定乗数法

拘束条件があるときの極値の求め方であるラグランジュの未定乗数法を示します。
「・」は時間微分です。

物体の運動に対して条件が付いている場合を考えます。例えば、単振り子には重りに繋がっている糸の長さ l の範囲での運動という制限があります。この条件は 2 次元での重りの位置を (x, y) とすれば $x^2+y^2-l^2=0$ と書けます。このような条件を拘束条件 (束縛条件、constraint) と言います。2 次元平面での位置を表すには 2 個の変数が必要ですが、単振り子では $x^2+y^2-l^2=0$ の拘束条件のために 1 個の変数で重りの位置を与えられます。このため、図 1 のように角度 θ を与えれば、単振り子の運動方程式は

$$\frac{d^2\theta}{dt^2}=-\frac{g}{l}\sin\theta$$

だけになります。言い換えると、拘束条件は 2 個の独立変数 x, y を 1 個の独立変数 θ へと変更します。これがオイラー・ラグランジュ方程式の導出に影響を与えます。

n 次元での一般化座標を $\boldsymbol{q}=(q_1, q_2, \ldots, q_n)$ として、作用 S とラグランジアン L は

$$S[q]=\int_{t_1}^{t_2} dt\ L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \qquad (\dot{\boldsymbol{q}}=\frac{d\boldsymbol{q}}{dt})$$

これの変分を取ることで

$$\delta S=\int_{t_1}^{t_2} dt\ \sum_{i=1}^{n}\delta q_i(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i})=0$$

δq_i は独立なので和の各項の括弧内はそれぞれ 0 とでき、$i=1,2,\ldots,n$ に対してそれぞれオイラー・ラグランジュ方程式が出てきます。しかし、例えば n 個の q_i に対して 1 個の拘束条件があると、q_i のどれか 1 個は独立ではなくなり、括弧内が 0 になるとは言えなくなります。例えば、$n=3$ とし $q_3=h(q_1, q_2)$ であるなら

$$\delta q_1(\frac{\partial L}{\partial q_1}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1})+\delta q_2(\frac{\partial L}{\partial q_2}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_2})+\delta h(q_1, q_2)(\frac{\partial L}{\partial q_3}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_3})=0$$

となり、第 3 項にも $\delta q_1, \delta q_2$ が現れるために各項の括弧内がそれぞれ 0 とは言えないです。これをどうにかします。

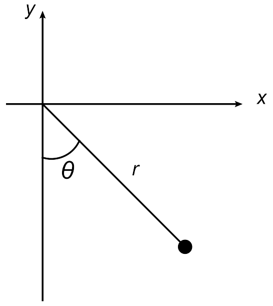

図 1

単語を与えておきます。拘束条件が関数 g によって

$$g(\boldsymbol{q}, t) = 0$$

と書けるときはホロノミック拘束条件 (holonomic constrain) と呼ばれ

$$g(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = 0$$

と書けるときは非ホロノミック拘束条件 (nonholonomic constrain) と呼ばれます。ほとんどの物理の問題ではどちらの場合でも同じようにできますが、非ホロノミック拘束条件の数学的な扱いの話が必要になるので、ここではホロノミック拘束条件として見ていきます。

q_i が独立な項と独立でない項が出てきているので、独立でない項を消せるように新しい項を加えます。m 個の拘束条件を関数 g_k $(k = 1, 2, \ldots, m)$ によって $g_k(\boldsymbol{q}, t) = 0$ とします。これに時間の関数 $\lambda_k(t)$ をかけても

$$\lambda_k(t) g_k(\boldsymbol{q}, t) = 0$$

時間で積分しても

$$\int_{t_1}^{t_2} dt\ \lambda_k(t) g_k(\boldsymbol{q}, t) = 0$$

これを作用に加えると

$$S = \int_{t_1}^{t_2} dt \big(L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \sum_{k=1}^{m} \lambda_k(t) g_k(\boldsymbol{q}, t) \big)$$

λ, g に対しても変分を取ることにして

$$\delta S = \int_{t_1}^{t_2} dt (\delta L + \sum_{k=1}^{m} \lambda_k \delta g_k + \sum_{k=1}^{m} \delta \lambda_k g_k) = 0$$

δg_k は、時間の変分は考えないので

$$\delta g_k = \frac{\partial g_k}{\partial q_i} \delta q_i$$

これを入れれば

$$\delta S = \int_{t_1}^{t_2} dt \big(\sum_{i=1}^{n} \delta q_i (\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial q_i}) + \sum_{k=1}^{m} \delta \lambda_k g_k \big) = 0$$

