Second Variation. One-variable problem

February 13, 2013

Contents

Stationary conditions point to a possibly optimal trajectory but they do not state that the trajectory corresponds to the minimum of the functional. A stationary solution can correspond to minimum, local minimum, maximum, local maximum, of a saddle point of the functional. In this chapter, we establish methods aiming to distinguish local minimum from local maximum or saddle. In addition to being a solution to the Euler equation, the true minimizer satisfies necessary conditions in the form of inequalities. We introduce variational tests, Weierstrass and Jacobi conditions, that supplement each other.

The conclusion of optimality of the tested stationary curve $u(x)$ is based on a comparison of the problem costs $I(u)$ and $I(u+\delta u)$ computed at u and any close-by admissible curve $u+\delta u$. The closeness of admissible curve is required to simplify the calculation and obtain convenient optimality conditions. The question whether or not two curves are close to each other or whether $v(x)$ is small depends on what curves we consider to be close. Below, we work out three tests of optimality using different definitions of closeness.

1 Local variations

1.1 Legendre Test

Consider again the simplest problem of the calculus of variations

$$\min_{u(x),x\in[a,b]} I(u), \quad I(u)=\int_a^b F(x,u,u')dx, \quad u(a)=u_a,\ u(b)=u_b \tag{1}$$

and function $u(x)$ that satisfies the Euler equation and boundary conditions,

$$\frac{\partial F}{\partial u}-\frac{d}{dx}\frac{\partial F}{\partial u'}=0, \quad u(a)=u_a,\ u(b)=u_b, \tag{2}$$

so that the first variation δI is zero.

Let us compute the increment $\delta^2 I$ of the objective caused by the variation

$$\delta u(x,x_0)=\begin{cases}\epsilon^2\phi\left(\frac{x-x_0}{\epsilon}\right) & \text{if } |x-x_0|<\epsilon\\ 0 & \text{if } |x-x_0|\geq\epsilon\end{cases} \tag{3}$$

where $\phi(x)$ is a function with the following properties:

$$\phi(-1)=\phi(1)=0, \quad \max_{x\in[-1,1]}|\phi(x)|\leq 1, \quad \max_{x\in[-1,1]}|\phi'(x)|\leq 1 \tag{4}$$

The magnitude of this *Legendre-type* variation tends to zero when $\epsilon\to 0$, and the magnitude of its derivative

$$\delta u'(x,x_0)=\begin{cases}-\epsilon\,\phi'\left(\frac{x-x_0}{\epsilon}\right) & \text{if } |x-x_0|<\epsilon\\ 0 & \text{if } |x-x_0|\geq\epsilon\end{cases}$$

tends to zero as well. Additionally, the variation is *local*: it is zero outside of the interval of the length 2ϵ. We use these features of the variation in the calculation of the increment of the cost.

Expanding F into Taylor series and keeping the quadratic terms, we obtain

$$\delta I = I(u+\delta u) - I(u) = \int_a^b (F(x, u+\delta u, u'+\delta u') - F(x,u,u'))dx$$
$$= \int_a^b \left(\left[\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u'} \right] \delta u + A\delta u^2 + 2B\delta u\, \delta u' + C(\delta u')^2 \right) dx + \left. \frac{\partial F}{\partial u'} \right|_{x=a}^{x=b}, \quad (5)$$

where

$$A = \frac{\partial^2 F}{\partial u^2}, \quad B = \frac{\partial^2 F}{\partial u \partial u'}, \quad C = \frac{\partial^2 F}{\partial (u')^2}$$

and all derivatives are computed at the point x_0 at the optimal trajectory $u(x)$. The term in the brackets in the integrand in the right-hand side of (5) is zero because the Euler equation is satisfied. Let us estimate the remaining terms

$$\int_a^b A(x)(\delta u)^2 dx = \int_{x_0-\varepsilon}^{x_0+\varepsilon} A(x)(\delta u)^2 dx$$
$$\leq \varepsilon^4 \int_{x_0-\varepsilon}^{x_0+\varepsilon} A(x)dx = A(x_0)\,\varepsilon^5 + o(\varepsilon^5)$$

Indeed, the variation δu is zero outside of the interval $[x-\varepsilon, x+\varepsilon]$, has magnitude of the order of ε^2 in this interval, and A(x) is assumed to be continuous at the trajectory. Similarly, we estimate

$$\int_a^b B(x)\delta u\, \delta u'\, dx \leq \varepsilon^3 \int_{x_0-\varepsilon}^{x_0+\varepsilon} B(x)dx = B(x_0)\,\varepsilon^4 + o(\varepsilon^4)$$
$$\int_a^b C(x)(\delta u')^2 dx \leq \varepsilon^2 \int_{x_0-\varepsilon}^{x_0+\varepsilon} C(x)dx = C(x_0)\,\varepsilon^3 + o(\varepsilon^3)$$

Its derivative's magnitude $\delta u'$ is of the order of ε, therefore $|\delta u'| \gg |\delta u|$ as $\varepsilon \to 0$; we conclude that the last term in the integrand in the right-hand side of (5) dominates. The inequality $\delta I > 0$ implies inequality

$$\frac{\partial^2 F}{\partial (u')^2} \geq 0 \quad (6)$$

which is called *Legendre condition* or *Legendre test.*

Remark 1.1 Here, it is assumed that $\frac{\partial^2 F}{\partial (u')^2} \neq 0$. If $\frac{\partial^2 F}{\partial (u')^2} = 0$, the Legendre test is inconclusive, and more sophisticated and sensitive variations most be used. An example is Kelly variation

$$v = \varepsilon^2 \begin{cases} \phi\left(\frac{x-x_0-\varepsilon}{\varepsilon}\right) & \text{if } x \in [x_0 - 2\varepsilon, x_0) \\ \phi\left(\frac{x-x_0+\varepsilon}{\varepsilon}\right) & \text{if } x \in [x_0 + 2\varepsilon, x_0) \\ 0 & \text{elsewhere.} \end{cases}$$

The corresponding condition for the minimum is []

$$\frac{d}{dx^2}\frac{\partial^2 F}{\partial (u')^2} \leq 0$$

It is obtained by the same method. This time, four terms in the Taylor expansion is kept.

