Purpose: This document is a compilation of notes generated for discussion in MATH 146 with reference credit due to J. David Logan’s text *Applied Mathematics*. 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

1 Introduction 1

2 Review Material 1

3 Examples with Functionals $J : V \to \mathbb{R}$ 7
  3.1 Simple Methods for Identifying Lower Bounds/Infimums . . . . . . . . . . . . . . . . . . . . 8
  3.2 The Gâteaux Derivative and Its Applications . . . . . . . . . . . . . . . . . . . . . . . . . . 12
Section 1: 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 discussion. And, the solutions to problems provided here illustrate the level of rigor desired from students this quarter.

Section 2: Review Material

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

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

and define

\[
f'(x) := \lim_{t \to x} \phi(t),
\]

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

\[
f'(1) = \lim_{t \to 1} \frac{f(t) - f(1)}{t - 1}
\]

\[
= \lim_{h \to 0} \frac{f(1+h) - f(1)}{(1+h) - 1}
\]

\[
= \lim_{h \to 0} \frac{(1+h)^2 - 1}{h}
\]

\[
= \lim_{h \to 0} \frac{1 + 2h + h^2 - 1}{h}
\]

\[
= \lim_{h \to 0} \frac{2h + h^2}{h}
\]

\[
= \lim_{h \to 0} 2 + h
\]

\[
= 2 + 0
\]

\[
= 2.
\]

Last Modified: 10/12/2017
Taylor’s Theorem: Let $I \subset \mathbb{R}$ be a neighborhood of $x_0$ and $n$ be a nonnegative integer. Suppose the function $f : I \to \mathbb{R}$ has $n+1$ derivatives. Then, for each point $x \neq x_0$ in $I$, there is a point $\xi$ strictly between $x$ and $x_0$ such that

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

(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

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

(5)

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

$$f(x+h) = f(x) + f’(x)h + \varepsilon(h),$$

(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

$$f’(x) = \lim_{h \to 0} f’(x) = \lim_{h \to 0} \left( \frac{f(x+h) - f(x) - \varepsilon(h)}{h} \right) = \lim_{h \to 0} \left( \frac{f(x+h) - f(x) - \varepsilon(h)}{h} \right) = f’(x) - \lim_{h \to 0} \frac{\varepsilon(h)}{h}. $$

(7)

Thus $\lim_{h \to 0} \varepsilon(h)/h = 0$. 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 $x \to x^*$ provided

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

(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

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

(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 $f : \mathbb{R} \to \mathbb{R}$ 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

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

But, using the reverse triangle inequality, we see

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

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

$$f(z) = f(\bar{x}) + f'(\xi)(z - \bar{x}) = f(\bar{x}) + f'(\xi)|z - \bar{x}| < f(\bar{x}) - \frac{|f'(\bar{x})|}{2}|z - \bar{x}| < f(\bar{x}),$$ (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 \( x \) to be a local minimizer of \( f \) is that \( f'(x) = 0 \). Below we provide several examples illustrating the use and limitations of this theorem.

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

\[
\min_{x \in \mathbb{R}} f(x),
\]

using only the above theorem and definition of a minimizer.

Solution:

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

\[
f'(x) = 2(x - 3) + 5 + 0 = 2x - 1. \tag{15}
\]

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

\[
f(z) = f\left(\frac{1}{2} + \delta \right) = \left(\frac{1}{2} + \delta \right)^2 - \left(\frac{1}{2} + \delta \right) + 12
\]

\[
= \left(\frac{1}{4} + \delta + \delta^2 \right) - \left(\frac{1}{2} + \delta \right) + 12
\]

\[
= \left(\frac{1}{4} - \frac{1}{2} + 12 \right) + \delta^2 \tag{16}
\]

\[
= f\left(\frac{1}{2} \right) + \delta^2
\]

\[
\geq f\left(\frac{1}{2} \right).
\]

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.
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, $f(0) = 0 > -\epsilon^3 = f(-\epsilon)$ for every $\epsilon > 0$ 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 $\mathbb{R}$. □

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. ◇

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 $f$. □
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,

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

Thus a condition for the theorem does not hold and so it cannot be applied. □
Section 3: Examples with Functionals $J : V \to \mathbb{R}$

**Definition:** A *functional* $J : V \to F$ is a mapping from a vector space $V$ to a field of scalars $F$, e.g., the real numbers $\mathbb{R}$.

