Variational Methods & Optimal Control

lecture 29

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

April 14, 2016

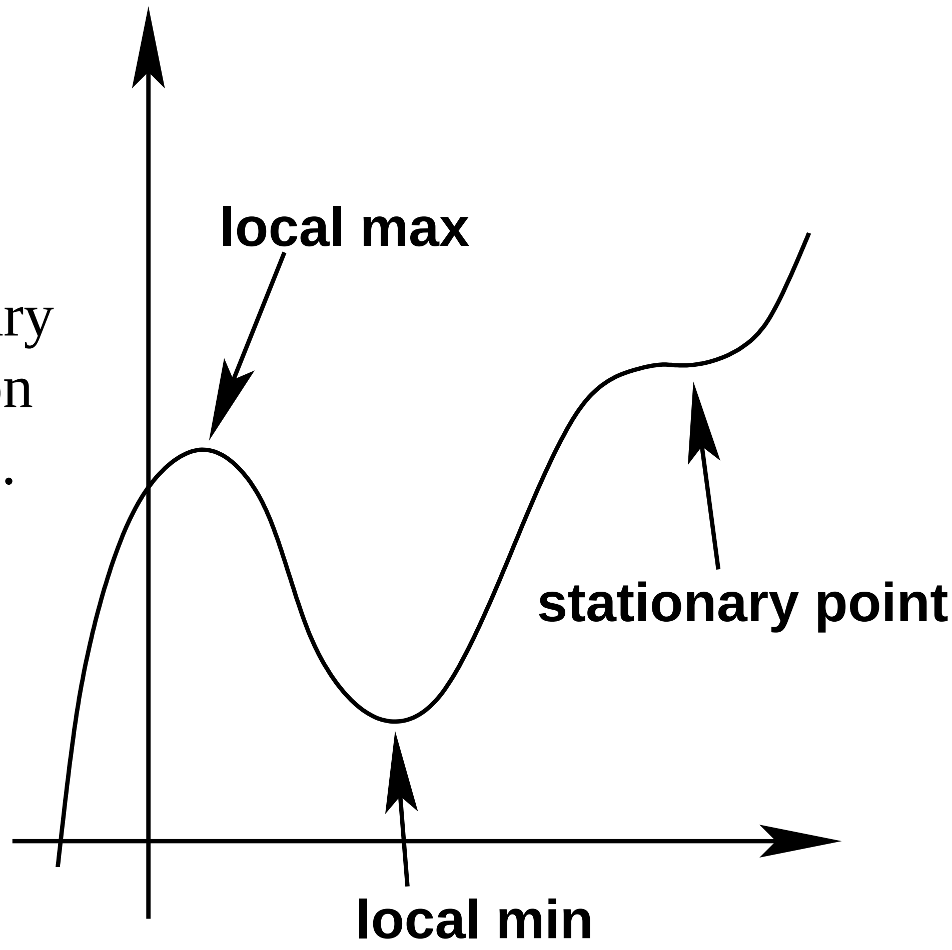
Classification of extrema

We have so far typically ignored the issue of classification of extrema, but remember that for simple stationary points we need to look at higher derivatives to see if a stationary point is a maximum, minimum or point of inflection. We need an analogous process for extremal curves as well.

Classification of extrema

Local extrema have $f'(x) = 0$

- $f''(x) > 0$ local minima
- $f''(x) < 0$ local maxima
- $f''(x) = 0$ it might be a stationary point of inflection, depending on higher order derivatives, e.g. x^4.



E-L solutions

- the E-L equations are a necessary condition
- the E-L equations are not sufficient
- along the extremal curve, the functional might have
  - a min, max, or stationary point
  - it might be global or local
- we really need to classify extremals
  - until now we have
    - just assumed it was the minima
    - used physical insight to understand the solution
    - tested it by inspection
  - we could also compare to alternative curves

Examples

- **Physical intuition:** Brachystochrone (or geodesic): we look for a minimum time path. So we can see that **physically** there can't be a maximum.
- **Examine the solution:** e.g. consider the functional

$$F\{y\} = \int_0^1 y'^2 \, dx$$

conditioned on $y(0) = y(1) = 0$.