Example 1.1 Legendre test is always satisfied in the geometric optics. The Lagrangian depends on the derivative as $F = \frac{\sqrt{1+y'^2}}{v(y)}$ and its second derivative

$$\frac{\partial^2 F}{\partial y'^{\,2}} = \frac{1}{v(y)(1+y'^2)^{\frac{3}{2}}}$$

is always nonnegative if $v > 0$. It is physically obvious that the fastest path is stable to short-term perturbations.

Example 1.2 Legendre test is also always satisfied in the Lagrangian mechanics. The Lagrangian $F = T - V$ depends on the derivatives of the generalized coordinates through the kinetic energy $T = \frac{1}{2}\dot{q}R(q)\dot{q}$ and its Hessian

$$\frac{\partial^2 F}{\partial q'^{\,2}} = R$$

is equal to the generalized inertia R which is always positive definite.

Physically speaking, inertia does not allow for infinitesimal oscillations because they always increase the kinetic energy while potential energy is insensitive to them.

Example 1.3 (Two-well Lagrangian) Consider the Lagrangian

$$F(u, u') = [(u')^2 - u^2]^2$$

for a simples variational problem with fixed boundary data $u(0) = a_0,\ u(1) = a_1$. The Legendre test is satisfied is the inequality is valid:

$$\frac{\partial^2 F}{\partial (u')^2} = 4(3u'^2 - u^2) \geq 0.$$

Consequently, the solution u of Euler equation

$$[(u')^3 - u^2 u']' + u(u')^2 - u^3 = 0, \quad u(0) = a_0,\ u(1) = a_1 \tag{7}$$

might correspond to a local minimum of the functional if, in addition, the inequality $3u'^2 - u^2 \geq 0$ is satisfied in all points $x \in (0, 1)$. Later we show how to transform (relax) the problem, if its solution does not satisfy this condition.

1.2 Weierstrass Test

The Weierstrass test detects optimality the trajectory by checking its stability against strong local perturbations. It is also local: it compares trajectories that coincide everywhere except a small interval where their derivatives significantly differ.

Suppose that u is the minimizer of the variational problem (1) that satisfies the Euler equation (2). Consider a variation that is shaped as an infinitesimal triangle supported on the interval $[x_0, x_0 + \varepsilon]$, where $x_0 \in (a,\, b)$ (see **??**):

$$\Delta u(x) = \begin{cases} 0 & \text{if} \;\; x \notin [x_0, x_0 + \varepsilon], \\ v_1(x - x_0) & \text{if} \;\; x \in [x_0, x_0 + \alpha\varepsilon], \\ v_2(x - x_0) - \alpha\varepsilon(v_1 - v_2) & \text{if} \;\; x \in [x_0 + \alpha\varepsilon, x_0 + \varepsilon] \end{cases}$$

where v_1 and v_2 are two real numbers and $\alpha \in (0, 1)$. Parameters α ($0 < \alpha < 1$), v_1 and v_2 are related

$$\alpha v_1 + (1 - \alpha)v_2 = 0 \tag{8}$$

to provide the continuity of $u + \Delta u$ at the point $x_0 + \varepsilon$, or equality

$$\Delta u(x_0 + \varepsilon - 0) = 0.$$

Condition (8) can be rewritten as

$$v_1 = (1 - \alpha)v, \quad v_2 = -\alpha v = 0, \tag{9}$$

where v is an arbitrary real number.

The considered variation (the Weierstrass variation) is localized and has an infinitesimal absolute value (if $\varepsilon \to 0$), but, unlike the Legendre variation, its derivative $(\Delta u)'$ is finite:

$$(\Delta u)' = \begin{cases} 0 & \text{if} \;\; x \notin [x_0, x_0 + \varepsilon], \\ v_1 & \text{if} \;\; x \in [x_0, x_0 + \alpha\varepsilon], \\ v_2 & \text{if} \;\; x \in [x_0 + \alpha\varepsilon, x_0 + \varepsilon]. \end{cases} \tag{10}$$

The increment is

$$\delta I = I(u + \delta u) - I(u) = \int_a^b (F(x, u + \delta u, u' + \Delta u') - F(x, u, u'))dx \tag{11}$$

is computed by splitting the first term in the integrant into two parts

$$\delta I = \int_{x_0}^{x_0+\alpha\epsilon} F(x, u + \delta u, u' + v_1)dx + \int_{x_0+\alpha\epsilon}^{x_0-\epsilon} F(x, u + \delta u, u' + v_2)dx - \int_{x_0}^{x_0+\epsilon} F(x, u, u')dx \tag{12}$$

and rounding integrands up to ε as follows

$$F(x, u(x) + \Delta u, u(x)' + v_1) = F(x_0, u(x_0), u'(x_0) + v_1) + O(\epsilon), \quad \forall x \in [x_0, x_0 + \epsilon]$$

and similarly for $F(x, u(x) + \Delta u, u(x)' + v_2)$. The smallness of the variation δu follows from to smallness $[x_0, x_0 + \epsilon]$ of the interval of the variation and finiteness of the variation of the derivative there.

With these simplifications, we compute the main term of the increment as

$$\delta I(u, x_0) = \\ \varepsilon[\alpha F(x_0, u, u' + v_1) + (1-\alpha) F(x_0, u, u' + v_2) - F(x_0, u, u')] + o(\varepsilon) \quad (13)$$

Repeating the variational arguments and using the arbitrariness of x_0, we find that an inequality holds

$$\delta I(u, x) \geq 0 \quad \forall x \in [a, b] \quad (14)$$

for a minimizer u. The last expression results in the Weierstrass necessary condition.