**Definition:** We define the space $C^\alpha[a, b]$ to be the set of all functions $f : [a, b] \to \mathbb{R}$ such that $f^{(n)}$ is continuous.

**Remark 4:** We take the following approach in the examples below.

**Step 1:** Pick an arbitrary $y \in A$ and try to find a reasonable lower bound for $J(y)$ from the definition of $J$ and information about $y$ from $A$.

**Step 2a:** If we are able to find $f \in A$ such that $J(f)$ equals this lower bound, then $f$ is a minimizer and $J(f)$ is the minimum.

**Step 2b:** If a minimizer does not exist in $A$, then we look for a sequence $\{f_n\}$ contained in $A$ such that $J(f_n)$ converges to the lower bound. If this is can be done, then the lower bound is, in fact, an infimum for $J$.

**Remark 5:** We will have more sophisticated methods at our disposal later. The examples below are given to familiarize students with the notions of the admissibility class $A$, a mapping $J : A \to \mathbb{R}$, and the existence and values of minimums/infimums and minimizers.

**Remark 6:** Note the infimum of a set is the greatest lower bound of that set. Consider the set $S := (0, 1]$. Here $\inf S = 0$. Note, however, that 0 is a lower bound for $S$ and $\alpha$ is also a lower bound for $S$ for each $\alpha < 0$. 
3.1 – Simple Methods for Identifying Lower Bounds/Infimums:

**Example 7:** Define the admissibility class \( A := \{ f \in C[a,b] : f(x) \geq 5 \} \) and let \( J : C[a,b] \to \mathbb{R} \) be the functional defined by

\[
J(y) := \int_a^b y(x)^2 - 8y(x) + 20 \, dx. \tag{19}
\]

Find the minimum of \( J(y) \) for \( y \in A \). What is the minimizer?

**Solution:**

Let \( y \in A \). Then

\[
J(y) = \int_a^b y(x)^2 - 8y(x) + 20 \, dx
\]

\[
= \int_a^b (y(x)^2 - 8y(x) + 16) + 4 \, dx
\]

\[
= \int_a^b (y(x) - 4)^2 + 4 \, dx
\]

\[
\geq \int_a^b (5 - 4)^2 + 4 \, dx
\]

\[
= \int_a^b 5 \, dx
\]

\[
= 5(b - a). \tag{20}
\]

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 \( f \in A \) such that \( J(f) = 5(b - a) \). This is accomplished if and only if the inequality in (20) is a strict equality. The only candidate is \( f(x) = 5 \). Since \( f \) is continuous on \([a,b]\) and \( f \geq 5 \), we see \( f \in A \). Thus we conclude \( f(x) = 5 \) is the minimizer of \( J(y) \) over \( A \) and \( 5(b - a) \) is the minimum of \( J(y) \) over \( A \). \( \square \)

**Remark 7:** Note in the above example we say \( f(x) = 5 \) is “the” minimizer. This is because the is the only function in \( A \) that gives \( J(f) = 5(b - a) \). In the next example, multiple minimizers exist. \( \diamond \)

**Remark 8:** Note \( J \) is not a functional in this case because \( A \) does not form a vector space. This follows from the fact it is not closed under scalar multiplication. For example, if \( f \in A \), then \(-f \notin A \). \( \diamond \)
Example 8: Define the admissibility class \( \mathcal{A} := \{ f \in C[0,1] : f(x) \geq x^2 - 10x + 28 \} \). Then let \( J : C[0,1] \to \mathbb{R} \) be the functional defined by

\[
J(f) := \inf_{x \in [0,1]} f(x).
\] (21)

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

**Solution:**

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

\[
f(x) \geq x^2 - 10x + 28 = (x^2 - 10x + 25) + 3 = (x - 5)^2 + 3.
\] (22)

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 fast, we see

\[
J(f) = \inf_{x \in [0,1]} f(x) \geq \inf_{x \in [0,1]} g(x) = g(1) = (1 - 5)^2 + 3 = 19.
\] (23)

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

\[
J(f) \geq J(g) = 19 \quad \forall \ f \in \mathcal{A}.
\] (24)