The E-L equations give straight line solutions, e.g. $y = c_1 x + c_2$, and the boundary conditions imply $c_1 = c_2 = 0$, so $y' = 0$. Clearly then $F\{y\} = 0$, which is the minimum possible value, for an integral of a non-negative function like y'^2.

Examples

- **Compare with alternative curves:** For the functional

$$F\{y\} = \int_0^1 \left(xy' + y'^2\right)\, dx$$

conditioned on $y(0) = 0$ and $y(1) = 1$.

The E-L equations give

$$y = -\frac{1}{4}x^2 + c_1 x + c_2$$

and the boundary conditions give $c_1 = 5/4$, $c_2 = 0$, so the solution is

$$y = \frac{5}{4}x - \frac{1}{4}x^2$$

Examples

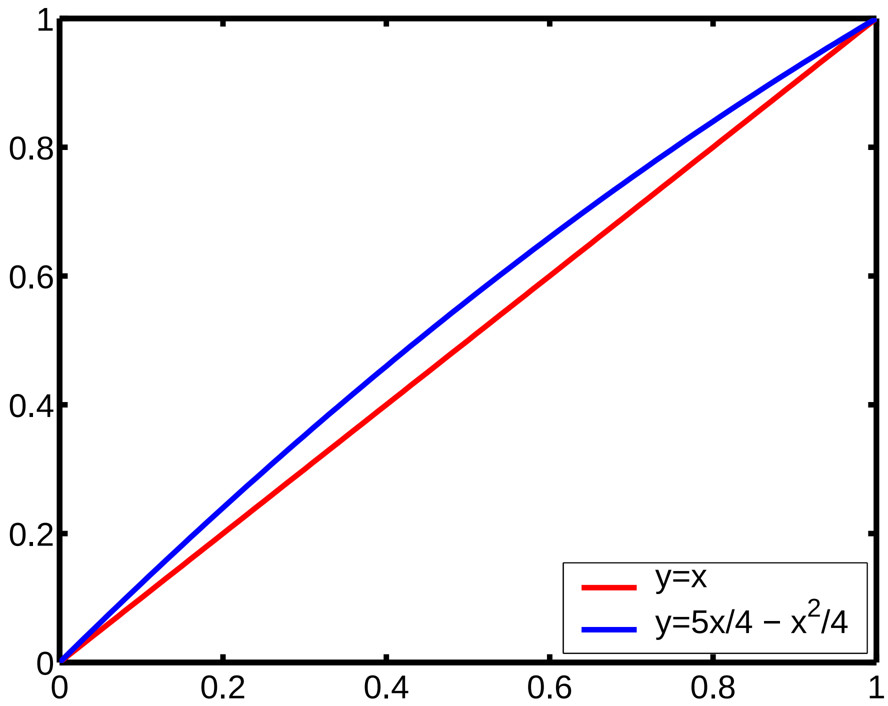
For $y = \frac{5}{4}x - \frac{1}{4}x^2$, we have $y' = \frac{5}{4} - \frac{1}{2}x$, so the function is

$$
\begin{aligned}
F\{y\} &= \int_0^1 \left[x\left(\frac{5}{4} - \frac{1}{2}x\right) + \left(\frac{5}{4} - \frac{1}{2}x\right)^2 \right] dx \\
&= \int_0^1 \left[\frac{25}{16} - \frac{1}{4}x^2 \right] dx \\
&= \left[\frac{25}{16}x - \frac{1}{12}x^3 \right]_0^1 \\
&= \frac{25}{16} - \frac{1}{12} \\
&= \frac{71}{48}
\end{aligned}
$$

Examples

For the curve $y(x) = x$, the $y' = 1$, so the functional is

$$\begin{aligned} F\{y\} &= \int_0^1 (x+1)\, dx \\ &= \left[x^2/2 + x\right]_0^1 \\ &= 3/2 \end{aligned}$$



Now $\frac{3}{2} > \frac{71}{48}$, so we should be looking at a local min.

