Lecture 1: Introductory Ideas

Goals:

1. Define what a variational problem is.
2. Discuss problems on finite dimensional spaces ($\mathbb{R}^n$).
3. Meat: Examples of ill posed problems in infinite dimensional spaces.
   a.) non-uniqueness
   b.) non-existence
   c.) non-smooth solutions.

Key Points:

1. coercivity
2. compactness
3. lower semi-continuity
4. minimizing sequences.

1. Definitions:

The calculus of variations is concerned with problems of the following form:

$$\inf_{f \in \mathcal{A}} I[f].$$

a.) $I : \mathcal{A} \to \mathbb{R}$ is called a functional.
b.) $\mathcal{A}$ is called the admissable set.

*The reason we say inf instead of min is the minimum may not exist.

example:

$\mathcal{A} = \mathbb{R}$ and $I[x] = e^{-x^2}$. Clearly $\inf_{x \in \mathbb{R}} I[x] = 0$. but $I[x] > 0$.

[redacted]

This issue is not so trivial on general function spaces.

2. Finite Dimensional Spaces.

Briefly we will look at functions $f: \mathbb{R}^n \to \mathbb{R}$.

Classic Approach:

1. If f is smooth solve $\nabla f = 0$ for $\vec{x}^*$.
2. Need to check convexity condition:

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \geq 0$$

meaning $\vec{x}^T \cdot \nabla^2 f \cdot \vec{x} \big|_{x^*} \geq 0$. Reason: in a neighbourhood around $\vec{x}^*$ we have that

$$f(x) = f(\vec{x}^*) + \frac{1}{2} \vec{x}^T \cdot \nabla^2 f\big|_c \cdot \vec{x} \Rightarrow f(x) \geq f(x^*).$$

This approach assumes a lot of smoothness of the data.

Direct Method:

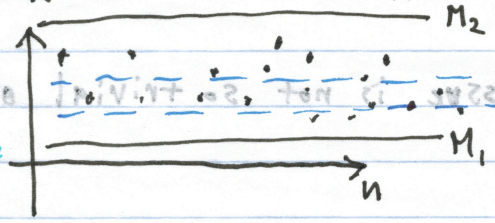
In terms of actually minimizing a function the above method is very impractical. (Solve n nonlinear equations). Here I will illustrate an important concept the direct method.

1. Let $\vec{x}_n$ be a minimizing sequence, i.e. a sequence satisfying $f(\vec{x}_n) \to \inf_{\vec{x} \in \mathbb{R}^n} f(\vec{x})$. We can abstractly consider this existence since $f: \mathbb{R}^n \to \mathbb{R}$. We would like to show two things:
   a.) There exists $\vec{x}^*$ such that $\vec{x}_n \to \vec{x}^*$.
   b.) For all $\vec{x} \in \mathbb{R}^n$, $f(\vec{x}^*) \leq f(\vec{x})$.

$|\vec{x}| = \sqrt{x_1^2 + \ldots + x_n^2}$

2. If f is coercive meaning $|\vec{x}| \to \infty \Rightarrow |f(\vec{x})| \to \infty$ then there exists $M \in \mathbb{R}^n$ such that $|\vec{x}_n| \leq M$. By the Bolzano-Weirstrass Theorem (Compactness) there exists a subsequence $\vec{x}_{n_k}$ and $\vec{x}^*$ such that $\vec{x}_{n_k} \to \vec{x}^*$.

(Proof of Bolzano-Weirstrass = keep dividing up sequence into terms with infinite number)



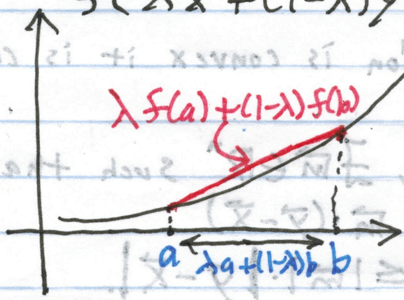
3. A function is lower semicontinuous if $x_n \to x$ implies $f(x) \le \liminf f(x_n)$.

Summary: Coercivity + lower semicontinuity $\Rightarrow$ existence of minimum.

These are the basic conditions we need for a well posed minimization problem on $\mathbb{R}^n$.