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

Remark 9: 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,

\[
q'(x) = g'(x) + 2(x - 1) = 2(x - 5) + 2(x - 1) \leq 2(x - 5) + 0 < 0 \quad \forall \ x \in [0,1].
\] (25)

This shows \( q \) is strictly decreasing on \([0,1]\). Thus

\[
J(q) = \inf_{x \in [0,1]} q(x) = q(1) = g(1) + (1 - 1)^2 = g(1) = 19.
\] (26)

This shows \( g \) and \( q \) are minimizers of \( J \) over \( \mathcal{A} \). \( \diamond \)
Example 9: Define the function \( h : \mathbb{R} \to \mathbb{R} \) by
\[
h(x) := \begin{cases} 
0 & \text{if } |x| < 1, \\
1 & \text{if } |x| \geq 1.
\end{cases}
\] (27)

Define the admissibility class \( \mathcal{A} := \{ f \in C^1(\mathbb{R}) : f(x) \geq h(x) \} \). Then let \( J : \mathcal{A} \to \mathbb{R} \) be the mapping
\[
J(y) := \int_{-1}^{1} y(x) \, dx.
\] (28)

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 \( y \in \mathcal{A} \),
\[
J(y) = \int_{-1}^{1} y(x) \, dx \geq \int_{-1}^{1} h(x) \, dx = \int_{-1}^{1} 0 \, dx = 0.
\] (29)

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 \( J(f_n) \to 0 \). 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
\[
J(f_n) = \int_{-1}^{1} f_n(x) \, dx = \int_{-1}^{1} x^{2n} \, dx = 2 \int_{0}^{1} x^{2n} \, dx = 2 \left( \frac{1^{2n+1}}{2n+1} \right) = \frac{2}{2n+1} \leq \frac{1}{n} \quad (30)
\]

Taking the limit as \( n \to \infty \), we see
\[
0 \leq \lim_{n \to \infty} J(f_n) \leq \lim_{n \to \infty} \frac{1}{n} = 0.
\] (31)

Thus \( \lim_{n \to \infty} J(f_n) = 0 \), 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 10: 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 \(|x|^n \longrightarrow 0\) as \( n \longrightarrow \infty \) (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 \( n = 1, 3, 10 \).](image)

Remark 11: In class, I made a typo in (30), writing \( x^{2n+1} \) instead of \( 1^{2n+1} \). It has been corrected in this set of notes. For those of you still wondering about the limit I took there, however, I prove the following lemma.

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

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

\[
\alpha = \lim_{n \to \infty} c^n = \lim_{n \to \infty} c^{n+1} = c \lim_{n \to \infty} c^n = ca.
\]

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

11 Last Modified: 10/12/2017
3.2 – The Gâteaux Derivative and Its Applications:

**Definition:** We say \( y_0 \in A \) is a local minimizer for \( J \) over \( A \) if there exists \( \varepsilon > 0 \) such that \( J(y) \geq J(y_0) \) for all \( y \in A \) satisfying \( \|y - y_0\| \leq \varepsilon \).

**Definition:** Suppose \( V \) is a vector space. The Gâteaux derivative, denoted \( DJ(y_0)h \), of \( J : V \to \mathbb{R} \) at \( y \in V \) in the direction of \( h \) is defined as a linear map \( DJ(y_0)h : V \to \mathbb{R} \) such that

\[
DJ(y_0)h := \lim_{\varepsilon \to 0} \frac{J(y + \varepsilon h) - J(y)}{\varepsilon} = \frac{d}{d\varepsilon} [J(y + \varepsilon h)]_{\varepsilon=0},
\]

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

**Remark 12:** Here we list the steps for computing the Gâteaux derivative.

1. Identify \( J \) and \( A \).
2. Fix \( y \in A \). Let \( h \in V \) such that \( y + \varepsilon h \in A \) for all \( \varepsilon \) with \( |\varepsilon| \) sufficiently small.
3. Define \( f(\varepsilon) : \mathbb{R} \to \mathbb{R} \) by \( f(\varepsilon) := J(y + \varepsilon h) \).
4. Compute \( f'(\varepsilon) \).
5. Evaluate \( f'(0) \), which equals \( DJ(y_0)h \).
6. Check that \( DJ(y_0)h \) is linear in \( h \).

If each of these steps are complete and the last statement is in the affirmative, then \( DJ(y_0)h \) is the Gâteaux derivative of \( J \) at \( y \) in the direction \( h \).