Any minimizer $u(x)$ of (1) satisfies the inequality

$$\alpha F(x, u, u' + (1-\alpha)v) + (1-\alpha) F(x, u, u' - \alpha v) - F(x, u, u') \geq 0 \\ \forall v, \ \forall \alpha \in [0, 1] \quad (15)$$

The reader may recognize in this inequality the definition of convexity, or the condition that the graph of the function $F(., ., z)$ (considered as a function of the third argument $z = u'$) lies below the the chord supported by points $z_1 = u' + (1-\alpha)v$ and $z_2 = u' - \alpha v$ in the interval $[z_1, z_2]$ between there points.

The Weierstrass condition requires *convexity of the Lagrangian* $F(x, y, z)$ *with respect to its third argument* $z = u'$. The first two arguments $x, \ = u$ are determined from the equation of the tested trajectory. Recall that the tested minimizer $u(x)$ is a solution to the Euler equation.

The Weierstrass test is stronger than the Legendre test because convexity implies nonnegativity of the second derivative. It compares the optimal trajectory with larger set of admissible trajectories.

Example 1.4 (Two-well Lagrangian. II) Consider again the Lagrangian discussed in Example 1.3

$$F(u, u') = \left((u')^2 - u^2\right)^2$$

$F(u, v)$ is convex as a function of v if $|v| \geq |u|$. Indeed, take

$$v_1 = u - u', \quad v_1 = -u - u', \quad \alpha = \frac{u + u'}{2u}$$

and apply formula (13). We have $F(u, u' + v_1) = F(u, u' + v_2) = 0$ and

$$\alpha F(u, u' + v_1) + (1-\alpha) F(u, u' + v_2) - F(u, u') < 0, \text{ if } \alpha \in [0, 1] \text{ or } u' \in [-u, u]$$

The Weierstrass test $u'^2 \geq u^2$ is stronger than Legendre test, $u'^2 \geq \frac{1}{3} u^2$ The stationary solution u (see (7)) may correspond to a local minimum of the functional if, the inequality $|u'(x)| \geq |u(x)|$ is satisfied in all points $x \in (0, 1)$.

Figure 1: The construction of Weierstrass $\mathcal{E}$-function. The graph of a convex function and its tangent plane.

Weierstrass $\mathcal{E}$-function Weierstrass suggested a convenient test for convexity of Lagrangian, the so-called $\mathcal{E}$-function equal to the difference between the value of Lagrangian $L(x,u,\hat{z})$ in a trial point $u, z = z'$ and the tangent hyperplane $L(x,u,u') - (\hat{z}-u')^T \frac{\partial L(x,u,u')}{\partial u'}$ to the optimal trajectory at the point u, u':

$$\mathcal{E}L(x,u,u',\hat{z}) = L(x,u,\hat{z}) - L(x,u,u') - (\hat{z}-u')\frac{\partial L(x,u,u')}{\partial u'} \tag{16}$$

Function $\mathcal{E}L(x,u,u',\hat{z})$ vanishes together with the derivative $\frac{\partial \mathcal{E}(L)}{\partial \hat{z}}$ when $\hat{z} = u'$:

$$\mathcal{E}L(x,u,u',\hat{z})|_{\hat{z}=u'} = 0, \quad \frac{\partial}{\partial \hat{z}}\mathcal{E}(L(x,u,u',\hat{z})|_{\hat{z}=u'} = 0.$$

According to the basic definition of convexity, the graph of a convex function is greater than or equal to a tangent hyperplane. Thereafter, the Weierstrass condition of minimum of the objective functional can be written as the condition of positivity of the Weierstrass $\mathcal{E}$-function for the Lagrangian,

$$\mathcal{E}(L(x,u,u',\hat{z}) \geq 0 \quad \forall \hat{z}, \ \forall x, u(x)$$

where $u(x)$ is the tested trajectory.

Example 1.5 Check the optimality of Lagrangian

$$L = u'^4 - \phi(u,x)u'^2 + \psi(u,x)$$

where ϕ and ψ are some functions of u and x using Weierstrass $\mathcal{E}$-function.

The Weierstrass $\mathcal{E}$-function for this Lagrangian is

$$\begin{aligned}&\mathcal{E}L(x,u,u',\hat{z}) = \left[\hat{z}^4 - \phi(u,x)\hat{z}^2 + \psi(u,x)\right] \\ &- \left[u'^4 - \phi(u,x)u'^2 + \psi(u,x)\right] - (\hat{z}-u')(4u'^3 - 2\phi(u,x)u).\end{aligned}$$

or

$$\mathcal{E}L(x,u,u',\hat{z}) = (\hat{z}-u')^2\left(\hat{z}^2 + 2\hat{z}u' - \phi + 3u'^2\right).$$

As expected, $\mathcal{E}L(x,u,u',\hat{z})$ is independent of an additive term ψ and contains a quadratic coefficient $(\hat{z}-u')^2$. It is positive for any trial function $\hat{z}$ if the quadratic

$$\pi(\hat{z}) = -\hat{z}^2 - 2u'\,\hat{z} + (\phi - 3u'^2)$$

does not have real roots, or if discriminant is negative:

$$4(u')^2 - \phi(u,x) \leq 0$$

If this condition is violated at a point of an optimal trajectory $u(x)$, the trajectory is nonoptimal.

1.3 Vector-Valued Minimizer

Legendre test The Legendre and Weierstrass conditions and can be naturally generalized to the problem with the vector-valued minimizer. If the Lagrangian is twice differentiable function of the vector $u' = z$, the Legendre condition becomes

$$He(F, z) \geq 0 \tag{17}$$

(see Section **??**) where $He(F, z)$ is the Hessian

$$He(F, z) = \begin{pmatrix} \frac{\partial^2 F}{\partial z_1^2} & \dots & \frac{\partial^2 F}{\partial z_1 \partial z_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 F}{\partial z_1 \partial z_n} & \dots & \frac{\partial^2 F}{\partial z_n^2} \end{pmatrix}$$

and inequality in (17) means that the matrix is nonnegative definite (all its eigenvalues are nonnegative). The Weierstrass test requires convexity of $F(x, y, z)$ with respect to the third vector argument. If the minimal eigenvalue of H is zero, more complex variations are needed to select minimum.

