Math 521 Lecture #30
§4.2: Necessary Conditions for Extrema, Part I

In extending the methods of calculus of one real variable to that of extrema of functionals, we require normed linear spaces in which to define the notion of a derivative.

§4.2.1: Normed Linear Spaces.

A necessary condition for a differentiable real-valued function $f(x)$ of a real variable to have an extrema at x_0 is that $f'(x_0) = 0$.

Recall that the derivative of a function at a point x_0 is defined by

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

provided the limit exists.

For a functional $J : A \to \mathbb{R}$, if we replace $f(x)$ with $J(y)$ and replace the numbers x_0 and Δx by functions y_0 and $\Delta y = y - y_0$ in A, the definition for the derivative becomes

$$J'(y_0) = \lim_{\Delta y \to 0} \frac{J(y_0 + \Delta y) - J(y_0)}{\Delta y}.$$

Does this make sense? No because we do not know what $\Delta y \to 0$ means, nor do we know if $y + \Delta y$ is in A, nor do we know what by dividing by Δy means.

We will overcome this problem by first resolving what we mean by $\Delta y \to 0$, or when are two functions "close."

To do this we need to impose a linear structure (to define the subtraction $y - y_0$) and a distance (or metric to determine how "close" two functions are) on A.

We want A to be a subset of a **normed linear space** V where there is defined an addition and scalar multiplication on V that satisfy the axioms of a real vector space, and a norm that induces a metric V.

We refer to a normed linear space of functions as a **function space**, which in the Calculus of Variations are typically infinite dimensional.

A **norm** on a vector space V is a function $\|\cdot\|$ from V to $[0, \infty)$ such that

(i) $\|v\| = 0$ if and only if $v = 0$,

(ii) $\|\alpha v\| = |\alpha|\ \|v\|$ for all $\alpha \in \mathbb{R}$ and all $v \in V$, and

(iii) $\|y_1 + y_2\| \leq \|y_1\| + \|y_2\|$ for all $y_1, y_2 \in V$.

Any norm $\|\cdot\|$ on a vector space V induces a **metric** on d on V by

$$d(y_1, y_2) = \|y_1 - y_2\|.$$

Any metric d on V satisfies

(i) $d(y_1, y_2) = 0$ if and only if $y_1 = y_2$,

(ii) $d(y_1, y_2) = d(y_2, y_2)$ for all $y_1, y_2 \in V$, and

(iii) $d(y_1, y_3) \leq d(y_1, y_2) + d(y_2, y_3)$ for all $y_1, y_2, y_3 \in V$.

For a finite dimensional vector space V, any two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are **equivalent** in that there exists a constant $C > 0$ such that for all $v \in V$ there holds

$$C^{-1}\|v\|_1 \leq \|v\|_2 \leq C\|v\|_1.$$

This means that the norms define through their induced metrics, the same open sets (and hence the same topology) so that two vectors are close in all of the metrics.

However, for an infinite dimensional vector space V there may exists many non-equivalent norms which induces different collections of open sets (and hence different topologies) so that two vectors may not be close in all the metrics.

Example 4.8. A infinite dimensional vector space is $C[a,b]$, the set of continuous functions with domain $[a,b]$ for $-\infty < a < b < \infty$.

We know from Calculus that the sum and scalar multiple of continuous functions are continuous.

The zero vector is the zero function, and the additive inverse of $f \in C[a,b]$ is the scalar multiple $-f$.

The rest of the axioms of a vector space are easily shown to hold.

A norm on $C[a,b]$ is the **strong** or **max norm** defined by

$$\|f\|_M = \max_{x\in[a,b]} |f(x)|.$$

Property (i) of a norm is satisfied for the max norm because $f = 0$ implies $\|f\|_M = 0$, and $\|f\|_M = 0$ implies that $|f(x)| \leq 0$ which implies that $f(x) = 0$ for all $x \in [a,b]$.

