Homework #6 Solutions

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

1 For submission to Thayer Anderson

Problem 1.1. Akin to the tensor product of functionals described in class, you can also define a tensor product of vectors: $V \otimes W$ is the vector space populated by formal sums of elements of the form $v \otimes w$, subject to the relations

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w, \qquad (kv) \otimes w = k(v \otimes w) = v \otimes (kw),$$
$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

(The tensor described in class is thus this definition of tensor, applied to the dual space $\mathbb{R}^*$.) Similarly, you can also build a wedge product of vectors $v_1, \ldots, v_k \in V$ as a particular kind of tensor:

$$v_1 \wedge \cdots \wedge v_k = \frac{1}{k!} \sum_{\sigma \text{ a permutation of } \{1,\ldots,k\}} \operatorname{sign}(\sigma) \cdot (v_{\sigma 1} \otimes \cdots \otimes v_{\sigma k}),$$

considered as a vector in $(\mathbb{R}^n)^{\otimes k}$. (Again, the wedge product defined in class is thus this definition of wedge product, applied to the dual space $\mathbb{R}^*$.)

1. Show that a linear map $V \otimes W \to U$ is identical information to a bilinear function $V \times W \to U$.
2. Let $e_1, \ldots, e_n$ be the standard basis for $\mathbb{R}^n$, and let $\varphi_1, \ldots, \varphi_n$ be the dual basis under the standard inner product. Demonstrate the identity

   $$(\varphi_{j_1} \wedge \cdots \wedge \varphi_{j_k})(e_{j_1} \wedge \cdots \wedge e_{j_k}) = 1$$

   by considering the input as a vector in $(\mathbb{R}^n)^{\otimes k}$ and the function as a vector in $((\mathbb{R}^n)^*)^{\otimes k} = ((\mathbb{R}^n)^{\otimes k})^*$.
3. Remark on the role of the binomial/factorial coefficient in the definition of the wedge product. If that factor were omitted, what would the above pairing evaluate to instead? Why?
4. Note that if $\{v_1, \ldots, v_k\}$ and $\{w_1, \ldots, w_\ell\}$ are basis of V and W respectively, then $\{v_i \otimes w_j\}$ forms a basis of $V \otimes W$. Conclude more generally that if $v_1, \ldots, v_k$ is a k–tuple of vectors in $\mathbb{R}^n$ and $\psi_1, \ldots, \psi_k$ is a k–tuple of linear functionals on $\mathbb{R}^n$, then the following two values agree:

   $$(\psi_1 \wedge \cdots \wedge \psi_k)(v_1, \ldots, v_k) = (\psi_1 \wedge \cdots \wedge \psi_k)(v_1 \wedge \cdots \wedge v_k).$$

Solution. 1. Given a linear map $g \colon V \otimes W \to U$, define a function $f \colon V \times W \to U$ by $f(u, w) = g(u \otimes w)$. In the other direction, a function $g \colon V \times W \to U$ defines a linear function from the space of all symbols $v \otimes w$ *without any relations imposed* to U, and then this function descends to a linear functon on the quotient space $V \otimes W$ because the bilinearity of g causes it to vanish on the subspace of symbols generated by the relations specified above.

2. This is a rather manual computation.

$$(\varphi_{j_1} \wedge \cdots \wedge \varphi_{j_k})(e_{j_1} \wedge \cdots \wedge e_{j_k}) = \left(\frac{k!}{k!} \sum_\eta \operatorname{sign}(\eta)(\varphi_{\eta 1} \otimes \cdots \otimes \varphi_{\eta k}) \right) \left(\frac{1}{k!} \sum_\sigma \operatorname{sign}(\sigma)(v_{\sigma 1} \otimes \cdots \otimes v_{\sigma k}) \right)$$
$$= \frac{1}{k!} \sum_{\sigma,\eta} (\varphi_{\eta 1} \otimes \cdots \otimes \varphi_{\eta k})(v_{\sigma 1} \otimes \cdots \otimes v_{\sigma k}).$$

This sum is nonzero if and only if $\sigma = \eta$, in which case is gives $\frac{1}{k!}$, and this happens $k!$ many times.