Convexity: A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$ and each $\lambda \in [0,1]$:

$$f(\lambda \vec{x} + (1-\lambda)\vec{y}) \le \lambda f(\vec{x}) + (1-\lambda) f(\vec{y})$$

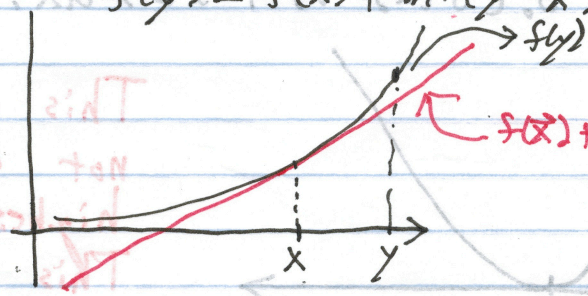


Line connecting two points lies "above" the function.

Caution: Convexity in terms of functions only makes sense on convex domains.

Theorem: Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is convex, then for each $x \in \mathbb{R}$ there exists ~~m~~ $\vec{m} \in \mathbb{R}$ such that $\forall \vec{y} \in \mathbb{R}^n$

$$f(\vec{y}) \ge f(\vec{x}) + \vec{m} \cdot (\vec{y} - \vec{x}).$$



Function lies above it supporting hyperplane.

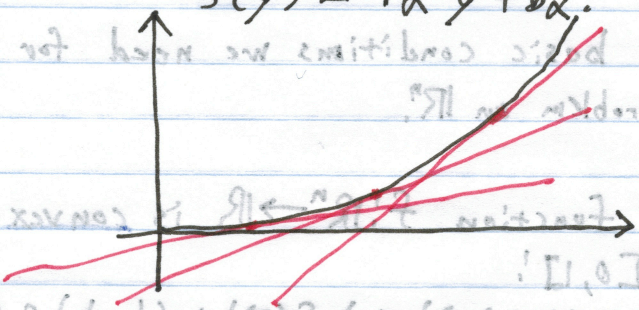
* If f is differentiable then $\vec{m} = \nabla f$.

Theorem: A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if it can be written as the supremum of bounding hyperplanes:

$$f(\vec{x}) = \sup_{\alpha} \vec{r}_\alpha^{\,T} \cdot \vec{x} + b_\alpha$$

where for all α in some indexing set and all $\vec{y} \in \mathbb{R}^n$

$$f(\vec{y}) \geq \vec{r}_\alpha^{\,T} \vec{y} + b_\alpha.$$



This theorem connects convexity with linear functions.

Theorem: If a function is convex it is continuous.

~~proof~~:

Let $\vec{x} \in \mathbb{R}^n$. $\forall \vec{y} \in \mathbb{R}^n$, $\exists \vec{m} \in \mathbb{R}^n$ such that

$$f(\vec{y}) \geq f(\vec{x}) + \vec{m} \cdot (\vec{y} - \vec{x})$$

$$\Rightarrow |f(\vec{y}) - f(\vec{x})| \leq |\vec{m}| \cdot |\vec{y} - \vec{x}|.$$

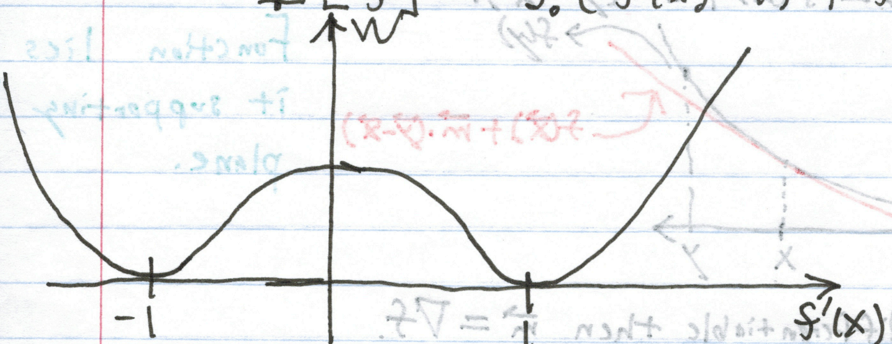
■

Hence convexity $\Rightarrow$ continuous $\Rightarrow$ lower semicontinuous.