Weierstrass test To derive Weierstrass test for ?????? ?????????, consider the variation of the type

$$\delta u(x, x_0) = \begin{cases} \epsilon\phi\left(\frac{x-x_0}{\epsilon}\right) & \text{if } |x - x_0| < \epsilon \\ 0 & \text{if } |x - x_0| \geq \epsilon \end{cases} \tag{18}$$

where $\phi(-1) = \phi(1) = 0$, $\max_{x\in[-1,1]} |\phi(x)| \leq 1$. Its derivative

$$\delta u'(x, x_0) = \begin{cases} \phi'\left(\frac{x-x_0}{\epsilon}\right) & \text{if } |x - x_0| < \epsilon \\ 0 & \text{if } |x - x_0| \geq \epsilon \end{cases}$$

is finite and its magnitude is independent of ϵ in the interval $|x - x_0| < \epsilon$. Let us call $v(x) = \phi'\left(\frac{x-x_0}{\epsilon}\right)$. Notice that

$$\int_{x_0-\epsilon}^{x_0+\epsilon} v(x)\, dx = 0 \tag{19}$$

The perturbed function is approximated as

$$F\left(x, u(x) + \delta u(x), u'(x) + v(x)\right) = F\left(x_0, u(x_0), u'(x_0) + v(x)\right) + O(\epsilon)$$

The main term of the variation

$$\Delta I = \epsilon\left[\frac{1}{2\epsilon}\int_{x_0-\epsilon}^{x_0+\epsilon} F\left(x_0, u(x_0), u'(x_0) + v(x)\right) dx - F\left(x_0, u(x_0), u'(x_0)\right)\right]$$

is positive for all x_0, if $u(x)$ is a minimizer.

Let us find and optimal (most sensitive) variation $v_{\text{opt}}(x)$

$$\Delta I(v_{\text{opt}}) = \min_{v(x) \text{ as in } (19)} \Delta I \quad \forall\, x_0$$

For briefness, we call

$$\Phi\left(u'(x_0)+v(x)\right)=\left(F\left(x_0,u(x_0),u'(x_0)+v(x)\right)\right)$$

Recall that convex envelope $\mathcal{C}\Phi(u')$ of a function $\Phi(u')$ and the point $x=x_0$ is defined as:

$$\mathcal{C}\Phi(u'(x_0))=\min_{v(x)\text{ as in (19)}}\frac{1}{2\epsilon}\int_{x_0-\epsilon}^{x_0+\epsilon}\Phi\left(u'(x_0)+v(x)\right)dx$$

We see that increment $\Delta I_{(\mathrm{opt})}$ is the difference between the convex envelope $\mathcal{C}\Phi(u'(x_0))$ and function $\Phi(u'(x_0))$ itself. By Caratheodory theorem, the convex envelope of Φ is supported by no more that $n+1$ points and is equal to:

$$\mathcal{C}\Phi(u'(x_0))=\min_{\{\alpha,v\}\in\mathcal{M}}\sum_{i=1}^{n+1}\alpha_i\Phi(u'(x_0)+\alpha_i v_i)$$

where

$$\mathcal{M}=\{\alpha,v\}:\ \left\{\sum_{i=1}^{n+1}\alpha_i=1,\ \sum_{i=1}^{n+1}v_i=0\ \alpha_i\geq 0\right\}.$$

Returning to original notations, we conclude that

1. F is convex with respect at the point $u'(x_0)$, then optimal $v(x)$ is zero, $v_{\mathrm{opt}}(x)=0$, and $\Delta_{\mathrm{opt}}I=0$. The extremal u satisfies the Weierstrass test in the point $x=x_0$. If F is nonconvex, then $\Delta_{\mathrm{opt}}I\leq 0$ and the trajectory fails the test and is non optimal.

2. The most sensitive Weierstrass variation is a continuous piece-wise linear function with the piece-wise constant slope that vanishes at x_0 and $x_0+\epsilon$. Only the values of its derivative and measures of the intervals of the constancy affect the increment.

Remark 1.2 Convexity of the Lagrangian does not guarantee the existence of a solution to a variational problem. It states only that a differentiable minimizer (if it exists) is optimal with fine-scale perturbations. However, the minimum may not exist at all or be unstable to other variations.

If the solution of a variational problem fails the Weierstrass test, then its cost can be decreased by adding infinitesimal centered wiggles to the solution. The wiggles are the Weierstrass trial functions, which decrease the cost. In this case, we call the variational problem ill-posed, and say that the solution is unstable against fine-scale perturbations.

1.4 Null-Lagrangians and convexity

Find the Lagrangian cannot be uniquely reconstructed from its Euler equation. Similarly to antiderivative, it is defined up to some term called null-Lagrangian.

Definition 1.1 The Lagrangians $\phi(x, u, u')$ for which the operator $S(\phi, u)$ of the Euler equation (**??**) identically vanishes

$$S(\phi, u) = 0 \quad \forall u$$

are called *Null-Lagrangians*.