But this isn't very formal, or rigorous!

Classification of extrema

- Above methods either
  - Aren’t very formal or rigorous
  - Aren’t easy to generalize
- Need to develop a means of formal classification
- The secret is by analogy to classification for functions of several variables
  - We need to look at second derivatives
  - Positive definiteness of the Hessian
- The analogy to second derivatives is called the **second variation**

Classification of extrema

Classification of extrema of functions (see Lecture 2)

Use Taylor's theorem in N-D

$$f(\mathbf{x}+\delta\mathbf{x}) = f(\mathbf{x}) + \delta\mathbf{x}^T \nabla f(\mathbf{x}) + \frac{1}{2}\delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} + O(\delta\mathbf{x}^3)$$

Where $H(\mathbf{x})$ is the Hessian matrix

$$H(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Maxima of N variables

If a smooth function $f(\mathbf{x})$ has a local extrema at $\mathbf{x}$ then $\nabla f(\mathbf{x}) = 0$, and so we can rewrite Taylor's theorem for small $\delta\mathbf{x}$ as

$$f(\mathbf{x}+\delta\mathbf{x}) - f(\mathbf{x}) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x}/2$$

A sufficient condition for the extrema $\mathbf{x}$ to be a local minimum is for the quadratic form

$$Q(\delta x_1, \ldots, \delta x_n) = \delta\mathbf{x}^T H(\mathbf{x})\delta\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\delta x_i \delta x_j$$

to be strictly positive definite.

Quadratic forms

A quadratic form

$$Q(\mathbf{x}) = \sum_{i,j} a_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x}$$

is said to be positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$.

A quadratic form is positive definite iff every eigenvalue of A is greater than zero.

A quadratic form is positive definite if all the principal minors in the top-left corner of A are positive, in other words

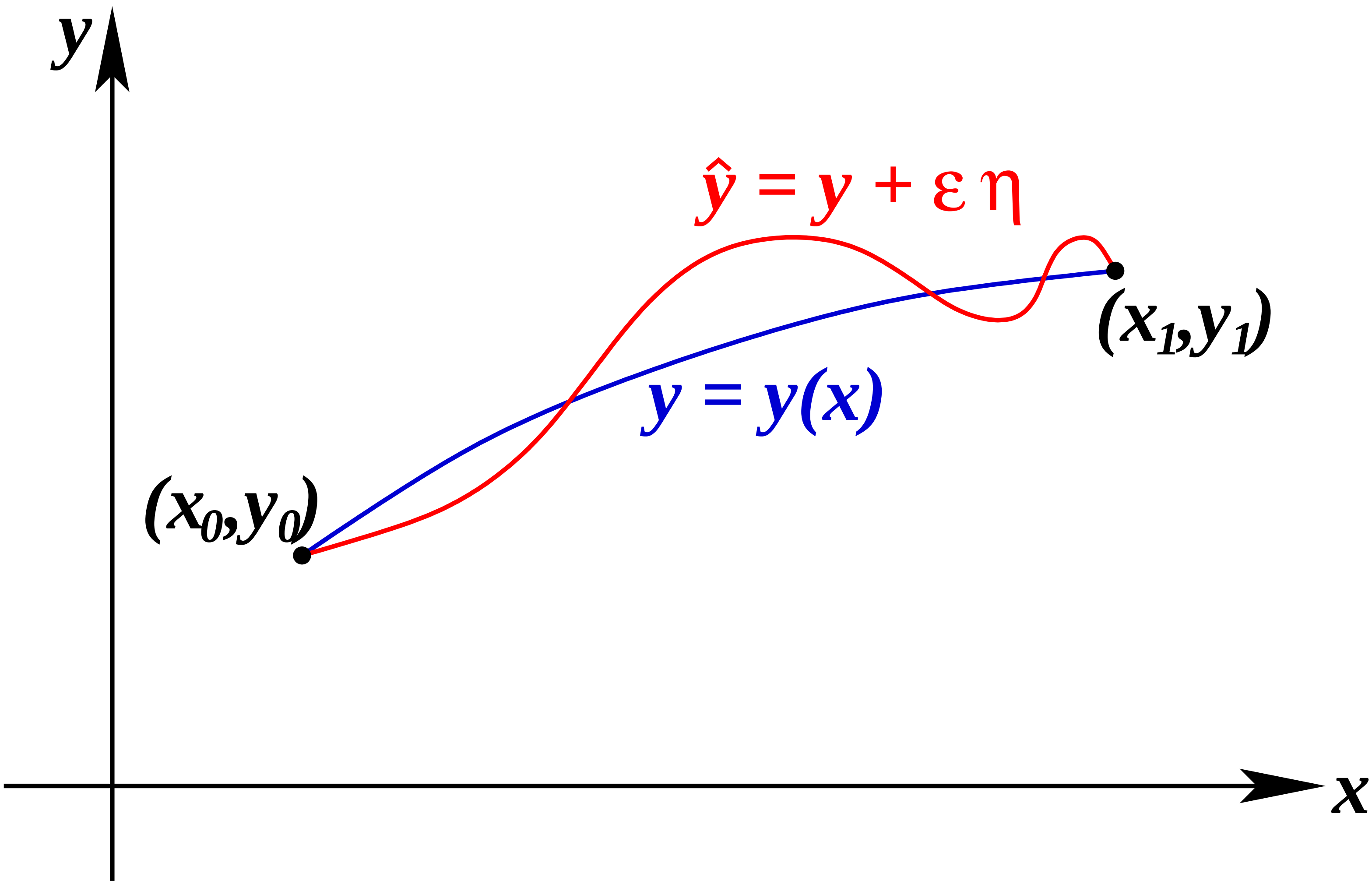
$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \cdots$$