となるので

$$\sum_{i=1}^{n}\delta q_i\big(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i}\big)+\sum_{k=1}^{m}\delta\lambda_k g_k=0$$

拘束条件 $g_k(\boldsymbol{q},t)=0$ から最後の項は消えるので

$$\sum_{i=1}^{n}\delta q_i\big(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k(t)\frac{\partial g_k}{\partial q_i}\big)=0 \tag{1}$$

拘束条件のために全ての q_i が独立になっていなく、m 個の拘束条件から独立なのは $n-m$ 個です。

今は λ_k の項がいるので、これを利用します。$i=1,2,\ldots,n-m$ を独立な部分、$i=n-m+1,n-m+2,\ldots,n$ を独立でない部分とします。独立でない部分では括弧内が 0 になるように λ_k を選びます。つまり

$$\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k(t)\frac{\partial g_k}{\partial q_i}=0 \quad (i=n-m+1,n-m+2,\ldots,n) \tag{2}$$

となるように選ぶことで、(1) の和におけるこれらの項を 0 にします。この方程式の数は m 個なので、原理的には m 個の λ_k について解けます。そして、残っている $n-m$ 個の独立な q_i の部分は

$$\sum_{i=1}^{n-m}\delta q_i(\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k(t)\frac{\partial g_k}{\partial q_i})=0$$

これの δq_i は独立なので何も気にすることなく

$$\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k(t)\frac{\partial g_k}{\partial q_i}=0 \quad (i=1,2,\ldots,n-m) \tag{3}$$

とできます。そうすると、(2),(3) は同じ形をしているので、全ての i に対して

$$\frac{\partial L}{\partial q_i}-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}+\sum_{k=1}^{m}\lambda_k(t)\frac{\partial g_k}{\partial q_i}=0 \tag{4}$$

という式が成立します。これは n 個の式ですが、式には $n+m$ 個の変数 (n 個の q_i と m 個の λ_k) があります。なので、解くには m 個の $g_k(\boldsymbol{q},t)=0$ を含めた $n+m$ 個の式が必要で、これらによって運動方程式と同じになります。このようにして拘束条件があるときの解（極値を与える関数 $\boldsymbol{q}$）を求める方法をラグランジュの未定乗数法 (method of Lagrange multiplier, method of Lagrange undetermined multiplier) と言います。λ_k はラグランジュ乗数 (Lagrange multiplier) や乗数と呼ばれます。

今のように $g(\boldsymbol{q},t)$ であるなら

$$F=L+\sum_{k=1}^{m}\lambda_k(t)g_k(\boldsymbol{q},t)$$

とすることで

$$\frac{\partial F}{\partial q_i}-\frac{d}{dt}\frac{\partial F}{\partial \dot{q}_i}=0 \tag{5}$$

として、オイラー・ラグランジュ方程式の形にできます。また、λ_k もオイラー・ラグランジュ方程式に従っているとして

$$\frac{\partial F}{\partial \lambda_k}-\frac{d}{dt}\frac{\partial F}{\partial \dot{\lambda}_k}=0$$

とすれば、$g_k(\boldsymbol{q},t)=0$ になるので拘束条件を導きます。
後で見るように、ラグランジュの未定乗数法は関数 f を条件 g とラグランジュ乗数 λ によって $f+\lambda g$ とすることで、条件のもとでの f の極値を求めるという方法です。

(4) を

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}=\frac{\partial L}{\partial q_i}+\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i}$$

と変形して、一般化運動量

$$p_i=\frac{\partial L}{\partial \dot{q}_i}$$

を使うと

$$\frac{dp_i}{dt}=\frac{\partial L}{\partial q_i}+\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i}$$

ラグランジアン L は運動エネルギーとポテンシャルの差なので、ポテンシャル V にしか q_i の依存性がないとして

$$\frac{dp_i}{dt}=-\frac{\partial V}{\partial q_i}+\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i}=-\frac{\partial}{\partial q_i}(V-\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i})$$

左辺は運動方程式の左辺、右辺は力です。そうすると、保存力 $F_i=-\partial V/\partial q_i$ に

$$F_i^{\mathrm{ex}}=\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial q_i}$$

という力が q_i 方向に加わっているように見えます。この力が拘束力 (束縛力、constrain force) に対応します。このため、λ_k を解くことは拘束力を見つけることに対応します。

