Math 118C HW1

Zih-Yu Hsieh

April 14, 2025

1

Question 1 *Rudin Pg. 239 Problem 1:*
If S is a nonempty subset of a vector space X. prove that the span of S is a vector space.

Pf:

(Remark: The notation $\mathbb{F}$ denotes the base field of the vector space X).

Let S' be the span of the set S. Then, S' is a collection of all arbitrary linear combinations of vectors in any finite subcollection of S.

Hence, for all $x \in S'$, there exists $x_1, ..., x_n \in S$, and $a_1, ..., a_n \in \mathbb{F}$, where $x = \sum_{k=1}^n a_k x_k$.

Which, the zero vector $\bar{0} \in S'$, since $0 = 0x$ for all $x \in S$.

For all $x, y \in S'$, there exists $x_1, ..., x_n, y_1, ..., y_m \in S$, and $a_1, ..., a_n, b_1, ..., b_m \in \mathbb{F}$, where $x = \sum_{k=1}^n a_k x_k$, and $y = \sum_{j=1}^m b_j y_j$. THen, the sum $x + y = \sum_{k=1}^n a_k x_k + \sum_{j=1}^m b_j y_j \in S'$, since it is a linear combination of $x_1, ..., x_n, y_1, ..., y_m \in S$.

Finally, for any $\lambda \in \mathbb{F}$, given $x \in S'$ above, $\lambda x \in S'$, since $\lambda x = \lambda \sum_{k=1}^n a_k x_k = \sum_{k=1}^n (\lambda a_k) x_k$, where each index $k \in \{1, ..., n\}$ satisfies $\lambda a_k \in \mathbb{F}$. Hence, λx is again a linear combination of $x_1, ..., x_n \in S$, showing that $\lambda x \in S'$.

Since the zero vector $\bar{0} \in S'$, S' is closed under addition (all $x, y \in S'$ has $x + y \in S'$), and it's closed under scalar multiplication (all $x \in S'$ and $\lambda \in \mathbb{F}$ satisfies $\lambda x \in S'$), hence S' (the span of S) is a vector space.

2

Question 2 *Rudin Pg. 239 Problem 4:*
Prove that null spaces and ranges of linear transformations are vector spaces.

Pf:

Let $\mathbb{F}$ be an arbitrary field, and V, W be arbitrary two vector spaces over base field $\mathbb{F}$, and $T \in \mathcal{L}(V, W)$ (an arbitrary linear transformation from V to W).

Null Space is a vector space:

The null space of T, $null(T) \subseteq V$ satisfies the following properties:

- $\bar{0}_V \in null(T)$: By definition, since $T\bar{0}_V = \bar{0}_W$, then $\bar{0}_V \in null(T)$.
- $null(T)$ is closed under addition: For all $u, v \in null(T)$, since $Tu, Tv = \bar{0}_W$, then $T(u+v) = Tu+Tv = \bar{0}_W + \bar{0}_W = \bar{0}_W$, hence $u + v$ also got mapped to $\bar{0}_W$, showing that $u + v \in null(T)$.
- $null(T)$ is closed under scalar multiplication: For all $v \in null(T)$ and $\lambda \in \mathbb{F}$, since $Tv = \bar{0}_W$, then $T(\lambda v) = \lambda Tv = \lambda \cdot \bar{0}_W = \bar{0}_W$, showing that λv also got mapped to $\bar{0}_W$, hence $\lambda v \in null(T)$.

With the above three conditions, $null(T)$ the null space of T, is a vector space.

Range is a vector space:

The range of T, $range(T) \subseteq W$ satisfies the following properties:

- $\bar{0}_W \in range(T)$: By definition, since $T\bar{0}_V = \bar{0}_W$, then $\bar{0}_W \in range(T)$.
- $range(T)$ is closed under addition: For all $u, v \in range(T)$, there exists $x, y \in V$, such that $Tx = u$, and $Ty = v$. Then, $T(x + y) = Tx + Ty = u + v$, showing that $u + v \in range(T)$.
- $range(T)$ is closed under scalar multiplication: For all $v \in range(T)$ and $\lambda \in \mathbb{F}$, since there exists $x \in V$, such that $Tx = v$, then $T(\lambda x) = \lambda(Tx) = \lambda v$, showing that $\lambda v \in range(T)$.

Again, with the above three conditions, $range(T)$ is a vector space.