Null-Lagrangians in variational problems with one independent variable are linear functions of u'. Indeed, the Euler equation is a second-order differential equation with respect to u:

$$\frac{d}{dx}\left(\frac{\partial}{\partial u'}\phi\right) - \frac{\partial}{\partial u}\phi = \frac{\partial^2 \phi}{\partial (u')^2} \cdot u'' + \frac{\partial^2 \phi}{\partial u' \partial u} \cdot u' + \frac{\partial^2 \phi}{\partial u \partial x} - \frac{\partial \phi}{\partial u} \equiv 0. \qquad (20)$$

The coefficient of u'' is equal to $\frac{\partial^2 \phi}{\partial (u')^2}$. If the Euler equation holds identically, this coefficient is zero, and therefore $\frac{\partial \phi}{\partial u'}$ does not depend on u'. Hence, ϕ linearly depends on u':

$$\begin{aligned} \phi(x, u, u') &= u' \cdot A(u, x) + B(u, x); \\ A = \tfrac{\partial^2 \phi}{\partial u' \partial u}&, \quad B = \tfrac{\partial^2 \phi}{\partial u \partial x} - \tfrac{\partial \phi}{\partial u}. \end{aligned} \qquad (21)$$

Additionally, if the following equality holds

$$\frac{\partial A}{\partial x} = \frac{\partial B}{\partial u}, \qquad (22)$$

then the Euler equation vanishes identically. In this case, ϕ is a null-Lagrangian.

We notice that the Null-Lagrangian (21) is simply a full differential of a function $\Phi(x, u)$:

$$\phi(x, u, u') = \frac{d}{dx}\Phi(x, u) = \frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial u} u';$$

equations (22) are the integrability conditions (equality of mixed derivatives) for Φ. The vanishing of the Euler equation corresponds to the Fundamental theorem of calculus: The equality

$$\int_a^b \frac{d\Phi(x, u)}{dx} dx = \Phi(b, u(b)) - \Phi(a, u(a)).$$

that does not depend on $u(x)$ only on its end-points values.

Example 1.6 Function $\phi = u\, u'$ is the null-Lagrangian. Indeed,we check

$$\frac{d}{dx}\left(\frac{\partial}{\partial u'}\phi\right) - \frac{\partial}{\partial u}\phi = u' - u' \equiv 0.$$

Null-Lagrangians and Convexity The convexity requirements of the Lagrangian F that follow from the Weierstrass test are in agreement with the concept of null-Lagrangians (see, for example [?]).

Consider a variational problem with the Lagrangian F,

$$\min_u \int_0^1 F(x,u,u')dx.$$

Adding a null-Lagrangian ϕ to the given Lagrangian F does not affect the Euler equation of the problem. The family of problems

$$\min_u \int_0^1 \left(F(x,u,u') + t\phi(x,u,u')\right) dx,$$

where t is an arbitrary number, corresponds to the same Euler equation. Therefore, each solution to the Euler equation corresponds to a family of Lagrangians $F(x,u,z) + t\phi(x,u,z)$, where t is an arbitrary real number. In particular, a Lagrangian cannot be uniquely defined by the solution to the Euler equation.

The stability of the minimizer against the Weierstrass variations should be a property of the Lagrangian that is independent of the value of the parameter t. It should be a common property of the family of equivalent Lagrangians. On the other hand, if $F(x,u,z)$ is convex with respect to z, then $F(x,u,z)+t\phi(x,u,z)$ is also convex. Indeed, $\phi(x,u,z)$ is linear as a function of z, and adding the term $t\phi(x,u,z)$ does not affect the convexity of the sum. In other words, convexity is a characteristic property of the family. Accordingly, it serves as a test for the stability of an optimal solution.

2 Nonlocal conditions

2.1 Distance on a sphere: Columbus problem

This simple example illustrates the use of second variation without a single calculation. We consider the problem of geodesics (shortest path) on a sphere.

Stationarity Let us prove that a geodesics is a part of the great circle. Suppose that geodesics is a different curve, or that it exists an arc C, C' that is a part of the geodesics but does not coincide with the arc of the great circle. Let us perform a variation: Replace this arc with its mirror image – the reflection across the plane that passes through the ends C, C' of this arc and the center of the sphere. The reflected curve has the same length of the path and it lies on the sphere, therefore the new path remains a geodesics. On the other hand, the new path is broken in two points C and C' , and therefore cannot be the shortest path. Indeed, consider a part of the path in an infinitesimal circle around the point C of breakage and fix the points A and B where the path crosses that circle. This path can be shorten by a arc of a great circle that passes through the points A and B. To demonstrate this, it is enough to imagine a human-size scale on Earth: The infinitesimal part of the round surface becomes flat and obviously the shortest path correspond to a straight line and not to a zigzag line with an angle.

Second variations The same consideration shows that the length of geodesics is no larger than π times the radius of the sphere, or it is shorter than the great semicircle. Indeed, if the length of geodesics is larger than the great semicircle one can fix two opposite points – the poles of the sphere – on the path and turn on an arbitrary angle the axis the part of geodesics that passes through these points. The new path lies on the sphere, has the same length as the original one, and is broken at the poles, thereby its length is not minimal. We conclude that the minimizer does not satisfy Jacobi test if the length of geodesics is larger than π times the radius of the sphere. Therefore, *geodesics on a sphere is a part of the great circle that joins the start and end points and which length is less that a half of the equator's length.*

Remark 2.1 The argument that the solution to the problem of shortest distance on a sphere bifurcates when its length exceeds a half of the great circle was famously used by Columbus who argued that the shortest way to India passes through the Western route. As we know, Columbus wasn't be able to prove or disprove the conjecture because he bumped into American continent discovering New World for better and for worse.

2.2 Sufficient condition for the weak local minimum

We assume that a trajectory $u(x)$ satisfies the stationary conditions and Legendre condition. We investigate the increment caused by a nonlocal variation δu of an infinitesimal magnitude:

$$|v| < \varepsilon, \quad |v'| < \varepsilon, \quad \text{variation interval is arbitrary.}$$

To compute the increment, we expand the Lagrangian into Taylor series keeping terms up to $O(\epsilon^2)$. Recall that the linear of ϵ terms are zero because the Euler equation $S(u, u') = 0$ for $u(x)$ holds. We have

$$\delta I = \int_0^r S(u, u')\delta u \, dx + \int_0^r \delta^2 F dx + o(\epsilon^2) \tag{23}$$

where

$$\delta^2 F = \frac{\partial^2 F}{\partial u^2}(\delta u)^2 + 2\frac{\partial^2 F}{\partial u \partial u'}(\delta u)(\delta u') + \frac{\partial^2 F}{\partial (u')^2}(\delta u')^2 \tag{24}$$

Because the variation is nonlocal, we cannot neglect v in comparison with v'.

