Math 521 Lecture #31
§4.2: Necessary Conditions for Extrema, Part II

§4.2.2: Derivatives of Functionals. To define what a local minimum for a functional J is, we need the set of admissible functions A to be a subset of a function space V with a norm $\|\cdot\|$.

We say $y_0 \in A$ is a **local minimizer** (or an **extremal** for J), and that J has a **local minimum** at y_0 (or is **stationary** at y_0), if there exists $\epsilon > 0$ such that $J(y_0) \leq J(y)$ for all $y \in A$ with $\|y - y_0\| < \epsilon$.

This notion of local minimizer does depend on the norm used.

If $A \subset C^1[a,b]$, we say that J has a strong local (or relative) minimum at y_0 if we use with the strong norm, while we say that J has a weak local (or relative) minimum at y_0 if we use the weak norm.

Example. The global minimizer of the arc length functional

$$J(y) = \int_0^1 \sqrt{1 + [y'(x)]^2}\, dx$$

over the admissible set

$$A = \{y \in C^1[0,1] : y(0) = 0,\ y(1) = 1\} \subset C^1[0,1]$$

is the straight line $y_0(x) = x$ for which $J(y_0) = \sqrt{2}$.

That is, we have

$$J(y_0) \leq J(y) \text{ for all } y \in A.$$

By choosing a norm on $C^1[0,1]$, we can view the global minimizer y_0 as a local minimizer for $J(y)$ relative to the norm.

In the strong norm

$$\|y\|_M = \max_{x \in [0,1]} |y(x)|$$

the open set

$$U_M = \{y \in C^1[a,b] : \|y - y_0\|_M < \eta\}$$

consists of curves close to y_0 but with arbitrarily large derivatives.

The strong local minimum $\sqrt{2}$ of J at y_0 is the smallest in comparison with the value $J(y)$ for $y \in U_M$ which can be quite large.

In the weak norm

$$\|y\|_w = \max_{x \in [0,1]} |y(x)| + \max_{x \in [0,1]} |y'(x)|$$

the open set

$$U_w = \{y \in C^1[0,1] : \|y - y_0\|_w < \eta\}$$

consists of curves close to y_0 with derivatives close to 1.

The weak local minimum $\sqrt{2}$ of J at y_0 is the smallest in comparison with the values of $J(y)$ for $y \in U_w$ which are all close to $\sqrt{2}$.

Note that U_w is a proper subset of U_M.

The weak local minimum is by comparison over a smaller subset of admissible functions, while the strong local minimum is by comparison over a larger subset of admissible functions.

In the Calculus of Variations, the choice of norm comes into play in the sufficient conditions for the existence of extrema.

That is, if $\{J(y) : y \in A\}$ is bounded below, then we can take a minimizing sequence $y_n \in A$ such that

$$\lim_{n\to\infty} J(y_n) = \inf\{J(y) : y \in A\},$$

and hope that in the norm chosen, a subsequence of $\{y_n\}$ converges to $y_0 \in A$, giving y_0 as a global minimizer.

In analogy with a real valued function, we expect that the derivative of J at y_0 should be 0.

But how do we define the derivative of a functional?

We take the definition of the derivative of a real-valued function $f(x)$,

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

and rewrite it in the form

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

from which we get

$$f(x_0+\Delta x) - f(x_0) = f'(x_0)\Delta x + o(\Delta x).$$

The **differential** of f at x_0, defined by $df(x_0, \Delta x) = f'(x_0)\Delta x$, is the linear part in the increment Δx of the total change $\Delta f = f(x_0 + \Delta x) - f(x)$.

This results in

$$\Delta f = df(x_0, \Delta x) + o(\Delta x).$$

For a functional $J : A \to \mathbb{R}$, where A is a subset of a normed function space V, an increment of $y_0 \in A$ has the form $y_0 + \epsilon h$ for $h \in V$ and small ϵ.

We require that $y_0 + \epsilon h$ be in A for all sufficiently small ϵ, so that we can evaluate $J(y_0 + \epsilon h)$.

We call the increment $\delta y_0 = (y_0 + \epsilon h) - y_0 = \epsilon h$ the **variation** of y_0.

The corresponding increment in J is given by

$$\Delta J = J(y_0 + \epsilon h) - J(y_0).$$

We want to find the linear part of the increment ΔJ which we do through the real-valued function

$$\mathcal{J}(\epsilon) = J(y_0 + \epsilon h)$$

which is defined on an open interval containing 0.

Assuming that $\mathcal{J}$ is sufficiently differentiable, we will have

$$\Delta J = \mathcal{J}(\epsilon) - \mathcal{J}(0) = \mathcal{J}'(0)\epsilon + o(\epsilon)$$

and so the differential of J is $\mathcal{J}'(0)\epsilon$.

The **first variation** or **Gâteaux derivative** of J at y_0 in the direction of $h \in V$ is

$$\delta J(y_0, h) = \mathcal{J}'(0) = \frac{d}{d\epsilon} J(y_0 + \epsilon h)\bigg|_{\epsilon=0},$$

provided the derivative exist.