単振り子を使って、どうなるのか見ていきます。まず、xy 平面上で質量 m の質点が重力加速度 g を $-y$ 方向に受けているとします。このときの運動エネルギー T とポテンシャル U は

$$T=\frac{1}{2}m(\dot{x}^2+\dot{y}^2)\ ,\ U=mgy$$

図 1 のように角度 θ と動径 r で書くことにすれば

$$x = r\sin\theta\ ,\ y = -r\cos\theta$$

$$\dot{x} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta\ ,\ \dot{y} = -\dot{r}\cos\theta + r\dot{\theta}\sin\theta$$

$$\dot{x}^2 + \dot{y}^2 = \dot{r}^2\sin^2\theta + (r\dot{\theta})^2\cos^2\theta + \dot{r}^2\cos^2\theta + (r\dot{\theta})^2\sin^2\theta = \dot{r}^2 + (r\dot{\theta})^2$$

から

$$L = T - U = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta$$

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)\ ,\ U = -mgr\cos\theta$$

これに拘束条件 $g(r) = r - l = 0$ を加えます。今の (4) は

$$\frac{\partial L}{\partial r} - \frac{d}{dt}\frac{\partial L}{\partial \dot{r}} + \lambda\frac{\partial g}{\partial r} = 0\ ,\ \frac{\partial L}{\partial \theta} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} + \lambda\frac{\partial g}{\partial \theta} = 0$$

$$\frac{\partial g}{\partial r} = 1$$

なので

$$mr\dot{\theta}^2 + mg\cos\theta - m\frac{d^2r}{dt^2} + \lambda = 0$$

$$-mgr\sin\theta - mr^2\frac{d^2\theta}{dt^2} = 0$$

解を求めるために拘束条件 $r = l$ を入れて

$$ml\dot{\theta}^2 + mg\cos\theta + \lambda = 0$$

$$-mgl\sin\theta - ml^2\frac{d^2\theta}{dt^2} = 0$$

上の式から λ は

$$\lambda = -ml\dot{\theta}^2 - mg\cos\theta$$

と求まり、拘束力として

$$\lambda\frac{\partial g}{\partial r} = -ml\dot{\theta}^2 - mg\cos\theta \tag{6}$$

が r 方向に働いていることが分かります。力学の見方をすれば、単振り子の r 方向には重力による $mg\cos\theta$、回転による向心力 $ml\dot{\theta}^2$、張力 S が働いており、これらによるつり合いの式 (r 方向の運動方程式) は

$$ml\dot{\theta}^2 = S - mg\cos\theta$$

なので、(6) は張力です。下の式は

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta$$

として単振り子の運動方程式になります。

ラグランジュの未定乗数法を作用に 0 の項を加えて求めましたが、変分を使った場合も見ておきます。分かりやすくするために 2 次元 (x, y) として、作用とラグランジアンを

$$S = \int_{t_1}^{t_2} dt\ L(x, y, \dot{x}, \dot{y}, t)$$

拘束条件は $g(x, y, t) = 0$ とします。通常通りなら、$\overline{x}, \overline{y}$ のときに極値を与えるとして、微小な ϵ と関数 η_x, η_y による変分で位置をズラして

$$x(t) = \overline{x}(t) + \epsilon\eta_x(t)\ ,\ y(t) = \overline{y}(t) + \epsilon\eta_y(t)$$

$$\eta_x(t_1) = \eta_x(t_2) = 0\ ,\ \eta_y(t_1) = \eta_y(t_2) = 0$$

しかし、今は拘束条件があるために x, y は独立ではないです。このことを加えます。まず、x ではそのまま

$$x(t) = \overline{x}(t) + \epsilon\eta_x(t)\ ,\ \eta_x(t_1) = \eta_x(t_2) = 0$$

拘束条件を満たす極値を求めたいので、$\overline{x}, \overline{y}$ は拘束条件を満たしているとします。そうすると、t_1, t_2 での拘束条件は

$$g(x(t_1) - \epsilon\eta_x(t_1), \overline{y}(t_1, \epsilon), t_1) = g(x(t_1), \overline{y}(t_1, \epsilon), t_1) = 0$$

$$g(x(t_2) - \epsilon\eta_x(t_2), \overline{y}(t_2, \epsilon), t_2) = g(x(t_2), \overline{y}(t_2, \epsilon), t_2) = 0$$

となるので

$$y(t) = \overline{y}(t) + \eta_y(t, \epsilon)$$

と書くことにすれば、$\eta_y(t_1, \epsilon) = \eta_y(t_2, \epsilon) = 0$ となります。$\epsilon = 0$ では

$$g(x, y - \eta_y(t, 0), t) = 0$$

なので、$\eta_y(t, 0) = 0$ です。

拘束条件がいても $\epsilon=0$ のときに極値を与えることは変わらないので、オイラー・ラグランジュ方程式に対応する (5) は

$$\frac{d}{d\epsilon}\int_{t_1}^{t_2} dt\ L(x,y,\dot{x},\dot{y},t)\Big|_{\epsilon=0}=0$$

