
Discussion Notes for Methods of Applied Math (MATH 146)

UCLA
WINTER 2018

CREATED BY
HOWARD HEATON

© Copyright 2018 Howard W. Heaton

Purpose: This document is a compilation of notes generated for discussion in MATH 146 with reference credit due John L. Troutman's text *Variational Calculus and Optimal Control* [3]. If the reader finds any errors/typos, please feel free to email me at heaton@math.ucla.edu and I will address these and post an updated set of notes to my [webpage](#).

CONTENTS

Introduction	1
Review Material	2
Simple Methods for Finding Minimizers of $J : V \rightarrow \mathbb{R}$	8
Min-Max Problems	14
The Gâteaux Derivative & Extremals	19
Euler-Lagrange Equations	32
Constrained Problems	40
Convexity	45
Legendre Transform	47
Midterm II Material	53
Problems for Students	54
References	57

INTRODUCTION

These notes are provided to compliment the TA discussion sessions on Thursdays for MATH 146. Typically, more detail is provided here than on the board during discussions since portions of solutions are given orally in class. The examples provided here are meant to be a constructive reference for students. These illustrate how to use set up the variational problems we will see this quarter, what details are important to include, and provide example of acceptable solution presentation. Before reading each solution, I highly encourage students to first seriously attempt the problems on their own. I cannot overstate the value of struggling through these problems before comparing your attempts to the example solutions.

These notes will be updated weekly (if not more often), reflecting the current discussion material.

REVIEW MATERIAL

The next definition follows that that of the analysis text often called *Baby Rudin* [1, Definition 5.1].

Definition: Define $f : [a, b] \rightarrow \mathbb{R}$. For any $x \in [a, b]$, define the quotient

$$\phi(t) := \frac{f(t) - f(x)}{t - x} \quad (a < t < b, t \neq x), \quad (1)$$

and define

$$f'(x) := \lim_{t \rightarrow x} \phi(t), \quad (2)$$

provided the limit exists. We associate the function f' with f at the points where the limit (2) exists. The function f' is called the *derivative* of f . If f' is defined at a point x , we say f is *differentiable* at x . And if f' is defined at every point in a set $I \subset [a, b]$, then we say f is differentiable on I . \triangle

Example 1: Use the above definition to compute $f'(1)$ for the function $f(x) = x^2$.

Solution:

Through direct computation, we find

$$\begin{aligned} f'(1) &= \lim_{t \rightarrow 1} \frac{f(t) - f(1)}{t - 1} \\ &= \lim_{h \rightarrow 0} \frac{f(1+h) - f(1)}{(1+h) - 1} \\ &= \lim_{h \rightarrow 0} \frac{(1+h)^2 - 1}{h} \\ &= \lim_{h \rightarrow 0} \frac{1 + 2h + h^2 - 1}{h} \\ &= \lim_{h \rightarrow 0} \frac{2h + h^2}{h} \\ &= \lim_{h \rightarrow 0} 2 + h \\ &= 2 + 0 \\ &= 2. \end{aligned} \quad (3)$$

□

Taylor's Theorem: Let $I \subset \mathbb{R}$ be a neighborhood of x_0 and n be a nonnegative integer. Suppose the function $f : I \rightarrow \mathbb{R}$ has $n + 1$ derivatives. Then for each point $x \neq x_0$ in I there is a point ξ strictly between x and x_0 such that

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n + 1)!} (x - x_0)^{n+1}. \tag{4}$$

△

REMARK 1: The second term in (4) is known as the *Lagrange Remainder*. ◇

Consider using Taylor's theorem when $n = 1$. That is, suppose f is twice differentiable at x and define

$$\varepsilon(h) := \frac{f^{(2)}(\xi(h))}{2} h^2 \tag{5}$$

where $\xi(h)$ is the point strictly between x and $x + h$ such that

$$f(x + h) = f(x) + f'(x)h + \varepsilon(h), \tag{6}$$

which we know exists by Taylor's theorem. This form of expansion will be useful for us to remember when we look at differentiation of more abstract quantities known as functionals. Furthermore, this shows

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \left(\frac{f(x+h) - f(x) - \varepsilon(h)}{h} \right) = \lim_{h \rightarrow 0} \left(\frac{f(x+h) - f(x)}{h} - \frac{\varepsilon(h)}{h} \right) = f'(x) - \lim_{h \rightarrow 0} \frac{\varepsilon(h)}{h}. \tag{7}$$

Thus $\lim_{h \rightarrow 0} \varepsilon(h)/h = 0$. Using little-oh notation (defined below), we write this as $\varepsilon(h) = o(h)$.

Definition: Assume $g(x)$ is nonzero. Then we say $f(x) = o(g(x))$ as $x \rightarrow x^*$ provided

$$\lim_{x \rightarrow x^*} \left| \frac{f(x)}{g(x)} \right| = 0. \tag{8}$$

This notation is referred to as *little-oh notation*. △

Example 2: Define $f(x) := x^2$. Express $f(x+h)$ explicitly in the form of (6).

Solution:

First observe $f'(x) = 2x$ and $f''(x) = 2$. Then we see

$$f(x+h) = (x+h)^2 = x^2 + 2xh + h^2 = f(x) + f'(x)h + \varepsilon(h) \quad (9)$$

where $\varepsilon(h) := h^2$. □

We now turn our attention to a necessary condition for a point \bar{x} to be a local minimizer of f .

Theorem: If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a continuously differentiable function and \bar{x} is a local minimizer of f , then $f'(\bar{x}) = 0$. △

Proof:

Let \bar{x} be a minimizer of f , i.e., there is a $\delta^* > 0$ such that $f(\bar{x}) \leq f(x)$ for all $x \in (\bar{x} - \delta^*, \bar{x} + \delta^*)$. We proceed by way of contradiction, i.e., suppose $f'(\bar{x}) \neq 0$. By hypothesis f' is continuous, and so there is a $\delta > 0$ such that

$$|z - \bar{x}| < \delta \quad \implies \quad |f'(z) - f'(\bar{x})| < \frac{|f'(\bar{x})|}{2}. \quad (10)$$

But, using the reverse triangle inequality, we see

$$|f'(\bar{x})| - |f'(z)| \leq |f'(z) - f'(\bar{x})| < \frac{|f'(\bar{x})|}{2} \quad \implies \quad \frac{|f'(\bar{x})|}{2} < |f'(z)|. \quad (11)$$

Suppose $f'(\bar{x}) > 0$ and pick $z \in (\bar{x} - \delta/2, \bar{x})$. Taylor's theorem asserts there is $\xi \in (z, \bar{x})$ such that

$$f(z) = f(\bar{x}) + f'(\xi)(z - \bar{x}) = f(\bar{x}) - f'(\xi)|z - \bar{x}| < f(\bar{x}) - \frac{|f'(\bar{x})|}{2}|z - \bar{x}| < f(\bar{x}). \quad (12)$$

This shows $f(z) < f(\bar{x})$ for all $z \in (\bar{x} - \delta/2, \bar{x})$. Thus \bar{x} cannot be a local minimizer of f , contradicting our initial assumption. Whence $f'(\bar{x}) \leq 0$. By analogous argument to above, if instead $f'(\bar{x}) < 0$, we pick $z \in (\bar{x}, \bar{x} + \delta/2)$ to deduce

$$f(z) = f(\bar{x}) + f'(\xi)(z - \bar{x}) = f(\bar{x}) + f'(\xi)|z - \bar{x}| < f(\bar{x}) - \frac{|f'(\bar{x})|}{2}|z - \bar{x}| < f(\bar{x}), \quad (13)$$

again giving a contradiction. This shows $f'(\bar{x}) \geq 0$. Therefore, combining our results, we conclude $f'(\bar{x}) = 0$, as desired. ■

REMARK 2: The above theorem shows that a necessary condition for \bar{x} to be a local minimizer of f is that $f'(\bar{x}) = 0$. Below we provide several examples illustrating the use and limitations of this theorem. \diamond

Example 3: Define $f(x) = (x - 3)^2 + 5x + 3$. Solve the optimization problem

$$\min_{x \in \mathbb{R}} f(x), \quad (14)$$

using only the above theorem and definition of a minimizer.

Solution:

First note f is continuously differentiable since it is a polynomial. And,

$$f'(x) = 2(x - 3) + 5 + 0 = 2x - 1. \quad (15)$$

The single critical point of f is at $x = 1/2$. The above theorem shows this is the only candidate solution to the optimization problem.

All that remains is to verify $x = 1/2$ is, in fact, a minimizer. We can rewrite f as $f(x) = x^2 - x + 12$. Pick any $z \in \mathbb{R}$ and set $\delta := z - 1/2$ so that $z = 1/2 + \delta$. Then

$$\begin{aligned} f(z) &= f\left(\frac{1}{2} + \delta\right) = \left(\frac{1}{2} + \delta\right)^2 - \left(\frac{1}{2} + \delta\right) + 12 \\ &= \left(\frac{1}{4} + \delta + \delta^2\right) - \left(\frac{1}{2} + \delta\right) + 12 \\ &= \left(\frac{1}{4} - \frac{1}{2} + 12\right) + \delta^2 \\ &= f\left(\frac{1}{2}\right) + \delta^2 \\ &\geq f\left(\frac{1}{2}\right). \end{aligned} \quad (16)$$

This shows $f(1/2) \leq f(z)$ for all $z \in \mathbb{R}$, i.e., $1/2$ is the global minimizer of f , and we are done. \square

Example 4: Define $f(x) = x^3$. Can the above theorem be applied to find a local minimum?

Solution:

Observe $f'(x) = 3x^2$ and so $f'(x) = 0$ if and only if $x = 0$. But, $f(0) = 0 > -\varepsilon^3 = f(-\varepsilon)$ for every $\varepsilon > 0$ and so 0 is not a local minimum of f . Thus the above theorem cannot be applied to find a local minimum. Moreover, because this was the only candidate for a minimizer, we are able to further conclude f has no global minimizer over \mathbb{R} . \square

REMARK 3: The above theorem shows that the condition $f'(x) = 0$ is necessary, but *not* sufficient. We illustrate this again with the following example. \diamond

Example 5: Define $f(x) = -x^2$. Can the above theorem be applied to find a local minimum?

Solution:

Observe $f'(x) = -2x$ and so $f'(x) = 0$ if and only if $x = 0$. But, $f(z) = -z^2 < 0 = f(0)$ for all $z \neq 0$. This shows 0 is not a local minimum of f . Thus the above theorem cannot be applied to find a local minimum. In fact, the above shows $x = 0$ is a global maximizer of f . \square

Example 6: Define $f(x) := 3|x - 5|$. What is the global minimizer of f and can the above theorem be applied? Explain.

