Solutions to Homework #1, Math116

Keziah Cook and Michael McElroy

October 26, 2003

Problem 2.3 (Luenberger)

Let M and N be subspaces of a vector space. Show that $[M \cup N] = M + N$.

Solution to 2.3

To show that two sets X and Y are equal, we first show that X contains Y and then that Y contains X.

- Let $x \in [M \cup N]$. Then x can be expressed as a linear combination of elements in $M \cup N$. We write

  $$x = \sum_{i=1}^{n} \alpha_i m_i + \sum_{j=1}^{m} \beta_j n_j$$

  for some $m_i \in M$, $n_i \in N$.

  Recall, α_i and β_j are scalars in the base field. Also, note the some of the m_i and n_j could be in $M \cap N$. In this case, we just assign them to either thesum of elements in M or the sum of element in N.

  Because M is a subspace and thus closed under addition and scalar multiplication, $\sum_{i=1}^{n} \alpha_i m_i \in M$. Likewise, $\sum_{j=1}^{m} \beta_j n_j \in N$. Hence, $x \in M + N$.

  This shows $[M \cup N] \subseteq M + N$.

- Let $y \in M + N$. $\exists m \in M$, $n \in N$ such that $y = m + n$ (by definition of $M + N$.) Because $m, n \in M \cup N$, a linear combination $m + n$ is an element of the span of $M \cup N$, $[M \cup N]$.

  Thus $M + N \subseteq [M \cup N]$.

Therefore $[M \cup N] = M + N$ as desired. □

Problem 2.4 (Luenberger)

A convex combination of the vectors $x_1, x_2, \ldots, x_n$ is a linear combination of the form $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ where $\alpha_i \geq 0$, for each i; and $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$. Given a set S in a vector space, let K be the set of vectors consisting of all convex combinations from S. Show that $K = \text{co}(S)$.

Solution to 2.4

We recall that $\text{co}(S)$ is the smallest convex set containing S or the intersection of all convex sets containing S . First we will show that K is a convex set containing S, which proves $\text{co}(S) \subseteq K$. Then we will show that any convex set containing S also contains K, which proves $K \subseteq \text{co}(S)$.

- $\text{co}(S) \subseteq K$

  To show that $S \subset K$, let $x \in K$. We simply note that for $\alpha = 1$, $x = \alpha x$ is a convex combination. Thus $x \in K$ and we see that K contains S.

  To see that K is convex, we take two elements $p, q \in K$. Each of these can be expressed as a convex combination of elements in S.

  $$p = \sum_{i=1}^{n} \alpha_i x_i$$

  $$q = \sum_{j=1}^{m} \beta_j y_j$$

  Thus for any γ such that $0 \leq \gamma \leq 1$, $z = \gamma p + (1 - \gamma) q$ is also a convex combination. To see this, we expand $z = \gamma \sum_{i=1}^{n} \alpha_i x_i + (1 - \gamma) \sum_{j=1}^{m} \beta_j y_j$ and note that the coefficients sum to 1.

  $$\gamma \sum_{i=1}^{n} \alpha_i + (1 - \gamma) \sum_{j=1}^{m} \beta_j = \gamma \cdot (1) + (1 - \gamma) \cdot 1 = 1$$

- $K \subseteq \text{co}(S)$

  Let X be a convex set containing S. Then X must also contain all elements of the form $\alpha x_1 + (1 - \alpha) x_2$, for all $x_l, x_2 \in X$. To see that every element $k \in K$ is in X, first express k as a convex combination $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ where $x_i \in S$, and proceed by induction of n (Note, n is the number of elements from S used in the linear combination k, not the cardinality of S.)

  - Base Case: If $n = 2$, then as X is convex, $k \in X$.
  - Inductive Hypothesis: Assume that convex combinations of $n - 1$ elements of S are in the convex set X.
  - Inductive Step: Let k be as above. For any value of n, we can rewrite k in the form $\alpha_1 x_1 + (1 - \alpha_1) l$ where $x \in S$ and $l \in K$ is a convex combination of exactly $n - 1$ elements of S. To do this, let $l = \beta_1 x_2 + \cdots + \beta_{n-1} x_n$ where $\beta_i = \frac{\alpha_{i+1}}{(1-\alpha_1)}$. Applying our inductive hypotheses, we see that l is in X. Recalling the definition of a convex set shows that X which contains l must also contain k. By induction this holds for all n, and thus for all $k \in K$.

  Therefore, $K \subseteq \text{co}(S)$

$\text{co}(S) \subseteq K$ and $K \subseteq \text{co}(S)$ imply $K = \text{co}(S)$ and the proof is completed.