から求められるはずです。ϵ と t は無関係なので積分の中に微分を入れて計算すれば

$$\begin{aligned}
0 =& \int_{t_1}^{t_2} dt\ (\frac{dx}{d\epsilon}\frac{\partial L}{\partial x}+\frac{d\dot{x}}{d\epsilon}\frac{\partial L}{\partial \dot{x}}+\frac{dy}{d\epsilon}\frac{\partial L}{\partial y}+\frac{d\dot{y}}{d\epsilon}\frac{\partial L}{\partial \dot{y}})\Big|_{\epsilon=0}\\
=& \int_{t_1}^{t_2} dt\ (\eta_x\frac{\partial L}{\partial x}+\dot{\eta}_x\frac{\partial L}{\partial \dot{x}}+\frac{\partial \eta_y}{\partial\epsilon}\frac{\partial L}{\partial y}+\frac{\partial\dot{\eta}_y}{\partial\epsilon}\frac{\partial L}{\partial \dot{y}})\Big|_{\epsilon=0}\\
=& \int_{t_1}^{t_2} dt\ (\eta_x\frac{\partial L}{\partial x}+\frac{d\eta_x}{dt}\frac{\partial L}{\partial \dot{x}}+\frac{\partial \eta_y}{\partial\epsilon}\frac{\partial L}{\partial y}+\frac{d}{dt}\frac{\partial\eta_x}{\partial\epsilon}\frac{\partial L}{\partial \dot{y}})\Big|_{\epsilon=0}\\
=& \int_{t_1}^{t_2} dt\ (\eta_x\frac{\partial L}{\partial x}-\eta_x\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}+\frac{\partial \eta_y}{\partial\epsilon}\frac{\partial L}{\partial y}-\frac{\partial\eta_x}{\partial\epsilon}\frac{d}{dt}\frac{\partial L}{\partial \dot{y}})\Big|_{\epsilon=0}\\
=& \int_{t_1}^{t_2} dt\big(\eta_x(\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}})+\frac{\partial \eta_y}{\partial\epsilon}(\frac{\partial L}{\partial y}-\frac{d}{dt}\frac{\partial L}{\partial \dot{y}})\big)\Big|_{\epsilon=0}
\end{aligned}\tag{7}$$

$\eta(t,\epsilon)$ なので偏微分にしています。$\partial\eta_y/\partial\epsilon$ は拘束条件から求められます。拘束条件を ϵ で微分すると

$$\frac{d}{d\epsilon}g(x,y,t)=\frac{dx}{d\epsilon}\frac{\partial g}{\partial x}+\frac{dy}{d\epsilon}\frac{\partial g}{\partial y}=\eta_x\frac{\partial g}{\partial x}+\frac{\partial\eta_y}{\partial\epsilon}\frac{\partial g}{\partial y}=0$$

これから

$$\begin{aligned}
\eta_x\frac{\partial g}{\partial x}+\frac{\partial\eta_y}{\partial\epsilon}\frac{\partial g}{\partial y}=0&\\
\frac{\partial\eta_y}{\partial\epsilon}=&\ -\eta_x(\frac{\partial g}{\partial y})^{-1}\frac{\partial g}{\partial x}
\end{aligned}$$

$\partial g/\partial y\neq 0$ を要求していることに注意してください。これを (7) の第 2 項に入れれば

$$\int_{t_1}^{t_2} dt\ \eta_x\big((\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}})-(\frac{\partial g}{\partial y})^{-1}\frac{\partial g}{\partial x}(\frac{\partial L}{\partial y}-\frac{d}{dt}\frac{\partial L}{\partial \dot{y}})\big)\Big|_{\epsilon=0}=0$$

となるので

$$(\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}})-(\frac{\partial g}{\partial y})^{-1}\frac{\partial g}{\partial x}(\frac{\partial L}{\partial y}-\frac{d}{dt}\frac{\partial L}{\partial \dot{y}})=0$$

これは

$$(\frac{\partial g}{\partial x})^{-1}(\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}})=(\frac{\partial g}{\partial y})^{-1}(\frac{\partial L}{\partial y}-\frac{d}{dt}\frac{\partial L}{\partial \dot{y}})$$

と書けば、両辺は x, y の入れ替えで同じ式になっているのが分かります。このため、x, y に依存しない t のみの関数でないといけないので、その関数を $\lambda(t)$ とすれば

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \lambda(t)\frac{\partial g}{\partial x}\ ,\ \frac{\partial L}{\partial y} - \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = \lambda(t)\frac{\partial g}{\partial y}$$