Solution:

The global minimizer is $x = 5$. Indeed,

$$f(5) = 0 \leq 3|x - 5| = f(x) \quad \forall x \in \mathbb{R}. \quad (17)$$

However, f is not continuously differentiable since f' is not continuous at $x = 5$. Indeed,

$$\lim_{x \rightarrow 5^-} f'(x) = -3 \neq 3 = \lim_{x \rightarrow 5^+} f'(x). \quad (18)$$

Thus a condition for the theorem does not hold and so it cannot be applied. \square

REMARK 4: Here we review integration by parts. Let $f, g \in C^1[a, b]$. Then using the product rule we write

$$\frac{d}{dx} [f(x)g(x)] = f'(x)g(x) + f(x)g'(x). \quad (19)$$

So,

$$\int_a^b \frac{d}{dx} [f(x)g(x)] \, dx = \int_a^b f'(x)g(x) + f(x)g'(x) \, dx. \quad (20)$$

But, the left hand side can be rewritten as

$$\int_a^b \frac{d}{dx} [f(x)g(x)] \, dx = \int_{f(a)g(a)}^{f(b)g(b)} d(fg) = [f(x)g(x)]_{x=a}^{x=b} = f(b)g(b) - f(a)g(a). \quad (21)$$

Thus the integration by parts formula becomes

$$\int_a^b f(x)g'(x) \, dx = - \int_a^b f'(x)g(x) \, dx + [f(x)g(x)]_{x=a}^{x=b}. \quad (22)$$

This will be especially useful tool for us and is important to have at our disposal. \diamond

SIMPLE METHODS FOR FINDING MINIMIZERS OF $J : V \rightarrow \mathbb{R}$

REMARK 5: For the following problems, we proceed roughly by taking the following steps.

1. Find a lower bound ℓ for $J(y)$ (the tightest lower bound we can establish).
2. Find a collection of candidates y for which $J(y)$ equals this lower bound, i.e., $J(y) = \ell$.
3. Any functions in the collection of candidates contained in \mathcal{A} are minimizers. If at least one candidate is contained in \mathcal{A} , then ℓ is the minimum.

This approach works when the minimizer is obtained. However, what can we do when it isn't obtained? In this case, we might attempt to do as follows.

1. Find a lower bound ℓ for $J(y)$ (the tightest lower bound we can establish).
2. Find a collection of candidates y for which $J(y)$ equals this lower bound.
3. If none of the candidates y is admissible (i.e., $y \notin \mathcal{A}$), we must find a sequence $\{y^n\}_{n=1}^{\infty}$ of functions in \mathcal{A} for which

$$\lim_{n \rightarrow \infty} J(y^n) = \ell. \tag{23}$$

Then ℓ is the infimum and it is *not* obtained.

◇

Definition: Let V be a vector space and $J : V \rightarrow \mathbb{R}$ be a mapping. Let \mathcal{A} be a subset of V , i.e., $\mathcal{A} \subset V$. Then we say $y \in \mathcal{A}$ is a *global minimizer* of J over \mathcal{A} provided

$$J(y) \leq J(z) \quad \forall z \in \mathcal{A}. \quad (24)$$

△

Example 7: Define the admissibility class $\mathcal{A} := C[0, 2]$ and let $J : C[0, 2] \rightarrow \mathbb{R}$ be the functional defined by

$$J(y) := \int_0^2 [y(x) - 9]^2 + 7 \, dx. \quad (25)$$

Find the minimum of J over \mathcal{A} . What is the minimizer?

Solution:

Let $y \in \mathcal{A}$. Then

$$J(y) = \int_0^2 [y(x) - 9]^2 + 7 \, dx \geq \int_0^2 0 + 7 \, dx = 14, \quad (26)$$

where the inequality holds since $[y(x) - 9]^2 \geq 0$ for all possible values of $y(x)$. This shows 14 is a lower bound for $J(y)$. To verify this is the minimum for $J(y)$, it suffices to find $f \in \mathcal{A}$ such that $J(f) = 14$. This is accomplished if and only if the inequality in (26) is a strict equality. The only candidate is $f(x) = 9$ since this would give

$$[f(x) - 9]^2 = [9 - 9]^2 = 0^2 = 0. \quad (27)$$

Indeed, this implies

$$J(y) \geq 14 = J(f) \quad \forall y \in \mathcal{A}. \quad (28)$$

Because f is constant, it is continuous on $[a, b]$, and so $f \in \mathcal{A}$. Thus we conclude $f(x) = 9$ is the minimizer of $J(y)$ over \mathcal{A} and 14 is the minimum of $J(y)$ over \mathcal{A} . □

Example 8: Define the admissibility class $\mathcal{A} := \{f \in C[a, b] : f(x) \geq 5\}$ and let $J : C[a, b] \rightarrow \mathbb{R}$ be the functional defined by

$$J(y) := \int_a^b y(x)^2 - 8y(x) + 20 \, dx. \quad (29)$$

Find the minimum of J over \mathcal{A} . What is the minimizer?

Solution:

Let $y \in \mathcal{A}$. Then

$$\begin{aligned} J(y) &= \int_a^b y(x)^2 - 8y(x) + 20 \, dx \\ &= \int_a^b (y(x)^2 - 8y(x) + 16) + 4 \, dx \\ &= \int_a^b (y(x) - 4)^2 + 4 \, dx \\ &\geq \int_a^b (5 - 4)^2 + 4 \, dx \\ &= \int_a^b 5 \, dx \\ &= 5(b - a). \end{aligned} \quad (30)$$

This shows $5(b - a)$ is a lower bound for $J(y)$. To verify this is the minimum for $J(y)$, it suffices to find $f \in \mathcal{A}$ such that $J(f) = 5(b - a)$. This is accomplished if and only if the inequality in (30) is a strict equality. The only candidate is $f(x) = 5$. Since f is continuous on $[a, b]$ and $f \geq 5$, we see $f \in \mathcal{A}$. Thus we conclude $\boxed{f(x) = 5}$ is the minimizer of $J(y)$ over \mathcal{A} and $\boxed{5(b - a)}$ is the minimum of $J(y)$ over \mathcal{A} . \square

REMARK 6: Note in the above example we say $f(x) = 5$ is “the” minimizer. This is because this is the only function in \mathcal{A} that gives $J(f) = 5(b - a)$. In the next example, multiple minimizers exist. \diamond

REMARK 7: Note \mathcal{A} does not form a vector space in the following example. This follows from the fact it is not closed under scalar multiplication. For example, if $f \in \mathcal{A}$, then $-f \notin \mathcal{A}$. \diamond

Example 9: Define the admissibility class $\mathcal{A} := \{f \in C[0, 1] : f(x) \geq x^2 - 10x + 28\}$. Then let $J : C[0, 1] \rightarrow \mathbb{R}$ be the functional defined by

$$J(f) := \inf_{x \in [0, 1]} f(x). \quad (31)$$

Find $\inf_{f \in \mathcal{A}} J(f)$. Does $J(f)$ attain its infimum?

Solution:

Let $f \in \mathcal{A}$. Then, for each $x \in [0, 1]$,

$$f(x) \geq x^2 - 10x + 28 = (x^2 - 10x + 25) + 3 = (x - 5)^2 + 3. \quad (32)$$

Set $g(x) := (x - 5)^2 + 3$. Also note $g'(x) = 2(x - 5) < 0$ for $x < 5$, and so g is strictly decreasing on $[0, 1]$. This implies $\inf_{x \in [0, 1]} g(x) = g(1)$. Using this fact, we see

$$J(f) = \inf_{x \in [0, 1]} f(x) \geq \inf_{x \in [0, 1]} g(x) = g(1) = (1 - 5)^2 + 3 = 19. \quad (33)$$

This shows $J(f) \geq 19$, i.e., 19 is a lower bound. Moreover, because g is a polynomial, it is continuous. Whence $g \in \mathcal{A}$ and

$$J(f) \geq J(g) = 19 \quad \forall f \in \mathcal{A}. \quad (34)$$

Thus g is a minimizer of J over \mathcal{A} and so $\boxed{\inf_{f \in \mathcal{A}} J(f) = 19}$. Yes, $J(f)$ attains its infimum. \square

REMARK 8: Note in the above example we say g is “a” minimizer. In general, there may be multiple minimizers. For instance, in the above example consider defining $q(x) := g(x) + (x - 1)^2$. Then $q \in C[0, 1]$ and $q(x) = g(x) + (x - 1)^2 \geq g(x)$, which implies $q \in \mathcal{A}$. Moreover,

$$q'(x) = g'(x) + 2(x - 1) = 2(x - 5) + 2(x - 1) \leq 2(x - 5) + 0 < 0 \quad \forall x \in [0, 1]. \quad (35)$$

This shows q is strictly decreasing on $[0, 1]$. Thus

$$J(q) = \inf_{x \in [0, 1]} q(x) = q(1) = g(1) + (1 - 1)^2 = g(1) = 19. \quad (36)$$

This shows g and q are minimizers of J over \mathcal{A} . \diamond

Example 10: Define the function $h : \mathbb{R} \rightarrow \mathbb{R}$ by

$$h(x) := \begin{cases} 0 & \text{if } |x| < 1, \\ 1 & \text{if } |x| \geq 1. \end{cases} \quad (37)$$

Define the admissibility class $\mathcal{A} := \{f \in C^1(\mathbb{R}) : f(x) \geq h(x)\}$. Then let $J : C^1(\mathbb{R}) \rightarrow \mathbb{R}$ be the functional

$$J(y) := \int_{-1}^1 y(x) \, dx. \quad (38)$$

Compute $\inf_{y \in \mathcal{A}} J(y)$. Does J attain its infimum?

Solution:

We proceed as follows. First we find a lower bound for J over \mathcal{A} . Then we show this is the greatest lower bound for J over \mathcal{A} . Lastly, we remark why J does *not* attain its infimum, i.e., there is no minimizer in \mathcal{A} . Note, for $y \in \mathcal{A}$,

$$J(y) = \int_{-1}^1 y(x) \, dx \geq \int_{-1}^1 h(x) \, dx = \int_{-1}^1 0 \, dx = 0. \quad (39)$$

This shows 0 is a lower bound for $J(y)$. We claim there is a sequence of functions $\{f_n\}_{n=1}^{\infty}$ contained in \mathcal{A} such that $J(f_n) \rightarrow 0$. This implies there is no lower bound greater than zero and, therefore, 0 must be the greatest lower bound for J . In other words, $0 = \inf_{y \in \mathcal{A}} J(y)$.

All that remains is to verify the claimed sequence $\{f_n\}_{n=1}^{\infty}$ exists. Define $f_n(x) := x^{2n}$ for $n \geq 1$. Then $f_n(x) = x^{2n} \geq 0 = h(x)$ for $|x| < 1$ and $f_n(x) = x^{2n} \geq 1^{2n} = 1 = h(x)$ for $|x| \geq 1$. Hence $f_n \geq h$ and, with the fact f is a polynomial (and thus smooth), we see $f_n \in \mathcal{A}$ for each n . Then computing $J(f_n)$ gives

$$J(f_n) = \int_{-1}^1 f_n(x) \, dx = \int_{-1}^1 x^{2n} \, dx = 2 \int_0^1 x^{2n} \, dx = 2 \left(\frac{1^{2n+1}}{2n+1} \right) = \frac{2}{2n+1} \leq \frac{1}{n} \quad (40)$$

Taking the limit as $n \rightarrow \infty$, we see

$$0 \leq \lim_{n \rightarrow \infty} J(f_n) \leq \lim_{n \rightarrow \infty} \frac{1}{n} = 0. \quad (41)$$

Thus $\lim_{n \rightarrow \infty} J(f_n) = 0$, as desired.

Lastly, we note J does not attain its infimum. This is because the infimum is obtained if and only if $y(x) = 0$ for $|x| < 1$. But, because we need $y(\pm 1) \geq 1$, such a minimizer would necessarily have a jump discontinuity, contradicting the fact $y(x)$ must be continuous to be in \mathcal{A} . \square

REMARK 9: After reading the above example, we may ask ourselves “But why did you pick $f_n(x) = x^{2n}$? How did you know to do that?”. I encourage the reader to draw a picture. A good picture can go a long way.