3 Examples

a.) Non-Convex

$$I[f] = \int_0^1 \underbrace{(f'(x)^2 - 1)^2}_{W(f'(x))} + f(x)^2 \, dx.$$

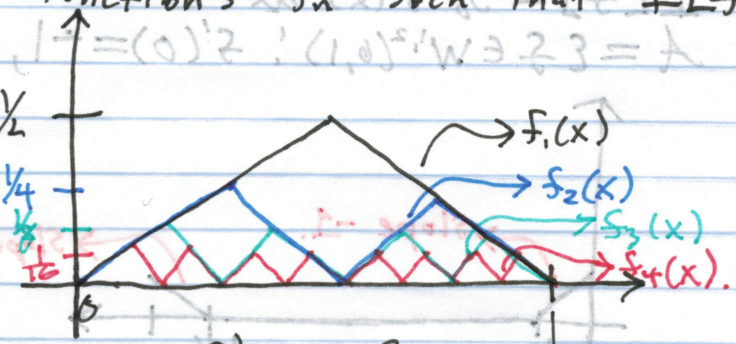


This functional is not convex in its highest derivative. This really means the integrand is not convex.

$$A = \{ f \in W_0^{1,4}(0,1) \}$$

we will talk about Sobolev spaces in more detail later.

Clearly, $I[f] \geq 0$. We construct a sequence of functions f_n such that $I[f_n] \to 0$.



(Oscillates to Zero)

$$I[f_n] = \int_0^1 f_n(x)^2 dx \leq \frac{1}{4n^2} \to 0.$$

However, if $I[f]=0$ then $f(x) \equiv 0$ a.e. and $f'(x) = \pm 1$ a.e. There is no minimum for this problem. $\blacksquare$

b.) Non-Coercive

$$I[f] = \int_0^1 \sqrt{f(x)^2 + f'(x)^2}\, dx.$$

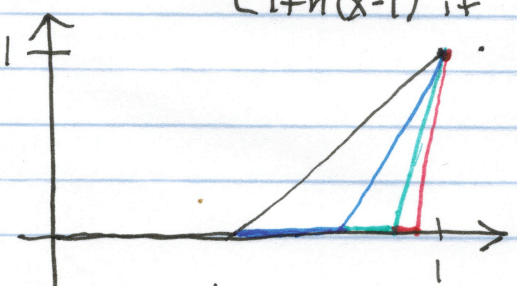
$$A = \{ f \in W^{1,1}(0,1) : f(0)=0 \text{ and } f(1)=1 \}.$$

We will show no minimum exists. Clearly,

$$I[f] = \int_0^1 \sqrt{f(x)^2 + f'(x)^2}\, dx \geq \int_0^1 |f'(x)|\, dx \geq \int_0^1 f'(x)\, dx = 1.$$

Now, construct a minimizing sequence as follows:

$$f_n(x) = \begin{cases} 0 & \text{if } x \in [0, 1-\frac{1}{n}] \\ 1+n(x-1) & \text{if } x \in [1-\frac{1}{n}, 1] \end{cases}$$



(derivative concentrates at a point).

$$I[f_n] = \int_{1-1/n}^1 \sqrt{(1+n(x-1))^2 + n^2}\, dx \leq \int_{1-1/n}^1 \sqrt{1+n^2}\, dx = \frac{1}{n}\sqrt{1+n^2}$$

Therefore,

$$1 \leq I[f_n] \leq \frac{1}{n}\sqrt{1+n^2} \Rightarrow \lim_{n\to\infty} I[f_n] = 1.$$

This implies that $\inf_{f \in A} I[f] = 0$. Now, suppose that $\exists f^* \in A$ such that $I[f^*]=1$. Then,

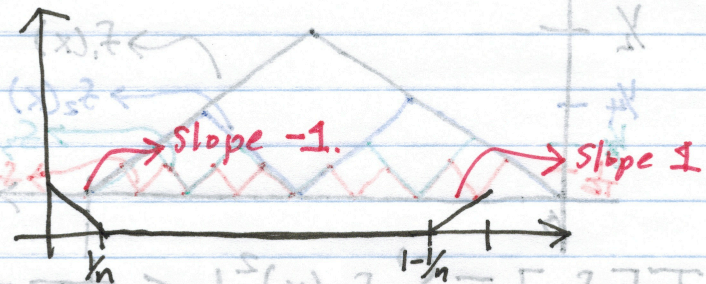
$$1 = \int_0^1 \sqrt{f^*(x)^2 + f^{*\prime}(x)^2}\, dx \geq \int_0^1 f'(x)\, dx = 1.$$

$\Rightarrow f^*(x) = 0$ which does not satisfy the boundary the boundary conditions. $\blacksquare$

C. Convex and Coercive (bad boundary conditions)

$$I[f] = \int_0^1 f'(x)^2\,dx$$

$$A = \{ f \in W^{1,2}(0,1) : f'(0) = -1,\ f'(0) = 1 \}.$$



Therefore $I[f_n] = \frac{1}{2n} \to 0$. However if $I[f] = 0$ then $f = \text{constant}$ which doesn't satisfy boundary conditions.