となり、(4) になります。そして

$$F = L(x, y, \dot{x}, \dot{y}, t) - \lambda(t)g(x, y, t)$$

とすれば

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

と書けます。これらと拘束条件を満たす解が極値となる $x = \overline{x}, y = \overline{y}$ です。というわけで、最初に求めた結果と同じものが出てきます。

ここからは解析力学というより数学での話になりますが、数学部分はかなり省きます。

作用 (汎関数) に対して見てきましたが、条件があるときの関数の極値を求める場合でもラグランジュの未定乗数法が出てきます。xy 平面上で言えば、関数 $f(x, y)$ と $g(x, y)$ があり、$g(x, y) = 0$ となる曲線上での $f(x, y)$ の極値を求めるということです。曲線は各点において微分ができるとし ($\nabla g(x, y) \neq 0$)、曲線上の x, y は $x(s), y(s)$ と書けるとします (曲線のベクトルを $\boldsymbol{r}(s)$ とすれば、$d\boldsymbol{r}/ds \neq 0$)。ここでの極値は、曲線上の (x_0, y_0) において $f(x_0, y_0)$ が最大値、もしくは最小値となるという意味です。

$f(x, y)$ の曲線上の極値を与えるのは

$$\frac{df(s)}{ds} = \frac{df(x(s), y(s))}{ds} = \frac{dx}{ds}\frac{\partial f}{\partial x} + \frac{dy}{ds}\frac{\partial f}{\partial y} = 0 \tag{8}$$

一方で、拘束条件の曲線 $g(x, y) = 0$ から

$$\frac{dg(x, y)}{ds} = \frac{dx}{ds}\frac{\partial g}{\partial x} + \frac{dy}{ds}\frac{\partial g}{\partial y} = 0 \tag{9}$$

$\partial g/\partial x, \partial g/\partial y \neq 0$ としているので

$$\frac{dy}{ds} = -(\frac{\partial g}{\partial y})^{-1}\frac{dx}{ds}\frac{\partial g}{\partial x}$$

これを (8) に入れれば

$$\frac{dx}{ds}\frac{\partial f}{\partial x} - (\frac{\partial g}{\partial y})^{-1}\frac{dx}{ds}\frac{\partial g}{\partial x}\frac{\partial f}{\partial y} = (\frac{\partial g}{\partial y})^{-1}\frac{dx}{ds}(\frac{\partial g}{\partial y}\frac{\partial f}{\partial x} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}) = 0$$

$dx/ds \neq 0$ なら

$$\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} = 0$$

これは $\nabla f, \nabla g$ を 3 次元ベクトルとして

$$\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, 0) \ , \ \nabla g = (\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, 0)$$

と与えれば

$$\nabla f \times \nabla g = 0$$

と書けます。ベクトル積が 0 のベクトルは平行になっているので、曲線上の $f(x, y)$ の極値となる地点において

$$\nabla f = \lambda \nabla g$$
$$\nabla (f - \lambda g) = 0$$

を満たす定数 λ が存在します。

まとめると、曲線 $g(x, y) = 0$ ($\nabla g \neq 0$) において $f(x, y)$ が (x_0, y_0) で極値を持つとき

$$F(x, y) = f(x, y) - \lambda g(x, y)$$
$$\frac{\partial F}{\partial x}\Big|_{x=x_0, y=y_0} = 0 \ , \ \frac{\partial F}{\partial y}\Big|_{x=x_0, y=y_0} = 0 \ , \ g(x_0, y_0) = 0 \tag{10}$$

となる定数 λ が存在するということです。λ がラグランジュ乗数です。これを使って、拘束条件 $g(x, y) = 0$ があるときの関数 $f(x, y)$ の極値を求めるのがラグランジュの未定乗数法です。これを汎関数で行っているのが上での話で、ラグランジュ乗数が定数でなく t を変数にする関数に変更されています。

変数と条件が増えても同じことができます。ついでなので、別の方法で導出します。極値の式 (8) は全微分 df で書くと

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0$$

拘束条件での (9) も全微分で書けば

$$dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy = 0$$

dg に λ をかけて、差を取れば

$$df - \lambda dg = (\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x}) dx + (\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y}) dy = 0$$

拘束条件で x, y は独立ではないですが

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

となるように λ を選べば (λ について解く)、第 2 項は

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0$$

となります。よって

$$F(x,y) = f(x,y) - \lambda g(x,y)\ ,\ dF = d(f - \lambda g)$$
$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0\ ,\ \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0\ \Rightarrow\ \nabla(f - \lambda g) = 0$$