Notes on maxima and minima

- maxima of $f(x)$ are minima of $-f(x)$.
- we need to generalize this to functionals
- we do this using the second variation
- note that even so, we only classify local min and max, the global min or max may occur at the boundary, or at one of several extrema.

The second variation

Once again consider the fixed end-point problem, with small perturbations about the extremal curve.



The second variation

Given the perturbation $\hat{y} = y + \varepsilon\eta$ we use Taylor's Theorem as when deriving the first variation, but this time we expand the $O(\varepsilon^2)$ terms as well, e.g.

$$
\begin{aligned}
f(x,\hat{y},\hat{y}') &= f(x,y,y') + \varepsilon\left[\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'}\right] \\
&\quad + \frac{\varepsilon^2}{2}\left[\eta^2\frac{\partial^2 f}{\partial y^2} + 2\eta\eta'\frac{\partial^2 f}{\partial y'\partial y} + \eta'^2\frac{\partial^2 f}{\partial y'^2}\right] + O(\varepsilon^3)
\end{aligned}
$$

$$
\begin{aligned}
F\{\hat{y}\} - F\{y\} &= \varepsilon\delta F(\eta,y) \\
&\quad + \frac{\varepsilon^2}{2}\int_{x_0}^{x_1}\left[\eta^2\frac{\partial^2 f}{\partial y^2} + 2\eta\eta'\frac{\partial^2 f}{\partial y\partial y'} + \eta'^2\frac{\partial^2 f}{\partial y'^2}\right]dx + O(\varepsilon^3)
\end{aligned}
$$

The second variation

$$\begin{aligned} F\{\hat{y}\} - F\{y\} &= \varepsilon \delta F(\eta, y) \\ &\quad + \frac{\varepsilon^2}{2} \int_{x_0}^{x_1} \left[\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'} \right] dx + O(\varepsilon^3) \\ &= \varepsilon \delta F(\eta, y) + \frac{\varepsilon^2}{2} \delta^2 F(\eta, y) + O(\varepsilon^3) \end{aligned}$$

Where we define the **Second Variation** by

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \left[\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'} \right] dx$$

Note for a stationary curve, we require $\delta F = 0$, so the behavior of $F\{\hat{y}\} - F\{y\}$ is captured in $\delta^2 F(\eta, y)$.