No variation of this kind can improve the stationary solution if the quadratic form

$$Q(u, u') = \begin{pmatrix} \frac{\partial^2 F}{\partial u^2} & \frac{\partial^2 F}{\partial u \partial u'} \\ \frac{\partial^2 F}{\partial u \partial u'} & \frac{\partial^2 F}{\partial (u')^2} \end{pmatrix}$$

is positively defined,

$$Q(u, u') > 0 \quad \forall x \text{ on the stationary trajectory } u(x) \tag{25}$$

This condition is called the *sufficient condition for the weak minimum.* It neglects the relation between δu and $\delta u'$ and treats them as two independent trial functions. If the sufficient condition is satisfied, no trajectory that is smooth and sufficiently close to the stationary trajectory can increase the objective functional of the problem compared with the objective at that tested stationary trajectory.

Notice that the first term $\frac{\partial^2 F}{\partial u'^2}$ is nonnegative because of the Legendre condition.

Problem 2.1 Show that the sufficient condition is satisfied for the Lagrangians

$$F_1 = \frac{1}{2}u^2 + \frac{1}{2}(u')^2 \text{ and } F_2 = \frac{1}{|u|}(u')^2$$

If the sufficient condition is not satisfied, we try to create a variation that improves the stationary solution. In the next sections, we examine two possibilities: a straightforward arbitrary construction of δu and investigation of the increment (23) using (24) to compute the increment, or finding of an optimal shape of such variation (Jacobi condition).

2.3 Nonlocal variations

Here we consider a nonlocal variation of small magnitude. This variation compliment Weierstrass test. As before, our goal is to find whether a particular variation decreases the cost functional below its stationary value. The technique is shown using a simplest example.

Consider the following problem:

$$I = \min_u \int_0^r \left(\frac{1}{2}(u')^2 - \frac{c^2}{2}u^2 \right) dx \quad u(0) = 0; \quad u(r) = A \tag{26}$$

where $c > 0$ is a constant. The first variation δI i

$$\delta I = \int_0^r \left(u'' + c^2 u \right) \delta u \, dx$$

s zero if $u(x)$ satisfies the Euler equation (that turns out to be the equation of oscillator)

$$u'' + c^2 u = 0, \quad u(0) = 0, u(r) = A. \tag{27}$$

The stationary solution $u(x)$ is

$$u(x) = \left(\frac{A}{\sin(cr)} \right) \sin(cx)$$

The Weierstrass test is satisfied, because the dependence of the Lagrangian on the derivative u' is convex, $\frac{\partial^2 L}{\partial u'^2} = 1$. The sufficient condition of local minimum is not satisfied because $\frac{\partial^2 L}{\partial u^2} = -c^2$.

Let us show that the stationary condition does not correspond to a minimum of I is the interval's length r is large enough. We simply demonstrate a variation that improves the stationary trajectory by decreasing cost of the problem. Compute the second variation (24):

$$\delta^2 I = \int_0^r \left(\frac{1}{2}(\delta u')^2 - \frac{c^2}{2}(\delta u)^2 \right) dx \tag{28}$$

Since the boundary conditions at the ends of the trajectory are fixed, the variation δu satisfies homogeneous conditions $\delta u(0) = \delta u(r) = 0$.

Let us choose the variation as follow:

$$\delta u = \begin{cases} \epsilon x(a-x), & 0 \le x \le a \\ 0 & x > a \end{cases}$$

where the interval of variation $[0, a]$ is not greater that $[0, r]$, $a \le r$. Computing the second variation of the goal functional from (28), we obtain

$$\delta^2 I(a) = \frac{\epsilon^2}{60} a^3 (10 - c^2 a^2), \quad a \le r$$

The increment $\delta^2 I$ is positive only if

$$a < r_{\text{crit}}, \quad r_{\text{crit}} = \frac{\sqrt{10}}{c} = \frac{3.16227}{c}$$

The most dangerous variation corresponds to the maximal value $a = r$. This increment is negative when r is sufficiently large,

$$r > r_{\text{crit}}.$$

In this case $\delta^2 I(a)$ is negative, $\delta^2 I(a) < 0$ We conclude that the stationary solution does not correspond to the minimum of I if the length of the trajectory is larger than r_{crit}.

If the length is smaller than r_{crit}, the situation is inconclusive. It could still be possible to choose another type of variation different from considered here that disproves the optimality of the stationary solution.

The general case is considered in the same manner. To examine a stationary solution $u(x)$, one chooses a nonlocal variation δu with the conditions $\delta u(\alpha) = \delta u(\beta) = 0$, where $\alpha \in [0, r]$ and $\beta \in [0, r]$ and compute the integral of the expression (23).

$$\delta^2 I = \int_0^r \delta^2 F \, dx$$

If we succeed to find a variation that makes $\delta^2 I$ negative, the stationary solution does not correspond to a minimum.

2.4 Jacobi variation

The Jacobi necessary condition chooses the most sensitive long and shallow variation and examines the increment caused by such variation. It complements the Weierstrass test that investigates stability of a stationary trajectory to strong localized variations. Jacobi condition tries to disprove optimality by testing stability against "optimal" nonlocal variations with small magnitude.