3. The fraction is there to handle this double-counting appearing from the double summation.

4. Each vector v_j and each functional φ_j decomposes as a sum of standard basis vectors, and then the above results apply. (ECP)

Problem 1.2. For $f, g\colon \mathbb{R}^n \to \mathbb{R}$, demonstrate the *Leibniz rule*

$$d(f \cdot g) = df \cdot g + f \cdot dg.$$

Solution. This is a totally formal consequence of the Leibniz rule for the full derivative of functions $\mathbb{R}^n \to \mathbb{R}$, solved way back in Problem 2.1.1.2. Recall $D_a(f \cdot g) = g \cdot D_a f + f \cdot D_a g$. By reading the coordinate functions of the 1–form $d(f \cdot g)$ out of $D(f \cdot g)$, we have

$$d(f \cdot g) = \sum_j [D(f \cdot g)]_j \, \mathrm{d}x_j = \sum_j \left(\frac{\partial f}{\partial x_j} \cdot g + f \cdot \frac{\partial g}{\partial x_j} \right) \mathrm{d}x_j = \mathrm{d}f \cdot g + f \cdot \mathrm{d}g. \qquad \text{(ECP)}$$

2 For submission to Davis Lazowski

Problem 2.1. Recall the definition of the cross-product in $\mathbb{R}^3$: for $v, w \in \mathbb{R}^3$, $v \times w$ is the vector representing the linear functional

$$u \mapsto \det \left(\begin{array}{c|c|c} u & v & w \end{array} \right)$$

under the standard inner product. Demonstrate the following truckload of identities:

1.

$$\begin{array}{lll} e_1 \times e_1 = 0, & e_1 \times e_2 = e_3, & e_1 \times e_3 = -e_2, \\ e_2 \times e_1 = -e_3, & e_2 \times e_2 = 0, & e_2 \times e_3 = e_1, \\ e_3 \times e_1 = e_2, & e_3 \times e_2 = -e_1, & e_3 \times e_3 = 0. \end{array}$$

2. $v \times w = (v_2 w_2 - v_3 w_2)e_1 + (v_3 w_1 - v_1 w_3)e_2 + (v_1 w_2 - v_2 w_1)e_3$.

3. $\|v \times w\| = \|v\| \cdot \|w\| \cdot \sin\theta$, where θ is the angle formed by v and w as rays intersecting at the origin. Conclude $\langle v \times w, v \rangle = 0$ and $\langle v \times w, w \rangle = 0$.

4. The juggling identities: $\langle v, w \times u \rangle = \langle w, u \times v \rangle = \langle u, v \times w \rangle$.

5. The associative identities: $v \times (w \times u) = \langle v, u \rangle w - \langle v, w \rangle u$ and $(v \times w) \times u) = \langle v, u \rangle w - \langle w, u \rangle v$.

6. $\|v \times w\| = \sqrt{\|v\|^2 \cdot \|w\|^2 - \langle v, w \rangle^2}$.

Solution. I will note two lemmas which make all the identities easy.

First, *the cross product is linear in both elements.* This is clear since $u \to \det \left(u \mid v \mid w \right)$ is a linear functional.

Second, *the cross product is rotation invariant*, in the sense that if R is a rotation, then $R(v \times W) = (Rv) \times (Rw)$.

As a proof, if v and w are not linearly independent, then Rv and Rw aren't either, so $v \times w = 0$ and $Rv \times Rw = 0 = R0$, so this is trivial.

Otherwise, v, w and $v \times w$ are linearly independent. (To prove this, observe that $\langle v \times w, v \times w \rangle$ is nonzero.) Therefore, Rv, Rw and $R(v \times w)$ are linearly independent. So we know that $Rv \times Rw \propto R(v \times w)$. So we just need to fix the proportionality constant, and show that it's one.

But since R is a rotation matrix,

$$\det \begin{pmatrix} R(v \times w) & | & Rv & | & Rw \end{pmatrix} = \det \begin{pmatrix} v \times w & | & v & | & w \end{pmatrix}$$

this constant is just one.

Let's use these to prove all the identities.

1. Clearly $e_j \times e_j = 0$, since the matrix $\begin{pmatrix} u & e_j & e_j \end{pmatrix}$ has less than full rank, so has zero determinant. Now suppose we have $k \neq i \neq j \neq k$. Then since $\begin{pmatrix} u & | & e_i & | & e_j \end{pmatrix}$ is only nonzero when u has a $\hat{e}_k$ component, so we know that $e_i \times e_j = \alpha_{ij} e_k$; we just need to fix α_{ij}.

   For $k = 1, i = 2, j = 3$, by inputting $u = e_1$, since the determinant is just the identity, we know that $1 = \alpha_{23} \langle e_1, e_1 \rangle = \alpha_{23}$. So $e_2 \times e_3 = e_1$.

   Also observe that $\det \begin{pmatrix} u & | & e_i & | & e_j \end{pmatrix}$ swaps sign under transposition of rows, since a row transposition matrix has determinant -1.

