Introduction to the Calculus of Variations
Lecture Notes
Lecture 6

Swarnendu Sil

Spring 2021, IISc

Classical Methods

Second Variation

So far in this chapter we were concerned with **any critical point**. Now we want to investigate necessary and sufficient conditions for a given critical point to be a minimizer of the functional. We begin with a simple result which calculates the second variation and gives a necessary criterion a given critical point to be a minimizer in terms of the second variation.

Theorem 1 (Second Variation)**.** *Let $f = f(t, u, \xi) \in C^3\left([a,b] \times \mathbb{R}^N \times \mathbb{R}^N\right)$, $\alpha, \beta \in \mathbb{R}^N$ be given and $X = \left\{u \in C^1\left([a,b];\mathbb{R}^N\right) : u(a) = \alpha,\ u(b) = \beta\right\}$. Consider the problem*

$$(P) \qquad \inf_{u \in X} \left\{ I(u) = \int_a^b f(t, u(t), \dot{u}(t))\ \mathrm{d}t \right\} = m.$$

If $\bar{u} \in X \cap C^2\left([a,b];\mathbb{R}^N\right)$ is a minimizer for (P), then the following integral

$$\int_a^b \left[\langle f_{uu}(t, \bar{u}, \dot{\bar{u}})\, \psi, \psi \rangle + 2 \left\langle f_{u\xi}(t, \bar{u}, \dot{\bar{u}})\, \psi, \dot{\psi} \right\rangle + \left\langle f_{\xi\xi}(t, \bar{u}, \dot{\bar{u}})\, \dot{\psi}, \dot{\psi} \right\rangle \right] \qquad (1)$$

is ***nonnegative*** *for any $\psi \in C_c^1\left([a,b];\mathbb{R}^N\right)$.*

Proof. As we did in deriving the EL equations, we take $\psi \in C_c^1\left([a,b];\mathbb{R}^N\right)$. Thus for any $h \in \mathbb{R}$, we have $\bar{u} + h\psi \in X$. Now we define the function $g : \mathbb{R} \to \mathbb{R}$ by $g(h) := I(\bar{u} + h\psi)$.

Then $g \in C^2(\mathbb{R})$ (Check!) and since $\bar{u}$ is a minimizer, g must have a local minima at 0. Thus we must have $g''(0) \geq 0$. But

$$g''(0) = \left. \frac{d^2}{dh^2} \left[I(\bar{u} + h\psi) \right] \right|_{h=0}.$$

The rest is a straight forward calculation. Note that since $f \in C^3$, thus in particular C^2 and thus $f_{u\xi}$, f_{uu} and $f_{\xi\xi}$ are all symmetric matrices. $\square$

Quadratic functional related to second variation

To understand the expression for the second variation better, we integrate by parts in the mixed term and obtain

$$2\int_a^b \left\langle f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\psi,\dot{\psi}\right\rangle = -\int_a^b \left\langle \frac{d}{dt}\left[f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\right]\psi,\psi\right\rangle.$$

Note that here again we used the fact that $f_{u\xi}$ takes values in the space of symmetric matrices. In view of this, we can rewrite the expression (1) as

$$\int_a^b \left[\left\langle\left[f_{uu}\left(t,\bar{u},\dot{\bar{u}}\right)-\frac{d}{dt}f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\right]\psi,\psi\right\rangle+\left\langle f_{\xi\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\dot{\psi},\dot{\psi}\right\rangle\right].$$

The important point here is that the matrix $f_{\xi\xi}$ plays the dominant role here in determining whether this quadratic form will be nonnegative or not.

The heuristic argument is, since ψ vanishes at the boundary, we have a Poincaré inequality. Roughly, since the function vanishes at the boundary, the value of the function itself can not be large while keeping its derivative small, since it has climb up from zero to the high values. But the converse is quite possible! The function can be small with large derivative! Why? It can oscillate a lot!

We now formalize the heuristic argument.