Assume that a trajectory $u(x)$ satisfies the stationary and Legendre conditions but does not satisfy the sufficient conditions for weak minimum, that is $Q(u, u')$ in (25) is not positively defined,

$$S(u, u') = 0, \quad \frac{\partial^2 F}{\partial (u')^2} > 0, \quad Q(u, u') \not> 0$$

To derive Jacobi condition, we consider an infinitesimal nonlocal variation: $\delta u = O(\epsilon) \ll 1$ and $\delta u' = O(\epsilon) \ll 1$ and examine the expression (24) for the second variation. When an infinitesimal nonlocal variation is applied, the increment increases because of assumed positivity of $\frac{\partial^2 F}{\partial (u')^2}$ and decreases because of assumed nonpositivity of the matrix Q. Depending on the length r of the interval of integration and of chosen form of the variation δu, one of these effects prevails. If the second effect is stronger, the extremal fails the test and is nonoptimal.

Jacobi conditions asks for the choice of the best δu of the variation. The expression (24) itself is a variational problem for δu which we rename here as v for short; the Lagrangian is quadratic of v and v' and the coefficients are functions of x determined at the stationary trajectory $u(x)$ which is assumed to be known:

$$\delta^2 I = \int_0^r \left[A v^2 + 2B\, v\, v' + C (v')^2 \right] dx, \quad v(0) = v(r) = 0 \tag{29}$$

where

$$A = \frac{\partial^2 F}{\partial u^2}, \quad B = \frac{\partial^2 F}{\partial u \partial u'}, \quad C = \frac{\partial^2 F}{\partial (u')^2}$$

The problem (29) is considered as a variational problem for the unknown variation v with fixed A, B and C,

$$\min_{v:\ v(0)=v(r)=0} \delta^2 I(v, v')$$

Its Euler equation:

$$\frac{d}{dx}(C v' + B v) - A v = 0, \quad v(r_0) = v(r_{\text{conj}}) = 0 \quad [r_0, r_{\text{conj}}] \subset [0, r] \tag{30}$$

is a solution to Sturm-Liouville problem The point r_0 and r_{conj} are called the conjugate points. The problem is homogeneous: If $v(x)$ is a solution and c is a real number, $cv(x)$ is also a solution.

Jacobi condition is satisfied if the interval does not contain conjugate points, that is if there is no nontrivial solutions to (30) on any subinterval of $[r_0, r_{\rm conj}] \subset [0, r]$, that is if there are no nontrivial solutions of (30) with boundary conditions $v(r_0) = v(r_{\rm conj}) = 0$.

If this condition is violated, than there exist a family of trajectories

$$U(x) = \begin{cases} u + \alpha v & \text{if} \quad x \in [r_0, r_{\rm conj}] \\ u & \text{if} \quad x \in [0, r]/[r_0, r_{\rm conj}] \end{cases}$$

that deliver the same value of the cost. Indeed, v is defined up to a multiplier: If v is a solution, αv is a solution too. These trajectories have discontinuous derivative at the points r_0 and $r_{\rm conj}$. Such discontinuity leads to a contradiction to the Weierstrass-Erdman condition which does not allow a broken extremal at these points.

Example 2.1 (Nonexistence of the minimizer: Blow up) Consider again problem (26)

$$I = \min_u \int_0^r \left(\frac{1}{2}(u')^2 - \frac{c^2}{2} u^2 \right) dx \quad u(0) = 0; \quad u(r) = A$$

The stationary trajectory and the second variation are given by formulas (27) and (28), respectively. Instead of arbitrary choosing the second variation (as we did above), we choose it as a solution to the homogeneous problem (30) for $v = \delta u$

$$v'' + c^2 v = 0, \quad r_0 = 0, \ u(0) = 0, \ u(r_{\rm conj}) = 0, \quad r_{\rm conj} \le r \tag{31}$$

This problem has a nontrivial solution $v = \epsilon \sin(cx)$ if the length of the interval is large enough to satisfy homogeneous condition of the right end. We compute $cr_{\rm conj} = \pi$ or

$$r(_{\rm conj}) = \frac{\pi}{c}$$

The second variation $\delta^2 I$ is positive when r is small enough,

$$\delta^2 I = \frac{1}{r} \epsilon^2 \left(\frac{\pi^2}{r^2} - c^2 \right) > 0 \quad \text{if } r < \frac{\pi}{c}$$

In the opposite case $r > \frac{\pi}{c}$, the increment is negative which shows that the stationary solution is not a minimizer.

To clarify this, let us compute the stationary solution (27). We have

$$u(x) = \left(\frac{A}{\sin(cr)} \right) \sin(cx) \quad \text{and } I(u) = -\frac{A^2}{\sin^2(cr)} \left(c^2 - \frac{\pi^2}{r^2} \right)$$

When r increases approaching the value $\frac{\pi}{c} - 0$, the magnitude of the stationary solution indefinitely grows, and the cost indefinitely decreases:

$$\lim_{r \to \frac{c}{\pi} - 0} I(u) = -\infty$$

On the other hand, the cost of the problem is monotonic function of the interval length r. To show this, it is enough to show that the problem cost for an interval $[0, r]$ can correspond to an admissible function defined at a larger interval $[0, r+d]$.

The admissible trajectory that correspond to the same cost is easily constructed. Let $u(x), x \in [0, r]$ be a minimizer (recall, that $u(0) = 0$) for the problem in $[0, r]$, and let the cost functional be I_r. In a larger interval $x \in [0, r+d]$, the admissible trajectory

$$\hat{u}(x) = \begin{cases} 0 & \text{if } 0 < x < d \\ u(x-d) & \text{if } d \le x \le r+d \end{cases}$$

corresponds to the same cost I_r. Therefore, the minimum I_{r+d} over $x \in [0, r+d]$ is not larger than I_r, or $I_{r+d} \le I_r$.

Obviously, this trajectory of the Euler equation is not a minimizer if $r > \frac{\pi}{c}$, because it corresponds to a finite cost $I(u) > -\infty$.