3

Question 3 *Rudin Pg. 239 Problem 5:*
Prove that to every $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ corresponds to a unique $y \in \mathbb{R}^n$, such that $Ax = x \cdot y$. Prove also that $\|A\| = |y|$.

Pf:

Existence of y:

If we pick the standard orthonormal basis $e_1, ..., e_n \in \mathbb{R}^n$, which for every $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, let $a_i = Ae_i \in \mathbb{R}$ for all index $i \in \{1, ..., n\}$.

Now, consider the vector $y = \sum_{i=1}^n a_i e_i$:

For any $x \in \mathbb{R}^n$, there exists unique $b_1, ..., b_n \in \mathbb{R}$, such that $x = \sum_{i=1}^{n} b_i e_i$. Then, when apply the transformation and the dot product, we get the following:

$$Ax = A\left(\sum_{i=1}^{n} b_i e_i\right) = \sum_{i=1}^{n} b_i (Ae_i) = \sum_{i=1}^{n} b_i a_i$$

$$x \cdot y = \left(\sum_{i=1}^{n} b_i e_i\right) \cdot \left(\sum_{j=1}^{n} a_j e_j\right) = \sum_{i=1}^{n} b_i \left(e_i \cdot \sum_{j=1}^{n} a_j e_j\right) = \sum_{i=1}^{n} b_i a_i$$

(Note: Since $e_1, ..., e_n \in \mathbb{R}^n$ is an orthonormal basis, then $e_i \cdot e_j = 1$ if $i = j$, and $e_i \cdot e_j = 0$ if $i \neq j$). Hence, $Ax = x \cdot y$, showing that there exists such $y \in \mathbb{R}^n$, with $Ax = x \cdot y$.

Uniqueness of y:

Suppose $y, z \in \mathbb{R}^n$ are two vectors satisfying $Ax = x \cdot y$ and $Ax = x \cdot z$ for all $x \in \mathbb{R}^n$. Then, by the bilinearity of real dot product, we have:

$$0 = Ax - Ax = (x \cdot y) - (x \cdot z) = x \cdot (y - z)$$

However, notice that the choice of x is arbitrary. In particular, we can choose $x = (y - z) \in \mathbb{R}^n$, and get the following:

$$0 = (y - z) \cdot (y - z)$$

By the property of dot product, any $x \in \mathbb{R}^n$ satisfies $x \cdot x \geq 0$, and $x \cdot x = 0$ iff $x = \bar{0}$, hence the above equality implies $(y - z) = \bar{0}$, or $y = z$. This proves the uniqueness of such corresponding vector y of A.

Norm of A:

First, we need to consider the special case where $A = 0$ as a linear functional: For all $x \in \mathbb{R}^n$, since $Ax = 0$, and $x \cdot \bar{0} = 0$, then the unique vector corresponding to $A = 0$ the zero map, is $\bar{0}$. In this case, all $x \in \mathbb{R}^n$ with $|x| = 1$ satisfies $|Ax| = 0 = |\bar{0}|$, hence $\|A\| = \sup_{|x|=1} |Ax| = 0 = |\bar{0}|$.

Now, suppose $A \neq 0$. For all $x \in \mathbb{R}^n$ with $|x| = 1$, based on Cauchy-Schwartz Inequality, we can get the following relationship:

$$|Ax| = |x \cdot y| \leq |x| \cdot |y| = |y|$$

Hence, $\|A\| = \sup_{|x|=1} |Ax| \leq |y|$.

On the other hand, since $A \neq 0$, then the corresponding vector $y \neq \bar{0}$ (or else all $x \in \mathbb{R}^n$ would satisfy $Ax = x \cdot \bar{0} = 0$, which is a contradiction). Then, $|y| > 0$, which we can define a unit vector $\hat{y} = \frac{y}{|y|}$ with $|\hat{y}| = 1$. Because Cauchy-Schwartz Inequality achieves an equality when the two vectors are scalar multiple of each other, then since $\hat{y}$ is a scalar multiple of y, we get the following:

$$|A\hat{y}| = |\hat{y} \cdot y| = |\hat{y}| \cdot |y| = |y|$$

Hence, $|A\hat{y}| = |y| \leq \|A\|$.

The above two inequalities show that $\|A\| = |y|$.