□

Problem 2.5 (Luenberger)

Let C and D be convex cones in a vector space. Show that $C \cap D$ and $C + D$ are convex cones.

Solution to 2.5

As the problem does not mention the vertices of C and D, we will assume both have the origin as a vertex.

- $C \cap D$ is a convex cone.

  By Proposition 2, $C \cap D$ is convex, thus it suffices to show $x \in C \cap D$ implies $\alpha x \in C \cap D$ for all $\alpha \geq 0$. Fix $x \in C \cap D$. Clearly, $x \in C$ and $x \in D$. Since C and D are both cones, we know that $\alpha x \in C$ and $\alpha x \in D$. Thus, $\alpha x \in C \cap D$ as desired.

- $C + D$ is a convex cone.

  By Proposition 1, $C + D$ is convex. Let $x \in C + D$. Write $x = c + d$, for $c \in C$, $d \in D$. Then

  $$\alpha x = \alpha (c + d) = \alpha c + \alpha d$$

  As C and D are cones, $\alpha c \in C$ and $\alpha d \in D$. Thus $\alpha x \in C + D$, as desired.

$\square$

Problem 2.7 (Luenberger)

Prove that the intersection of an arbitrary collection of closed sets is closed and that the union of a finite collection of closed sets is closed.

Solution to 2.7

Let $\{X_i \,|\, i \in \mathcal{I}\}$ be an infinite collection of closed sets. We want to prove that both $\bigcap_{i \in \mathcal{I}} X_i$ and $\bigcup_{j \in \mathcal{J}} X_j$ are closed, where $\mathcal{J} \subset \mathcal{I}$ is finite. Remember the notation that for any set S let $\overline{S}$ be its closure. We'll start with the former statement.

- We want to show that $\bigcap_{i \in \mathcal{I}} X_i = \overline{\bigcap_{i \in \mathcal{I}} X_i}$. Trivially we have $\bigcap_{i \in \mathcal{I}} X_i \subset \overline{\bigcap_{i \in \mathcal{I}} X_i}$. To show the other inclusion, let x be a point of closure for $\bigcap_{i \in \mathcal{I}} X_i$. Then for any $\delta > 0$ there is a $y \in \bigcap_{i \in \mathcal{I}} X_i$ such that $\|x - y\| < \delta$. Since y is in the intersection, we have $y \in X_i$ for all $i \in \mathcal{I}$, so x is a closure point for all X_i. Since the X_i are closed, we have $x \in X_i$ for all i, and so $x \in \bigcap_{i \in \mathcal{I}} X_i$. Hence $\overline{\bigcap_{i \in \mathcal{I}} X_i} \subset \bigcap_{i \in \mathcal{I}} X_i$.

- We want to show that $\overline{\bigcup_{j \in \mathcal{J}} X_j} = \bigcup_{j \in \mathcal{J}} X_j$. We will prove that for any pair of closed sets their union is closed, i.e. for all $k, \ell \in \mathcal{I}$ we have that $\overline{X_k \cup X_\ell} = X_k \cup X_\ell$. We can then build up to arbitrary finite unions one by one.

  Without loss of generality we consider X_1 and X_2. Showing equality means showing that the both are subsets of each other. It is trivial that $X_1 \cup X_2 \subset \overline{X_1 \cup X_2}$. Now we show that $\overline{X_1 \cup X_2} \subset X_1 \cup X_2$. We'll achieve this not by choosing $x \in \overline{X_1 \cup X_2}$ and showing that $x \in X_1 \cup X_2$; instead, we'll show $y \notin X_1 \cup X_2 \implies y \notin \overline{X_1 \cup X_2}$. Given $y \notin X_1 \cup X_2$, choose $\delta_1, \delta_2 > 0$ such that there is no $x_1 \in X_1$ and no $x_2 \in X_2$ such that $\|y - x_1\| < \delta_1$ and $|y - x_1\| < \delta_2$. We can do this because y is neither in X_1 or X_2, both of which are closed sets. Thus there is no $x_3 \in X_1 \cup X_2$ such that $\|y - x_3\| < \min(\delta_1, \delta_2)$ and hence $y \notin \overline{X_1 \cup X_2}$.

$\square$

Problem 2.8 (Luenberger)

Show that the closure of a set S in a normed space is the smallest closed set containing S.

