Necessary Conditions for Extrema - introduction

A necessary condition for a smooth function $f(x)$ to have a local minimum at x_0 is

(1) $\qquad f'(x_0) = 0$

where

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

We seek an analogous necessary condition for functionals $J : A \to \mathbb{R}$.

To even be able to define the notions of

(i) local min

(ii) derivatives

for a functional one must first have some sort of definition of distance between functions.

We can't just use (2) since we need to know what $\Delta x \to 0$ means when Δx is a function and not a real number as it is in (2).

Linear Spaces $\mathbb{V}$

Are sets with addition + and scalar multiplication definitions satisfying the following axioms

* (1) $u, v \in \mathbb{V} \Rightarrow u + v \in \mathbb{V}$

(2) $u + v = v + u$

(3) $u + (v + w) = (u + v) + w$

(4) $\exists 0 \in \mathbb{V}$ such that $u + 0 = u \quad \forall u \in \mathbb{V}$

(5) $\forall u \in \mathbb{V} \quad \exists (-u)$ s.t. $u + (-u) = 0$

(6) $\alpha u \in \mathbb{V}$ for all $\alpha \in \mathbb{R}$

(7) $\alpha(\beta u) = (\alpha\beta) u \qquad \alpha, \beta \in \mathbb{R}$

(8) $(\alpha + \beta) u = \alpha u + \beta v$

(9) $\alpha(u + v) = \alpha u + \alpha v$

(10) $1 u = u \qquad \forall u \in \mathbb{V}$

Examples of Linear Spaces

$V = \mathbb{R}$ (set of real numbers)

$\mathbb{V} = \mathbb{R}^{n \times n}$ (set of n by n matrices)

$\mathbb{V} = C[0,1]$ (set of cont fns on $[0,1]$)

Example of a space that is not linear

$$\mathbb{V} = A = \{ y : y \in C[0,1],\ y(0) = 0,\ y(1) = 3 \}$$

If $u(x) \in A$ and $v(x) \in A$ then $u + v \notin \mathbb{V}$ since $u(1) + v(1) = 6 \neq 3$. This set is not closed under addition

Normed Linear Spaces

Many function spaces are normed linear spaces which have a notion of distance associated with them

Defn A linear space V is a normed linear space if it has a norm $\|\cdot\|$ satisfying

(1) $\|y\| = 0 \iff y = 0$

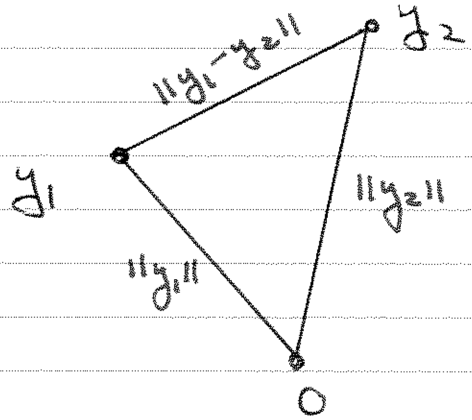
(2) $\|\alpha y\| = |\alpha| \, \|y\| \qquad \alpha \in \mathbb{R}, \ y \in V$

(3) $\|y_1 + y_2\| \leq \|y_1\| + \|y_2\| \qquad \forall y_1, y_2 \in V$

The last condition is called the triangle inequality and is used to define the distance between elements:

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

Using this defn of distance mimics that of $\mathbb{R}^3$



EXAMPLES OF NORMED LINEAR FUNCTION SPACES

Let $V = C[0,1]$. Then each of the following are norms for V

$$\|y\|_\infty = \max_{x \in [0,1]} |y(x)| \qquad \text{Strong norm}$$

$$\|y\|_1 = \int_0^1 |y(x)|\,dx \qquad L^1 \text{ norm}$$

$$\|y\|_2 = \left(\int_0^1 |y(x)|^2 dx\right)^{1/2} \qquad L^2 \text{ norm}$$

Note in particular that $\|y\| = 0$ in all the above implies $y(x) \equiv 0$, the zero function, since $y(x)$ must be continuous.

