Lagrange Multipliers and Variational Problems with Constraints

Integral Constraints. Consider the variational problem of finding the extremals for the functional

\[ J[y] = \int_a^b F(x, y, y') \, dx \]

with \( y(a) = A \) and \( y(b) = B \) subject to the additional integral constraint that

\[ K[y] = \int_a^b G(x, y, y') \, dx = \ell, \]

where \( \ell \) is a given constant. Suppose we have an extremal \( y = y(x) \) for this variational problem. To derive a necessary condition for the extremal, we embed it in a 2-parameter family \( y = y(x, s, t) \) so that the given extremal \( y = y(x) \) corresponds to \( (s, t) = (0, 0) \). The reason why we need a 2-parameter family instead of a 1-parameter family is that the family \( y = y(s, t) \) has to satisfy the integral constraint

\[ K[y] = \int_a^b G(x, y(x, s, t), \frac{\partial}{\partial x} y(x, s, t)) \, dx = \ell \]

for all \((s, t)\). Consider now \( J[y] \) as a function of two variables \((s, t)\) subject to the condition \( K[y] = \ell \). Since \( J[y] \) has a critical point at \((s, t) = (0, 0)\) subject to the condition that the point \((s, t)\) lies on the curve \( K[y] = \ell \) in the space of the two real variables \((s, t)\), it follows that the gradient of \( J[y] \) with respect to \((s, t)\) is proportional to the gradient of \( K[y] \) with respect to \((s, t)\) at the point \((s, t) = (0, 0)\). Thus there exists some real number \( \lambda \) such that

\[ \frac{\partial}{\partial s} J[y] = \lambda \frac{\partial}{\partial s} K[y], \]
\[ \frac{\partial}{\partial t} J[y] = \lambda \frac{\partial}{\partial t} K[y] \]

at the point \((s, t) = (0, 0)\). In other words, we have

\[ \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \left( \frac{\partial}{\partial s} y \right) \, dx = \lambda \int_a^b \left( G_y - \frac{d}{dx} G_{y'} \right) \left( \frac{\partial}{\partial s} y \right) \, dx, \]
\[ \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \left( \frac{\partial}{\partial t} y \right) \, dx = \lambda \int_a^b \left( G_y - \frac{d}{dx} G_{y'} \right) \left( \frac{\partial}{\partial t} y \right) \, dx, \]
The left-hand side
\[ \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \left( \frac{\partial}{\partial s} y \right) dx \]
of the first equation is the derivative of \( J \) with respect to the vector \( \frac{\partial}{\partial s} y \) in the space of functions. The left-hand side
\[ \int_a^b \left( F_y - \frac{d}{dx} F_{y'} \right) \left( \frac{\partial}{\partial t} y \right) dx \]
of the second equation is the derivative of \( J \) with respect to the vector \( \frac{\partial}{\partial t} y \) in the space of functions. The right-hand side
\[ \int_a^b \left( G_y - \frac{d}{dx} G_{y'} \right) \left( \frac{\partial}{\partial s} y \right) dx \]
of the first equation is the derivative of \( K \) with respect to the vector \( \frac{\partial}{\partial s} y \) in the space of functions. The right-hand side
\[ \int_a^b \left( G_y - \frac{d}{dx} G_{y'} \right) \left( \frac{\partial}{\partial t} y \right) dx \]
of the second equation is the derivative of \( K \) with respect to the vector \( \frac{\partial}{\partial t} y \) in the space of functions.

The constant \( \lambda \) is already determined by the first equation which says that the component of the gradient of \( J \) in the direction of the vector \( \frac{\partial}{\partial s} y \) in the space of functions is equal to \( \lambda \) times the component of the gradient of \( K \) in the direction of the vector \( \frac{\partial}{\partial s} y \) in the space of functions. The second equation says that as a result the component of the gradient of \( J \) in the direction of any other vector \( \frac{\partial}{\partial t} y \) in the space of functions is equal to the same constant \( \lambda \) times the component of the gradient of \( K \) in the direction of the vector \( \frac{\partial}{\partial t} y \) in the space of functions.

As a consequence we can say that the full gradient of \( J \) in the space of functions is equal to \( \lambda \) times the full gradient of \( K \) in the space of functions. In other words,
\[ \int_a^b \left( \left( F_y - \frac{d}{dx} F_{y'} \right) - \lambda \left( G_y - \frac{d}{dx} G_{y'} \right) \right) (\delta y) dx = 0 \]
for all $\delta y$ with $(\delta y)(a) = (\delta y)(b) = 0$. Therefore we get the Euler-Lagrange equation
\[
(*) \quad F_y - \frac{d}{dx} F_{y'} = \lambda \left( G_y - \frac{d}{dx} G_{y'} \right)
\]
for some constant $\lambda$ which is known as the Lagrange multiplier. Another way to write it is
\[
(F - \lambda G)_y - \frac{d}{dx} (F - \lambda G)_{y'} = 0.
\]
Besides the two initial conditions $y(a) = A$ and $y(b) = B$ to determine the two constant of integrations for the solution of the second-order differential equation $(*)$, we have an extra unknown $\lambda$ which will be determined by the integral constraint $\int_a^b G(x, y, y') \, dx = \ell$.

