Homework #2 Solutions

Thayer Anderson, Davis Lazowski, Handong Park, Rohil Prasad
Eric Peterson

1 For submission to Thayer Anderson

Problem 1.1. A function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l$ is bilinear when it satisfies:

$$f(cx, y) = f(x, cy) = cf(x, y)$$
$$f(x + x', y) = f(x, y) + f(x', y)$$
$$f(x, y + y') = f(x, y) + f(x, y')$$

1. Show that if f is bilinear, then
$$\lim_{(h,k)\to 0} \frac{||f(h, k)||}{||(h, k)||} = 0.$$
2. Prove $D_{(a,b)}f(x, y) = f(a, y) + f(x, b)$
3. Conclude from this the product rule for functions with target $\mathbb{R}$.

Solution. 1. Analogous to the Problem 1.1.2, we wish to first show that f is a bounded operator. (Although we will be more generous in this case and show that it behaves at worst quadratically). Take orthonormal bases $(e_1, \ldots, e_n)$ and $(g_1, \ldots, g_m)$ of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Suppose that $x = \sum_{i=1}^n a_i e_i + \sum_{i=1}^m b_i g_i \in \mathbb{R}^l$. Then

$$\begin{aligned} ||f(x)|| &\le \sum_{i=1}^n \sum_{j=1}^m |a_i| \cdot |b_j| ||f(e_i, g_j)|| \\ &\le |\max a_i| \cdot |\max b_i| \cdot nm \sum_{i,j} ||f(e_i, g_j)|| \\ &\le \left(nm \sum_{i,j} ||f(e_i, g_j)|| \right) \cdot (||u|| + ||v||)) \end{aligned}$$

where we let $x = (u, v)$ with $U \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$. Let $M = \Big(nm \sum_{i,j} ||f(e_i, g_j)||\Big)$. Then

$$||f(x)|| \le M \cdot (||u|| + ||v||) \le M \cdot 2(\max \{||u||, ||v||\}) \le 2M||x||$$

Then we see that f is bounded and retrieve the limiting behaviour identically to Problem 1.1.2.

2. We will test our putative derivative to verify that it is, in fact, the derivative:

$$\begin{aligned} &\lim_{(x,y)\to 0} \frac{||f(a + x, b + y) - (f(a, b) + f(a, y) + f(x, b))||}{||(x, y)||} \\ &= \lim_{(x,y)\to 0} \frac{||f(a + x, b + y) - (f(a + x, b + y) - f(x, y)))}{||(x, y)||} \\ &= \lim_{(x,y)\to 0} \frac{||f(x, y)||}{||(x, y)||} \end{aligned}$$

and by the previous problem this limit is 0 and the derivative is as expected.

3. Define the function $p : \mathbb{R}^2 \to \mathbb{R}$ as $p(x, y) = xy$. This function is evidently linear. Then the derivative of p at (a, b) is given by the previous part:

$$\left(D_{(a,b)}p\right)(x, y) = p(a, y) + p(x, b) = ay + xb$$

Take functions $f, g : \mathbb{R} \to \mathbb{R}$. Then:

$$\begin{aligned} D_a(fg) = D_a p(f, g) = (D_{(f(a),g(a))}p) \circ (D_a f, D_a g) \\ = (f(a)y + xg(a))(D_a f, D_a g) = f(a)(D_a g) + (D_a f)g(a) \end{aligned}$$

and we are done. (TA)

Problem 1.2. Here's the same problem with more indices. Consider the product $V = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ of Euclidean spaces of various dimensions. A function $f : V \to \mathbb{R}^l$ is called multilinear when for any choice of tuple of vectors $(v_1, \ldots, v_m)$ the function

$$\begin{aligned} f_j : \mathbb{R}^{n_j} \to \mathbb{R}^l \\ f_j(v) = f(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_m) \end{aligned}$$

describes a linear function.

1. For any point $(a_1, \ldots, a_m) \in V$, difference vector $(h_1, \ldots, h_m) \in V \simeq T_a V$ and pair of coordinates $i < j$, show the following:

$$\lim_{||h|| \to 0} \frac{||f(a_1, \ldots, a_{i-1}, h_i, a_{i+1}, \ldots, a_{j-1}, h_j, a_{j+1}, \ldots, a_m)}{||h||} = 0$$

2. Conclude that the derivative of f is given by

$$(D_a f)(h) = \sum_{j=1}^{m} f(a_1, \ldots, a_{j-1}, h_j, a_{j+1}, \ldots, a_m).$$

Solution. 1. Take the restriction of f to the arguments i and j of interest: $g : \mathbb{R}^{n_i} \times \mathbb{R}^{n_j} \to \mathbb{R}^l$ defined by

$$g(h_i, h_j) = f(a_1, \ldots, a_{i-1}, h_i, a_{i+1}, \ldots, a_{j-1}, h_j, a_{j+1}, \ldots, a_m)$$

g inherits bilinearity from the multi-linearity of f. Then we see:

$$\lim_{||h|| \to 0} \frac{||g(h_i, h_j)||}{||h||} \leq \lim_{||(h_i, h_j)|| \to 0} \frac{||g(h_i, h_j)}{||(h_i, h_j)||} = 0$$

after an application of 1.1.1.