We want a continuous function f with $f(-1) \geq 1$ and $f(1) \geq 1$, but approaches 0 for $|x| < 1$. To keep things simple, we may restrict our consideration to even functions. Perhaps an initial guess might be to use x^2 to get an even function with $(-1)^2 = 1 = 1^2$. Then because $|x| < 1$, we know $|x|^n \rightarrow 0$ as $n \rightarrow \infty$ (see Lemma below). So, we could try $(x^2)^n = x^{2n}$. Indeed, we see graphically below this does do the trick.

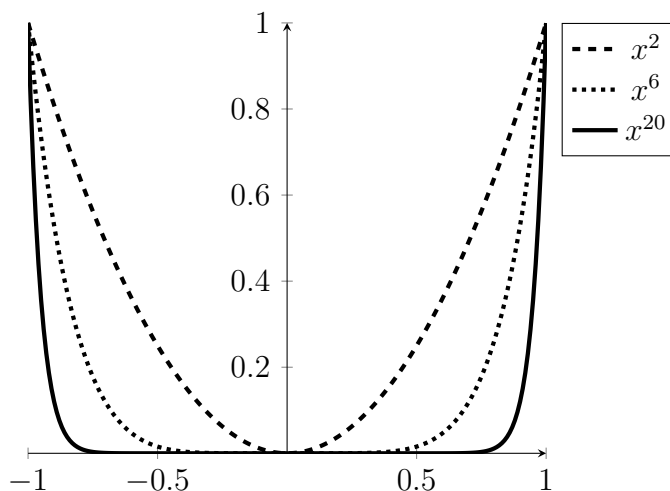


Figure 1: Plots of x^{2n} on $[-1, 1]$ for $n = 1, 3, 10$. \diamond

Lemma: Let $c \in (0, 1)$. The $\lim_{n \rightarrow \infty} c^n = 0$. \triangle

Proof:

Let $n \in \mathbb{N}$. Then $c^{n+1} = cc^n < 1c^n = c^n$. This shows the sequence $\{c^n\}_{n=1}^\infty$ is decreasing. And, the fact $c^n \geq 0^n = 0$ shows it is bounded from below. The Monotone Convergence Theorem then asserts $\{c^n\}_{n=1}^\infty$ converges to some limit $\alpha \in \mathbb{R}$. Observe

$$\alpha = \lim_{n \rightarrow \infty} c^n = \lim_{n \rightarrow \infty} c^{n+1} = c \lim_{n \rightarrow \infty} c^n = c\alpha. \tag{42}$$

Because $c \in (0, 1)$, the above can hold if and only if $\alpha = 0$. Thus $\lim_{n \rightarrow \infty} c^n = 0$. \blacksquare

MIN-MAX PROBLEMS

In this section, we discuss how a constrained minimization problem can be turned into an equivalent min-max problem. To illustrate this, we will first work with an example of a minimization problem in \mathbb{R} , and then we will take the ideas from there and apply them to finding minimizers in a space of functions (e.g., $C^2[a, b]$).

Example 11: Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $f(x) := 5(x - 7)^2$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $g(x) := x^3$. Express the constrained minimization problem

$$\min_{x \in \mathbb{R}} f(x) \quad \text{s.t.} \quad g(x) = 8 \tag{43}$$

as an unconstrained min-max problem.

Solution:

Observe our optimization problem can be rewritten as

$$\min_{x \in \mathbb{R}} f(x) \quad \text{s.t.} \quad g(x) = 8, \tag{44}$$

which is equivalent to

$$\min_{x \in \mathbb{R}} \begin{cases} 5(x - 7)^2 & \text{if } (x^3 - 8) = 0, \\ \infty & \text{otherwise.} \end{cases} \tag{45}$$

We can then express this problem as

$$\boxed{\min_{x \in \mathbb{R}} \max_{\lambda \in \mathbb{R}} 5(x - 7)^2 + \lambda(x^3 - 8)}. \tag{46}$$

□

REMARK 10: The step to rewrite the constrained problem as (45) initially seems as though we are moving backwards; however, this makes the following form in (46) more clear. Indeed, if $(x^3 - 8)$ is not zero, then we can pick λ to make this as big as we'd like. For example, if $x^3 - 8 = \alpha > 0$, then we may heuristically write

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} 5(x-7)^2 + \lambda(x^3 - 8) &= \lim_{\lambda \rightarrow \infty} 5(x-7)^2 + \lambda\alpha \\ &= 5(x-7)^2 + \lim_{\lambda \rightarrow \infty} \lambda\alpha \\ &= 5(x-7)^2 + \infty \\ &= \infty. \end{aligned} \tag{47}$$

We could do similarly taking $\lambda \rightarrow -\infty$ if $x^3 - 8 < 0$. This same idea will next be used for a constrained optimization problem using a functional. \diamond

REMARK 11: Suppose we wish to find the minimizer of J over a set X . Here we will consider the problem unconstrained when X is a vector space without any constraints other than those imposed on the smoothness. So, if

$$X := \left\{ y \in C^2[a, b] : \int_a^b y \, dx = 0 \right\}, \tag{48}$$

and we are given a functional $J : C^2[a, b] \rightarrow \mathbb{R}$, then the problem

$$\min_{y \in X} J(y) \tag{49}$$

is here considered a constrained problem. However,

$$\min_{y \in C^2[a, b]} J(y) \tag{50}$$

is here considered unconstrained. \diamond

Example 12: Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be continuous, i.e., $f \in C(\mathbb{R})$. Also assume $q : [0, 1] \rightarrow \mathbb{R}$ is continuous and define $J : C[0, 2] \rightarrow \mathbb{R}$ by

$$J(y) := \int_0^2 f(y(x)) \, dx \tag{51}$$

and set

$$X := \{y \in C^2[0, 2] : y''(x) = q(x) \, \forall x \in [0, 1]\}. \tag{52}$$

Rewrite the constrained optimization problem

$$\min_{y \in X} J(y) = \min_{y \in C^2[0, 2]} J(y) \text{ s.t. } y''(x) = q(x) \, \forall x \in [0, 1] \tag{53}$$

as an unconstrained min-max problem.

Solution:

Observe

$$\min_{y \in X} J(y) = \min_{y \in C^2[0, 2]} \begin{cases} J(y) & \text{if } y''(x) - q(x) = 0 \text{ for all } x \in [0, 1], \\ \infty & \text{otherwise.} \end{cases} \tag{54}$$

To ensure the constrained $y''(x) - q(x) = 0$ holds for all $x \in [0, 1]$, we need a function $\lambda(x) \in C[0, 1]$. Indeed, then (54) becomes

$$\boxed{\min_{y \in C^2[0, 2]} \max_{\lambda \in C[0, 1]} J(y) + \int_0^1 \lambda(x) [y''(x) - q(x)] \, dx.} \tag{55}$$

Alternatively, we could write this problem as

$$\boxed{\min_{y \in C^2[0, 2]} \max_{\lambda \in \mathbb{R}} J(y) + \lambda \int_0^1 [y''(x) - q(x)]^2 \, dx.} \tag{56}$$

□

REMARK 12: Why do we need λ to be a function in (55) rather than simply a number in \mathbb{R} ? The answer is this. If we merely impose that $\lambda \in \mathbb{R}$, then

$$0 = \int_0^1 \lambda y''(x) - q(x) \, dx = \lambda \int_0^1 y''(x) - q(x) \, dx, \tag{57}$$

which holds whenever the average of $y'' - q$ on $[0, 1]$ is zero. For example, if $y''(x) - q(x) = x - 1/2$, then for each λ we have

$$\lambda \int_0^1 y''(x) - q(x) \, dx = \lambda \int_0^1 x - 1/2 \, dx = \lambda \left[\frac{x^2}{2} - \frac{x}{2} \right]_0^1 = \lambda 0 = 0. \tag{58}$$

But, $x - 1/2$ is not identically zero for all $x \in [0, 1]$. This is why we must use a function $\lambda(x)$ in (55).

A simpler route, which does allow for λ to be a scalar is given in (56). For there the integral term with y'' and q is equal to zero if and only if $y''(x) = q(x)$ for all $x \in [0, 1]$. \diamond

REMARK 13: In the next example, we impose two constraints on a minimization problem. \diamond

Example 13: For a functional $J : C[0, 1] \rightarrow \mathbb{R}$, rewrite the constrained minimization problem

$$\min_{y \in C^1[0,1]} J(y) \quad \text{s.t.} \quad \int_0^1 y^2 \, dx = 5, \quad \int_0^1 y' \, dx = 0. \tag{59}$$

as an unconstrained problem.

Solution:

Here the optimization problem may be rewritten as

$$\min_{y \in C^1[0,1]} \max_{\lambda \in \mathbb{R}} J(y) + \lambda \left[\int_0^1 y^2 \, dx - 5 \right] \quad \text{s.t.} \quad \int_0^1 y' \, dx = 0. \tag{60}$$

We now have a min-max problem, but there is still a constraint involved. So, we add another parameter μ to obtain the unconstrained problem

$$\boxed{\min_{y \in C^1[0,1]} \max_{\lambda \in \mathbb{R}} \max_{\mu \in \mathbb{R}} J(y) + \lambda \left[\int_0^1 y^2 \, dx - 5 \right] + \mu \int_0^1 y' \, dx.} \tag{61}$$

\square

Example 14: Consider the optimization problem

$$\min_{y \in C^1[0,1]} \int_0^1 (y' - 2)^2 dx \quad \text{s.t.} \quad \int_0^1 y dx = 2. \quad (62)$$

- a) Rewrite the constrained problem as an unconstrained min-max problem.
- b) Solve the constrained minimization problem. What is the minimizer? What is the minimum?

Solution:

a) We can rewrite the constrained problem as

$$\min_{y \in C^1[0,1]} \begin{cases} \int_0^1 (y' - 2)^2 dx & \text{if } (\int_0^1 y dx - 2) = 0, \\ \infty & \text{otherwise,} \end{cases} \quad (63)$$

which in turn can be expressed as

$$\min_{y \in C^1[0,1]} \max_{\lambda \in \mathbb{R}} \int_0^1 (y' - 2)^2 dx + \lambda \left[\int_0^1 y dx - 2 \right], \quad (64)$$

and then simplified as

$$\boxed{\min_{y \in C^1[0,1]} \max_{\lambda \in \mathbb{R}} \int_0^1 (y' - 2)^2 + \lambda(y - 2) dx.} \quad (65)$$

b) First note 0 is a lower bound for our functional since

$$\forall y \in C^1[0, 1], \quad \int_0^1 (y' - 2)^2 dx \geq \int_0^1 0 dx = 0. \quad (66)$$

We claim 0 is the minimum and verify this as follows. Suppose $y \in C^1[0, 1]$. Then the inequality in (66) is an equality if and only if $y'(x) = 2$ for all x in $[0, 1]$, which implies $y(x) = 2x + c$ for some $c \in \mathbb{R}$. In order for y to satisfy the constraint, we need

$$2 = \int_0^1 y dx = \int_0^1 2x + c dx = [x^2 + cx]_0^1 = 1 + c \quad \Leftrightarrow \quad c = 1. \quad (67)$$

This shows the inequality in (66) holds if and only if $y(x) = 2x + 1$. All that remains is to note y is smooth since it is a polynomial, and so $y \in C^1[0, 1]$. Thus we conclude the minimizer is $\boxed{y(x) = 2x + 1}$ and the minimum is $\boxed{0}$.