**Example.** Given $a > 0$ and $\ell > 0$. Find an extremal for the variational problem

\[
J[y] = \int_{-a}^a y \, dx
\]

subject to $y(-a) = y(a) = 0$ and

\[
K[y] = \int_{-a}^a \sqrt{1 + y'^2} \, dx = \ell.
\]

**Solution.** The Euler-Lagrange equation is
\[
\frac{\partial}{\partial y} \left( y - \lambda \sqrt{1 + y^2} \right) - \frac{d}{dx} \left( \frac{\partial}{\partial y'} \left( y - \lambda \sqrt{1 + y^2} \right) \right) = 0
\]
for some Lagrange multiplier $\lambda$. We can rewrite it as
\[
1 + \lambda \frac{d}{dx} \frac{y'}{\sqrt{1 + y^2}} = 0.
\]
Integrating it once, we get
\[
x + \lambda \frac{y'}{\sqrt{1 + y^2}} = C_1.
\]
Squaring to remove the square root, we get
\[
(x - C_1)^2 = \lambda^2 \frac{y'^2}{1 + y'^2} = \lambda \left( 1 - \frac{1}{1 + y'^2} \right).
\]
As a result,

\[
\frac{(x - C_1)^2}{\lambda^2} = 1 - \frac{1}{1 + y'^2},
\]

\[
\frac{1}{1 + y'^2} = 1 - \frac{(x - C_1)^2}{\lambda^2},
\]

\[
1 + y'^2 = \frac{1}{1 - \frac{(x - C_1)^2}{\lambda^2}},
\]

\[
y'^2 = \frac{1}{1 - \frac{(x - C_1)^2}{\lambda^2}} - 1 = \frac{(x - C_1)^2}{\lambda^2} - \frac{(x - C_1)^2}{\lambda^2},
\]

\[
y' = \pm \int \frac{x - C_1}{\lambda \sqrt{1 - \frac{(x - C_1)^2}{\lambda^2}}} \, dx.
\]

Let \(\frac{x - C_1}{\lambda} = \cos \theta\). Then

\[
y' = \pm \int \frac{\cos \theta (-\sin \theta) \lambda d\theta}{\sqrt{1 - \cos^2 \theta}} = \mp \lambda \sin \theta + C_2.
\]

Eliminating \(\theta\), we get

\[
(x - C_1)^2 + (y - C_2)^2 = \lambda^2.
\]

The constants \(C_1, C_2\), and \(\lambda\) are to be determined by \(y(-a) = y(a) = 0\) and

\[
K[y] = \int_{-a}^{a} \sqrt{1 + y'^2} \, dx = \ell.
\]

**Pointwise Constraints.** Consider the variational problem of finding the extremals for the functional

\[
J[y, z] = \int_{a}^{b} F(x, y, z, y', z') \, dx
\]

with \(y(a) = A\) and \(y(b) = B\) subject to the additional pointwise constraint that \(g(x, y, z) = 0\). Suppose we have an extremal \(y = y(x), z = z(x)\) for this variational problem. To derive a necessary condition for the extremal, we embed it in a 1-parameter family \(y = y(x, t), z = z(x, t)\) which satisfy the pointwise constraint \(g(x, y(x, t), z(x, t)) \equiv 0\) for all \(x, t\) so that the given
extremal $y = y(x), z = z(x)$ corresponds to $t = 0$. Taking the first variation of $J$, we get

$$
(\dagger) \quad \delta J = \int_a^b \left( F_y - \frac{d}{dx} F'_y \right) (\delta y) \, dx + \left( F_z - \frac{d}{dx} F'_z \right) (\delta z) \, dx,
$$

where $\delta J = \left. \frac{dJ}{dt} \right|_{t=0}$ and $\delta y = \left. \frac{\partial y}{\partial t} \right|_{t=0}$ and $\delta z = \left. \frac{\partial z}{\partial t} \right|_{t=0}$. Differentiating $g(x, y(x, t), z(x, t)) \equiv 0$ with respect to $t$ yields

$$
g_y(\delta y) + g_z(\delta z) \equiv 0
$$

for all $x$. Solving for $\delta z$, we get

$$
\delta z = - \frac{g_y}{g_z} (\delta y).
$$

Putting it into (\dagger) gives us

$$
0 = \delta J = \int_a^b \left( F_y - \frac{d}{dx} F'_y \right) (\delta y) \, dx.
$$

By the arbitrariness of $\delta y$, we conclude that the integrand vanishes and get

$$
F_y - \frac{d}{dx} F'_y - \frac{g_y}{g_z} \left( F_z - \frac{d}{dx} F'_z \right) = 0
$$

which can be rewritten as

$$
\frac{F_y - \frac{d}{dx} F'_y}{g_y} = \frac{F_z - \frac{d}{dx} F'_z}{g_z}.
$$

Let $-\lambda(x)$ be either of the two equal sides of the above equation. We get

$$
(F - \lambda g)_y - \frac{d}{dx} (F - \lambda g)_y' = 0,
$$

$$
(F - \lambda g)_z - \frac{d}{dx} (F - \lambda g)_z' = 0.
$$

Heuristically, we can also interpret the pointwise condition as a 1-parameter family of integral constraint

$$
\int_a^b (\delta \xi)(x) g(x, y, z) \, dx = 0
$$

with parameter $\xi \in [a, b]$, where $\delta \xi(x)$ is the Dirac delta at $\xi$ so that we end up with one constant multiplier $\lambda(\xi)$ for each of such an integral constraint.