   Therefore $\alpha_{ij} = \text{sign}(\sigma)$, where σ is the permutation sending (k, i, j) to $(1, 2, 3)$. This demonstrates the remaining identities.

2. By linearity in both elements, we only need to calculate $v_i \times w_j$. By the previous part, this is $\alpha_{ij} v_i w_j e_k$. Summing all these individual parts, we get the required identity.

3. . Since rotation doesn't change absolute value, we might as well rotate so that v and w are on the same plane– i.e. $Rv = v_1 e_1$, and $Rw = w_1 e_1 + w_2 e_2$.

   Then

   $$Rv \times Rw = v_1 w_2 e_3$$

   So that

   $$\begin{aligned} \frac{||v \times w||}{||v||||w||} \\ = \frac{||Rv \times Rw||}{||Rv||||Rw||} \\ = \frac{|v_1||w_2|}{|v_1|\sqrt{w_1^2 + w_2^2}|} \\ = \frac{|w_2|}{\sqrt{w_1^2 + w_2^2}} \\ = \sin\theta \end{aligned}$$

   as desired.

4. This is equivalent to the statement that

   $$\det \begin{pmatrix} v & | & w & | & u \end{pmatrix} = \det \begin{pmatrix} w & | & u & | & v \end{pmatrix} = \det \begin{pmatrix} u & | & v & | & w \end{pmatrix}$$

   Which is true because each matrix is related by two row transpositions, and the determinant of the matrix which performs two row transpositions is 1 (since it is the product of two transposition matrices of determinant -1.)

5. First, observe that (with i, j, k not necessarily distinct)

$$e_i \times (e_j \times e_k) = \langle e_i, e_k \rangle e_j - \langle e_i, e_j \rangle e_k$$

directly by computing the cross product.

Now

$$\begin{aligned} e_i \times (e_j \times u) &= e_i \times (e_j \times (u_1 e_1 + u_2 e_2 + u_3 e_3)) \\ &= e_i \times [u_1(e_j \times e_1) + u_2(e_j \times e_2) + u_3(e_j \times e_3)] \\ &= \sum_{k=1}^{3} u_k e_i \times (e_j \times e_k) \\ &= \sum_{k=1}^{3} \langle e_i, u_k e_k \rangle e_j - \langle e_i, e_j \rangle (u_k e_k) \\ &= \langle e_i, u \rangle e_j - \langle e_i, e_j \rangle u \end{aligned}$$

Now use linearity the exact same way to replace e_i and e_j with v and w. Prove the second associative identity the exact same way.

6. Again, rotate so that $Rv = v_1 e_1$. and $Rw = w_1 e_1 + w_2 e_2$. Then we want to show

$$\begin{aligned} ||Rv \times Rw|| &= |v_1 w_2| \\ \sqrt{v_1^2(w_1^2 + w_2^2) - (v_1 w_1)^2} &= |v_1 w_2| \end{aligned}$$

which is true. Then since $||Rv|| = v$, $\langle Rv, Rw \rangle = \langle v, w \rangle$, done. (DL)

Problem 2.2. Recall the polar coordinate transformation $x(r, \theta) = r \cos \theta$ and $y(r, \theta) = r \sin \theta$, defined for $0 < \theta < 2\pi$ and $r > 0$. Prove that where θ is defined as a function of x and y, we have

$$d\theta = \frac{-y}{x^2 + y^2}\, \mathrm{d}x + \frac{x}{x^2 + y^2}\, \mathrm{d}y.$$

Solution. Recall that we have a lemma saying

$$u^*(dx_i) = \sum_{j=1}^{m} \frac{\partial u_i}{\partial x_j} dx_j$$

Let u be the map $(x, y) \to (\theta, r)$, i.e. $(x, y) \to (\arctan \frac{y}{x}, \sqrt{x^2 + y^2})$.
Then this lemma applied to dx says

$$d\theta = \frac{\partial}{\partial x}[\arctan(y/x)]dx + \frac{\partial}{\partial y}[\arctan(y/x)]dy$$

Computing the derivatives, we have the requisite

$$d\theta = -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

(DL)

3 For submission to Handong Park

Problem 3.1. For $f\colon \mathbb{R}^n \to \mathbb{R}$, we define a *vector field* ∇f by the formula

$$\nabla f\colon\ (a \in \mathbb{R}^n) \mapsto \begin{pmatrix} \left.\frac{\partial f}{\partial x_1}\right|_{x=a} \\ \vdots \\ \left.\frac{\partial f}{\partial x_n}\right|_{x=a} \end{pmatrix} \in T_a\mathbb{R}^n.$$

Recall also the directional derivative from Homework #3: given a tangent vector $v \in T_a\mathbb{R}^n$, we set

$$\mathbb{D}_a^v f = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}.$$

Conclude $\mathbb{D}_a^v(f) = \langle v, \nabla f(a)\rangle$, and hence that $\nabla f(a)$ is the direction of greatest ascent[1].