として、同じ結果が出てきます。これは全微分を使っているので、変数を $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ としても同じことができます。さらに、条件を $g_k(\boldsymbol{x})$ $(k = 1, 2, \ldots, m)$ としても

$$\begin{aligned}
df &= \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n = 0 \\
dg_1 &= \frac{\partial g_1}{\partial x_1}dx_1 + \frac{\partial g_1}{\partial x_2}dx_2 + \cdots + \frac{\partial g_1}{\partial x_n}dx_n = 0 \\
dg_2 &= \frac{\partial g_2}{\partial x_1}dx_1 + \frac{\partial g_2}{\partial x_2}dx_2 + \cdots + \frac{\partial g_2}{\partial x_n}dx_n = 0 \\
&\vdots
\end{aligned}$$

となるだけなので、独立でない k 個の項が 0 になるように λ_k を選べば

$$F(\boldsymbol{x}) = f(\boldsymbol{x}) - \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x})$$
$$\nabla(f(\boldsymbol{x}) - \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x})) = 0$$

となります。∇ は n 次元です。

例として $f(x,y) = x^2 - y^2,\ g(x,y) = x^2 + y^2 - a^2 = 0$ (a は定数) としてみます。条件は半径 a の円です。なので、この円周上での最大値か最小値を求めることになります。単純に求めるなら、$f(x,y)$ に条件の式を入れて

$$x^2 - y^2 = 2x^2 - a^2$$

とすれば、最大値は $x = \pm a, y = 0$ での a^2、最小値は $x = 0, y = \pm a$ での $-a^2$ と分かります。

ラグランジュの未定乗数法を使ってみます。(10) に入れれば

$$\frac{\partial}{\partial x}(x^2 - y^2 - \lambda(x^2 + y^2 - a^2)) = 0\ ,\ \frac{\partial}{\partial y}(x^2 - y^2 - \lambda(x^2 + y^2 - a^2)) = 0$$

なので

$$x - \lambda x = 0 \ , \ y + \lambda y = 0$$

x の式から $x = 0$ か $\lambda = 1$ です。$x = 0$ を条件 g に入れれば $y = \pm a$ です。$\lambda = 1$ を y の式に入れれば $y = 0$ となり、条件 g に入れれば $x = \pm a$ となります。よって、最大値と最小値は $(0, \pm a)$ と $(\pm a, 0)$ の地点での $\pm a^2$ と求まります。

さらについでの話として、等周問題 (isoperimetric problem) にも触れておきます。細かいことは無視していきます。汎関数 I が

$$I[y] = \int_{x_1}^{x_2} dx \ f(x, y, y') \qquad (y' = \frac{dy}{dx})$$

$$y(x_1) = y(x_2) = 0$$

となっているとして、ここに拘束条件として

$$K[y] = \int_{x_1}^{x_2} dx \ g(x, y, y') = L \qquad (y' = \frac{dy}{dx}) \tag{11}$$

を加えたときの $I[y]$ の極値を求めるのが等周問題です。拘束条件はある関数 g による積分が定数 L になるというものです。等周問題は簡単に言えば、xy 平面における面の周長 (perimeter) に条件を与えたとき、その面はどうなるのかという問題です。この拘束条件で問題になるのは、$\overline{y}$ で極値になるとして y の変分を

$$y(x) = \overline{y}(x) + \epsilon \eta(x) \ , \ \eta(x_1) = \eta(x_2) = 0$$

と与えたとき、I だけでなく拘束条件である K にも $\epsilon \eta$ が現れることです。このため、通常の手順のように η を任意とできないです。これの対処として、新しくさらに変分の項を加えることで任意性を与えます。

というわけで、η_1 は任意とし、η_2 は拘束条件を満たす関数であるとして

$$y(x) = \overline{y}(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x)$$

$$\eta_1(x_1) = \eta_1(x_2) = 0 \ , \ \eta_2(x_1) = \eta_2(x_2) = 0$$

ϵ_1, ϵ_2 は微小量です。これで η_1 が実際に任意なのかどうかの話は省き (必ず任意になるわけではない)、任意にできているとします。汎関数 I, K を ϵ_1, ϵ_2 の関数と見ることにして

$$I(\epsilon_1, \epsilon_2) = \int_{x_1}^{x_2} dx \ f(x, y, y')$$

$$K(\epsilon_1, \epsilon_2) = \int_{x_1}^{x_2} dx \ g(x, y, y') = L$$

変数が変わっていますが、同じ I, K の表記を使ってしまいます。$\epsilon_1 = \epsilon_2 = 0$ で、拘束条件を満たす I の極値となります。拘束条件は $K(\epsilon_1, \epsilon_2) - L = 0$ と書けるので

$$J(\epsilon_1, \epsilon_2) = I(\epsilon_1, \epsilon_2) - \lambda(K(\epsilon_1, \epsilon_2) - L)$$