Lemma 2. *If the following inequality*

$$\int_a^b \left[\left\langle\left[f_{uu}\left(t,\bar{u},\dot{\bar{u}}\right)-\frac{d}{dt}f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\right]\psi,\psi\right\rangle+\left\langle f_{\xi\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\dot{\psi},\dot{\psi}\right\rangle\right]\ \mathrm{d}t\ \geq\ 0,$$

holds for every $\psi\in C_c^1\left([a,b];\mathbb{R}^N\right)$, then the matrix $f_{\xi\xi}\left(t,\bar{u},\dot{\bar{u}}\right)$ is nonnegative definite for every $t\in(a,b)$.

Proof. If $f_{\xi\xi}<0$ for some $t_0\in(a,b)$, this means there exist a $\zeta\in\mathbb{R}^N$ and $\beta>0$ such that

$$\left\langle f_{\xi\xi}\left(t_0,\bar{u}\left(t_0\right),\dot{\bar{u}}\left(t_0\right)\right)\zeta,\zeta\right\rangle<-\beta.$$

By continuity of $f_{\xi\xi}$, we can assume there exists $\alpha>0$ such that $a<t_0-\alpha<t_0+\alpha<b$ and we have

$$\left\langle f_{\xi\xi}\left(t,\bar{u}\left(t\right),\dot{\bar{u}}\left(t\right)\right)\zeta,\zeta\right\rangle<-\beta \qquad \text{for all } t\in[t_0-\alpha,t_0+\alpha].$$

Choose

$$\psi\left(t\right)=\begin{cases}\alpha\sin^2\left[\dfrac{\pi\left(t-t_0\right)}{\alpha}\right]\zeta & \text{if } t\in[t_0-\alpha,t_0+\alpha]\\ 0 & \text{otherwise.}\end{cases}$$

Clearly $\psi \in C_c^1\left([a,b];\mathbb{R}^N\right)$ and plugging it, we obtain

$$\pi^2 \int_{t_0-\alpha}^{t_0+\alpha} \sin^2\left[\frac{2\pi(t-t_0)}{\alpha}\right]\left\langle f_{\xi\xi}\left(t,\bar{u}(t),\dot{\bar{u}}(t)\right)\zeta,\zeta\right\rangle \, \mathrm{d}t$$
$$+\alpha^2 \int_{t_0-\alpha}^{t_0+\alpha} \sin^4\left[\frac{\pi(t-t_0)}{\alpha}\right]\left\langle\left[f_{uu}\left(t,\bar{u},\dot{\bar{u}}\right)-\frac{d}{dt}f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\right]\zeta,\zeta\right\rangle \, \mathrm{d}t \geq 0$$

But this implies

$$2M\alpha^3 - 2\beta\pi^2\alpha \geq 0,$$

where

$$M = \max_{t\in[t_0-\alpha,t_0+\alpha]}\left|\left\langle\left[f_{uu}\left(t,\bar{u},\dot{\bar{u}}\right)-\frac{d}{dt}f_{u\xi}\left(t,\bar{u},\dot{\bar{u}}\right)\right]\zeta,\zeta\right\rangle\right|.$$

But this means

$$\beta \leq \frac{M}{\pi^2}\alpha^2,$$

which we can easily contradict by letting $\alpha \to 0$. So we deduce

$$\left\langle f_{\xi\xi}\left(t,\bar{u}(t),\dot{\bar{u}}(t)\right)\zeta,\zeta\right\rangle \geq 0 \qquad \text{for all } \zeta \in \mathbb{R}^N, \text{ for all } t \in (a,b).$$

This proves the lemma. $\square$

This condition is known as the **Legendre condition**. This is implied by **convexity** of the map $\xi \mapsto f(t,u,\xi)$. If $n>1$, $N>1$, then the corresponding condition is called the **Legendre-Hadamard condition**.