2. We test our candidate derivative in the limit:

$$\lim_{||h|| \to 0} \frac{||f(a_1 + h_1, \ldots, a_m + h_m) - f(a_1, \ldots, a_m) - \sum_{j=1}^{m} f(a_1, \ldots, a_{j-1}, h_j, a_{j+1}, \ldots, a_m)||}{||h||}$$

After applying multilinearity to the value of $f(a + h)$ (with vectors a and h) we are left with a sum over terms with every combination of an a or an h in each entry. The candidate derivative cancels all the terms with exactly one h. The value of $f(a)$ cancels the term with no hs. Now it remains to be shown that those terms with 2 or more hs go to 0 in the limit.

Consider the quantity

$$\lim_{||h|| \to 0} \frac{||f(a_1, \ldots, h_{i_1}, \ldots, h_{i_2}, \ldots, h_{i_k}, \ldots, a_m)||}{||h||}$$

Where there is some enumeration i_j of the k-many indices containing an h instead of an a. Then:

$$\begin{aligned}
\lim_{||h|| \to 0} \frac{||f(a_1, \ldots, h_{i_0}, \ldots, h_{i_1}, \ldots, h_{i_k}, \ldots, a_m)||}{||h||} &= \lim_{||h|| \to 0} \prod_{j=3}^{k} |h_j| \cdot \frac{||f(a_1, \ldots, h_{i_1}, \ldots, h_{i_2}, \ldots, 1, \ldots, a_m)||}{||h||} \\
&\leq \lim_{(h_i, h_j) \to 0} \prod_{j=3}^{k} |h_j| \cdot \frac{||f(a_1, \ldots, h_{i_1}, \ldots, h_{i_2}, \ldots, 1, \ldots, a_m)||}{||h||}
\end{aligned}$$

which is equal to 0 by Problem 1. And we are done. (TA)

2 For submission to Davis Lazowski

Problem 2.1. In this problem, we show that C^∞ functions are generally *not* determined by their Taylor expansion.[1]

1. Define $f_1 \colon \mathbb{R} \to \mathbb{R}$ by the following piecewise formula:

$$f_1(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

   Show that f_1 is a C^∞ function and that $f_1^{(j)}(0) = 0$ for all orders j.

2. Now define $f_2 \colon \mathbb{R} \to \mathbb{R}$ by

$$f_2(x) = \begin{cases} e^{-(x-1)^{-2}} \cdot e^{-(x+1)^{-2}} & \text{if } x \in (-1, 1), \\ 0 & \text{otherwise.} \end{cases}$$

   Show that f_2 is a C^∞ function which is positive on $(-1, 1)$ and zero elsewhere.

3. For any $\varepsilon > 0$, show that there is a C^∞ function $g \colon \mathbb{R} \to \mathbb{R}$ such that $g(x \leq 0) = 0$, $g(x \geq \varepsilon) = 1$, and $0 \leq g(x) \leq 1$. (If you want, you're welcome to use the single-variable fundamental theorem of calculus here. You're not obligated, of course.)

Solution. 1. Clearly, by the chain rule and $1/x^2, e^x$ infinitely differentiable away from zero, the function is infinitely differentiable everywhere except zero. It's enough to show compute $f_1^{(j)}(0)$ to all orders.

We can observe that at any order, where it exists,

$$f^{(j)}(x) = e^{-\frac{1}{x^2}} \sum_{i=0}^{m_j} \frac{c_{i,j}}{x^{a_{i,j}}}$$

Where $a_{i,j}$ is some positive integer, possibly zero.

Observe that

$$\left| \frac{e^{-1/x^2}}{x^n} \right| \leq \left| \frac{e^{-1/x}}{x^n} \right|$$

[1] This stronger condition is sometimes called *analyticity*, and the function is called *analytic*.

and we can change variables and take

$$\lim_{y\to\infty} e^{-y}y^n$$

Which goes to zero, for example by L'Hospitale's.

2. As a product of C^∞ functions, it is C^∞ directly – we can see the individual C^∞ functions as $f_1(x-1)$ as $f_1(x+1)$. It is positive over $(-1,1)$ with e^x, for $x \in \mathbb{R}$, and 0 elsewhere by definition.

