Legendre's sufficient condition of extremum

Francis J. Narcowich

5 апреля 2023 г.

1. Variational Calculus
2. Legendre's trick
3. Proof of local minimum

What we know from variational calculus

Consider an optimisation problem

$$J[x] = \int_a^b F(t, x(t), \dot{x}(t)) \to \mathrm{extr},$$

$$x = x(t) \in C^1[a,b]; \;\; x(a) = A, \;\; x(b) = B$$

What we know from variational calculus

Consider an optimisation problem

$$J[x] = \int_a^b F(t, x(t), \dot{x}(t)) \to \mathrm{extr},$$

$$x = x(t) \in C^1[a, b]; \;\; x(a) = A, \;\; x(b) = B$$

Fix $x(t)$ and $h(t)$, where $h(a) = h(b) = 0$.
Observe Taylor's formula:

$$J[x + \varepsilon h] = J[x] + \varepsilon \delta J[x; h] + \frac{1}{2}\varepsilon^2 \delta^2 J[x; h] + o(\varepsilon^2),$$

What we know from variational calculus

Consider an optimisation problem

$$J[x] = \int_a^b F(t, x(t), \dot{x}(t)) \to \text{extr},$$

$$x = x(t) \in C^1[a, b]; \ \ x(a) = A, \ x(b) = B$$

Fix $x(t)$ and $h(t)$, where $h(a) = h(b) = 0$.
Observe Taylor's formula:

$$J[x + \varepsilon h] = J[x] + \varepsilon \delta J[x; h] + \frac{1}{2}\varepsilon^2 \delta^2 J[x; h] + o(\varepsilon^2),$$

where we put

$$\delta J[x; h] = \frac{d}{d\varepsilon} J[x + \varepsilon h]\bigg|_{\varepsilon=0}, \quad \delta^2 J[x; h] = \frac{d^2}{d\varepsilon^2} J[x + \varepsilon h]\bigg|_{\varepsilon=0}$$

If x is an extremum, then it is necessary that $\delta J[x; h] = 0$ for all h. After some machinery, the Euler-Lagrange equation is deduced from this:

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0.$$

If x is an extremum, then it is necessary that $\delta J[x; h] = 0$ for all h. After some machinery, the Euler-Lagrange equation is deduced from this:

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0.$$

The Euler-Lagrange is a necessary condition, since it only implies $\delta J[x; h] = 0$. What about $\delta^2 J[x; h]$?

If x is an extremum, then it is necessary that $\delta J[x;h] = 0$ for all h. After some machinery, the Euler-Lagrange equation is deduced from this:

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0.$$

The Euler-Lagrange is a necessary condition, since it only implies $\delta J[x;h] = 0$. What about $\delta^2 J[x;h]$?
There are a number of **sufficient conditions** for extremum in variational calculus. Naturally they have to do something with $\delta^2 J[x;h]$: if it is positive for all $h \neq 0$, then x is a local minimum.

If x is an extremum, then it is necessary that $\delta J[x;h] = 0$ for all h. After some machinery, the Euler-Lagrange equation is deduced from this:

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0.$$

The Euler-Lagrange is a necessary condition, since it only implies $\delta J[x;h] = 0$. What about $\delta^2 J[x;h]$?
There are a number of **sufficient conditions** for extremum in variational calculus. Naturally they have to do something with $\delta^2 J[x;h]$: if it is positive for all $h \neq 0$, then x is a local minimum.
Today's goal: study Legendre's sufficient condition for a local minimum.

Legendre's trick

Let's calculate $\delta^2 \mathcal{J}$ explicitly.

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \frac{d^2}{d\varepsilon^2}\int_a^b F(t, x+\varepsilon h, \dot{x}+\varepsilon \dot{h})dt$$

Legendre's trick

Let's calculate $\delta^2 J$ explicitly.

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \frac{d^2}{d\varepsilon^2}\int_a^b F(t, x+\varepsilon h, \dot{x}+\varepsilon \dot{h})dt$$

$$= \frac{d}{d\varepsilon}\int_a^b (F_x h + F_{\dot{x}} \dot{h})dt$$

Legendre's trick

Let's calculate $\delta^2 J$ explicitly.