$$\left\langle f_{\xi\xi}\left(x,\bar{u}(x),D\bar{u}(x)\right)a\otimes b, a\otimes b\right\rangle \geq 0$$

for all $a \in \mathbb{R}^n$, $b \in \mathbb{R}^N$ and for all $x \in \Omega$. This is **weaker** than the Legendre condition in that case, which would read

$$\left\langle f_{\xi\xi}\left(x,\bar{u}(x),D\bar{u}(x)\right)\xi,\xi\right\rangle \geq 0$$

for all $\xi \in \mathbb{R}^{n\times N}$ and for all $x \in \Omega$. Another way to see that these two conditions are different in the vectorial case $(n,N>1)$ is to notice that the Legendre-Hadamard is implied by convexity only along rank one matrices, which is in general significantly weaker than convexity.

Towards a sufficient condition

Possible candidate for sufficiency Can $f_{\xi\xi} \geq 0$ be a sufficient condition? **Clearly not!** Just think of $f(x)=x^3$, $x \in \mathbb{R}$. $x=0$ is a critical point and the second derivative vanishes, but it is not a minima! Then, can $f_{\xi\xi}>0$, i.e. positive definite instead of nonnegative definite, be a sufficient condition? This looks more promising, as this would be enough for finding minima of functionals. However, somewhat surprisingly, the answer is still **No!**

Understanding the trouble The reason is that the condition is purely local, whereas being a minimizer is not really a local property. We go back to geodesics. Think of the unit sphere in $\mathbb{R}^3$ centered at the origin and consider the points $A = (1, 0, 0)$, $B = (0, 1, 0)$ and $C = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$. All three points lie on the circle $\left\{(x, y, 0) : x^2 + y^2 = 1\right\}$, which being a great circle is a geodesic on the sphere. Now, the part of the circle going from A to B is a minimizing path with lenth $\pi/2$ and so is the part of the circle going from B to C, which has length $3\pi/4$. However, clearly the part of the circle going from A to C can not be minimizing, since its length is $\pi/2 + 3\pi/4 = 5\pi/4$, whereas the part of the circle going from C to A is definitely shorter, with length $2\pi - 5\pi/4 = 3\pi/4$.

Jacobi theory and Legendre method

We now consider the second variation itself as an integral functional

$$J[\psi] := \int_a^b \left[\left\langle P\dot{\psi}, \dot{\psi}\right\rangle + \langle Q\psi, \psi\rangle\right] \,\mathrm{d}t, \qquad \psi \in C^1, \psi(a) = \psi(b) = 0,$$

where

$$P := f_{\xi\xi}(t, \bar{u}, \dot{\bar{u}}) \qquad \text{and} \qquad Q := \left[f_{uu}(t, \bar{u}, \dot{\bar{u}}) - \frac{d}{dt} f_{u\xi}(t, \bar{u}, \dot{\bar{u}})\right]$$

Note that it is not difficult to establish that if

$$J[\psi] > c \int_a^b \left|\dot{\psi}\right|^2, \qquad \text{for all } \psi \in C^1, \psi \not\equiv 0 \text{ with } \psi(a) = \psi(b) = 0, \tag{2}$$

for some $c > 0$, then $\bar{u}$ is a minimizer. (See the problem sheet for a detailed proof). Note that

$$J[\psi] > 0, \qquad \text{for all } \psi \in C^1, \psi \not\equiv 0 \text{ with } \psi(a) = \psi(b) = 0,$$

is not sufficient, as can be seen in the following example.