Solution to 2.8

Consider the set of closed sets containing S: $\mathcal{C} \equiv \left\{C_i \,|\, S \subset C_i = \overline{C_i}\right\}$. The logical diagram for what we want to prove is

$$\left\{\forall C_i \in \mathcal{C},\, \overline{S} \subset C_i\right\} \iff \left\{x \in \overline{S} \implies \forall C_i \in \mathcal{C},\, x \in C_i\right\} \iff \left\{x \notin \overline{S} \Longleftarrow \forall C_i \in \mathcal{C},\, x \notin C_i\right\}.$$

We will prove the statement on the right. Consider any $C_i \in \mathcal{C}$: let some $x \notin C_i$ be given. Then $\exists \delta > 0$ s.t. there does not exist $y \in C_i$ s.t. $\|x - y\| < \delta$; alternatively we can say $\exists \delta$ s.t. $B(x, \delta) \cap C_i = \emptyset$. Since $S \subset C_i$, we have $B(x, \delta) \cap S = \emptyset$, and x is not a closure point of S. Hence $x \notin \overline{S}$. $\square$

Problem 2.9 (Luenberger)

Let X be a normed linear space and let $x_1, x_2, \ldots, x_n$ be linearly independent vectors from X. For fixed $y \in X$, show that there are coefficients $a_1, a_2, \ldots, a_n \in \mathbb{R}$ minimizing $\|y - a_1x_1 - a_2x_2 \cdots - a_nx_n\|$.

Solution to 2.9

We will solve the problem by invoking the Weierstrass theorem. I start with two wrong ideas, which many students tried.

Wrong Idea 1 One may choose a basis for every given vector space. We want a basis for X that includes $\{x_1, \ldots, x_n\}$: choose an independent set $\{z_i \,|\, i \in \mathcal{I}\}$ independent of $\{x_1, \ldots, x_n\}$ s.t. $\{x_1, \ldots, x_n\} \cup \{z_i \,|\, i \in \mathcal{I}\}$ is a basis for X. Now express y in terms of the basis: $y = \sum_{j=1}^n b_jx_j + \sum_{i \in \mathcal{I}} b_iz_i$, the b's being real numbers.

$$\left\| y - \sum_{j=1}^n a_jx_j \right\| = \left\| \sum_{j=1}^n (b_j - a_j)x_j + \sum_{i \in \mathcal{I}} b_iz_i \right\| \le \left\| \sum_{j=1}^n (b_j - a_j)x_j \right\| + \left\| \sum_{i \in \mathcal{I}} b_iz_i \right\|$$

So does choosing $a_1 = b_1$, $a_2 = b_2$, ..., $a_n = b_n$ minimize $\|y - a_1x_1 - \ldots - a_nx_n\|$? No. This is not implied by the above inequality (which is true). See the example.

Example. Consider the simplest example: $X = \mathbb{R}^2$, $y = (12, 5)$, $x_1 = (1, 0)$ and $z_1 = (1, 1)$. Now $y = 7x_1 + 5z_1$, but $\|y - 7x_1\| = \|5z_1\| = 5\sqrt{2}$, while $\|y - 12x_1\| = \| - 5x_1 + 5z_1\| = 5$. So in general it is not the case that $a_1 = b_1, \ldots, a_n = b_n$ minimizes the norm.

Wrong idea 2 We are choosing $\{a_1, a_2, \ldots, a_n\}$ from $\mathbb{R}^n$, which is a finite dimensional space and therefore complete. Hence any convergent sequence in $\mathbb{R}^n$ will converge in $\mathbb{R}^n$. For any $\vec{a} = \{a_1, a_2, \ldots, a_n\} \in \mathbb{R}^n$, define $f(\vec{a}) = \|y - \sum_i a_ix_i\|$. So $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous function (see below, also proved in class, that the norm is a continuous function). We want to prove that f has a minimum. If f did not have a minimum, then we could choose a sequence of $\vec{a}_j$'s such that $f(\vec{a}_j)$ was a decreasing sequence. Since f is bounded below, we have that the sequence $\{f(\vec{a}_j)\}$ converges. However, this does not mean that the $\vec{a}_j$'s converge, which is what we needed. You see, if we have a continuous map $T : A \to B$, and a convergent sequence $\{a_i\} \subset A$, then the sequence $\{T(a_i)\} \subset B$ must converge within B. The converse is not generally true, however.

Example. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \begin{cases} 1 & |x| \le 1 \\ |1/x| & 1 < |x| \end{cases}$. f is bounded below by 0 and is continuous. Clearly we may choose a sequence $(x, y)_i = (i, 0)$, and $f(i, 0) = 1/i$ converges to 0 (in the range space $\mathbb{R}$), but the sequence $(x, y)_i = (i, 0)$ does not converge (in the domain space $\mathbb{R}^2$).