□

THE GÂTEAUX DERIVATIVE & EXTREMALS

In this section, we effectively discuss how to take “derivatives” of functionals. First we provide the definition of a Gâteaux derivative. Then we discuss admissibility classes, provide a couple examples of these, outline a path for computing Gâteaux derivatives, and then provide more examples.

Definition: Suppose V is a vector space. The **Gâteaux derivative**, denoted $J(y, v)$, of $J : V \rightarrow \mathbb{R}$ at $y \in V$ in the direction of v is defined as a mapping $\delta J(y, v)$ from V into \mathbb{R} such that

$$\delta J(y, v) := \lim_{\varepsilon \rightarrow 0} \frac{J(y + \varepsilon h) - J(y)}{\varepsilon} = \frac{d}{d\varepsilon} [J(y + \varepsilon h)]_{\varepsilon=0}, \quad (68)$$

provided the limit exists. If the limit exists for all $h \in V$, then we say J is **Gâteaux differentiable** at y . △

Definition: Let \mathcal{A} be a subset of a vector space V , written $\mathcal{A} \subseteq V$. When solving a variational problem, the set of feasible solutions (i.e., those satisfying the constraints) are defined by the **admissibility class** \mathcal{A} . △

REMARK 14: Sometimes in this course the admissibility class is expressed by X rather than \mathcal{A} . ◇

Definition: Suppose $y \in \mathcal{A}$. We call a function v an **admissible variation** provided there is an interval $(-\varepsilon, \varepsilon)$ such that $y + \varepsilon v \in \mathcal{A}$. △

Example 15: Let $\mathcal{A} = C([0, 1])$. Then for each $y, v \in \mathcal{A}$ and $\varepsilon \in \mathbb{R}$ we have $y + \varepsilon v \in \mathcal{A}$ since the sum of continuous functions is continuous and scalar multiples of continuous functions are continuous.

Example 16: Find the set of all admissible variations when $\mathcal{A} = \{y \in C([0, 1]) : y(0) = 1\}$.

Solution:

Let $y \in \mathcal{A}$ and $\varepsilon \in \mathbb{R}$. Also let $h \in C([0, 1])$. Since the the sum of continuous functions is continuous and scalar multiples of continuous functions are continuous, $y + \varepsilon h$ is always continuous. All that remains is to identify a condition on h so that $y(0) + \varepsilon h(0) = 0$. Observe this implies

$$0 = y(0) + \varepsilon h(0) = 0 + \varepsilon h(0) \implies h(0) = 0, \tag{69}$$

since we may take $\varepsilon \neq 0$. Thus the set of admissible variations is given by

$$\boxed{\{v \in C([0, 1]) : v(0) = 0\}}. \tag{70}$$

□

REMARK 15: Computing the Gâteaux derivative:

1. Identify J and \mathcal{A} .
2. Fix $y \in \mathcal{A}$. Let $v \in V$ such that $y + \varepsilon v \in \mathcal{A}$ for all ε with $|\varepsilon|$ sufficiently small.¹
3. Compute $\frac{d}{d\varepsilon} [J(y + \varepsilon v)]$ and then evaluate the result at $\varepsilon = 0$.

◇

Theorem: Suppose \bar{y} is a local minimizer of a functional $J : V \rightarrow \mathbb{R}$ over an open set contained in \mathcal{A} . Then $\delta J(\bar{y}, v) = 0$ for every admissible variation v . △

REMARK 16: We can state the result of the above theorem more intuitively, using knowledge from calculus and earlier theorem. Fix v to be any admissible variation. Then define the function $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(\varepsilon) := J(\bar{y} + \varepsilon v)$. Since \bar{y} is a minimizer for J , it follows that 0 is a local minimizer of f . Therefore $f'(0) = 0$. In other words,

$$0 = f'(0) = \frac{d}{d\varepsilon} [J(\bar{y} + \varepsilon v)]_{\varepsilon=0} = \lim_{\varepsilon \rightarrow 0} \frac{J(\bar{y} + \varepsilon v) - J(\bar{y})}{\varepsilon} = \delta J(\bar{y}, v). \tag{71}$$

◇

¹Note the choice of \mathcal{A} may impose restrictions on v .

Lagrange's Lemma: Suppose $f \in C[a, b]$. If for all $v \in C[a, b]$ we have

$$0 = \int_a^b f(x)v(x) \, dx, \quad (72)$$

then $f(x) = 0$ for all $x \in [a, b]$, i.e., f is identically zero. \triangle

REMARK 17: The above lemma will be of significant importance to use when trying to determine what differential equation a minimizer y of a functional J must satisfy. \diamond

Definition: We say that y is an **extremal** of J provided $\delta J(y, v) = 0$ for arbitrary v . \triangle

REMARK 18: From the above definition and our theorem giving a necessary condition for minimizers, we know every minimizer is an extremal. However, an extremal need not necessarily be a minimizer (e.g., it could be a maximizer as is the case in Example 24). \diamond

REMARK 19: Below we list steps for finding extremals of $J : V \rightarrow \mathbb{R}$ over \mathcal{A} , assuming \mathcal{A} is nonempty.

Finding Extremals:

1. Let $y \in \mathcal{A}$ and find conditions on $v \in V$ such that $y + \varepsilon v \in \mathcal{A}$ when $|\varepsilon|$ is sufficiently small.
2. Compute $\delta J(y, v) = \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0}$ and simplify using information known from \mathcal{A} .
3. Set $\delta J(y, v) = 0$ and use our lemma to obtain an equation for each solution \bar{y} to this equation.
4. Find the solutions to this equation to obtain each possible extremal $\bar{y} \in \mathcal{A}$.

If, in addition, we would like to verify an extremal y_0 is a *local minimizer* of J over \mathcal{A} , then we must show there is $\varepsilon > 0$ such that $J(y_0) \leq J(y)$ for all $y \in \mathcal{A}$ satisfying $\|y - y_0\| < \varepsilon$ for some appropriate norm $\|\cdot\|$. If $J(y_0) \leq J(y)$ for all $y \in \mathcal{A}$, then it is a *global minimizer*. Note every global minimizer is also a local minimizer. \diamond

Example 17: Define $J : C^1([0, 1]) \rightarrow \mathbb{R}$ by

$$J(y) := \frac{1}{2} \int_0^1 (y')^2 + y^2 + 2ye^x \, dx. \quad (73)$$

Let $v \in C^1([0, 1])$ and compute $\delta J(y, v)$.

Solution:

Through direction computation we find

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\frac{1}{2} \int_0^1 (y' + \varepsilon v')^2 + (y + \varepsilon v)^2 + 2(y + \varepsilon v)e^x \, dx \right]_{\varepsilon=0} \\ &= \left[\int_0^1 (y' + \varepsilon v')v' + (y + \varepsilon v)v + 2ve^x \, dx \right]_{\varepsilon=0} \\ &= \boxed{\int_0^1 y'v' + yv + 2ve^x \, dx.} \end{aligned} \quad (74)$$

□

Example 18: Define $J : C^1([0, 2]) \rightarrow \mathbb{R}$ by

$$J(y) := \int_0^2 yy' + \cos(y)dx. \quad (75)$$

Let $v \in C^1([0, 2])$ with $v(0) = v(2) = 0$. Compute $\delta J(y, v)$ and simplify as much as possible.

Solution:

Through direction computation we find

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\int_0^2 (y + \varepsilon v)(y' + \varepsilon v') + \cos(y + \varepsilon v) dx \right]_{\varepsilon=0} \\ &= \left[\int_0^2 v(y' + \varepsilon v') + (y + \varepsilon v)v' - \sin(y + \varepsilon v)v dx \right]_{\varepsilon=0} \\ &= \int_0^2 vy' + yv' - \sin(y)v dx \\ &= \int_0^2 (y' - y' - \sin(y))v dx + \cancel{[yv]_0^2}^0 \\ &= \boxed{- \int_0^2 \sin(y)v dx.} \end{aligned} \quad (76)$$

□

Example 19: Define $J : C^1([0, 1]) \rightarrow \mathbb{R}$ by

$$J(y) := \int_0^1 \sqrt{1 + (y')^2} dx. \quad (77)$$

Let $v \in C^1([0, 1])$ and compute $\delta J(y, v)$.

Solution:

Through direction computation we find

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\int_0^1 \sqrt{1 + (y' + \varepsilon v')^2} dx \right]_{\varepsilon=0} \\ &= \left[\int_0^1 \frac{1}{2} (1 + (y' + \varepsilon v')^2)^{-1/2} \cdot 2(y' + \varepsilon v') \cdot v' dx \right]_{\varepsilon=0} \\ &= \left[\int_0^1 \frac{(y' + \varepsilon v')v'}{\sqrt{1 + (y' + \varepsilon v')^2}} dx \right]_{\varepsilon=0} \\ &= \boxed{\int_0^1 \frac{y'v'}{\sqrt{1 + (y')^2}} dx}. \end{aligned} \quad (78)$$

□

Example 20: Define $J : C^1([1, 2]) \rightarrow \mathbb{R}$ by

$$J(y) := \int_1^2 3x^2(y')^2 - e^y \, dx \quad (79)$$

Let $v \in C^1([1, 2])$ and compute $\delta J(y, v)$.

Solution:

Through direct computation we find

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\int_1^2 3x^2(y' + \varepsilon v')^2 - e^{y+\varepsilon v} \, dx \right]_{\varepsilon=0} \\ &= \left[\int_1^2 6x^2(y' + \varepsilon v')v' - v e^{y+\varepsilon v} \, dx \right]_{\varepsilon=0} \\ &= \boxed{\int_1^2 6x^2 y' v' - v e^y \, dx.} \end{aligned} \quad (80)$$

□

REMARK 20: With some practice of taking Gâteaux derivatives, we next move toward using these to find minimizers of a few functionals. ◇

Example 21: Define the admissibility class $\mathcal{A} := C[a, b]$. Then let $J : \mathcal{A} \rightarrow \mathbb{R}$ be the functional

$$J(y) = \int_a^b (y(x) - 4)^2 \, dx. \tag{81}$$

Find a minimizer of J over \mathcal{A} using the Gâteaux derivative. You may suppose a minimizer $y_0 \in \mathcal{A}$ exists.

Solution:

Fix $y, v \in \mathcal{A}$. Then $y + \varepsilon v \in \mathcal{A}$ for all $\varepsilon \in \mathbb{R}$ since \mathcal{A} is a vector space. We expand $J(y + \varepsilon v)$ to find

$$\begin{aligned} J(y + \varepsilon v) &= \int_a^b [y(x) + \varepsilon v(x) - 4]^2 \, dx \\ &= \int_a^b [y(x) + \varepsilon v(x)]^2 - 8[y(x) + \varepsilon v(x)] + 16 \, dx \\ &= \int_a^b y(x)^2 + 2\varepsilon y(x)v(x) + \varepsilon^2 v(x)^2 - 8[y(x) + \varepsilon v(x)] + 16 \, dx \\ &= \int_a^b y(x)^2 - 8y(x) + 16 \, dx + 2\varepsilon \int_a^b [y(x) - 4]v(x) \, dx + \varepsilon^2 \int_a^b v(x)^2 \, dx \\ &= J(y) + 2\varepsilon \int_a^b [y(x) - 4]v(x) \, dx + \varepsilon^2 \int_a^b v(x)^2 \, dx. \end{aligned} \tag{82}$$

Thus

$$\begin{aligned} \delta J(y, v) &= \lim_{\varepsilon \rightarrow 0} \frac{J(y + \varepsilon v) - J(y)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \left(2 \int_a^b [y(x) - 4]v(x) \, dx + \varepsilon \int_a^b v(x)^2 \, dx \right) \\ &= 2 \int_a^b [y(x) - 4]v(x) \, dx. \end{aligned} \tag{83}$$

By our theorem, we know that if \bar{y} is a minimizer of J over \mathcal{A} , then

$$0 = \delta J(\bar{y}, v) = 2 \int_a^b [\bar{y}(x) - 4]v(x) \, dx \tag{84}$$

for all admissible variations v . This implies the only candidate minimizer is $\bar{y}(x) = 4$. Since $J(y) \geq 0$ for all $y \in \mathcal{A}$ and $J(\bar{y}) = 0$, we conclude $\boxed{\bar{y}(x) = 4}$ is the global minimizer of J over \mathcal{A} . \square

REMARK 21: The above computations may seem quite tedious. This is because they are. A more elegant approach for computing $\delta J(\bar{y}, v)$ is given in the following reworking of the above example, utilizing our knowledge of derivatives for real valued functions. \diamond

Example 22: Repeat the previous example, making use of derivatives of real-valued functions.