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \frac{d^2}{d\varepsilon^2}\int_a^b F(t, x+\varepsilon h, \dot{x}+\varepsilon \dot{h})dt$$

$$= \frac{d}{d\varepsilon}\int_a^b (F_x h + F_{\dot{x}}\dot{h})dt = \int_a^b (F_{xx}h^2 + 2F_{x\dot{x}}h\dot{h} + F_{\dot{x}\dot{x}}\dot{h}^2)dt$$

Legendre's trick

Let's calculate $\delta^2 J$ explicitly.

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \frac{d^2}{d\varepsilon^2}\int_a^b F(t, x+\varepsilon h, \dot{x}+\varepsilon \dot{h})dt$$

$$= \frac{d}{d\varepsilon}\int_a^b (F_x h + F_{\dot{x}}\dot{h})dt = \int_a^b (F_{xx}h^2 + 2F_{x\dot{x}}h\dot{h} + F_{\dot{x}\dot{x}}\dot{h}^2)dt$$

Write $\int_a^b 2F_{x\dot{x}}h\dot{h}dt = \int_a^b F_{x\dot{x}}d(h^2) = -\int_a^b \frac{d}{dt}F_{x\dot{x}}h^2 dt$ and get

Legendre's trick

Let's calculate $\delta^2 J$ explicitly.

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\Big|_{\varepsilon=0} = \frac{d^2}{d\varepsilon^2}\int_a^b F(t, x+\varepsilon h, \dot{x}+\varepsilon \dot{h})dt$$

$$= \frac{d}{d\varepsilon}\int_a^b (F_x h + F_{\dot{x}}\dot{h})dt = \int_a^b (F_{xx}h^2 + 2F_{x\dot{x}}h\dot{h} + F_{\dot{x}\dot{x}}\dot{h}^2)dt$$

Write $\int_a^b 2F_{x\dot{x}}h\dot{h}dt = \int_a^b F_{x\dot{x}}d(h^2) = -\int_a^b \frac{d}{dt}F_{x\dot{x}}h^2 dt$ and get

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\Big|_{\varepsilon=0} = \int_a^b \Big(F_{xx} - \frac{d}{dt}F_{x\dot{x}}\Big)h^2 + F_{\dot{x}\dot{x}}\dot{h}^2 dt$$

$$= \int_a^b P\dot{h}^2 + Qh^2 dt$$

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2)dt$$

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

$$\frac{d^2}{d\varepsilon^2} J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2)dt$$

$$= \int_a^b P\dot{h}^2 + 2wh\dot{h} + (Q+\dot{w})h^2 dt$$

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2)dt$$

$$= \int_a^b \underline{P\dot{h}^2 + 2wh\dot{h}} + (Q+\dot{w})h^2 dt$$

Complete the square: $P\dot{h}^2 + 2wh\dot{h} = P(\dot{h} + \frac{w}{P}h)^2 - \frac{w^2}{P}h^2$,

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2)dt$$

$$= \int_a^b P\dot{h}^2 + 2wh\dot{h} + (Q+\dot{w})h^2 dt$$

Complete the square: $P\dot{h}^2 + 2wh\dot{h} = P(\dot{h} + \frac{w}{P}h)^2 - \frac{w^2}{P}h^2$, so we have

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P(\dot{h} + \frac{w}{P}h)^2 + (\dot{w} + Q - \frac{w^2}{P})h^2 dt$$

Consider arbitrary function $w(t)$. Add $\int_a^b \frac{d}{dt}(wh^2)dt$, which is zero by Fund. T. of Calculus, to the integral:

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P\dot{h}^2 + Qh^2 + \frac{d}{dt}(wh^2)dt$$

$$= \int_a^b P\dot{h}^2 + 2wh\dot{h} + (Q+\dot{w})h^2 dt$$

Complete the square: $P\dot{h}^2 + 2wh\dot{h} = P(\dot{h} + \frac{w}{P}h)^2 - \frac{w^2}{P}h^2$, so we have

$$\frac{d^2}{d\varepsilon^2}J[x+\varepsilon h]\bigg|_{\varepsilon=0} = \int_a^b P(\dot{h} + \frac{w}{P}h)^2 + (\dot{w} + Q - \frac{w^2}{P})h^2 dt$$

Now assume $P > 0$ and $\dot{w} + Q - \frac{w^2}{P} = 0$. If it is so, then trivially $\delta^2 J[x;h] > 0$ for all h. One might think that such w can always be found, but that's not true.

Now consider this equation (Riccati's equation):

$$\dot{w} + Q - \frac{w^2}{P} = 0$$

Now consider this equation (Riccati's equation):

$$\dot{w} + Q - \frac{w^2}{P} = 0$$

Substitute $w = -(\dot{u}/u)P$ to get Jacobi's equation:

$$-\frac{d}{dt}\left(P\frac{du}{dt}\right) + Qu = 0$$

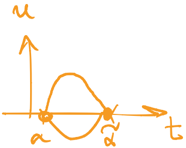
Now consider this equation (Riccati's equation):

$$\dot{w} + Q - \frac{w^2}{P} = 0$$

Substitute $w = -(\dot{u}/u)P$ to get Jacobi's equation:

1) $u(0) \neq 0$
2) $u(0) = 0$ but $\neq 0$ on $[a,b]$

$$-\frac{d}{dt}\left(P\frac{du}{dt}\right) + Qu = 0$$



The central notion here is **conjugate points**:

Definition

$t = \alpha$ and $t = \tilde{\alpha}$ are said to be **conjugate for Jacobi's equation**, if there is a solution u for which $u(\alpha) = u(\tilde{\alpha}) = 0$ and $u(x) \neq 0$ between α and $\tilde{\alpha}$.

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows.

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows. Let u_0 and u_1 be solutions to the Jacobi's equation such that:

$$u_0(a) = 0, \ \dot{u}_0(a) = 1, \ u_1(a) = 1, \ \dot{u}_1(a) = 1.$$

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows. Let u_0 and u_1 be solutions to the Jacobi's equation such that:

$$u_0(a) = 0,\ \dot{u}_0(a) = 1,\ u_1(a) = 1,\ \dot{u}_1(a) = 1.$$

Observe that u_0 is positive on $(a, b]$, and u_1 is positive on some segment $[a, c) \subset [a, b]$, Then it is possible to choose a linear combination $m_0 u_0 + m_1 u_1$ to be strictly positive on $[a, b]$.

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows. Let u_0 and u_1 be solutions to the Jacobi's equation such that:

$$u_0(a) = 0,\ \dot{u}_0(a) = 1,\ u_1(a) = 1,\ \dot{u}_1(a) = 1.$$

Observe that u_0 is positive on $(a, b]$, and u_1 is positive on some segment $[a, c) \subset [a, b]$, Then it is possible to choose a linear combination $m_0 u_0 + m_1 u_1$ to be strictly positive on $[a, b]$.
We have thus solved the Riccati equation $\dot{w} + Q - \frac{w^2}{P} = 0$, and the second variation becomes

$$\delta^2 J[x; h] = \int_a^b P \left(\dot{h} + \frac{w}{P} h\right)^2 dt$$

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows. Let u_0 and u_1 be solutions to the Jacobi's equation such that:

$$u_0(a) = 0,\ \dot{u}_0(a) = 1,\ u_1(a) = 1,\ \dot{u}_1(a) = 1.$$

Observe that u_0 is positive on $(a, b]$, and u_1 is positive on some segment $[a, c) \subset [a, b]$, Then it is possible to choose a linear combination $m_0 u_0 + m_1 u_1$ to be strictly positive on $[a, b]$.
We have thus solved the Riccati equation $\dot{w} + Q - \frac{w^2}{P} = 0$, and the second variation becomes

$$\delta^2 J[x; h] = \int_a^b P\left(\dot{h} + \frac{w}{P}h\right)^2 dt$$

We assumed that $P > 0$ and $h(a) = h(b) = 0$. Can $\delta^2 J[x; h] = 0$ for some $h \neq 0$?