**Theorem:** Suppose \( y_0 \) is a local minimizer of a functional \( J : V \to \mathbb{R} \) over an open set contained in \( A \). Then \( DJ(y_0)h = 0 \) for every admissible variation \( h \).
Remark 13: We can state the result of the above theorem more intuitively, using knowledge from calculus. Fix $h$ to be any admissible variation. Then define the function $f : \mathbb{R} \to \mathbb{R}$ by $f(\varepsilon) := J(y_0 + \varepsilon h)$. Since $y_0$ is a minimizer for $J$, it follows that $0$ is a local minimizer of $f$. Therefore $f'(0) = 0$. In other words,

$$0 = f'(0) = \frac{d}{d\varepsilon} [J(y_0 + \varepsilon h)]_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{J(y_0 + \varepsilon h) - J(y_0)}{\varepsilon} = DJ(y_0)h. \quad (34)$$

Lemma: Suppose $f \in C[a,b]$. Then

$$0 = \int_a^b f(x)h(x) \, dx \quad \forall \ h \in C[a,b] \implies f(x) = 0. \quad (35)$$

Remark 14: Here we list steps for finding minimizers of a functional $J : V \to \mathbb{R}$ over $A$.

1. Show $A$ is nonempty.

2. Define $f : \mathbb{R} \to \mathbb{R}$ by $J(y + \varepsilon h)$ for $y \in A$ and $y + \varepsilon h \in A$, where $h$ is an admissible variation.

3. Compute $f'(\varepsilon)$ and then $DJ(y)h = f'(0)$ and show $DJ(y)h$ is linear in $h$.

4. Use our theorem to assert that if there is a local minimizer $y_0 \in A$, then $DJ(y_0)h = 0$ for all admissible variations $h$.

5. Use our lemma to obtain an equation for $y_0$.

6. Use this equation to solve for each possible candidate minimizer $y_0$.

7. Show that $y_0 \in A$ and $y_0$ is a local minimizer. This requires showing there is $\varepsilon > 0$ such that $J(y_0) \leq J(y)$ for all $y \in A$ satisfying $\|y - y_0\| < \varepsilon$.

$\diamond$
Example 10: Define the admissibility class \( \mathcal{A} := C[a,b] \). Then let \( J : \mathcal{A} \to \mathbb{R} \) be the functional

\[
J(y) = \int_a^b (y(x) - 4)^2 \, dx.
\]

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

\[
J(y + \varepsilon h) = \int_a^b [y(x) + \varepsilon h(x) - 4]^2 \, dx
= \int_a^b [y(x) + \varepsilon h(x)]^2 - 8 [y(x) + \varepsilon h(x)] + 16 \, dx
= \int_a^b y(x)^2 + 2\varepsilon y(x) h(x) + \varepsilon^2 h(x)^2 - 8 [y(x) + \varepsilon h(x)] + 16 \, dx
= \int_a^b y(x)^2 - 8y(x) + 16 \, dx + 2\varepsilon \int_a^b [y(x) - 4] h(x) \, dx + \varepsilon^2 \int_a^b h(x)^2 \, dx
= J(y) + 2\varepsilon \int_a^b [y(x) - 4] h(x) \, dx + \varepsilon^2 \int_a^b h(x)^2 \, dx.
\]

Thus

\[
DJ(y)h = \lim_{\varepsilon \to 0} \frac{J(y + \varepsilon h) - J(y)}{\varepsilon}
= \lim_{\varepsilon \to 0} 2 \int_a^b [y(x) - 4] h(x) \, dx + \varepsilon \int_a^b h(x)^2 \, dx
= 2 \int_a^b [y(x) - 4] h(x) \, dx
\]