Solution:

Fix $y, v \in \mathcal{A}$. Then $y + \varepsilon v \in \mathcal{A}$ for all $\varepsilon \in \mathbb{R}$ since \mathcal{A} is a vector space. Through direct computation we find

$$\frac{d}{d\varepsilon} J(y + \varepsilon v) = \frac{d}{d\varepsilon} \int_a^b [y(x) + \varepsilon v(x) - 4]^2 dx = \int_a^b 2[y(x) + \varepsilon v(x) - 4] v(x) dx. \quad (85)$$

Thus

$$\delta J(y, v) = \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} = \int_a^b 2[y(x) - 4] v(x) dx. \quad (86)$$

If \bar{y} is a minimizer of J , then it is an extremal of J , and so $\delta J(\bar{y}, v) = 0$. Since this result holds for arbitrary admissible variations v , our lemma states the only candidate minimizer is $\bar{y}(x) = 4$. Since $J(y) \geq 0$ for all $y \in \mathcal{A}$ and $J(\bar{y}) = 0$, we conclude $\boxed{\bar{y}(x) = 4}$ is the global minimizer of J over \mathcal{A} . \square

Example 23: Define the functional $J : C^2[0, 1] \rightarrow \mathbb{R}$ by

$$J(y) := \int_0^1 \frac{1}{2} m \dot{y}^2 - mgy \, dt, \tag{87}$$

here using the dot notation for time derivatives. Let $\mathcal{A} := \{y \in C^2[0, 1] : y(0) = \alpha_1, y(1) = \alpha_2\}$. Find an extremal $y_0 \in \mathcal{A}$ of J .

Solution:

Pick $y \in \mathcal{A}$. Since the end points of $y \in \mathcal{A}$ are fixed, if $h \in V$ and $y_0 + \varepsilon h \in \mathcal{A}$ for any nonzero ε , then

$$\alpha_1 = y(0) + \varepsilon h(0) = \alpha_1 + \varepsilon h(0) \implies h(0) = 0. \tag{88}$$

Similarly, $h(1) = 0$. Fixing h , we compute

$$\frac{d}{d\varepsilon} [J(y + \varepsilon h)] = \frac{d}{d\varepsilon} \int_0^1 \frac{1}{2} m (\dot{y} + \varepsilon \dot{h})^2 - mg(y + \varepsilon h) \, dt = \int_0^1 m (\dot{y} + \varepsilon \dot{h}) \dot{h} - mgh \, dt. \tag{89}$$

To make this more useful, we use integration by parts with the first to rewrite the above in a more useful form. That is,

$$\frac{d}{d\varepsilon} [J(y + \varepsilon h)] = \int_0^1 -m(\ddot{y} + \varepsilon \ddot{h})h - mgh \, dt = -m \int_0^1 (\ddot{y} + \varepsilon \ddot{h} + g) h \, dt. \tag{90}$$

Evaluating the above at $\varepsilon = 0$, we deduce

$$\delta J(y, h) = -m \int_0^1 (\ddot{y} + g)h \, dt. \tag{91}$$

Thus for a minimizer y_0 of J we see $\delta J(y, h) = 0$. Since this holds for an arbitrary admissible variation h , by our lemma we deduce the only candidate minimizer satisfies $\ddot{y}_0 = -g$. Then

$$\dot{y}_0 = -gt + c_1 \implies y_0 = -\frac{1}{2}gt^2 + c_1t + c_2 \tag{92}$$

for some $c_1, c_2 \in \mathbb{R}$. Using the fact $y_0(0) = \alpha_1$, we know $c_2 = \alpha_1$. Then

$$\alpha_2 = y_0(1) = -\frac{1}{2}g + c_1 + \alpha_1 \implies c_1 = \alpha_2 - \alpha_1 + \frac{g}{2}, \tag{93}$$

and we conclude

$$y_0(t) = -\frac{1}{2}gt^2 + \left(\alpha_2 - \alpha_1 + \frac{g}{2}\right)t + \alpha_1. \quad (94)$$

□

REMARK 22: From classical mechanics in physics, we know the kinetic energy of a ball of mass m is given by $T = \frac{1}{2}m\dot{y}^2$ and its potential energy is $U = mgh$. In the above, we are minimizing $T - U$ over a time interval. It turns out this is associated with *Hamilton's Principle* and it gives the same result as would be obtained using Newton's second law of motion. ◇

Example 24: Define $J : C[0, 1] \rightarrow \mathbb{R}$ by

$$J(y) = \int_0^1 -y(x)^2 + 6y(x) + 10 \, dx. \quad (95)$$

Find $y_0 \in C[0, 1]$ such that $\delta J(y, v) = 0$ for all $C[0, 1]$. Is y_0 a minimizer of J over $C[0, 1]$?

Solution:

Let $y, h \in \mathcal{A}$ and $\varepsilon \in \mathbb{R}$. Differentiating, we see

$$\frac{d}{d\varepsilon} J(y + \varepsilon v) = \frac{d}{d\varepsilon} \int_0^1 -[y + \varepsilon v]^2 + 6[y + \varepsilon v] + 10 \, dx = \int_0^1 -2[y + \varepsilon v]v + 6h \, dx. \quad (96)$$

This implies

$$\delta J(y, v) = \int_0^1 -2yv + 6v \, dx = \int_0^1 (-2y + 6)v \, dx. \quad (97)$$

So, taking $y_0(x) = 3$, we obtain $\delta J(y_0, v) = 0$. Now observe

$$J(y) = \int_0^1 -(y^2 - 6y + 9) + 19 \, dx = \int_0^1 -(y - 3)^2 + 19 \, dx \leq \int_0^1 19 \, dx = 19. \quad (98)$$

This shows 19 is an upper bound for $J(y)$. However,

$$J(y_0) = \int_0^1 -(3 - 3)^2 + 19 \, dx = 19. \quad (99)$$

Moreover, if $y \neq 3 = y_0$, then $J(y) < J(y_0)$. Thus y_0 is **not** a minimizer of J . In fact, this shows y_0 is a maximizer of J . □

REMARK 23: The above example shows the condition $\delta J(y_0, v) = 0$ is **not** a sufficient condition to conclude y_0 is a minimizer of J . \diamond

Example 25: Now consider a bounded subset $\Omega \subset \mathbb{R}^n$. Let $f : \Omega \rightarrow \mathbb{R}$ be continuous and $y \in C^3(\Omega)$. Then define $J : C^3(\Omega) \rightarrow \mathbb{R}$ by

$$J(y) = \frac{1}{2} \int_{\Omega} (f - y)^2 + \lambda(\Delta y)^2 \, dx. \tag{100}$$

Compute the Gâteaux derivative of $\delta J(y, v)$, assuming all boundary terms vanish.

Solution:

Observe

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} \left[\frac{1}{2} \int_{\Omega} (f - y - \varepsilon v)^2 + \lambda(\Delta y + \varepsilon \Delta v)^2 \, dx \right]_{\varepsilon=0} \\ &= \left[\int_{\Omega} -(f - y - \varepsilon v)v + \lambda(\Delta y + \varepsilon \Delta v)\Delta v \, dx \right]_{\varepsilon=0} \\ &= \int_{\Omega} -(f - y)v + \lambda \Delta y \Delta v \, dx \\ &= \int_{\Omega} -(f - y)v + \lambda \left(\sum_{i=1}^n y_{x_i x_i} \right) \left(\sum_{j=1}^n v_{x_j x_j} \right) \, dx \\ &= \int_{\Omega} -(f - y)v + \lambda \sum_{i,j=1}^n y_{x_i x_i} v_{x_j x_j} \, dx \\ &= \int_{\Omega} -(f - y)v - \lambda \sum_{i,j=1}^n \partial_{x_j} y_{x_i x_i} v_{x_j} \, dx \\ &= \int_{\Omega} -(f - y)v + \lambda \sum_{i,j=1}^n \partial_{x_j x_j} y_{x_i x_i} v \, dx \\ &= \int_{\Omega} -(f - y)v + \lambda \sum_{j=1}^n \partial_{x_j x_j} \left(\sum_{i=1}^n y_{x_i x_i} \right) v \, dx \\ &= \int_{\Omega} -(f - y)v + \lambda \sum_{j=1}^n \partial_{x_j x_j} \Delta y v \, dx \\ &= \int_{\Omega} (-(f - y) + \lambda \Delta \Delta y) v \, dx. \end{aligned} \tag{101}$$

Then if y is an extremal, Lagrange's Lemma asserts y satisfies

$$-(f - y) + \lambda \Delta \Delta y = 0 \text{ in } \Omega. \tag{102}$$

□

EULER-LAGRANGE EQUATIONS

Suppose we have a function $L(\dot{y}, y, x)$ and $J : C^2[a, b] \rightarrow \mathbb{R}$ given by

$$J(y) := \int_a^b L(\dot{y}, y, x) \, dx. \quad (103)$$

Then for each $v \in C^2[a, b]$

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} \left[\int_a^b L(\dot{y} + \varepsilon \dot{v}, y + \varepsilon v, x) \right]_{\varepsilon=0} \\ &= \left[\int_a^b L_{\dot{y}}(\dot{y} + \varepsilon \dot{v}, y + \varepsilon v, x) \dot{v} + L_y(\dot{y} + \varepsilon \dot{v}, y + \varepsilon v, x) v \, dx \right]_{\varepsilon=0} \\ &= \int_a^b L_{\dot{y}}(\dot{y}, y, x) \dot{v} + L_y(\dot{y}, y, x) v \, dx. \end{aligned} \quad (104)$$

For notational compactness, we henceforth suppress the arguments of $L(\dot{y}, y, x)$ and simply write L . Integrating by parts yields

$$\begin{aligned} \delta J(y, v) &= \int_a^b L_{\dot{y}} \dot{v} + L_y v \, dx \\ &= \int_a^b -\frac{d}{dx} [L_{\dot{y}}] v + L_y v \, dx + [L_{\dot{y}} v]_a^b \\ &= \int_a^b \left(L_y - \frac{d}{dx} [L_{\dot{y}}] \right) v \, dx + [L_{\dot{y}} v]_a^b. \end{aligned} \quad (105)$$

If y is an extremal of J , then $\delta J(y, v) = 0$ for each $v \in C^2[a, b]$. Consequently, the integral term must equal zero. Applying Lagrange's Lemma yields

$$0 = L_y - \frac{d}{dx} [L_{\dot{y}}] \quad \text{for } x \in [a, b]. \quad (106)$$

We call this equation the **Euler-Lagrange equation**. Furthermore, for an extremal y the boundary terms must vanish, i.e., we need

$$0 = L_{\dot{y}} v|_{x=a} \quad \text{and} \quad 0 = L_{\dot{y}} v|_{x=b}. \quad (107)$$

These are known as **natural boundary conditions**. If $y(a)$ is fixed, then $v(a) = 0$ and so the first natural boundary condition holds automatically. Similarly applies if $y(b)$ is fixed. However, if the

endpoint $y(a)$ is not fixed and so the admissible variation v can take on any value at a , then we see

$$0 = L_{\dot{y}}|_{x=a}. \quad (108)$$

Similarly applies for $y(b)$. For more reading, we encourage the reader to see the set of [notes](#). (Unfortunately, it is not clear who created these notes.)