Remark 2.2 Comparing the critical length $r_{\text{conj}} = \frac{\Pi}{c}$ with the critical length $r_{\text{crit}} = \frac{\sqrt{10}}{c}$ found in Example (2.1) by a guessed (nonoptimal) variation, we see that an optimal choice of variation improved the length of the critical interval at only 0.65%.

2.5 Nature does not minimize action

The next example deals with a system of multiple degrees of freedom. Consider the variational problem with the Lagrangian

$$L = \sum_{i=1}^{n} \frac{1}{2} m \dot{u}_i^2 - \frac{1}{2} C (u_i - u_{i-1})^2, \quad u(0) = u_0$$

We will see later in Chapter **??** that this Lagrangian describes the *action* of a chain of particles with masses m connected by springs with constant C. In turn, the chain models an elastic continuum.

Stationarity is the solution to the system

$$m_i \ddot{u}_i + C(-u_{i+1} + 2u_i - u_{i-1}), \quad u(0) = u_0$$

That describes dynamics of the chain. The continuous limit of the chain dynamics is the dynamics of an elastic rod.

The second variation (here we also use the notation $v = \delta u$)

$$\delta^2 L = \sum_{i=1}^{n} \frac{1}{2} m \dot{v_i}^2 - \frac{1}{2} C (v_i - v_{i-1})^2, \quad v_0 = 0, \quad v_n = 0$$

corresponds to the Euler equation – the eigenvalue problem

$$m \ddot{v} = \frac{C}{m} A v$$

where $v(t) = [v_1(t), \dots, v_n(t)]$ is the vector of variations and

$$A = \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -2 \end{pmatrix}.$$

The problem has a solution

$$v(t) = \sum \alpha_k v_k \sin \omega_k t \quad v(0) = v(T_{\text{conj}}) = 0, \quad T_{\text{conj}} \le T$$

where v_k are the eigenvectors, α are coefficients found from initial conditions, and ω_k are the square roots of eigenvalues of the matrix A. Solving the characteristic equation for eigenvalues $\det(A - \omega^2 I) = 0$ we find that these eigenvalues are

$$\omega_k = 2\sqrt{\frac{C}{m}} \sin^2 \left(\sqrt{\frac{C}{m}} \frac{\pi k}{n} \right), \quad k = 1, \dots n$$

The Jacobi condition is violated if $v(t)$ is consistent with the homogeneous initial and final conditions that is if the time interval is short enough. Namely, the condition is violated when the duration T is larger than

$$T \ge \frac{\pi}{\max(\omega_k)} \approx 2\pi \sqrt{\frac{m}{C}}$$

.

The continuous limit of the chain is achieved when the number N of masses indefinitely growth and each mass decreases correspondingly as $m(N) = \frac{m(0)}{N}$. The distance between masses decreases, the stiffness of one link increases as $C(N) = C(0)N$ as it become N times shorter. Correspondingly,

$$\sqrt{\frac{C(N)}{m(N)}} = N \sqrt{\frac{C(0)}{m(0)}}$$

and the maximal eigenvalue ω_N tends to infinity as $N \to \infty$. This implies Jacobi condition is violated at any finite time interval, or that action J of the continuous system is not minimized at any finite time interval.

What is minimized in classical mechanics? Lagrangian mechanics states that differential equations of Newtonian mechanics correspond to the stationarity of action: the integral of difference between kinetic T and potential V energies. Kinetic energy is a quadratic form of velocities $\dot{q}_i$ of particles, and potential energy depends only on positions (generalized coordinates) q_i of them

$$T(q, \dot{q}) = \frac{1}{2} \sum_i^n \dot{q}^T R(q) \dot{q} \quad V = V(q)$$

We assume that T is a convex function of q and $\dot{q}$, and V is a convex function of q.

As we have seen at the above examples, action $L = T - V$ does not satisfy Jacobi condition because kinetic and potential energies, which both are convex functions or q and $\dot{q}$, enter the action with different signs. Generally, the action is a saddle function of q and $\dot{q}$. The notion that Newton mechanics is not equivalent to minimization of a universal quantity, had significant philosophical implications, it destroyed the hypothesis about universal optimality of the world.

The minimal action principle can be made a minimal principle, in the Minkovski space. Formally, we replace time t with the imaginary variable $t = i\tau$ and use the second-order homogeneity of T:

$$T(q, \dot{q}) = \frac{1}{2}\dot{q}^T R(q)\dot{q} = -{q'_\tau}^T R(q) q'_\tau$$

The Lagrangian, considered as a function of q and q'_τ instead of q and $\dot{q}$, become a negative of a convex function if potential energy V and inertia $R(q)$ are convex. It become formally equal to the first integral (the energy)

$$L(q, q'_\tau) = -{q'_\tau}^T R(q) q'_\tau - V(q)$$

that is conserved in the original problem.

The local maximum of the variational problem, or

$$J = -\min_{q(\tau)} \int_{t_0}^{t} (-L(q, q'_\tau)) d\tau$$

does exist, since the Lagrangian $-L(q, q'_\tau)$ is convex with respect to q and q'_τ.

Example 2.2 The Lagrangian L for an oscillator

$$L = \frac{1}{2}\left(m\dot{u}^2 - Cu^2\right)$$

becomes

$$\hat{L} = -\frac{1}{2}\left(m{u'}^2 + Cu^2\right).$$

The Euler Equation for $\hat{L}$

$$m\, u'' - Cu = 0$$

corresponds to the solution

$$u(\tau) = A\cosh(\omega\tau) + B\sinh(\omega\tau), \quad \omega = \sqrt{\frac{C}{m}}$$

The stationary solution satisfies Weierstrass and Jacobi conditions. Returning to original notations $t = i\tau$ we obtain

$$A\cos(\omega t) + B\sin(\omega t)$$

the correct solution of the original problem. Remarkable, that this solution is unstable, but its transform to Minkovski space is stable.

These ideas have been developed in the special theory of relativity (world lines).