としてみます。$\epsilon_1=\epsilon_2=0$ で拘束条件を満たす I の極値になるとしているので

$$\frac{\partial}{\partial\epsilon_1}J(\epsilon_1,\epsilon_2)\big|_{\epsilon=0}=\frac{\partial}{\partial\epsilon_1}(I-\lambda(K-L))\big|_{\epsilon=0}=0$$
$$\frac{\partial}{\partial\epsilon_2}J(\epsilon_1,\epsilon_2)\big|_{\epsilon=0}=\frac{\partial}{\partial\epsilon_2}(I-\lambda(K-L))\big|_{\epsilon=0}=0$$

$\epsilon_1=\epsilon_2=0$ を $\epsilon=0$ と省略して書いてます。ϵ_1 での偏微分は

$$\begin{aligned}\frac{\partial}{\partial\epsilon_1}J(\epsilon_1,\epsilon_2)\big|_{\epsilon=0}=&\int_{x_1}^{x_2}dx\frac{\partial}{\partial\epsilon_1}f(x,y,y')\big|_{\epsilon=0}-\lambda\int_{x_1}^{x_2}dx\frac{\partial}{\partial\epsilon_1}g(x,y,y')\big|_{\epsilon=0}\\
=&\int_{x_1}^{x_2}dx(\frac{\partial y}{\partial\epsilon_1}\frac{\partial f}{\partial y}+\frac{\partial y'}{\partial\epsilon_1}\frac{\partial f}{\partial y'})\big|_{\epsilon=0}-\lambda\int_{x_1}^{x_2}dx(\frac{\partial y}{\partial\epsilon_1}\frac{\partial g}{\partial y}+\frac{\partial y'}{\partial\epsilon_1}\frac{\partial g}{\partial y'})\big|_{\epsilon=0}\\
=&\int_{x_1}^{x_2}dx\big(\eta_1\frac{\partial f}{\partial y}+\eta_1'\frac{\partial f}{\partial y'}-\lambda(\eta_1\frac{\partial g}{\partial y}+\eta_1'\frac{\partial g}{\partial y'})\big)\big|_{\epsilon=0}\\
=&\int_{x_1}^{x_2}d\big(\eta_1\frac{\partial f}{\partial y}+\frac{d\eta_1}{dx}\frac{\partial f}{\partial y'}-\lambda(\eta_1\frac{\partial g}{\partial y}+\frac{d\eta_1}{dx}\frac{\partial g}{\partial y'})\big)\big|_{\epsilon=0}\\
=&\int_{x_1}^{x_2}dx\ \eta_1\big(\frac{\partial}{\partial y}(f-\lambda g)-\frac{d}{dx}\frac{\partial}{\partial y'}(f-\lambda g)\big)\big|_{\epsilon=0}\end{aligned}$$

よって

$$\frac{\partial}{\partial\epsilon_1}J(\epsilon_1,\epsilon_2)\big|_{\epsilon=0}=\int_{x_1}^{x_2}dx\ \eta_1(\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'})\big|_{\epsilon=0}=0\quad(F=f-\lambda g)$$

η_1 は任意なので

$$\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'}=0\tag{12}$$

として、F によるオイラー・ラグランジュ方程式になり、ラグランジュの未定乗数法と同じ結果になります。また、ϵ_2 の偏微分でも同じように

$$\frac{\partial}{\partial\epsilon_2}J(\epsilon_1,\epsilon_2)|_{\epsilon=0}=\int_{x_1}^{x_2}dx\ \eta_2(\frac{\partial F}{\partial y}-\frac{d}{dx}\frac{\partial F}{\partial y'})|_{\epsilon=0}$$

となりますが、これは (12) から 0 になるので成立しています (η_2 は任意ではない)。よって、拘束条件 (**??**) があるときの極値を与える関数 y は (12) から求まります。

等周問題の例として、ディド (Dido) の問題を扱います。x 軸上の $-a$ から $+a$ の直線と、$-a\leq x\leq a$ の範囲で与えられている長さ L の曲線 $y(x)$ があるとします。曲線の両端は $x=-a,a$ の位置として ($y(-a)=y(a)=0$)、この直線と曲線によって囲まれた面の面積を最大にするのはどのような曲線なのか求める、というのがディドの問題です。ちなみに、ディドはギリシャ神話に登場する女王で、その話の中でこの問題が出てきています。