REMARK 24: Now that we have more powerful tools available, we can almost immediately identify the differential equation satisfied by extremals to functionals of the form in (103). The steps here are roughly:

1. Identify the Lagrangian.
2. Write the Euler-Lagrange equation.
3. Identify the appropriate boundary conditions.
4. Solve the differential equation with the boundary conditions to obtain the extremals.

◇

Example 26: Find the form of extremals of $J : C^2[0, 1] \rightarrow \mathbb{R}$ defined by

$$J(y) := \frac{1}{2} \int_0^1 y^2 + \dot{y}^2 + 2ye^x \, dx. \quad (109)$$

Solution:

Here the Lagrangian is $L = \frac{1}{2}(y^2 + \dot{y}^2) + ye^x$. Using the Euler-Lagrange equation, we know any extremal y satisfies

$$0 = L_y - \frac{d}{dx} L_{\dot{y}} = (y + e^x) - \frac{d}{dx} [\dot{y}] = y + e^x - \ddot{y} \quad \forall x \in [0, 1]. \quad (110)$$

Since this is a linear ODE, we may write $y = y_H + y_P$ where $y_H - \ddot{y}_H = 0$ and $y_P - \ddot{y}_P = e^x$. The general solution to the homogeneous equation is

$$y_H = c_1 e^x + c_2 e^{-x} \quad (111)$$

for some scalars $c_1, c_2 \in \mathbb{R}$. The particular solution is $y_P = -\frac{x}{2}e^x$. Thus the extremal is of the form

$$\boxed{y = \left(c_1 - \frac{x}{2}\right) e^x + c_2 e^{-x}.} \quad (112)$$

□

REMARK 25: Using the natural boundary condition $L_{\dot{y}} = 0$ at $x = 0$ and $x = 1$, we would obtain two boundary conditions, which could be used to explicitly solve for c_1 and c_2 . ◇

REMARK 26: The solution to the above problem was *much simpler* than that obtained using the method of previous examples where we computed the Gâteaux derivative. ◇

Example 27: Find the extremal of $J : C^2[0, 1] \rightarrow \mathbb{R}$ given by

$$J(y) = \frac{1}{2} \int_0^\pi y^2 - \dot{y}^2 \, dx \quad (113)$$

subject to the condition $y(0) = 3i$

Solution:

Here the Lagrangian is $L = \frac{1}{2}(y^2 - \dot{y}^2)$. Then any extremal y of J satisfies the Euler-Lagrange equation

$$0 = L_y - \frac{d}{dx} L_{\dot{y}} = y - \frac{d}{dx} [-\dot{y}] = y + \ddot{y}. \quad (114)$$

Furthermore, since y is not fixed at $x = \pi$, we have the natural boundary condition

$$0 = L_{\dot{y}}|_{x=\pi} = -\dot{y}(\pi). \quad (115)$$

The general solution to the ODE $y + \ddot{y} = 0$ is $y = c_1 \sin(x) + c_2 \cos(x)$ for scalars $c_1, c_2 \in \mathbb{R}$. Then the natural boundary condition implies

$$0 = c_1 \cos(\pi) - c_2 \sin(\pi) = -c_1 - 0 \quad \implies \quad c_1 = 0. \quad (116)$$

Thus $y = c_2 \cos(x)$ and the condition on $y(0)$ yields

$$3 = y(0) = c_2 \cos(0) = c_2 \quad \implies \quad \boxed{y = 3 \cos(x)}. \quad (117)$$

□

Example 28: Find the extremal of J over $\mathcal{A} := \{y \in C^2[0, 1] : y(0) = 2\}$ where

$$J(y) := \frac{1}{2} \int_0^1 y^2 + \dot{y}^2 \, dx. \quad (118)$$

Solution:

Let $y \in \mathcal{A}$. Then if $v \in C^2[0, 1]$ is an admissible variation, we need $v(0) = 0$. Also, here the Lagrangian is $L = \frac{1}{2} (y^2 + \dot{y}^2)$. Then any extremal y satisfies the Euler-Lagrange equation

$$0 = L_y - \frac{d}{dx} L_{\dot{y}} = y - \frac{d}{dx} [\dot{y}] = y - \ddot{y} \quad \forall x \in [0, 1]. \quad (119)$$

The general solution to this ODE is $y = c_1 e^x + c_2 e^{-x}$ for scalars $c_1, c_2 \in \mathbb{R}$. Since v can take on any value at $x = 1$, we have the natural boundary condition

$$0 = L_{\dot{y}}|_{x=1} = \dot{y}(1) = [c_1 e^x - c_2 e^{-x}]_{x=1} = c_1 e - c_2 e^{-1} \quad \implies \quad c_1 = \frac{c_2}{e^2}. \quad (120)$$

Then the initial condition yields

$$2 = y(0) = c_1 e^0 + c_2 e^{-0} = c_1 + c_2 = c_2 \left(\frac{1}{e^2} + 1 \right) \quad \implies \quad c_2 = \frac{2}{1 + 1/e^2} = \frac{2e^2}{1 + e^2}. \quad (121)$$

Back substitution then reveals $c_1 = 2/(1 + e^2)$. Combining our results, we conclude

$$\boxed{y = \frac{2}{1 + e^2} (e^x + e^{2-x})}. \quad (122)$$

□

Example 29: Find the form of extremals of

$$J(y) = \int_a^b x^2 \dot{y}^2 + y^2 \, dx. \quad (123)$$

Solution:

Here the Lagrangian is $L = x^2 \dot{y}^2 + y^2$. Then the Euler-Lagrange equation reveals any extremal y satisfies

$$0 = L_y - \frac{d}{dx} L_{\dot{y}} = 2y - \frac{d}{dx} [2x^2 \dot{y}] = 2y - 4x\dot{y} - 2x^2 \ddot{y} = 2(y - 2x\dot{y} - x^2 \ddot{y}) \quad \forall x \in [a, b]. \quad (124)$$

Thus

$$x^2 \ddot{y} + 2x\dot{y} - y = 0 \quad \forall x \in [a, b]. \quad (125)$$

This is a second order Cauchy-Euler equation, which has solutions of the form

$$\boxed{y = c_1 x^{(\sqrt{5}-1)/2} + c_2 x^{-(\sqrt{5}+1)/2}}, \quad (126)$$

for scalars $c_1, c_2 \in \mathbb{R}$. □

Example 30: Find the form of extremals of J over all $y \in \mathcal{A} = \{y \in C^2[0, 1] : y(0) = 0, y(\pi) = 2\pi\}$ where

$$J(y) = \frac{1}{2} \int_0^\pi 9y^2 - \dot{y}^2 - 36xy \, dx. \quad (127)$$

Solution:

Here the Lagrangian is $L = \frac{1}{2}(9y^2 - \dot{y}^2 - 36xy)$. Then any extremal y of J satisfies the Euler-Lagrange equation

$$0 = L_y - \frac{d}{dx} L_{\dot{y}} = (9y - 36x) - \frac{d}{dx} [-\dot{y}] = 9y - 36x + \ddot{y} \quad \forall x \in [a, b]. \quad (128)$$

Thus y is a solution of the ODE $\ddot{y} + 9y = 18x$. Since this ODE is linear, we may write $y = y_H + y_P$ where y_H is the solution to the associated homogeneous problem and y_P is a particular solution. The general solution to the homogeneous equation is

$$y_H = c_1 \cos(3x) + c_2 \sin(3x) \quad (129)$$

for some scalars $c_1, c_2 \in \mathbb{R}$. Note $y_P = 2x$ by inspection. Then the first boundary condition implies

$$0 = y(0) = [c_1 \cos(3x) + c_2 \sin(3x) + 2x]_{x=0} = c_1 + 0 + 0 \implies c_1 = 0. \quad (130)$$

Thus $y = c_2 \sin(3x) + 2x$. The second boundary condition implies

$$2\pi = y(\pi) = 0 + 2\pi, \quad (131)$$

which provides no new information. Consequently, the form of each extremal is

$$\boxed{y = c_2 \sin(3x) + 2x.} \quad (132)$$

□

REMARK 27: Finding the extremals of functionals is quite useful for identifying the motion of objects in classical mechanics. **Hamilton's Principle** states that if we let $\mathcal{L} = T - U$ where T is the kinetic energy of a system and U is its potential energy, then the extremal of J as defined in (103) gives the motion of the system over a given period of time, swapping time t for the variable x . ◇

(MORE PHYSICS EXAMPLES WILL BE ADDED SOON.)

CONSTRAINED PROBLEMS

Suppose we seek to solve

$$\min_y \int_a^b \mathcal{L}(\dot{y}, y, x) \, dx \quad \text{s.t.} \quad \int_a^b G(\dot{y}, y, x) \, dx = 0. \quad (133)$$

Here we define the augmented Lagrangian $\mathcal{L}^* = \mathcal{L} + \lambda G$ for some scalar $\lambda \in G$. Then the extremals satisfy the Euler-Lagrange equation

$$0 = \mathcal{L}_y^* - \frac{d}{dx} \mathcal{L}_y^* \quad \forall x \in [a, b]. \quad (134)$$

An alternative approach to using the augmented Lagrangian \mathcal{L}^* is to rewrite the problem as a min-max problem. Then we swap the order of the min and max, find a solution to the min problem, and then use the constraint to identify the variable λ . This is illustrated in the following example.

Example 31: Find the local extrema over $C[0, 1]$ of²

$$J(y) = \int_0^1 xy \, dx \quad \text{subject to} \quad \int_0^1 y^2 \, dx = \frac{1}{3}. \quad (135)$$

Solution:

The constrained problem may be rewritten as the min-max problem

$$\min_y \max_\lambda \int_0^1 xy + \lambda y^2 - \frac{\lambda}{12} \, dx = \max_\lambda \min_y \int_0^1 xy + \lambda y^2 - \frac{\lambda}{12} \, dx, \quad (136)$$

where λ is a real scalar. Note for each $y, v \in C[0, 1]$ and $\varepsilon \in \mathbb{R}$ we have $y + \varepsilon v \in C[0, 1]$. Then taking the Gâteaux derivative for such y and v yields

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\int_0^1 x(y + \varepsilon v) + \lambda(y + \varepsilon v)^2 - \frac{\lambda}{12} \, dx \right]_{\varepsilon=0} \\ &= \left[\int_0^1 vx + 2\lambda(y + \varepsilon v)v - 0 \, dx \right]_{\varepsilon=0} \\ &= \int_0^1 (x + 2\lambda y) v \, dx. \end{aligned} \quad (137)$$

Consequently, if y is an extrema, then applying Lagrange's Lemma with the fact $0 = \delta J(y, v)$ yields

$$0 = x + 2\lambda y \quad \implies \quad y = -\frac{x}{2\lambda}. \quad (138)$$

Then plugging this into our constraint yields

$$\frac{1}{3} = \int_0^1 y^2 \, dx = \int_0^1 \left(-\frac{x}{2\lambda}\right)^2 \, dx = -\frac{1}{4\lambda^2} \int_0^1 x^2 \, dx = -\frac{1}{12\lambda^2}. \quad (139)$$

This shows $3 = 12\lambda^2$, from which we deduce $\lambda = \pm 1/2$. Whence this fact together with (140) shows the potential local extrema are given by $\boxed{y = \pm x}$. \square

²This problem is modeled after an exercise in [2].