As an example of computing distances.

$$y_1(x) = x \qquad y_2(x) = -x$$

then $y_1(x) - y_2(x) = 2x$ and

$$\|y_1 - y_2\|_\infty = \max_{x \in [0,1]} |2x| = 2$$

$$\|y_1 - y_2\|_1 = \int_0^1 2x\,dx = 1$$

$$\|y_1 - y_2\|_2 = \left(\int_0^1 (2x)^2 dx\right)^{1/2} = \frac{2}{\sqrt{3}}$$

Continuous Functionals

Defn A real valued function $f(x)$ is continuous at x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

The definition for functionals on normed linear spaces is similar

Defn Let $A \subset V$ where V is a normed linear space. Let $J : A \to \mathbb{R}$. The functional J is continuous at $y_0 \in A$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\| y - y_0 \| < \delta \Rightarrow | J(y) - J(y_0) | < \varepsilon$$

EXAMPLE $A = C[0,1]$, $J(y) = \int_0^1 y(x)\,dx$

Is continuous functional using the strong norm

$$
\begin{aligned}
|J(y) - J(y_0)| &= \left| \int_0^1 y(x) - y_0(x)\,dx \right| \\
&\le \int_0^1 |y(x) - y_0(x)|\,dx \\
&\le \| y - y_0 \|_\infty \equiv \max_{x \in [0,1]} |y(x) - y_0(x)|
\end{aligned}
$$

Hence if $\delta = \varepsilon$ we'll have $\| y - y_0 \|_\infty < \varepsilon \Rightarrow$

$$| J(y) - J(y_0) | < \varepsilon$$

Local and Global Minima for $J(y)$

Let $J : \mathcal{A} \to \mathbb{R}$ for some admissible set $\mathcal{A}$.

Defn $\bar{y}$ is a global minima for J on $\mathcal{A}$ iff

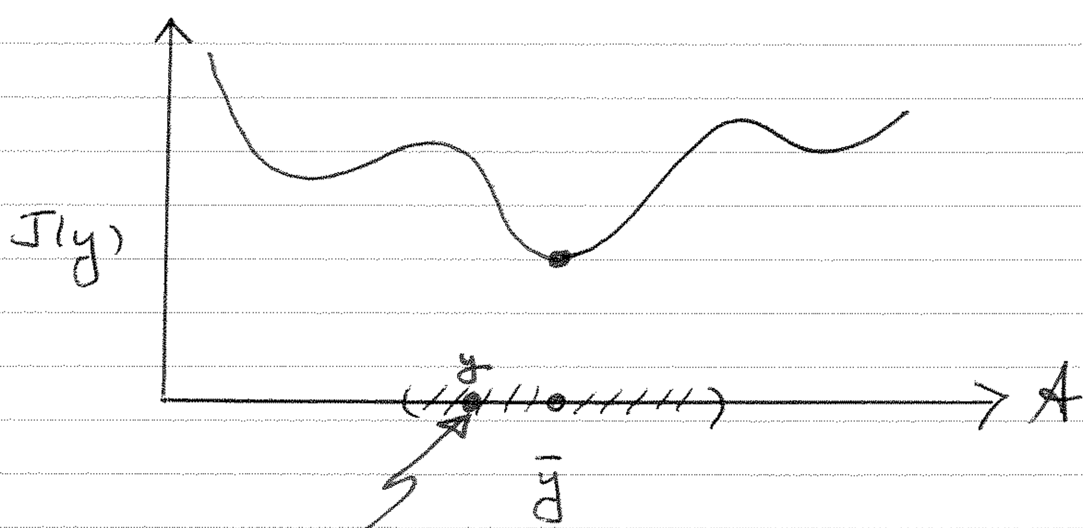
$$J(\bar{y}) \leq J(y) \qquad \forall y \in \mathcal{A}.$$

This definition does not require any notion of distance between y and $\bar{y}$.

Defn Let $\mathcal{A} \subset V$ where V is a normed linear space and $J : \mathcal{A} \to \mathbb{R}$. $\bar{y}$ is a local minima for J if there exists a $\delta > 0$ such that