Let \( g \) be another admissible variation and \( \alpha, \beta \in \mathbb{R} \). Then observe

\[
DJ(y_0)(\alpha h + \beta g) = 2 \int_a^b [y_0(x) - 4] (\alpha h(x) + \beta g(x)) \, dx
= \alpha 2 \int_a^b [y_0(x) - 4] h(x) \, dx + \beta 2 \int_a^b [y_0(x) - 4] g(x) \, dx
= \alpha DJ(y_0)h + \beta DJ(y_0)g,
\]

which shows \( DJ(y_0)h \) is linear in \( h \). Whence \( DJ(y_0)h \) is the Gâteaux derivative at \( y_0 \) in the
direction $h$. By our theorem, we know that if $y_0$ is a minimizer of $J$ over $A$, then

$$0 = DJ(y_0)h = 2 \int_a^b [y_0(x) - 4] h(x) \, dx$$

(40)

for all admissible variations $h$. This implies the only candidate minimizer is $y_0(x) = 4$. Since $J(y) \geq 0$ for all $y \in A$ and $J(y_0) = 0$, we conclude $y_0(x) = 4$ is the global minimizer of $J$ over $A$. \qed

**Remark 15:** The above computations may seem quite tedious. This is because they are. It is important to note, if $g(\varepsilon) := J(y_0 + \varepsilon h)$, then the statement $q'(0) = 0$ is equivalent to saying $DJ(y_0)h = 0$ since

$$q'(0) = \lim_{\varepsilon \to 0} \frac{q(\varepsilon) - q(0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{J(y_0 + \varepsilon h) - J(y_0)}{\varepsilon} = DJ(y_0)h.$$ 

(41)

So, a more elegant approach for computing $DJ(y)h$ is given in the following reworking of the above example, utilizing our knowledge of derivatives for real valued functions.
Example 11: Repeat the previous example, making use of derivatives of real-valued functions.

Solution:
Fix \( y, h \in \mathcal{A} \). Then \( y + \varepsilon h \in \mathcal{A} \) for all \( \varepsilon \in \mathbb{R} \) since \( \mathcal{A} \) is a vector space. Define a function \( q : \mathbb{R} \to \mathbb{R} \) by \( q(\varepsilon) := J(y + \varepsilon h) \). Through direct computation we find

\[
q'(\varepsilon) = \frac{d}{d\varepsilon} J(y + \varepsilon h) = \frac{d}{d\varepsilon} \int_a^b \left[ y(x) + \varepsilon h(x) - 4 \right]^2 \, dx = \int_a^b \left[ 2y(x) + 2\varepsilon h(x) - 4 \right] h(x) \, dx. \tag{42}
\]

Thus

\[
q'(0) = \int_a^b 2[y(x) - 4] h(x) \, dx. \tag{43}
\]

If \( y = y_0 \), then \( q \) is minimized when \( \varepsilon = 0 \), thereby implying \( DJ(y_0) h = q'(0) = 0 \). That is,

\[
0 = q'(0) = \int_a^b 2[y(x) - 4] h(x) \, dx. \tag{44}
\]

Since this result holds for arbitrary admissible variations \( h \), our lemma states the only candidate minimizer is \( y_0(x) = 4 \). Let \( g \) be another admissible variation and \( \alpha, \beta \in \mathbb{R} \). Then observe

\[
DJ(y_0)(\alpha h + \beta g) = 2 \int_a^b [y_0(x) - 4] (\alpha h(x) + \beta g(x)) \, dx
= \alpha \int_a^b [y_0(x) - 4] h(x) \, dx + \beta \int_a^b [y_0(x) - 4] g(x) \, dx
= \alpha DJ(y_0) h + \beta DJ(y_0) g,
\]

which shows \( DJ(y_0) h \) is linear in \( h \). Whence \( DJ(y_0) h \) is the Gâteaux derivative at \( y_0 \) in the direction \( h \). By our theorem, we know that if \( y_0 \) is a minimizer of \( J \) over \( \mathcal{A} \), then

\[
0 = DJ(y_0) h = 2 \int_a^b [y_0(x) - 4] h(x) \, dx \tag{46}
\]

for all admissible variations \( h \). This implies the only candidate minimizer is \( y_0(x) = 4 \). Since \( J(y) \geq 0 \) for all \( y \in \mathcal{A} \) and \( J(y_0) = 0 \), we conclude \( \boxed{y_0(x) = 4} \) is the global minimizer of \( J \) over \( \mathcal{A} \).
Example 12: Let $V := C^1[0,1]$ and define the functional $J : V \to \mathbb{R}$ by

$$J(y) := \int_0^1 \frac{1}{2} m(y'(t))^2 - mgy(t) \, dt.$$  \hspace{1cm} (47)

Let $A := \{ y \in V : y(0) = \alpha_1, \; y(1) = \alpha_2 \}$. Find $y_0$ that minimizes $J$ over $A$. Here just find $y_0$ such that $DJ(y_0)h = 0$. You do not need to verify this $y_0$ is, in fact, a local minimizer of $J$.