Example 32: Here we present an alternative solution to the previous problem.

Solution:

First we identify the augmented Lagrangian $\mathcal{L}^* = xy + \lambda y^2$. Then we use the fact any local extrema satisfies the Euler-Lagrange equation to write

$$0 = \mathcal{L}_y^* - \frac{d}{dx} [\mathcal{L}_{y'}^*] = (x + 2\lambda y) - \frac{d}{dx} [0] = x + 2\lambda y \quad \implies \quad y = -\frac{x}{2\lambda}. \quad (140)$$

Then plugging this into our constraint yields

$$\frac{1}{3} = \int_0^1 y^2 dx = \int_0^1 \left(-\frac{x}{2\lambda}\right)^2 dx = -\frac{1}{4\lambda^2} \int_0^1 x^2 dx = -\frac{1}{12\lambda^2}. \quad (141)$$

This shows $3 = 12\lambda^2$, from which we deduce $\lambda = \pm 1/2$. Whence this fact together with (140) shows the potential local extrema are given by $\boxed{y = \pm x}$. \square

Example 33: Find the local minimizer of³

$$J(y) = \int_0^1 y \, dx \quad (142)$$

subject to the constraint

$$\int_0^1 y^2 + xy \, dx = \frac{47}{12}. \quad (143)$$

You may assume the local minimizer is one of the extremals.

Solution:

First we rewrite the problem as

$$\min_y \max_\lambda \int_0^1 y + \lambda y^2 + \lambda xy \, dx - \frac{47}{12} \lambda = \max_\lambda \min_y \int_0^1 y + \lambda y^2 + \lambda xy \, dx - \frac{47}{12} \lambda. \quad (144)$$

where λ is a real scalar. Note for each $y, v \in C[0, 1]$ and $\varepsilon \in \mathbb{R}$ we have $y + \varepsilon v \in C[0, 1]$. Then taking the Gâteaux derivative for such y and v yields

$$\begin{aligned} \delta J(y, v) &= \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left[\int_0^1 (y + \varepsilon v) + \lambda(y + \varepsilon v)^2 + \lambda x(y + \varepsilon v) \, dx - \frac{47}{12} \lambda \right]_{\varepsilon=0} \\ &= \left[\int_0^1 v + 2\lambda(y + \varepsilon v)v + \lambda xv \, dx - \frac{47}{12} \lambda \right]_{\varepsilon=0} \\ &= \int_0^1 (1 + 2\lambda y + \lambda x) v \, dx. \end{aligned} \quad (145)$$

Consequently, if y is an extrema, then applying Lagrange's Lemma with the fact $0 = \delta J(y, v)$ yields

$$0 = 1 + 2\lambda y + \lambda x \quad \implies \quad y = -\frac{1}{2} \left[x + \frac{1}{\lambda} \right]. \quad (146)$$

³This example is due to [2].

Then plugging this into our constraint yields

$$\begin{aligned}
 \frac{47}{12} &= \int_0^1 y^2 + xy \, dx = \int_0^1 \left(-\frac{1}{2} \left[x + \frac{1}{\lambda} \right] \right)^2 + x \left(-\frac{1}{2} \left[x + \frac{1}{\lambda} \right] \right) \, dx \\
 &= \frac{1}{4} \int_0^1 \frac{1}{\lambda^2} - x^2 \, dx \\
 &= \frac{1}{4} \left[\frac{1}{\lambda^2} - \frac{1}{3} \right] \\
 &= \frac{1}{4\lambda^2} - \frac{1}{12}.
 \end{aligned} \tag{147}$$

Thus

$$\frac{47}{12} = \frac{1}{4\lambda^2} - \frac{1}{12} \implies 4 = \frac{1}{4\lambda^2} \implies \lambda^2 = \frac{1}{16} \implies \lambda = \pm \frac{1}{4}. \tag{148}$$

This gives two candidate local minimizers. Observe

$$J \left(-\frac{1}{2} [x \pm 4] \right) = \int_0^1 -\frac{1}{2} [x \pm 4] \, dx = -\frac{1}{4} \mp 2. \tag{149}$$

Consequently, we conclude $y = -\frac{1}{2} [x + 4]$ is the local minimizer. \square

REMARK 28: If you were to do the previous problem with an augmented Lagrangian, you'd have $\mathcal{L}^* = y + \lambda(y^2 + xy)$. Then

$$0 = \mathcal{L}_y - \frac{d}{dx} [\mathcal{L}_{y'}] = 1 + \lambda(2y + x) \implies y = -\frac{1}{2} \left[x + \frac{1}{\lambda} \right]. \tag{150}$$

\diamond

CONVEXITY

Definition: Let V be a vector space. We say a functional $J : V \rightarrow \mathbb{R}$ is convex provided for each $y_1, y_2 \in V$ and $t \in (0, 1)$ we have

$$J(ty_1 + (1 - t)y_2) \leq tJ(y_1) + (1 - t)J(y_2). \quad (151)$$

△

REMARK 29: Note the above definition is quite general. This same definition can be applied to the special case of a function mapping $f : \mathbb{R} \rightarrow \mathbb{R}$. ◇

Example 34: Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and convex and define $J : C[0, 1] \rightarrow \mathbb{R}$ by $J(y) := \int_0^1 f(y(x)) \, dx$. Show J is convex.⁴

Solution:

Let $y_1, y_2 \in C[0, 1]$ and $t \in (0, 1)$. Then observe the convexity of f implies

$$\begin{aligned} J(ty_1 + (1 - t)y_2) &= \int_0^1 f(ty_1 + (1 - t)y_2) \, dx \\ &\leq \int_0^1 tf(y_1) + (1 - t)f(y_2) \, dx \\ &= t \int_0^1 f(y_1) \, dx + (1 - t) \int_0^1 f(y_2) \, dx \\ &= tJ(y_1) + (1 - t)J(y_2). \end{aligned} \quad (152)$$

Thus we conclude J is convex. □

REMARK 30: Since functions like $|x|$ and x^2 and e^x are convex, this shows

$$J(y) = \int_0^1 |x| \, dx, \quad J(y) = \int_0^1 x^2 \, dx, \quad \text{and} \quad J(y) = \int_0^1 e^x \, dx \quad (153)$$

are convex. ◇

⁴This example is due to Professor Leger.

REMARK 31: One result is that J is convex provided $J(y+v) - J(y) \geq \delta J(y, v)$ for each admissible direction v . We use this in the following example ◇

Example 35: Show $J : C^1[0, 1] \rightarrow \mathbb{R}$ defined by

$$J(y) = \frac{1}{2} \int_a^b \dot{y}^2 + y^2 \, dx \tag{154}$$

is convex.

Solution:

For each $y, v \in C^1[0, 1]$ observe

$$\delta J(y, v) = \frac{d}{d\varepsilon} [J(y + \varepsilon v)]_{\varepsilon=0} = \frac{d}{d\varepsilon} \left[\frac{1}{2} \int_0^1 (y + \varepsilon v)^2 + (\dot{y} + \varepsilon \dot{v})^2 \, dx \right]_{\varepsilon=0} = \int_0^1 \dot{v}^2 + v^2 \, dx. \tag{155}$$

Then observe

$$\int_0^1 (y + v)^2 - y^2 - 2yv \, dx = \int_0^1 (y^2 + 2yv + v^2) - y^2 - 2yv \, dx = \int_0^1 v^2 \, dx. \tag{156}$$

A similar result holds, substituting \dot{y} and \dot{v} for y and v , respectively. Consequently,

$$\begin{aligned} J(y + v) - J(y) - \delta J(y, v) &= \int_0^1 [(\dot{y} + \dot{v})^2 + (y + v)^2] - [\dot{y}^2 + y^2] - 2[\dot{y}\dot{v} + yv] \, dx \\ &= \int_0^1 \dot{v}^2 + v^2 \, dx \\ &\geq 0. \end{aligned} \tag{157}$$

Thus we conclude J is convex. □

LEGENDRE TRANSFORM

The following definition follows that on the [Wikipedia page](#) for Legendre transforms.

Definition: Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be convex. Then the *dual* f^* is given by

$$f^*(p) := \max_{x \in I} px - f(x). \quad (158)$$

△

Example 36: Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(x) = \alpha x^2$ for some $\alpha > 0$. Find f^* .

Solution:

Through the definition of the dual we see

$$f^*(p) := \max_{x \in \mathbb{R}} px - f(x) = \max_{x \in \mathbb{R}} px - \alpha x^2. \quad (159)$$

Differentiating, we see

$$\frac{d}{dx} [px - \alpha x^2] = p - 2\alpha x, \quad (160)$$

and so the single critical point of the expression to be maximized occurs at $x = p/2\alpha$. This point is a local maximum since f is convex; indeed,

$$\forall x \in \mathbb{R}, \quad \frac{d^2}{dx^2} [px - \alpha x^2] = \frac{d}{dx} [p - 2\alpha x] = -2\alpha < 0. \quad (161)$$

Thus

$$\boxed{f^*(p) = p \left(\frac{p}{2\alpha} \right) - \alpha \left(\frac{p}{2\alpha} \right)^2 = \frac{p^2}{4\alpha}.} \quad (162)$$

□

Example 37: Let $f(x) = 5x^2 + 3x$. Find f^* .

Solution:

Through the definition of the dual we see

$$f^*(p) := \max_{x \in \mathbb{R}} px - f(x) = \max_{x \in \mathbb{R}} px - (5x^2 + 3x) \quad (163)$$

Differentiating, we see

$$\frac{d}{dx} [px - (5x^2 + 3x)] = p - 10x - 3, \quad (164)$$

and so the single critical point is $x = (p - 3)/10$. Consequently,

$$f^*(p) = p \left(\frac{p-3}{10} \right) - \left(5 \left(\frac{p-3}{10} \right)^2 - 3 \left(\frac{p-3}{10} \right) \right) = \frac{1}{10} [2p^2 - 3p + 18], \quad (165)$$

from which we conclude

$$\boxed{f^*(p) = \frac{1}{10} [2p^2 - 3p + 18]}. \quad (166)$$

□

Example 38: Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = |x|$. Find $f^*(p)$.