The second variation

Note that

$$2\eta\eta' = \frac{d}{dx}(\eta^2)$$

So we can write

$$\begin{aligned}
\int_{x_0}^{x_1} 2\eta\eta' f_{yy'}\, dx &= \int_{x_0}^{x_1} \frac{d(\eta^2)}{dx} f_{yy'}\, dx \\
&= \left[\eta^2 f_{yy'}\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta^2 \frac{d f_{yy'}}{dx}\, dx \\
&= -\int_{x_0}^{x_1} \eta^2 \frac{d f_{yy'}}{dx}\, dx
\end{aligned}$$

using integration by parts and the fact that $\eta(x_0) = \eta(x_1) = 0$.

The second variation

So we can write the second variation as

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) + \eta'^2 f_{y'y'} \, dx$$

This form has the advantage that

- $\eta^2 \geq 0$
- $\eta'^2 \geq 0$
- after solving E-L equations we know f and its derivatives

Classifying extrema

For an extremal curve y to be a local minima, we require

$$\delta^2 F(\eta, y) \geq 0$$

for all valid perturbation curves η. Likewise we get a maxima if $\delta^2 F(\eta, y) \leq 0$ for all η and a stationary curve if the second variation changes sign.

- Note that we have already solved the E-L equations, and so we know y. Hence we can calculate f_{yy}, $f_{yy'}$, and $f_{y'y'}$ explicitly.
- we still need to ensure $\delta^2 F(\eta, y) \geq 0$ for all possible η.

Legendre condition

The **Legendre condition** is a **necessary** condition for a local minima.

If y is a local minima of a functional $F\{y\} = \int f(x,y,y')\,dx$, then

$$f_{y'y'}(x,y,y') \geq 0$$

along the extremal curve y.

Legendre condition

Sketch proof: Remember that f and y are known functions (now), so we know f_{yy}, $f_{yy'}$ and $f_{y'y'}$, explicitly as functions of x, and hence we can write the second variation

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 B(x) + \eta'^2 A(x)\, dx$$

where

$$\begin{aligned} A(x) &= f_{y'y'} \\ B(x) &= \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) \end{aligned}$$

Legendre condition

Sketch proof: The proof relies on the fact that we can find functions η such that $|\eta|$ is small, but $|\eta'|$ is large.

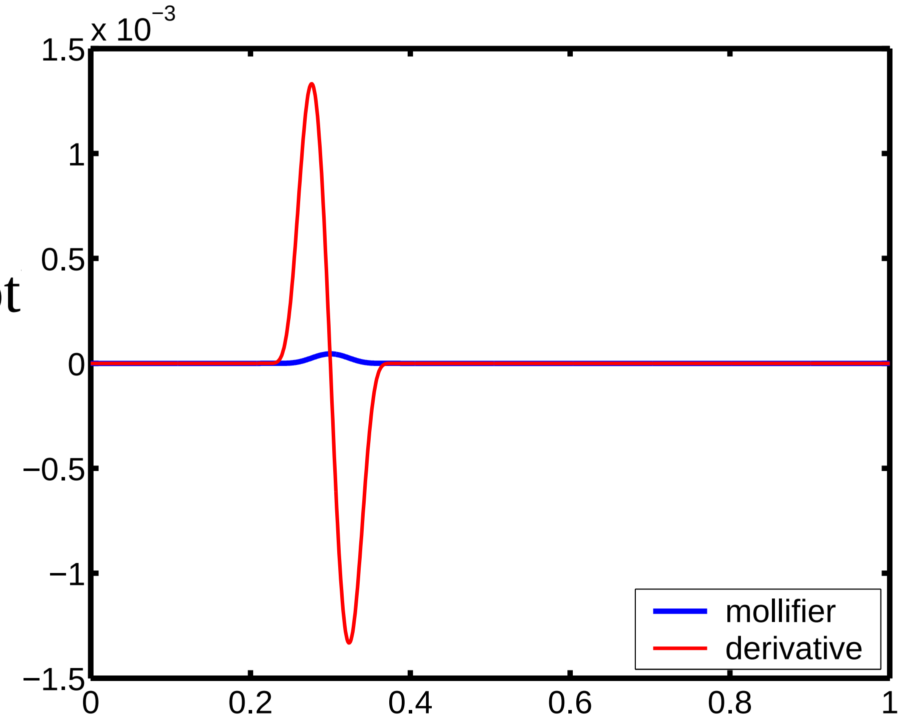
Note we cannot do the opposite, because $|\eta'|$ small, implies that η is smooth, which given the end conditions implies that $|\eta|$ will be small.