Such an $h \in V$ for which $\delta J(y_0, h)$ exists is called an **admissible variation** at y_0, i.e., $h \in V$ satisfies $y_0 + \epsilon h \in A$ for all sufficiently small ϵ and $\delta J(y_0, h)$ exists.

The first variation is analogous with the directional derivative for a function of several variables, and when written in limit form is

$$\delta J(y_0, h) = \lim_{\epsilon \to 0} \frac{J(y_0 + \epsilon h) - J(y_0)}{\epsilon}.$$

4.2.3. Necessary Conditions.

Equipped with the first variation of a functional, we can now state a necessary condition for the existence of a local minimum.

Theorem 4.11. For a functional $J : A \to \mathbb{R}$ with $A \subset V$, if y_0 is a local minimum for J relative to the norm $\|\cdot\|$ on V, then for all admissible $h \in V$ at y_0, there holds

$$\delta J(y_0, h) = 0.$$

Proof. Suppose that $J(y_0)$ is a local minimum of J relative to the norm $\|\cdot\|$.

For an admissible variation $h \in V$ at y_0, we have that $y_0 + \epsilon h \in A$ is "close" to y_0 because

$$\|(y_0 + \epsilon h) - y_0\| = \|\epsilon h\| = |\epsilon| \; \|h\|$$

can be made sufficiently small by choosing ϵ sufficiently small.

This means that $y_0 + \epsilon h$ is in the open set $U = \{y \in V : \|y - y_0\| < \eta\}$

The real-valued function $\mathcal{J}(\epsilon) = J(y_0 + \epsilon h)$ has a local minimum at $\epsilon = 0$, and hence its derivative is 0, implying that $\delta J(y_0, h) = 0$. □

The vanishing of the first variation for all admissible variations often permits the elimination of admissible variation from the necessary condition to get a differential equation that y_0 must satisfy (another necessary condition).

Because we are dealing with a necessary condition, the y_0 we find from the differential equation may not be a minimizer, but could be a "saddle point."

Example. For the arc length functional

$$J(y) = \int_0^1 \sqrt{1 + [y'(x)]^2}\, dx$$

on the set of admissible functions

$$A = \{y \in C^2[0,1] : y(0) = 0,\ y(1) = 1\},$$

an admissible variation $h \in C^2[0,1]$ satisfies $h(0) = 0$ and $h(1) = 0$.

[Notice that we have switched to $C^2[0,1]$ from $C^1[0,1]$. We will see why in a minute.]

This means that $y(0) + \epsilon h(0) = 0$ and $y(1) + \epsilon h(1) = 1$ for all ϵ.

For $y_0 \in A$ we have that

$$J(y_0 + \epsilon h) = \int_0^1 \sqrt{1 + [y_0'(x) + \epsilon h'(x)]^2}\, dx.$$

Then

$$\begin{aligned} \frac{d}{d\epsilon} J(y_0 + \epsilon h) &= \frac{d}{d\epsilon} \int_0^1 \sqrt{1 + [y_0'(x) + \epsilon h'(x)]^2}\, dx \\ &= \int_0^1 \frac{\partial}{\partial \epsilon} \sqrt{1 + [y_0'(x) + \epsilon h'(x)]^2}\, dx \\ &= \int_0^1 \frac{[y_0'(x) + \epsilon h'(x)] h'(x)}{\sqrt{1 + [y_0'(x) + \epsilon h'(x)]^2}}\, dx \end{aligned}$$

and so by setting $\epsilon = 0$ we get

$$\delta J(y_0, h) = \int_0^1 \frac{y_0'(x) h'(x)}{\sqrt{1 + [y_0'(x)]^2}}\, dx.$$

If we now integrate by parts, with

$$u = \frac{y_0'(x)}{\sqrt{1 + [y_0'(x)]^2}},\ \ dv = h'(x) dx,$$

we get

$$\delta J(y_0, h) = \left. \frac{y_0'(x) h(x)}{\sqrt{1 + [y_0'(x)]^2}} \right|_0^1 - \int_0^1 \frac{d}{dx} \left[\frac{y_0'(x)}{\sqrt{1 + [y_0'(x)]^2}} \right] h(x)\, dx.$$

Because $h(0) = 0$ and $h(1) = 0$, the first part vanishes.

So if y_0 is a minimizer of J, then for all admissible variations h we have that

$$0 = \delta J(y_0, h) = -\int_0^1 \frac{d}{dx} \left[\frac{y_0'(x)}{\sqrt{1 + [y_0'(x)]^2}} \right] h(x)\, dx.$$

Because $h(x)$ is an arbitrary admissible variation, we conclude that y_0 satisfies the second-order differential equation

$$\frac{d}{dx} \left[\frac{y_0'(x)}{\sqrt{1 + [y_0'(x)]^2}} \right] = 0.$$