囲まれた面の面積は

$$I[y]=\int_{-a}^{a}dx\ y(x)\quad(y(-a)=y(a)=0)$$

と与えられ ($y(x)\Delta x$ を足していった面積の極限)、これが最大になる関数 y を求めればいいです。ただし、今は曲線の長さに L という条件がついています (曲線は x 軸上の直線より長くないと囲めないので $L > 2a$)。曲線の微小な長さは

$$\sqrt{(\Delta x)^2 + (\Delta y)^2} = \Delta x\sqrt{1 + (\Delta y/\Delta x)^2}$$

これを $x = -a$ から $+a$ の範囲で足したら L になるという条件なので

$$L = \int_{-a}^{a} dx\sqrt{1 + (\frac{dy}{dx})^2}$$

これを拘束条件として (12) に入れれば

$$\begin{aligned}
&\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \\
&F = y - \lambda\sqrt{1 + y'^2} \qquad (y' = \frac{dy}{dx})
\end{aligned}$$

これの解が面積を最大にします。微分を行って、x で積分すると

$$\begin{aligned}
1 + \lambda\frac{d}{dx}\frac{y'}{\sqrt{1 + y'^2}} &= 0 \\
\lambda\frac{y'}{\sqrt{1 + y'^2}} &= -x + C_1
\end{aligned}$$

C_1 は積分定数です。変形して

$$\begin{aligned}
\lambda^2\frac{{y'}^2}{1 + y'^2} &= (x - C_1)^2 \\
(\lambda^2 - (x - C_1)^2)y'^2 &= (x - C_1)^2 \\
y'^2 &= \frac{(x - C_1)^2}{\lambda^2 - (x - C_1)^2} \\
y' &= \pm\frac{x - C_1}{\sqrt{\lambda^2 - (x - C_1)^2}}
\end{aligned}$$

x で積分して

$$y = \pm\int dx\frac{x - C_1}{\sqrt{\lambda^2 - (x - C_1)^2}} = \pm\int dz\frac{z}{\sqrt{\lambda^2 - z^2}} \qquad (z = x - C_1)$$

曲線を表す y は実数でないといけないので、$\lambda > z$ として積分を行うことにして

$$
\begin{aligned}
y =& \pm\lambda \int d\theta \frac{\sin\theta\cos\theta}{\sqrt{1-\cos^2\theta}} \qquad (z = \lambda\cos\theta,\ dz = \lambda\cos\theta d\theta) \\
=& \pm\lambda \int d\theta \cos\theta \\
=& \pm\lambda\sin\theta + C_2 \\
=& \pm\sqrt{\lambda^2 - z^2} + C_2 \\
=& \pm\sqrt{\lambda^2 - (x-C_1)^2} + C_2
\end{aligned}
$$

C_2 は積分定数です。さらに変形すると

$$
\begin{aligned}
(y-C_2)^2 = \lambda^2 - (x-C_1)^2 \\
(x-C_1)^2 + (y-C_2)^2 = \lambda^2
\end{aligned}
$$

これは xy 平面において (C_1, C_2) を中心とする半径 $|\lambda|$ の円の式です。なので、この円と x 軸で囲まれている面が最大の面積を与えます。

$y(-a) = y(a) = 0$ から

$$
(-a-C_1)^2 + C_2^2 = \lambda^2\ ,\ (a-C_1)^2 + C_2^2 = \lambda^2
$$

となるので、$C_1 = 0$ です。円の半径は $\lambda^2 = a^2 + C_2^2$ で与えられるために、円の中心が x 軸より上にいると円周の一部は $x = -a, a$ の外側にいけてしまい、曲線は $x = -a$ から a の範囲としている状況に合わなくなります。なので、C_2 を $-C_2$ $(C_2 \geq 0)$ にして

$$
x^2 + (y+C_2)^2 = \lambda^2\ ,\ a^2 + C_2^2 = \lambda^2 \tag{13}
$$

とします。円の中心 $(0, -C_2)$ と $(a, 0)$ を繋いだ直線は円の半径なので長さは $|\lambda|$ です。この直線と y 軸との間の角度を ϕ とすれば、$(-a, 0)$ から $(a, 0)$ への円弧の長さは

$$
2\pi|\lambda|\frac{2\phi}{2\pi} = 2|\lambda|\phi
$$

これが囲っている曲線の長さなので $L = 2\lambda\phi$ となり、C_2 で書くと

$$
L = 2|\lambda|\phi = 2\sqrt{a^2 + C_2^2}\arctan\frac{a}{C_2}
$$

これによって与えられた L, a に対して C_2 が決まります。C_2 が決まれば (13) から円の半径 $|\lambda|$ が分かり、面積が最大になる面が作れます。