$$\| y - \bar{y} \| < \delta \quad \Rightarrow \quad J(\bar{y}) \leq J(y)$$

Here $J(y)$ is larger for y "near" $\bar{y}$

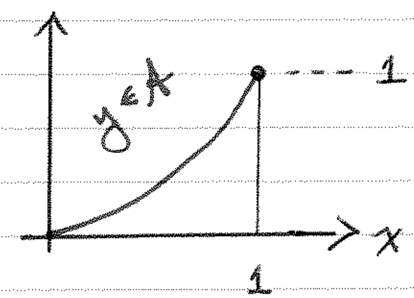


Some y near $\bar{y}$, i.e. $\| y - \bar{y} \| < \delta$

EXAMPLE   Explicit Local Min Example

$$J(y) \equiv 30 \int_0^1 (1 + y^2 + y'^2)\,dx$$

where the admissible set will be quadratics thru $(0,0)$ and $(1,1)$



$$\mathcal{A} = \{ y : y(x) = a x(x-1) + x \}$$

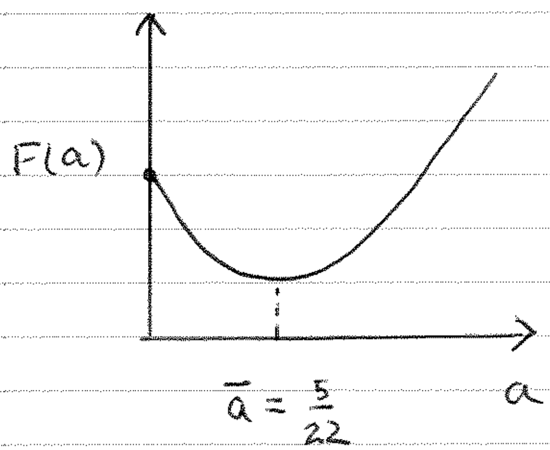
Note that for any $a \in \mathbb{R}$

$$y(0) = 0 \qquad y(1) = 1$$

This is a one parameter family of admissible functions. $\mathcal{A}$ is a very restrictive and small set of functions. Nevertheless we can illustrate the local min idea.

$$J(y) = 11a^2 - 5a + 70 = F(a), \quad y \in \mathcal{A}$$

which we can plot noting $F'(a) = 22a - 5$



At least over the $y \in \mathcal{A}$, the smallest $J(y)$ can be is

$$J(\bar{y}) = \frac{3055}{44}$$

for $a = \bar{a} = 5/22$.

Admissible Variations

Suppose $\bar{y} \in \mathcal{A}$ is a candidate for a local min or max of $J(y)$.

One might test this by asking if

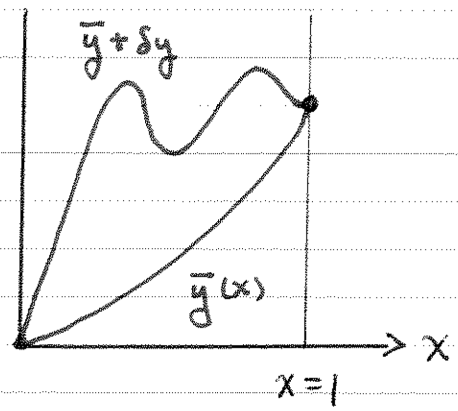
$$J(\bar{y}) \leq J(\bar{y} + \delta y)$$

where $\delta y \in V$ is some admissible variation. δy is admissible variation only if

$$\bar{y} + \delta y \in \mathcal{A}$$

Not all variations are admissible.

EX $\quad \mathcal{A} = \{ y : y \in C[0,1],\ y(0) = 0,\ y(1) = 1 \}$



Here the admissible variation is a function $\delta y(x)$

Given $\bar{y}(1) = 1$, any admissible variation must have $\delta y(1) = 0$ so that $\bar{y}(1) + \delta y(1) = 1$, and that $y + \delta y \in \mathcal{A}$.

In fact the set $\mathcal{A}^*$ of all admissible variations is

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

$\mathcal{A}$ and $\mathcal{A}^*$ are not the same sets!

First Variation (Gateaux Derivative)

Let $A \subset V$ where V is a normed linear space. Also let $J : A \to \mathbb{R}$.

Further let

$$\bar{y}(x) \in A$$

and that for some fixed $h(x) \in V$

$$\bar{y}(x) + \varepsilon h(x) \in A \qquad \text{(Admissible)}$$

for ε small.

Define

$$F(\varepsilon) = J(\bar{y} + \varepsilon h)$$

If $F(\varepsilon)$ is sufficiently smooth in ε

$$\Delta J \equiv F(\varepsilon) - F(0) = F'(0)\,\varepsilon + o(\varepsilon)$$

since it has a Taylor Series in ε.

We use this to define the "derivative" or first variation of J at $\bar{y}$

Definition Let $J : A \to \mathbb{R}$, $\bar{y} \in A$ and suppose $\bar{y} + \varepsilon h \in A$ for all $\varepsilon \in \mathbb{R}$ sufficiently small.

The first variation $\delta J(\bar{y}, h)$ of J at $\bar{y}$ in the direction h is

$$\delta J(\bar{y}, h) = F'(0) = \frac{d}{d\varepsilon} J(\bar{y} + \varepsilon h)\Big|_{\varepsilon = 0}$$

(if it exists).