Example 3. *The Lagrangian*

$$I[u] = \int_{-1}^1 \left[t^2 (\dot{u}(t))^2 + t(\dot{u}(t))^3\right] \,\mathrm{d}t,$$

has $u \equiv 0$ as an extremal where the second variation is positive for every non-trivial test function, but $u \equiv 0$ is not a minimizer. Indeed, at $u \equiv 0$, the second variation is

$$J[\psi] = \int_{-1}^1 t^2 \left(\dot{\psi}(t)\right)^2 \,\mathrm{d}t,$$

which is clearly positive for every $\psi \in C^1$, $\psi \not\equiv 0$ with $\psi(a) = \psi(b) = 0$. Now consider the family of functions

$$u_\varepsilon(t) := \begin{cases} \varepsilon\left(\frac{3}{4}\varepsilon + t\right) & \text{for } -\frac{3}{4}\varepsilon \leq t \leq 0, \\ \varepsilon\left(\frac{3}{4}\varepsilon - t\right) & \text{for } 0 \leq t \leq \frac{3}{4}\varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Now we have $I[u_\varepsilon] = -\frac{9}{32}\varepsilon^5 < 0 = I[0]$. We can easily round off the corners of u_ε, making it C^1 and still keep the condition that the functional is strictly negative on these functions. Moreover, clearly these modified C^1 functions converge in C^1 to $u \equiv 0$ as $\varepsilon \to 0$. So $u \equiv 0$ does not minimize the functional among C^1 functions in its C^1 neighborhood.

However, $P > 0$, i.e. positive definiteness of P for all $t \in (a, b)$ is not enough to obtain (2). So what other condition is needed to ensure this? To find out, Legendre wanted to 'complete the square' by adding a **null Lagrangian**, i.e. *integral functionals which always equate to zero irrespective of the argument in the class of admissible functions.*

Legendre method Let W be an arbitrary differentiable symmetric matrix. Then

$$0 = \int_a^b \frac{d}{dt}\left[\langle W\psi, \psi\rangle\right]\, \mathrm{d}t \qquad \text{for all } \psi \text{ with } \psi(a) = \psi(b) = 0.$$

Thus

$$\frac{d}{dt}\left[\langle W\psi, \psi\rangle\right] \text{ is a null lagrangian for any } W.$$

Hence adding such a term does not alter the value of $J[\psi]$. So we get

$$\begin{aligned} J[\psi] &= J[\psi] + \int_a^b \frac{d}{dt}\left[\langle W\psi, \psi\rangle\right]\, \mathrm{d}t \\ &= \int_a^b \left[\left\langle P\dot{\psi}, \dot{\psi}\right\rangle + 2\left\langle W\psi, \dot{\psi}\right\rangle + \left\langle \left(Q + \dot{W}\right)\psi, \psi\right\rangle\right]\, \mathrm{d}t \end{aligned}$$

When can we make this a perfect square?

Riccati equation

Proposition 4. *Suppose W is a solution of the following* matrix Riccati equation*,*

$$\dot{W} = -Q + WP^{-1}W. \tag{3}$$

Then we have

$$\left[\left\langle P\dot{\psi}, \dot{\psi}\right\rangle + 2\left\langle W\psi, \dot{\psi}\right\rangle + \left\langle \left(Q + \dot{W}\right)\psi, \psi\right\rangle\right] = \left|P^{\frac{1}{2}}\dot{\psi} + P^{-\frac{1}{2}}W\psi\right|^2.$$

Remark 5. *Here $P^{\frac{1}{2}}$ is the square root of P. Note that since P is symmetric and positive definite, the square root $P^{\frac{1}{2}}$ is well defined and is itself symmetric and positive definite.*