Suppose there are no conjugate points to a in the interval $[a, b]$. Then we can construct a strictly positive solution u as follows. Let u_0 and u_1 be solutions to the Jacobi's equation such that:

$$u_0(a) = 0,\ \dot{u}_0(a) = 1,\ u_1(a) = 1,\ \dot{u}_1(a) = 1.$$

Observe that u_0 is positive on $(a, b]$, and u_1 is positive on some segment $[a, c) \subset [a, b]$, Then it is possible to choose a linear combination $m_0 u_0 + m_1 u_1$ to be strictly positive on $[a, b]$.
We have thus solved the Riccati equation $\dot{w} + Q - \frac{w^2}{P} = 0$, and the second variation becomes

$$\delta^2 J[x; h] = \int_a^b P\left(\dot{h} + \frac{w}{P}h\right)^2 dt$$

We assumed that $P > 0$ and $h(a) = h(b) = 0$. Can $\delta^2 J[x; h] = 0$ for some $h \neq 0$? No:

$$\delta^2 J = 0 \iff \boxed{\dot{h} + \frac{w}{P}h = 0} \iff h \equiv 0$$

We have arrived at the following result.

Theorem

If $P > 0$ *and there are no conjugate points to a in* $[a, b]$*, then* $\delta^2 J[x; h] > 0$ *for all non-zero h.*

We have arrived at the following result.

Theorem

If $P > 0$ and there are no conjugate points to a in $[a, b]$, then $\delta^2 J[x; h] > 0$ for all non-zero h.

Note that it is not a sufficient condition for extremum yet. We need something in the form of

$$\delta^2 J[x, h] \geq c\|h\|^2_{C^1[a,b]}.$$

This is due to Taylor's formula:

$$J[x + h] = J[x] + \delta J[x; h] + \frac{1}{2}\delta^2 J[x; h] + o(\|h\|^2).$$

We have arrived at the following result.

Theorem

If $P > 0$ and there are no conjugate points to a in $[a, b]$, then $\delta^2 J[x; h] > 0$ for all non-zero h.

Note that it is not a sufficient condition for extremum yet. We need something in the form of

$$\delta^2 J[x, h] \geq c\|h\|^2_{C^1[a,b]}.$$

This is due to Taylor's formula:

$$J[x + h] = J[x] + \delta J[x; h] + \frac{1}{2}\delta^2 J[x; h] + o(\|h\|^2).$$

It is achievable under our current assumptions: $P > 0$ and no points conjugate to a. It just needs a little tweak.

Proof of local minimum

Take the second variation in one of the forms:

$$\delta^2 J[x;h] = \int_a^b \left(P\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2 \right) dt$$

Proof of local minimum

Take the second variation in one of the forms:

$$\delta^2 J[x;h] = \int_a^b \left(P\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2\right) dt$$

Let $0 < \sigma < \min P$ be a constant. We add and subtract $\sigma \dot{h}^2$:

$$\delta^2 J[x;h] = \int_a^b \left((P-\sigma)\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2\right) dt + \sigma \int_a^b \dot{h}^2 dt.$$

Proof of local minimum

Take the second variation in one of the forms:

$$\delta^2 J[x;h] = \int_a^b \left(P\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2 \right) dt$$

Let $0 < \sigma < \min P$ be a constant. We add and subtract $\sigma \dot{h}^2$:

$$\delta^2 J[x;h] = \int_a^b \left((P-\sigma)\dot{h}^2 + 2wh\dot{h} + (\dot{w} + Q)h^2 \right) dt + \sigma \int_a^b \dot{h}^2 dt.$$