Solution Let a be the n-vector $[a_1, a_2, \ldots, a_n] \in \mathbb{R}^n$. Let's define a function f to minimize: $f : \mathbb{R}^n \times X \to \mathbb{R}$, where for $z \in X$, $a \in \mathbb{R}^n$ we have $f(a, z) = \|z - \sum_{i=1}^n a_ix_i\|$. Our problem is to show that $f(a, y)$ has a minimum over a for fixed y.

We want to use the Weierstrass theorem, that a continuous function on a compact space achieves a maximum and a minimum on the compact space. First we need to show that $f(a, z)$ is continuous over $\mathbb{R}^n$ for fixed $z \in X$. For $a, b \in \mathbb{R}^n$ we have

$$\begin{aligned}
f(a,z) - f(b,z) &= \left\| z - \sum_{i=1}^n a_i x_i \right\| - \left\| z - \sum_{i=1}^n b_i x_i \right\| \\
&\leq \left\| \left(z - \sum_{i=1}^n a_i x_i \right) - \left(z - \sum_{i=1}^n b_i x_i \right) \right\| = \left\| \sum_{i=1}^n (b_i - a_i) x_i \right\| \\
&\leq \sum_{i=1}^n |a_i - b_i| \cdot \|x_i\| \\
&\leq \max_i \|x_i\| \cdot \sum_{i=1}^n |a_i - b_i| = \max_i \|x_i\| \cdot \|a - b\|_1 \\
&\leq \sqrt{n} \max_i \|x_i\| \cdot \|a - b\|_2
\end{aligned}$$

So we have related the difference between function values of f to distance between a and b in Euclidean space. We see that f is upper-semicontinuous (over $\mathbb{R}^n$), and switching a with b gives that f is lower semicontinuous. So $f(a,z)$ is continuous for fixed z, and in particular, for $z = y$.

Now we need to work with a compact space, but only closed and bounded subsets of $\mathbb{R}^n$ are compact. Our intuition tells us that for $\|a\|_2$ really big, $f(a,y)$ will be big[1]. So we only want to consider a ball about 0 in $\mathbb{R}^n$. How big should this ball be to guarantee that the minimum point of $f(a,y)$ will be inside? We know that $f(0,y) = \|y\|$, so this is an upper bound for any minimum of $f(a,y)$.

The sphere in n dimensions, S^{n-1}, is compact, and so $f(\cdot, \theta)$ has a minimum on it,[2] say, at a^*. Choose $N > 0$ s.t. $N \cdot f(a^*, \theta) > 2\|y\|$. Then we have for any a such that $\|a\|_2 > N$

$$\begin{aligned}
f(a,y) &= \left\| y - \sum_i a_i x_i \right\| \geq \left\| \sum_i a_i x_i \right\| - \|y\| = \|a\|_2 \cdot \left\| \sum_i \frac{a_i}{\|a\|_2} x_i \right\| - \|y\| \\
&\geq N \cdot \left\| \sum_i a_i^* x_i \right\| - \|y\| > 2\|y\| - \|y\| = \|y\|.
\end{aligned}$$

Now we are done, because we consider $B(0,N) \subset \mathbb{R}^n$, which is compact, and so $f(a,y)$ achieves a minimum on $B(0,N)$, which will also be a minimum for all of $\mathbb{R}^n$.

□

Problem 1 (Writing)

Find or think up another theoretical or applied problem related to convexity and optimization of interest to you and write up a brief description of it in the styple of section 1.2 of Luenberger.

Solution to 1

There are many, many acceptable responses. One problem I've encountered recently in the area of public economics is that of finding the optimal non-linear tax schedule for a country. If you view a tax schedule as a continuous function from income to tax payments, then the problem is that of maximizing a specified social welfare over the set of possible tax schedules.

[1] Observe that we use the Euclidena norm $\|\cdot\|_2$ in $\mathbb{R}^n$, which should not be confused with the given norm $\|\cdot\|$ on X.

[2] Here θ is the zero element in X

Please feel free to post your answers to this question on the course website, and to comment on the problems posed by your peers. □

Problem 2 (Writing)

> Find and describe another situation where a vector space other than $\mathbb{R}^n$ is important. Explain and give references. How do you know it's different?

Solution to 2

Consider the L_2 functions on $\mathbb{R}$. This means all functions $f : \mathbb{R} \to \mathbb{R}$ s.t. $\int_{-\infty}^{\infty} x^2(t)\,\mathrm{d}t < \infty$. In Fourier analysis[3] you learn that you can write any such function as an infinite sum of sines and cosines. This is handy for solving partial differential equations, like the heat equation. Fourier transforms are ubiquitous in electrical engineering. The space $L_2(\mathbb{R})$ is infinite dimensional, so it is very different than the finite dimensional Euclidean space $\mathbb{R}^n$. □

[3]See the book by Tom K orner