Propoerty (ii) of a norm is satisfies for the max norm because for $\alpha \in \mathbb{R}$ we have

$$\|\alpha f\|_M = \max_{x\in[a,b]} |\alpha f(x)| = \max_{x\in[a,b]} |\alpha|\ |f(x)| = |\alpha| \max_{x\in[a,b]} |f(x)| = |\alpha|\ \|f\|_M.$$

To show that property (iii) holds for the max norm, we are required to show that

$$\max_{x\in[a,b]} |f(x) + g(x)| \leq \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |g(x)|.$$

We can do this through Calculus: because $|f(x) + g(x)|$ is continuous on $[a,b]$, there exists x_0 such that

$$\max_{x\in[a,b]} |f(x) + g(x)| = |f(x_0) + g(x_0)|.$$

The absolute value function is a norm on $\mathbb{R}$ so that

$$|f(x_0) + g(x_0)| \leq |f(x_0)| + |g(x_0)|.$$

Since $|f(x_0)|$ is just one the values of $|f(x)|$ we have by the definition of the max norm that

$$|f(x_0)| \leq \max_{x\in[a,b]} |f(x)|.$$

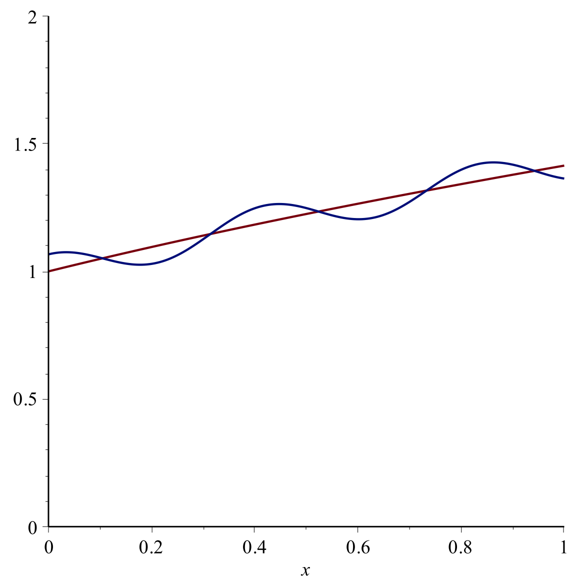
An identical statement holds for $|g(x_0)|$, so that we get

$$\max_{x\in[a,b]} |f(x)+g(x)| \leq |f(x_0)| + |g(x_0)| \leq \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |g(x)|.$$

This gives property (iii) for the max norm.

Two functions $f, g \in C[a,b]$ are "close" in $C[a,b]$ if $\|f-g\|_M$ is small.

Here is an example of two functions in $C[0,1]$ are are close in the max norm.



Another norm on $C[a,b]$ is the L^1**-norm** (or **strong norm**) defined by

$$\|f\|_1 = \int_a^b |f(x)|\ dx.$$

The L^1-norm satisfies property (i) because $f(x)=0$ implies $\|f\|_1 = 0$, and $\|f\|_1 = 0$ implies $f(x) = 0$ for all $x \in [a,b]$ (via the contrapositive of $f(x) \neq 0$ at some $x_0 \in [a,b]$, whence the continuity of f at x_0 implies the integral of $|f(x)|$ is bigger than 0).

It satisfies property (ii) because

$$\|\alpha f\|_1 = \int_a^b |\alpha f(x)|\ dx = |\alpha| \int_a^b |f(x)| dx = |\alpha|\ \|f\|_1.$$

It satisfies property (iii) because

$$\begin{aligned}\|f+g\|_1 &= \int_a^b |f(x)+g(x)|\ dx \\ &\leq \int_a^b \big(|f(x)| + |g(x)|\big)\ dx \\ &= \int_a^b |f(x)|\ dx + \int_a^b |g(x)|\ dx \\ &= \|f\|_1 + \|g\|_1.\end{aligned}$$

For the functions whose graphs are above, is the L^1-norm of their difference small? No, it is large compared to the max norm of their difference.

If the wavy curve oscillated more but maintained the same max-norm distance away from the mostly straight curve, the L^1-norm would be bigger.

If the nearly straight graph had a very tall spike over a very short subinterval, then the L^1-norm of the difference would be small, while the max norm would be big.

This says that the max-norm and the L^1-norm are not equivalent, that they define different topologies on $C[a,b]$.

Example 4.9. A common function space for the Calculus of Variations is $C^1[a,b]$, the space of continuously differentiable functions on $[a,b]$ (the derivative exists and is continuous).

The **weak norm** on $C^1[a,b]$ is defined by

$$\|f\|_w = \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |f'(x)|.$$

It is left to you to verify that this is indeed a norm. [We did most of the heavy lifting when showing that the strong or max norm is indeed a norm.]

What is the weak norm of the difference of the two functions in the graph on the previous page?

What happens to the norm of the difference when the wavy curve oscillates more and more but stays close to the nearly straight curve?

It goes to infinity so that the weak norm and the strong norm define different topologies on $C^1[a,b]$.

A similar weak type norm can be defined on the space $C^n[a,b]$ of functions n-times differentiable with continuous n^{th} derivative.