Proof. The proof is elementary calculation. Indeed, we have

$$\begin{aligned}
\left|P^{\frac{1}{2}}\dot{\psi}+P^{-\frac{1}{2}}W\psi\right|^2 &= \left\langle P^{\frac{1}{2}}\dot{\psi}+P^{-\frac{1}{2}}W\psi, P^{\frac{1}{2}}\dot{\psi}+P^{-\frac{1}{2}}W\psi\right\rangle \\
&= \left\langle P^{\frac{1}{2}}\dot{\psi}, P^{\frac{1}{2}}\dot{\psi}\right\rangle + \left\langle P^{\frac{1}{2}}\dot{\psi}, P^{-\frac{1}{2}}W\psi\right\rangle + \left\langle P^{-\frac{1}{2}}W\psi, P^{\frac{1}{2}}\dot{\psi}\right\rangle \\
&\qquad + \left\langle P^{-\frac{1}{2}}W\psi, P^{-\frac{1}{2}}W\psi\right\rangle \\
&= \left\langle P\dot{\psi}, \dot{\psi}\right\rangle + \left\langle \dot{\psi}, W\psi\right\rangle + \left\langle W\psi, \dot{\psi}\right\rangle + \left\langle P^{-1}W\psi, W\psi\right\rangle \\
&= \left\langle P\dot{\psi}, \dot{\psi}\right\rangle + 2\left\langle W\psi, \dot{\psi}\right\rangle + \left\langle WP^{-1}W\psi, \psi\right\rangle \\
&\overset{(3)}{=} \left\langle P\dot{\psi}, \dot{\psi}\right\rangle + 2\left\langle W\psi, \dot{\psi}\right\rangle + \left\langle \left(Q+\dot{W}\right)\psi, \psi\right\rangle.
\end{aligned}$$

This completes the proof. ∎

Now, to solve the Riccati equation

$$\dot{W} = -Q + WP^{-1}W,$$

let us substitute

$$W = -P\dot{\Psi}\Psi^{-1}.$$

Plugging it in the Riccati equation, we obtain

$$\frac{d}{dt}\left(P\dot{\Psi}\right) = Q\Psi. \tag{4}$$

Any solution Ψ of the above equation would furnish a solution W of the Riccati equation if Ψ is **invertible.**

Jacobi equation and Jacobi fields However, the equation above has another nice interpretation. We again consider the second variation itself as an integral functional

$$J[\psi] := \int_a^b \left[\left\langle P\dot{\psi}, \dot{\psi}\right\rangle + \langle Q\psi, \psi\rangle\right]\, \mathrm{d}t, \qquad \psi \in C^1, \psi(a) = \psi(b) = 0.$$

The Euler-Lagrange equation to this variational problem is

$$\frac{d}{dt}\left(P\dot{\psi}\right) = Q\psi. \tag{5}$$

This is called the **Jacobi equation** and its solutions (for a given u) is called a **Jacobi field along u.** Clearly, this equation looks very much the same as equation (4), except that here the unknown ψ is $\mathbb{R}^N$-valued, whereas Ψ in

(4) is an $N \times N$ matrix-valued function. However, since the form of both the equations are the same, the matrix formed by a system of N linearly independent solutions of (5) would solve (4). In fact, any N solutions of (5) would have the same property. However, they being linearly independent would mean that Ψ is invertible as well and thus would furnish a solution of the matrix Ricatti equation. Thus, linear independence is a crucially important property in this context and we would do well to be very interested in determining if and when this fails. In accordance with this line of thinking, we define the notion of conjugate points.

Definition 6 (Conjugate points)**.** *Let Ψ be the matrix of N solutions of the Jacobi equation, i.e.*

$$\Psi := \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix}.$$

where $\psi_1, \ldots, \psi_N$ solves the Jacobi equation and satisfies

$$\Psi(a) = 0 \quad \text{and} \quad \dot{\Psi}(a) = \mathbb{I}_N.$$

A point $\bar{a} \in (a, b]$ is called a ***conjugate to the point*** *a or simply a* ***conjugate point of*** *a if we have*

$$\det \Psi(\bar{a}) = 0.$$

At this stage, it should be clear what we are trying to achieve. If there are no conjugate points to a in $(a, b]$ for $J[\psi]$, then there would not be one as well for the functional

$$J_c[\psi] := J[\psi] - c \int_a^b \left|\dot{\psi}\right|^2,$$

by continuous dependence of solutions to ODEs on parameters. But then the corresponding Ψ would be invertible for every $t \in (a, b]$ and thus would furnish a solution to the corresponding Riccati equation. This in turn would imply that $J_c[\psi] > 0$ for all nontrivial $\psi \in C_c^1((a, b))$, which is sufficient for $\bar{u}$ to be a minimizer.