4

Question 4 *Rudin Pg 239 Problem 7:*
Suppose that f is a real-valued function defined in an open set $E \subseteq \mathbb{R}^n$, and that the partial derivatives $D_1 f, ..., D_n f$ are bounded in E. Prove that f is continuous in E.

Pf:

First, since $D_1 f, ..., D_n f$ are all bounded in E, and there are finitely many such partial derivatives functions, then there exists a universal $M > 0$, such that regardless of the index $i \in \{1, ..., n\}$, all $x \in E$ satisfies $|D_i f(x)| \leq M$. Second, for all $x \in E$, since E is open, there exists $r > 0$, with the neighborhood $B_r(x) \in E$.

Now, notice that if we fix arbitrary $a \in E$ (where $x = (a_1, .., a_n)$), given corresponding $r > 0$, for any index $i \in \{1, ..., n\}$, fixing all other entries except for the i^{th} entry to the coordinates of a, we get a single-variable function, with its derivative given by the partial derivative:

$$f(a_1, ..., x_i, ..., a_n) : (a_i - r, a_i + r) \to \mathbb{R}$$

$$\frac{d}{dx_i} f(a_1, ..., x_i, ..., a_n) = \lim_{h \to 0} \frac{f(a_1, ..., x_i + h, ..., a_n) - f(a_1, ..., x_i, ..., a_n)}{h} = D_i f(a_1, ..., x_i, ..., a_n)$$

Then, by Mean Value Theorem of differentiable real single-valued function, for any $|h| < r$, since $a_i + h \in (a_i - r, a_i + r)$, then there exists c_i strictly between a_i and $a_i + h$ (which $|a_i - c_i| < |h| < r$), such that:

$$f(a_1, ..., a_i + h, ..., a_n) - f(a_1, ..., a_i, ..., a_n) = D_i f(a_1, ..., c_i, ..., a_n)(a_i + h - a_i) = D_i f(a_1, ..., c_i, ..., a_n) \cdot h$$

Hence, the following is true:

$$|f(a_1, ..., a_i + h, ..., a_n) - f(a)| = |D_i f(a_1, ..., c_i, ..., a_n)| \cdot |h| \leq M \cdot |h|$$

This shows the Lipschitz Continuity of the function f when only varying one coordinate.

Continuity of f:

To prove continuity, we'll go through an iterative process, by varying only one coordinate at a time:

Given any $a \in E$ (where $a = (a_1, ..., a_n)$), assume afterward we're working in an open neighborhood of a contained in E. For all $\epsilon > 0$ (which $\frac{\epsilon}{n} > 0$), choose $\delta = \frac{\epsilon}{nM} > 0$. Then, for all $h \in \mathbb{R}^n$ (where $h = (h_1, ..., h_n)$) with $|h| < \delta = \frac{\epsilon}{nM}$, each entry $|h_i| < \frac{\epsilon}{nM}$. Now, consider the vector $a + h = (a_1 + h_1, ..., a_n + h_n)$:

1. First, from the formula before, we know if we vary only the first entry, we get:

$$|f(a_1 + h_1, ..., a_n + h_n) - f(a_1, a_2 + h_2, ..., a_n + h_n)| \leq M \cdot |h_1| < M \cdot \frac{\epsilon}{nM} = \frac{\epsilon}{n}$$

2. Then, for the second point $(a_1, a_2 + h_2, ..., a_n + h_n)$, if we vary only the second entry, we get:

$$|f(a_1, a_2 + h_2, ..., a_n + h_n) - f(a_1, a_2, a_3 + h_3, ..., a_n + h_n)| \leq M \cdot |h_2| < M \cdot \frac{\epsilon}{nM} = \frac{\epsilon}{n}$$

i. At the i^{th} step (where $3 \leq i \leq n$), since in the previous steps, we've varied the first $(i-1)^{th}$ entries (starting with $(a_1, ..., a_{i-1}, a_i + h_i, ..., a_n + h_n)$), if only vary the i^{th} entry, we get:

$$|f(a_1, ..., a_{i-1}, a_i + h_i, ..., a_n + h_n) - f(a_1, ..., a_i, a_{i+1} + h_{i+1}, ..., a_n + h_n)| \leq M \cdot |h_i| < M \cdot \frac{\epsilon}{nM} = \frac{\epsilon}{n}$$