Solution:

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

$$\alpha_1 = y(0) + \varepsilon h(0) = \alpha_1 + \varepsilon h(0) \Rightarrow h(0) = 0. \hspace{1cm} (48)$$

Similarly, $h(1) = 0$. Now, fixing $h$, define $g : \mathbb{R} \to \mathbb{R}$ by $f(\varepsilon) := J(y + \varepsilon h)$. We compute the derivative of $f$, which reveals

$$f'(\varepsilon) = \frac{d}{d\varepsilon} \int_0^1 \frac{1}{2} m(y' + \varepsilon h')^2 - mgy + \varepsilon h \, dt = \int_a^b m(y' + \varepsilon h') h' - mgh \, dt. \hspace{1cm} (49)$$

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

$$f'(\varepsilon) = \int_0^1 -m(y'' + \varepsilon h'')h - mgh \, dt = -m \int_a^b (y'' + \varepsilon h'' + g) h \, dt. \hspace{1cm} (50)$$

This shows

$$DJ(y)h = f'(0) = -m \int_a^b (y'' + g) h \, dt. \hspace{1cm} (51)$$

Let $\alpha, \beta \in \mathbb{R}$ and $r$ be another admissible variation. Then, by linearity of the integral,

$$DJ(y)(\alpha h + \beta r) = -m \int_a^b (y'' + \varepsilon h'' + g) (\alpha h + \beta r) \, dt$$

$$= -\alpha m \int_a^b (y'' + \varepsilon h'' + g) h \, dt - \beta m \int_a^b (y'' + \varepsilon h'' + g) r \, dt$$

$$= \alpha DJ(y)h + \beta DJ(y)r, \hspace{1cm} (52)$$

and so $DJ(y)h$ is linear in $h$. Thus $DJ(y)h$, as given above, is the Gâteaux derivative of $J$ at $y$ in the direction $h$. Now, if $y = y_0$, then we know 0 is a minimizer of $f$ and so $DJ(y_0)h = f'(0) = 0$. Thus

$$0 = f'(0) = -m \int_0^1 (y''_0 + g) h \, dt. \hspace{1cm} (53)$$
Since this holds for an arbitrary admissible variation $h$, by our lemma we deduce the only candidate minimizer satisfies $y_0'' = -g$. Then

$$y_0' = -gt + c_1 \implies y_0 = -\frac{1}{2}gt^2 + c_1t + c_2$$

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

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

(55) and we conclude

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

(56) \qed
Example 13: Let $V := C[0,1]$ and $A := V$. Define $J : V \to \mathbb{R}$ by

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

(57)

Find $y_0 \in A$ such that $DJ(y_0)h = 0$. Is $y_0$ a minimizer of $J$ over $A$?

Solution:

Let $y, h \in A$ and $\varepsilon \in \mathbb{R}$. Then define $g(\varepsilon) := J(y + \varepsilon h)$. We must find $y$ such that $g'(0) = 0$.

Computing the derivative, we see

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

(58)

This implies

$$DJ(y)h = g'(0) = \int_0^1 -2y h + 6h \, dx = \int_0^1 (-2y + 6) \, h \, dx.$$  

(59)

So, taking $y_0 = 3$, we obtain $DJ(y_0)h = 0$. Now observe

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

(60)

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

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

(61)

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 $J$. $\square$

Remark 16: The above example shows the condition $DJ(y_0)h = 0$ is not a sufficient condition to conclude $y_0$ is a minimizer of $J$. $\diamond$
Example 14: Let $V := C^2[0, 1]$ and $\mathcal{A} := \{y \in V : y(1) = \alpha_1, \ y'(0) = \alpha_2\}$. Define $J : V \to \mathbb{R}$ by

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

(62)

Show any minimizer of $J$ over $\mathcal{A}$ is a solution $y$ to the ODE

$$y'' = y, \quad y(1) = \alpha_1, \quad y'(0) = \alpha_2.$$  

(63)

You may assume a minimizer $y_0 \in \mathcal{A}$ exists.

Solution:

This is currently left as an exercise for the reader.

A solution will be posted on Monday (October 16). □