We then add $\int_a^b \frac{d}{dt}(wh^2)dt = 0$ and complete the square just the same way as earlier. The result is only changed in P, which has become $P - \sigma$:

$$\delta^2 J[x;h] = \int_a^b (P-\sigma)\left(\dot{h} + \frac{w}{P-\sigma} h \right)^2 dt$$
$$+ \int_a^b \left(\dot{w} + Q - \frac{w^2}{P-\sigma} \right) h^2 dt + \sigma \int_a^b \dot{h}^2 dt$$

Continuing as earlier, we request

$$\dot{w} + Q - \frac{w^2}{P - \sigma} = 0,$$

Continuing as earlier, we request

$$\dot{w} + Q - \frac{w^2}{P - \sigma} = 0,$$

which with substitution $w = -(\dot{u}/u)P$ leads to a modified Jacobi equation

$$-\frac{d}{dt}\left((P - \sigma)\frac{du}{dt}\right) + Qu = 0$$

Continuing as earlier, we request

$$\dot{w} + Q - \frac{w^2}{P - \sigma} = 0,$$

which with substitution $w = -(\dot{u}/u)P$ leads to a modified Jacobi equation

$$-\frac{d}{dt}\left((P - \sigma)\frac{du}{dt}\right) + Qu = 0$$

We assumed that there were no conjugate points to a for the original Jacobi equation $-\frac{d}{dt}\left(P\frac{du}{du}\right) + Qu = 0$. Since this is a question of solution existence, for small σ there are no conjugate points to a for the modified equation too. So we fix some small enough σ.

Proceeding solving appropriately the Jacobi equation and thus the Riccati equation, we nullify one of the terms in $\delta^2 J$, and so we get

$\sigma < P$

$$\delta^2 J[x;h] = \int_a^b (P-\sigma)\left(\dot{h} + \frac{w}{P-\sigma}h\right)^2 + \sigma \int_a^b \dot{h}^2 dt$$

$$\geq \sigma \int_a^b \dot{h}^2 dt. \geq C\|h\|^2_{C^1}$$

$\|h\|_{C^1} = \int h^2 + \int \dot{h}^2$

We're actually done here, since by Friedrichs' inequality,

$$\int_a^b h^2 dt \leq \frac{(b-a)^2}{2} \int_a^b \dot{h}^2 dt$$

we have

$$\|h\|^2_{C^1} = \int_a^b h^2 dt + \int_a^b \dot{h}^2 dt \leq C \int_a^b \dot{h}^2 dt \implies \delta^2 J[x;h] \geq C\|h\|^2_{C^1}$$

Proceeding solving appropriately the Jacobi equation and thus the Riccati equation, we nullify one of the terms in $\delta^2 J$, and so we get

$$\delta^2 J[x;h] = \int_a^b (P-\sigma)\left(\dot h + \frac{w}{P-\sigma}h\right)^2 + \sigma\int_a^b \dot h^2 dt$$
$$\geq \sigma \int_a^b \dot h^2 dt.$$

We're actually done here, since by Friedrichs' inequality,

$$\int_a^b h^2 dt \leq \frac{(b-a)^2}{2}\int_a^b \dot h^2 dt$$

If 1) $P > 0$ $\forall t \in [a,b]$
2) there are no conj pts to a
$\Rightarrow$ minimum

Proceeding solving appropriately the Jacobi equation and thus the Riccati equation, we nullify one of the terms in $\delta^2 \mathcal{J}$, and so we get

$$\delta^2 \mathcal{J}[x;h] = \int_a^b (P-\sigma)\left(\dot{h} + \frac{w}{P-\sigma}h\right)^2 + \sigma \int_a^b \dot{h}^2 dt$$
$$\geq \sigma \int_a^b \dot{h}^2 dt.$$

We're actually done here, since by Friedrichs' inequality,

$$\int_a^b h^2 dt \leq \frac{(b-a)^2}{2} \int_a^b \dot{h}^2 dt$$

we have

$$\|h\|^2_{C^1} = \int_a^b h^2 dt + \int_a^b \dot{h}^2 dt \leq C \int_a^b \dot{h}^2 dt \implies \delta^2 \mathcal{J}[x;h] \geq C\|h\|^2_{C^1}$$