Then, from the above process, we get the following inequality;

$$|f(a+h)-f(a)| = \left|\sum_{i=1}^{n} f(a_1,...,a_{i-1},a_i+h_i,...,a_n+h_n) - f(a_1,...,a_i,a_{i+1}+h_{i+1},...,a_n+h_n)\right|$$

$$\leq \sum_{i=1}^{n} |f(a_1,...,a_{i-1},a_i+h_i,...,a_n+h_n) - f(a_1,...,a_i,a_{i+1}+h_{i+1},...,a_n+h_n)|$$

$$< \sum_{i=1}^{n} \frac{\epsilon}{n} = \epsilon$$

(Note: For each index i, we compare the difference of the function by removing the difference of the i^{th} entry, and each time the function is bounded by $\frac{\epsilon}{n}$, which is proven above).

So, the above process proves that for all $|h| < \delta = \frac{\epsilon}{nM}$, we have $|f(a+h)-f(a)| < \epsilon$, which proves that f is continuous at a. Then, since this choise of $a \in E$ is arbitrary, f is in fact continuous in E.

5

Question 5 *Rudin Pg. 239 Problem 8:*
Suppose that f is a differentiable real function in an open set $E \subseteq \mathbb{R}^n$, and that f has a local maximum at a point $x \in E$. Prove that $f'(x) = Df(x) = 0$.

Pf:

First, since f has a local maximum at x, then there exists a $r > 0$, such that any $y \in B_r(x)$ (a small open neighborhood of x), satisfies $f(y) \leq f(x)$.

Then, since f is differentiable implies the existence of all partial derivative and the uniqueness of the differential $Df(x)$, we know it is given as follow:

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x)\right)$$

So, to prove that $Df(x) = 0$, it suffices to prove that each partial derivative is 0 at x.

Let $x = (a_1, ..., a_n) \in \mathbb{R}^n$. For each $i \in \{1, ..., n\}$, the partial derivative is given as follow:

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \frac{f(a_1,...,a_i+h,...,a_n) - f(a_1,...,a_i,...,a_n)}{h}$$

Now, if we consider any $0 < |h| < r$, since $|(a_1,...,a_i+h,...,a_n)-(a_1,...,a_i,...,a_n)| = |(0,...,h,...,0)| = |h| < r$, then the vector $(a_1,...,a_i+h,...,a_n) \in B_r(x)$. Hence, $f(a_1,...,a_i+h,...,a_n) \leq f(x) = f(a_1,...,a_i,...,a_n)$, so the difference $f(a_1,...,a_i+h,...,a_n) - f(a_1,...,a_i,...,a_n) \leq 0$.

Then, there are two cases to consider:

- For all $h > 0$, the following is true:

$$\frac{f(a_1,...,a_i+h,...,a_n) - f(a_1,...,a_i,...,a_n)}{h} \leq 0 \implies \lim_{h \to 0} \frac{f(a_1,...,a_i+h,...,a_n) - f(a_1,...,a_i,...,a_n)}{h} \leq 0$$

- Else, for all $h < 0$, the following is true:

$$\frac{f(a_1, ..., a_i + h, ..., a_n) - f(a_1, ..., a_i, ..., a_n)}{h} \geq 0 \implies \lim_{h\to 0} \frac{f(a_1, ..., a_i + h, ..., a_n) - f(a_1, ..., a_i, ..., a_n)}{h} \geq 0$$

(Note: the above two inequalities are followed by the properties of limit).

Then, we can conclude the following:

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h\to 0} \frac{f(a_1, ..., a_i + h, ..., a_n) - f(a_1, ..., a_i, ..., a_n)}{h} = 0$$

So, because each partial derivative is 0, the differential $Df(x) = 0$.

Therefore, f is differentiable over E and $x \in E$ is a local maximum, implies that $Df(x) = 0$.

6

Question 6 *Rudin Pg. 239 Problem 11:*
If f and g are differentiable real functions in $\mathbb{R}^n$, prove that

$$D(fg) = f(Dg) + g(Df)$$

and that $D(1/f) = -f^{-2}(Df)$ wherever $f \neq 0$.