Example: **mollifier**

$$\eta(x) = \begin{cases} \exp\left(-\frac{\gamma}{\gamma^2-(x-c)^2}\right), & \text{if } x \in [c-\gamma, c+\gamma] \\ 0, & \text{otherwise} \end{cases}$$

Mollifier

$$\eta(x) = \begin{cases} \exp\left(-\frac{\gamma}{\gamma^2-(x-c)^2}\right), \\ \qquad \text{if } x \in (c-\gamma, c+\gamma) \\ 0, & \text{ot} \end{cases}$$



$$\eta'(x) = \begin{cases} -\frac{2\gamma(x-c)}{(\gamma^2-(x-c)^2)^2}\exp\left(-\frac{\gamma}{\gamma^2-(x-c)^2}\right), & \text{if } x \in (c-\gamma, c+\gamma) \\ 0, & \text{otherwise} \end{cases}$$

Ratio of derivative to function is larger for smaller γ.

Legendre condition

Sketch proof: Given $|\eta|$ small, we can essentially ignore the η^2 terms, and we get only the term

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta'^2 A(x)\, dx$$

If A changes sign, then we could choose η to be a mollifier such that it is localized in the part where A is positive, and a mollifier such that it is localized in the part of A which is negative. The two would produce integrals with different signs, and so we would get a change of sign of $\delta^2 F(\eta, y)$, which is what we are trying to avoid.

Example

Find the minimum of

$$F\{y\} = \int_0^1 \left(xy' + y'^2\right)\, dx$$

conditioned on $y(0) = 0$ and $y(1) = 1$.

The solution is

$$y = \frac{5}{4}x - \frac{1}{4}x^2$$

Then (from earlier)

$$\begin{aligned} f(x,y,y') &= xy' + y'^2 \\ &= \frac{25}{16} - \frac{1}{4}x^2 \end{aligned}$$

Example

$$
\begin{aligned}
f(x,y,y') &= xy' + y'^2 \\
f_{y'} &= x + 2y' \\
f_{y'y'} &= 2 \\
&> 0
\end{aligned}
$$

Hence Legendre's condition is satisfied, so this **could** be a local minimum.

Sufficient conditions

- various approaches to sufficient conditions
- problem is that we have to get away from point-wise conditions
  - like the Legendre condition
  - point-wise conditions couldn't classify which of two possible arcs of a great circle is the shortest path between two points on a sphere.
- a sufficient condition is the Jacobi condition, but there are others (van Brunt, 10.4, or Cragg's p.37, or Bliss, p.37)
- still mostly only conditions for local minima, so need to do more work

All is not lost

Example: Find the minimum of

$$F\{y\} = \int_0^1 \left(xy' + y'^2\right)\, dx$$

So

$$\begin{aligned} f_{y'y'} &= 2 \\ f_{yy'} &= 0 \\ f_{yy} &= 0 \end{aligned}$$

So the second variation

$$\delta^2 F(\eta, y) = \int_{x_0}^{x_1} \eta^2 \left(f_{yy} - \frac{d}{dx} f_{yy'} \right) + \eta'^2 f_{y'y'}\, dx = 2\int_{x_0}^{x_1} \eta'^2\, dx \geq 0$$

for all η so we have a local minimum!