Solution:

Through the definition of the dual we see

$$f^*(p) := \max_{x \in \mathbb{R}} px - f(x) = \max_{x \in \mathbb{R}} px - |x|. \quad (167)$$

Then note

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0, \end{cases} \quad (168)$$

which implies

$$\frac{d}{dx} [px - f(x)] = \begin{cases} p - 1 & \text{if } x > 0, \\ p + 1 & \text{if } x < 0. \end{cases} \quad (169)$$

Now suppose $p > 1$. Then the right hand side of (169) is always positive, which implies the expression $px - |x|$ to be maximized is strictly increasing with x . Thus for $p > 1$

$$f^*(p) = \lim_{x \rightarrow \infty} px - f(x) = \lim_{x \rightarrow \infty} px - |x| = \lim_{x \rightarrow \infty} (p - 1)x = \infty. \quad (170)$$

Similarly, if $p < -1$, then the right hand side of (169) is always negative, thereby implying the expression to be maximized is strictly decreasing. Whence for $p < -1$

$$f^*(p) = \lim_{x \rightarrow -\infty} px - |x| = \lim_{x \rightarrow -\infty} -|x|(p - 1) = \infty, \quad (171)$$

where we note $x = -|x|$ when $x < 0$. Lastly, consider the case where $-1 \leq p \leq 1$. Then the derivative of $px - f(x)$ is nonpositive for $x > 0$ and nonnegative for $x < 0$. Thus the maximum occurs at $x = 0$, for which we see

$$f^*(p) = p0 - f(0) = 0 - 0 = 0. \quad (172)$$

Combining our results, we conclude

$$f^*(p) = \begin{cases} 0 & \text{if } |p| \leq 1, \\ \infty & \text{if } |p| > 1. \end{cases} \quad (173)$$

□

Example 39: Define $f : (0, \infty) \rightarrow \mathbb{R}$ by $f(x) = -\ln(x)$. Find f^* and f^{**} .

Solution:

By definition,

$$f^*(p) = \max_{x \in (0, \infty)} px - f(x) = \max_{x \in (0, \infty)} px + \ln(x). \quad (174)$$

Through differentiation we see

$$\frac{d}{dx} [px + \ln(x)] = p + \frac{1}{x}, \quad (175)$$

which implies the single critical point occurs at $x = -1/p = -p^{-1}$. And, this corresponds to a max since

$$\frac{d^2}{dx^2} [px + \ln(x)] = \frac{d}{dx} \left[p + \frac{1}{x} \right] = -\frac{1}{x^2} < 0. \quad (176)$$

Thus

$$f^*(p) = p(-p^{-1}) + \ln(-p^{-1}) = -[1 + \ln(-p)]. \quad (177)$$

Since the domain of the logarithm is $(0, \infty)$, we conclude the dual is $f^* : (-\infty, 0) \rightarrow \mathbb{R}$ defined by

$$\boxed{f^*(p) = -[1 + \ln(-p)]}. \quad (178)$$

Now observe

$$f^{**}(x) = \max_{p \in (-\infty, 0)} xp - f^*(p) = \max_{p \in (-\infty, 0)} xp + 1 + \ln(-p). \quad (179)$$

Differentiating reveals

$$\frac{d}{dp} [xp + 1 + \ln(-p)] = x + \frac{1}{p}, \quad (180)$$

which implies the single critical point is at $p = -x^{-1}$. And, this is a max since

$$\frac{d^2}{dp^2} [xp + 1 + \ln(-p)] = \frac{d}{dp} \left[x + \frac{1}{p} \right] = -\frac{1}{p^2} < 0. \quad (181)$$

Thus

$$f^{**}(x) = x(-x^{-1}) + 1 + \ln(-(-x)^{-1}) = -1 + 1 + \ln(x^{-1}) = -\ln(x), \quad (182)$$

and so $\boxed{f^{**} = f}$. □

Now let $\mathcal{L}(v, x, t)$ be some Lagrangian. Note here we use $x = x(t)$ instead of $y = y(t)$ and we will typically take $v = \dot{x}$. Then using the Legendre transform we can define a new quantity as follows.

Definition: Let $\mathcal{L}(v, x, t)$ be a Lagrangian. Then we define the corresponding *Hamiltonian* $H(p, x, t)$ is defined by

$$H(p, x, t) := \mathcal{L}^*(v, x, t) := \max_v pv - \mathcal{L}(v, x, t). \tag{183}$$

△

REMARK 32: Note the function H is *not* a function of v . Also, there are now two uses of the term \mathcal{L}^* in these notes. Earlier, we used this to describe an augmented Lagrangian; however, here we use this to describe the dual of a Lagrangian. It may always be inferred from the context which description is meant. ◇

Example 40: Consider a particle in space with its potential energy given by some function $U(x)$. Find the Hamiltonian for the system.

Solution:

The Lagrangian \mathcal{L} is given by $\mathcal{L} = T - U$ where $T = \frac{1}{2}m\dot{x}^2$ is the kinetic energy. Letting $v = \dot{x}$, we see

$$H = \mathcal{L}^* = \max_v pv - \mathcal{L} = \max_v pv - \frac{1}{2}mv^2 + U(x) = U(x) + \max_v pv - \frac{1}{2}mv^2. \tag{184}$$

Referring to Example 36, we may take $\alpha = \frac{1}{2}m$ to deduce

$$H = U(x) + \frac{p^2}{4(\frac{1}{2}m)} = U(x) + \frac{p^2}{2m}. \tag{185}$$

□

Some useful relations are that

$$p = \mathcal{L}_{\dot{x}} \quad \text{and} \quad \dot{p} = \mathcal{L}_x. \tag{186}$$

Then Hamilton's equations of motion are given by

$$H_x = -\dot{p} \quad \text{and} \quad H_p = \dot{x}. \tag{187}$$

REMARK 33: A curious property of the Hamiltonian should be noted. When a system follows Hamilton's equations of motion, we see the Hamiltonian is conserved when it does not explicitly depend on time. Indeed, observe

$$\dot{H} = H_x \dot{x} + H_p \dot{p} + H_t = H_x H_p + H_p (-H_x) + H_t = H_t. \quad (188)$$

So, if t is not explicitly given in H , then $\dot{H} = 0$. In many classical mechanics problems, we may verify the Hamiltonian equals the total energy and thus deduce energy is conserved when H does not depend on t . \diamond

Example 41: The potential for an anharmonic oscillator is $U = kx^2/2 + bx^4/4$ where k and b are constants. Find Hamilton's equations of motion.

Solution:

The Lagrangian \mathcal{L} for this system is given by

$$\mathcal{L} = T - U = \frac{1}{2}m\dot{x}^2 - k\frac{x^2}{2} - b\frac{x^4}{4}.$$

This implies the momentum is given by

$$p = \mathcal{L}_{\dot{x}} = m\dot{x}$$

so that the Hamiltonian H is

$$H = p\dot{x} - \mathcal{L} = m\dot{x}^2 - \left[\frac{1}{2}m\dot{x}^2 - k\frac{x^2}{2} - b\frac{x^4}{4} \right] = \frac{1}{2}m\dot{x}^2 + k\frac{x^2}{2} + b\frac{x^4}{4} = \frac{p^2}{2m} + k\frac{x^2}{2} + b\frac{x^4}{4}.$$

Consequently, Hamilton's equations of motion are

$$\boxed{\dot{x} = \frac{\partial H}{\partial p} = \frac{p_x}{m}} \quad \text{and} \quad \boxed{\dot{p}_x = -\frac{\partial H}{\partial x} = -kx - bx^3}.$$

\square

MIDTERM II MATERIAL

The second midterm will include the following material:

- i) Finding local minima without constraints
- ii) Finding local minima with constraints
- iii) Finding admissible directions/variatioins
- iv) Showing a simple function/functional is convex

REMARK 34: As mentioned earlier, it may be useful to refer to these [notes](#) for discussion about boundary conditions (e.g., see the list of cases on page 138). \diamond

PROBLEMS FOR STUDENTS

REMARK 35: Below we provide a few problems for students to try and rewrite in min-max form.

◇

Example 42: Rewrite the constrained optimization problem

$$\min_{y \in C^1[0,2]} \frac{1}{2} \int_0^2 y^2 + (y')^2 dx \quad \text{s.t.} \quad \int_0^1 y dx = 12 \quad (189)$$

as an unconstrained min-max problem. Then find a solution to the problem.

Example 43: Rewrite the constrained optimization problem

$$\min_{y \in C^2[0,2]} \frac{1}{2} \int_0^2 y^2 + 2yy' + (y')^2 dx \quad \text{s.t.} \quad y''(x) - y'(x) = 5x \quad \forall x \in [0, 1] \quad (190)$$

as an unconstrained min-max problem.

Example 44: Rewrite the constrained optimization problem

$$\min_{y \in C[0,2]} y(0)^2 - 3y(0) + 7 \quad \text{s.t.} \quad \int_0^1 y' dx = 0. \quad (191)$$

as an unconstrained min-max problem.

Example 45: Rewrite the constrained optimization problem

$$\min_{y \in C^1[0,2]} \int_0^2 (y - 2)^2 + (y - 4)^2 dx \quad \text{s.t.} \quad y'(x) = 1 \quad \forall x \in [0, 2]. \quad (192)$$

as an unconstrained min-max problem. Also find the minimizer and the minimum for the constrained problem.

Example 46: Rewrite the constrained optimization problem

$$\min_{y \in C^1[0,2]} \int_0^2 (y-2)^2 + (y-4)^2 dx \quad \text{s.t.} \quad y'(x) = 1 \quad \forall x \in [0,2]. \quad (193)$$

as an unconstrained min-max problem. Also find the minimizer and the minimum for the constrained problem.

Example 47: Find *all* solutions to the minimization problem

$$\min_{y \in C[0,1]} \int_0^1 \sin(y(x)) dx \quad \text{s.t.} \quad \int_0^1 y' dx = 0. \quad (194)$$

Since we have not yet covered how to handle natural boundary conditions, in each of the examples below you can simply find $\delta J(y, v)$ for appropriate v . At a later date, we can return and try to fully tackle these problems.

Example 48: Find the extremals for $J : C^1([0, 1]) \rightarrow \mathbb{R}$ defined by

$$J(y) := \int_0^1 (y')^2 + y^2 dx, \quad (195)$$

taking $y(0) = 2$ and $y(1)$ to be free.

Example 49: Find the extremals for $J : C^1([0, 4]) \rightarrow \mathbb{R}$ defined by

$$J(y) := \int_0^4 e^{2x} [(y')^2 - y^2] dx \quad (196)$$

taking $y(0) = 1$ and $y(4)$ to be free.

Example 50: Find the extremals for $J : C([0, 1]) \rightarrow \mathbb{R}$ defined by

$$J(y) := y(1)^2. \quad (197)$$

(Here $y(1)$ does denote the function y evaluated at 1.)

Example 51: Find the extremals for $J : C([0, 1]) \rightarrow \mathbb{R}$ defined by

$$J(y) := \int_0^1 4^y \, dx. \quad (198)$$

Example 52: Use the limit definition of the Gâteaux derivative to compute $\delta J(y, v)$ where $J : C([a, b]) \rightarrow \mathbb{R}$ is defined by

$$J(y) := \int_a^b y^2 - 6y \, dx. \quad (199)$$

Example 53: Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ are convex. Assume g is nondecreasing. Show the composition $f \circ g$ is convex.

Example 54: Show the functional $J : C[0, 1] \rightarrow \mathbb{R}$ defined by

$$J(y) := \int_a^b e^{y^2} \, dx \quad (200)$$

is convex.

Hint: Use the result of the previous example.

REFERENCES

- [1] Rudin, W., et al.: Principles of mathematical analysis, vol. 3. McGraw-hill New York (1964)
- [2] Smith, D.R.: Variational Methods in Optimization. Prentice-Hall (1974)
- [3] Troutman, J.L.: Variational calculus and optimal control: optimization with elementary convexity. Springer Science & Business Media (2012)