Pf:

Product Rule:

Given f, g differentiable real functions in $\mathbb{R}^n$, hence the differential Df, Dg are defined, such that for all $x \in \mathbb{R}^n$, there exists $\delta > 0$, with $|h| < \delta$ implies the following:

$$f(x+h) - f(x) = Df(x)(h) + o_f(h), \quad g(x+h) - g(x) = Dg(x)(h) + o_g(h)$$

$$\lim_{h\to 0} \frac{|o_f(h)|}{|h|} = 0, \quad \lim_{h\to 0} \frac{|o_g(h)|}{|h|} = 0$$

Then, if we consider the following for $0 < |h| < \delta$, we get:

$$|f(x+h)g(x+h) - f(x)g(x) - (f(x)(Dg(x)) + g(x)(Df(x)))h|$$

$$= |f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x) - (f(x)(Dg(x)) + g(x)(Df(x)))h|$$

$$\leq |f(x+h)(g(x+h) - g(x)) - f(x)(Dg(x))(h)| + |g(x)(f(x+h) - f(x)) - g(x)(Df(x))(h)|$$

$$= |f(x+h)((Dg(x))(h) + o_g(h)) - f(x)(Dg(x))(h)| + |g(x)((Df(x))(h) + o_f(h)) - g(x)(Df(x))(h)|$$

$$\leq |f(x+h) - f(x)| \cdot |Dg(x)(h)| + |f(x+h)| \cdot |o_g(h)| + |g(x)| \cdot |o_f(h)|$$

Now, since $Dg(x)$ is a linear transformation, then for all $h \in \mathbb{R}^n$, we have $|Dg(x)(h)| \leq \|Dg(x)\| \cdot |h|$. Which, we further get the following:

$$0 \leq \frac{|f(x+h)g(x+h) - f(x)g(x) - (f(x)(Dg(x)) + g(x)(Df(x)))h|}{|h|}$$

$$\leq \frac{|f(x+h) - f(x)| \cdot |Dg(x)(h)|}{|h|} + \frac{|f(x+h)| \cdot |o_g(h)|}{|h|} + \frac{|g(x)| \cdot |o_f(h)|}{|h|}$$

$$\leq |f(x+h)-f(x)| \cdot \frac{\|Dg(x)\| \cdot |h|}{|h|} + |f(x+h)| \cdot \frac{|o_g(h)|}{|h|} + |g(x)| \cdot \frac{|o_f(h)|}{|h|}$$

$$= |f(x+h)-f(x)| + |f(x+h)| \cdot \frac{|o_g(h)|}{|h|} + |g(x)| \cdot \frac{|o_f(h)|}{|h|}$$

Then, since f is differentiable, which implies f is continuous, hence $\lim_{h\to 0} f(x+h) = f(x)$. Then, taking the limit, we get:

$$0 \leq \lim_{h\to 0} \frac{|f(x+h)g(x+h) - f(x)g(x) - (f(x)(Dg(x)) + g(x)(Df(x)))h|}{|h|}$$

$$\leq \lim_{h\to 0} |f(x+h)-f(x)| + |f(x+h)| \cdot \frac{|o_g(h)|}{|h|} + |g(x)| \cdot \frac{|o_f(h)|}{|h|} = 0$$

(Note: we have $lim_{h\to 0}|f(x+h)-f(x)| = 0$, $\lim_{h\to 0} |f(x+h)| \cdot \frac{|o_g(h)|}{|h|} = |f(x)| \cdot 0 = 0$ based on the definition of $o_g(h)$, and $\lim_{h\to 0} |g(x)| \cdot \frac{|o_f(h)|}{|h|} = 0$ again by the definition of $o_f(h)$).

Hence, since $A = f(x)Dg(x) + g(x)Df(x)$ is a linear transformation satisfying the following limit:

$$\lim_{h\to 0} \frac{|f(x+h)g(x+h) - f(x)g(x) - Ah|}{|h|} = 0$$

Then, fg is in fact differentiable at $x \in \mathbb{R}^n$. And, by the uniqueness of derivative, $A = f(x)Dg(x)+g(x)Df(x)$ is the derivative of fg at x. Therefore, the general formula of derivative is given by:

$$D(fg) = f(Dg) + g(Df)$$

Derivative of $1/f$:

Given f is differentiable, and $f(x) \neq 0$ for given $x \in \mathbb{R}^n$, then since $\frac{1}{f(x)}$ is defined, $1 = \frac{1}{f(x)} \cdot f(x)$, then $0 = D(1) = 1/f(Df) + f(D(1/f))$. Hence, the derivative $D(1/f)$ is given by:

$$f(D(1/f)) = -\frac{1}{f}(Df), \quad D(1/f) = -\frac{1}{f^2}(Df)$$