3. If we had g_ε for one such ε, then for ε' we can define

$$g_{\varepsilon'}(x) := g_\varepsilon(\frac{\varepsilon x}{\varepsilon'})$$

Which directly by the chain rule is smooth with g_ε. So it's enough to find it for one ε.

Now define

$$F(x) := \frac{\int_{-\infty}^{x} f_2(2t-1)dt}{\int_{\mathbb{R}} f_2(2t-1)dt}$$

This is C^∞ as the integral of a C^∞ function, 1 for $x \geq 1$ and 0 for $x \leq 0$. (DL)

Problem 2.2. We now use the functions from the previous problem to construct interesting functions on $\mathbb{R}^n$.

1. Let f_2 be as above, and select any $\varepsilon > 0$ and point $(a_1, \dots, a_n) \in \mathbb{R}^n$. Define a function $h\colon \mathbb{R}^n \to \mathbb{R}$ by the formula

$$h(x_1, \dots, x_n) = \prod_j f_2\left(\frac{x_j - a_j}{\varepsilon}\right).$$

Show that h is a C^∞ function which is positive on the open rectangle $\prod_j (a_j - \varepsilon, a_j + \varepsilon)$ and zero elsewhere.

2. For $A \subseteq \mathbb{R}^n$ open and $C \subseteq A$ compact, show there is a nonnegative C^∞ function $f\colon A \to \mathbb{R}$ such that $f(x \in C) > 0$, and f becomes the zero function outside of some closed set contained in A.

3. Show we can choose the f from the previous part so that $0 \leq f(x) \leq 1$ and and $f(x \in C) = 1$.

Solution.

1. It is C^∞ as the product of C^∞ functions, positive in the rectangle with each copy of f_2, and zero elsewhere with the copy of f_2 such that $|\frac{x_j - a_j}{\varepsilon}| \geq 1$.

2. Choose d by Problem 2.2.3 of Pset 1 so that if $x \in \mathbb{R}^n - A, y \in C$, then $||x-y|| \geq d$. Then let $\delta = \frac{d}{2}$. Cover C with an open cover of open squares of diagonal length δ or less. C can be covered with finitely many of these open squares, say $S_1 .... S_n$.

   Define h_j as in part 1) to be positive on S_j and 0 elsewhere. Then let $f(x) := \sum_{j=1}^{n} h_j(x)$, which is C^∞ as a finite sum of C^∞ functions.

   Now $\overline{S_j} \subset A$, because of the diagonal length we have chosen.

   So $\bigcup_{j=1}^{n} \overline{S_j}$ is a closed set outside of which f is the zero function.

3. By compactness, h from the last part has a minimum over C; say δ. Choose g as in part 2.1.3 so that $x \geq \delta \implies g(x) = 1$. Then $g \circ h$ is the desired function. (DL)

3 For submission to Handong Park

Let's keep it going from Thayer's part with some applications.

Problem 3.1. Recall that inner products give examples of bilinear functions. Define $IP\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by the following formula:

$$IP(x, y) = \langle x, y \rangle,$$

where $\langle -, - \rangle$ is the usual Euclidean inner product.

1. Describe $D_{(a,b)} IP$.
2. Given two differentiable functions $f, g\colon \mathbb{R} \to \mathbb{R}^n$, form the composite $h(t) = \langle f(t), g(t) \rangle$. For any $t_0 \in \mathbb{R}$, show
$$h'(t_0) = \langle (D_{t_0} f)(t), g(t_0) \rangle + \langle f(t_0), (D_{t_0} g)(t) \rangle,$$
where we are identifying the $(n \times 1)$–matrix $D_{t_0} f$ with a column vector, or an element in $\mathbb{R}^n$.
3. For $f\colon \mathbb{R} \to \mathbb{R}^n$ differentiable and $\|f(t)\| = 1$ for all $t \in \mathbb{R}$, conclude that for any $t_0 \in \mathbb{R}$ we have
$$\langle (D_{t_0} f)(t), f(t_0) \rangle = 0.$$

Solution.

1. The inner product is a bilinear function, hence we can use Problem 1.1.2 to show
$$(D_{(a,b)} IP)(h, k) = \langle a, k \rangle + \langle h, b \rangle.$$
2. I'm going to call the composite $p(t) = \langle f(t), g(t) \rangle$ instead, because I've decided I don't like the conflict of h–the–function and h–the–displacement.[2] This is then a matter of applying the chain rule:
$$\begin{aligned}(D_{t_0} p)(h) &= (D_{(f(t_0), g(t_0))} IP \circ D_{t_0}(f, g))(h) \\ &= D_{(f(t_0), g(t_0))} IP \circ ((D_{t_0} f)(h), (D_{t_0} g)(h)) \\ &= \langle f(t_0), (D_{t_0} g)(h) \rangle + \langle (D_{t_0} f)(h), g(t_0) \rangle.\end{aligned}$$
3. This looks like one half of the expression we were just dealing with, so we set $f = g$. In this case, the function $p(t) = \langle f(t), f(t) \rangle = \|f(t)\|^2 = 1$ becomes a constant function, hence its derivative vanishes. From the previous part, we also have a generic formula for the derivative of any such composite p, and marrying these formulas gives
$$0 = (D_{t_0} p)(h) = \langle f(t_0), (D_{t_0} f)(h) \rangle + \langle (D_{t_0} f)(h), f(t_0) \rangle = 2\langle (D_{t_0} f)(h), f(t_0) \rangle. \tag{ECP}$$