EXAMPLE Define $J(y) = y(0)^2$ $\qquad A = C[0,1]$

$$J(y + \varepsilon h) = (y(0) + \varepsilon h(0))^2$$

$$F(\varepsilon) = J(y + \varepsilon h) = y(0)^2 + 2\varepsilon y(0) h(0) + \varepsilon^2 h(0)^2$$

Thus

$$F'(\varepsilon) = 2y(0)h(0) + 2\varepsilon h(0)$$

Evaluating at $\varepsilon = 0$ we arrive at

$$\delta J(y, h) = 2y(0)h(0)$$

Is much like the directional derivative of $f : \mathbb{R}^n \to \mathbb{R}^n$. The "direction" is the function h.

<u>EXAMPLE</u> $J(y) = \int_0^1 (1 + y'(x)^2)\,dx$ on

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

Some admissible variations are

$$\delta y(x) \quad = \quad \varepsilon h(x) \quad = \varepsilon x(1-x)$$

so that for any $\bar{y} \in \mathcal{A}$, $\bar{y} + \delta y \in \mathcal{A}$ as well.

For clarity, suppose

$$\bar{y}(x) = x \qquad\qquad \delta y(x) = \varepsilon h(x)$$

Then

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

After some calculations

$$F(\varepsilon) = J(\bar{y} + \varepsilon h) = 2 + \tfrac{1}{3}\varepsilon^2$$

so that

$$\delta J(\bar{y}, h) = \frac{d}{d\varepsilon}\left(2 + \tfrac{1}{3}\varepsilon^2\right)\Big|_{\varepsilon=0} = 0$$

Necessary Conditions for Local Min (Max)

If $\bar{y}$ is a local min of J then

$$(1) \qquad \Delta J = J(\bar{y}+\varepsilon h) - J(\bar{y}) \geq 0$$

for all admissible variations $\delta y = \varepsilon h(x)$

In this setting $y = \bar{y} + \varepsilon h$ and $\bar{y}$ are near each other if ε is small since

$$\|y - \bar{y}\| = \|\varepsilon h\| = \varepsilon \|h\| \to 0 \qquad \text{as } \varepsilon \to 0.$$

Since

$$\Delta J = \delta J(\bar{y}, h)\,\varepsilon + o(\varepsilon)$$

a necessary condition for $\Delta J \geq 0$ is that δJ vanish (noting though ε issmall it can be positive or negative).

This argument is true for every $h(x)$ for which $\varepsilon h(x)$ is an admissible variation

<u>Theorem</u> $\bar{y}$ is a local min of $J : A \to \mathbb{R}$ only if

$$\delta J(\bar{y}, h) = 0$$

for all admissible variations $\delta y = \varepsilon h$.

The theorem is true for local max $\bar{y}$ as well.

Theorem Let $g(x) \in C[a,b]$ and suppore

$$(1) \qquad \int_a^b g(x)\,v(x)\,dx = 0 \qquad \forall v \in A^*$$

where A^* is any one of the following spaces

$$A^* = C[a,b]$$

$$A^* = C^n[a,b]$$

$$A^* = \{y : y \in C^n[a,b],\ y(a) = y(b) = 0\}$$

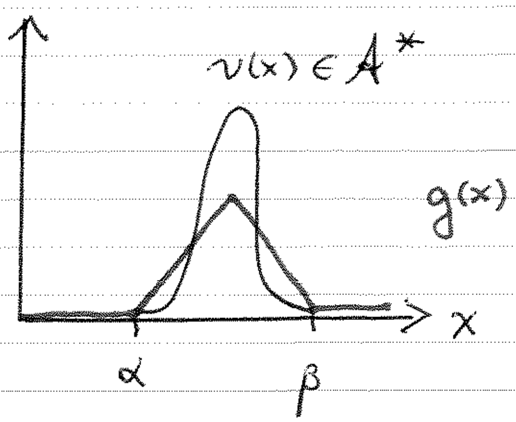
Then $g(x) \equiv 0$.

Proof Outline (By contradiction)

Suppose (1) is true by $g(x)$ is not identically zero. Since $g(x)$ is continuous there must be some interval $[\alpha, \beta] \in [0, 1]$ on which it is nonzero. Without loss of generality $g(x) > 0$ on $I = [\alpha, \beta]$.

Pick any $v \in A^*$ positive on I but zero elsewhere. Then

$$\int_a^b g(x)\,v(x)\,dx = \int_\alpha^\beta g(x)\,v(x)\,dx > 0$$



contradicts (1)