Solution. Our conclusion from Homework #3 was $\mathbb{D}_a^v f = (D_a f)(v)$. We know that $D_a f$ admits expression as a matrix:

$$D_a f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix},$$

and hence we have the evaluation formula

$$(D_a f)(v) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \langle v, \nabla f(a)\rangle.$$

By writing this as an inner product, we can appeal to Cauchy–Schwarz to see that the expression $\langle v, \nabla f(a)\rangle$ is maximized over unit-length tangent vectors v by $v = (\nabla f(a))/\|\nabla f(a)\|$. (ECP)

Problem 3.2. Let $f\colon U \to \mathbb{R}^n$ be a differentiable function with a differentiable inverse $f^{-1}\colon f(U) \to \mathbb{R}^n$. If every closed form on U is exact, show that the same is true for $f(U)$.

Solution. Let ω be a closed form defined on $f(U)$. Then $(f^{-1})^*\omega$ is a form on U, and because exterior differentiation commutes with pullback we have

$$d((f^{-1})^*\omega) = (f^{-1})^* d\omega = (f^{-1})^* 0 = 0$$

shows that $(f^{-1})^*\omega$ is closed as well. Since on U every closed form is exact, there exists some ξ with $d\xi = (f^{-1})^*\omega$. We claim that $f^*\xi$ witnesses ω as an exact form:

$$df^*\xi = f^* d\xi = f^*(f^{-1})^*\omega = (f^{-1}\circ f)^*\omega = \mathrm{id}^*\omega = \omega. \qquad \text{(ECP)}$$

4 For submission to Rohil Prasad

Problem 4.1. Let $c : [0,1] \to (\mathbb{R}^n)^n$ be a 1-parameter continuous family of families of n vectors in $\mathbb{R}^n$, and suppose that $c(t) = \{c_1(t), \ldots, c_n(t)\}$ is a basis of $\mathbb{R}^n$ for each $0 \le t \le 1$. Show that the orientation of each basis must be the same, i.e. the value

$$[c_1(t), \ldots, c_n(t)] = \mathrm{sign}(\det(c_1(t)|\ldots|c_n(t)))$$

is constant even as t varies.

Solution. Postcompose c with the determinant map to get a continuous map $\varphi : [0,1] \to \mathbb{R}$.

Since $c(t)$ is a basis of $\mathbb{R}^n$ for every t, this determinant must be nonzero for every t.

Now assume for the sake of contradiction that the orientation switches. This implies that there exists $s, t \in [0,1]$ such that one of $\varphi(s), \varphi(t)$ is less than 0 and one is greater than 0.

Applying the intermediate value theorem to φ, this tells us that there is some $a \in [s,t]$ such that $\varphi(a) = 0$. However, we assumed in the problem statement that φ is nonzero everywhere, so we arrive at a contradiction and the orientation remains constant. (RP)

[1] Or "direction of fastest change", if you prefer.

Problem 4.2. In class, we "proved" by example that any quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$ can be written in the form

$$Q = a_1^2 + \cdots + a_k^2 - b_1^2 - \ldots b_l^2$$

for a family of linearly independent linear functionals a_* and b_*. Complete our discussion by turning our examples into an honest proof. (Don't worry about the invariance of the signature; just work on this existence half.)

Solution. We can do this by induction on n.

For $n = 1$, $Q(x)$ is simply equal to Cx^2 for some constant C.

Now take a quadratic form $Q(x_1, \ldots, x_n)$. We can write it as $Cx_n^2 + x_n \cdot A(x_1, \ldots, x_{n-1}) + B(x_1, \ldots, x_{n-1})$. Here C is a constant, A is a linear functional, and B is a quadratic form.

Then we can set u to be the linear functional $\sqrt{C}x_n + A(x_1, \ldots, x_{n-1})/2\sqrt{C}$.

It follows that we can write $Q = u^2 + B(x_1, \ldots, x_{n-1}) - \frac{1}{4C}A(x_1, \ldots, x_{n-1})^2$. The term $B(x_1, \ldots, x_{n-1}) - \frac{1}{4C}A(x_1, \ldots, x_{n-1})^2$ is a quadratic form in $n-1$ variables, so we can apply our induction hypothesis and obtain the desired formula for Q. (RP)