4 For submission to Rohil Prasad

Problem 4.1. We now use this technology to calculate the derivative of the determinant. Use the columns of an $(n \times n)$ matrix M to present it as an element as

$$M \in \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n \text{ times}}.$$

1. Demonstrate $(D_a \det)(v) = \sum_{j=1}^{n} \det(a_1 | \ldots | a_{j-1} | v_j | a_{j+1} | \ldots | a_n)$.

[2] Yes, I wrote this assignment and so picked these bad original names myself.

2. Assemble a collection of differentiable functions $a_{ij}\colon \mathbb{R} \to \mathbb{R}$ into a matrix $(a_{ij})_{ij}$, and set $f(t) = \det(a_{ij}(t))$. Calculate its derivative to be

$$(D_{t_0}f)(h) = \sum_{j=1}^{n} \det \left(\begin{array}{c|c|c|c|c|c|c} a_{11}(t_0) & \cdots & a_{1(j-1)}(t_0) & a'_{1j}(t_0)\cdot h & a_{1(j+1)}(t_0) & \cdots & a_{1n}(t_0) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}(t_0) & \cdots & a_{n(j-1)}(t_0) & a'_{nj}(t_0)\cdot h & a_{n(j+1)}(t_0) & \cdots & a_{nn}(t_0) \end{array} \right).$$

3. Take f as in the previous part, and suppose $f(t) \neq 0$ for all t. Given $b_1, \ldots, b_n : \mathbb{R} \to \mathbb{R}$ differentiable, let $x_1, \ldots, x_n : \mathbb{R} \to \mathbb{R}$ be functions satisfying the matrix equation

$$(a_{ij}(t))_{ij} \cdot (x_i(t))_i = (b_i(t))_i$$

guaranteed by the invertibility of $(a_{ij}(t))_{ij}$. Show that the functions x_i are all differentiable and calculate their derivative.

Solution. 1. Recall a theorem of Axler that the determinant considered as a function on the columns of the matrix, i.e. as a map

$$\det : \overbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}^{n \text{ times}}.$$

is multilinear.

Therefore, by problem 1.2 of this problem set, we can plug the determinant function to get exactly

$$(D_a\det)(v) = \sum_{j=1}^{n} \det(a_1|\ldots|a_{j-1}|v_j|a_{j+1}|\ldots|a_n)$$

as desired.

2. Let $(a_{ij}(t))_{ij} : \mathbb{R} \to \mathbb{R}^{n^2}$ denote the function given by packing all the a_{ij}'s together in a matrix. Then, we find by definition $f = \det \circ (a_{ij}(t))_{ij}$. By the Chain rule, it follows that

$$D_{t_0}(f)(h) = D_{(a_{ij}(t_0))_{ij}}\det((a'_{ij}(t_0)(h))_{ij})$$

and therefore we can calculate by the first part the derivative in column notation:

$$D_{t_0}(f)(h) = \sum_{j=1}^{n} \det(a_1(t_0)|\ldots|a_{j-1}(t_0)|a'_j(t_0)\cdot h|a_{j+1}(t_0)|\ldots|a_n(t_0)).$$

which matches the matrix given in the problem statement.

3. Set $A(t) = (a_{ij}(t))_{ij}$, and $A_i(t)$ to be the matrix formed by replacing the ith column of $A(t)$ with the column vector $(b_k(t))_k$. By Cramer's rule, a solution to the system of equations is given by

$$x_i(t) = \frac{\det(A_i(t))}{\det(A(t))}$$

for every i.

Since by assumption we have $\det(A(t)) \neq 0$ for every t and both the numerator and the denominator of the fraction are differentiable (as shown in part 2) it follows that $x_i(t)$ is differentiable as well for every i. We then use the quotient rule to calculate its differential to be

$$D_{t_0}x_i = \frac{D_{t_0}\det(A_i(t))\cdot\det(A(t)) - \det(A_i(t))\cdot D_{t_0}\det(A(t))}{(\det(A(t)))^2}$$

which can be expanded out using the previous part if one desires. (RP)