EXTERIOR DERIVATIVE

BENJAMIN GUNBY

We now use the concept of exterior algebras to define more generally the notion of derivative. First, we will define the differential of a function df, where f is a function $\mathbb{R}^n \to \mathbb{R}$. (This aligns with our intuition as we are usually used to seeing expressions like dx_i, and x_i is indeed a function $\mathbb{R}^n \to \mathbb{R}$ given by just looking at one coordinate.)

At each point, this should align with the total derivative. So we consider $D_a f$, which is a map from $T_a(\mathbb{R}^n)$ to $T_{f(a)}(\mathbb{R})$, which is just isomorphic to $\mathbb{R}$. Therefore, $D_a f$ is an element of the dual space of $T_a(\mathbb{R}^n)$ (also known as the cotangent space).

Now, given that x_i is projection onto one coordinate, $D_a x_i$ (which we will call dx_i as above) will be the same. Thus $dx_i|_a$ is the dual of the unit vector $e_i \in T_a(\mathbb{R}^n)$–it sends e_i to 1 and e_j to 0 for all $j \neq i$.

Now, as $D_a f = \displaystyle\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) D_a x_i$ by the chain rule, we have $df|_a = \displaystyle\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i|_a$. Thus at each point a, we have a space of differentials at a, which is generated by $dx_1|_a, \ldots, dx_n|_a$. We refer to this space of differentials at a as $\Omega^1(T_a^*(\mathbb{R}^n))$.

Now, we expand beyond just one point a. After all, dx_i is defined at every point, not just a single one! This motivates the following definition.

Definition. A 1-*form* is an assignment of each $a \in \mathbb{R}$ to a differential at a (that is, an element of $\Omega^1(T_a^*(\mathbb{R}^n))$, or in other words a map $T_a(\mathbb{R}^n) \to \mathbb{R}$). Since at each a, $\Omega^1(T_a^*(\mathbb{R}^n))$ is generated by the $dx_i|_a$, every 1-form ω must be of the form $\displaystyle\sum_{i=1}^{n} g_i dx_i$ for some functions g_i. That is, $\omega|_a = \displaystyle\sum_{i=1}^{n} g_i(a) dx_i|_a$. We call the space of all 1-forms Ω^1 (or $\Omega^1(\mathbb{R}^n)$).

We say that ω is continuous/differentiable/C^n/etc if all g_i are.

There are several operations we can apply to 1-forms.

- First, given a 1-form ω on $\mathbb{R}^n$ and a function $g : \mathbb{R}^n \to \mathbb{R}$, we can multiply ω by g. As above, the multiplication is just done pointwise, so $(g\omega)|_a = g(a)\omega|_a$.
- Second, given a 1-form ω on $\mathbb{R}^n$ and a function $f : \mathbb{R}^m \to \mathbb{R}^n$, we can find the *pullback* of ω under f. This will be a 1-form ω' on $\mathbb{R}^m$, and it will be given by the property that $\omega'|_a : T_a\mathbb{R}^m \to \mathbb{R}$ is the map $\omega|_{f(a)} \circ D_a f$. Note that $D_a f$ maps $T_a\mathbb{R}^m$ to $T_{f(a)}\mathbb{R}^n$ and $\omega|_{f(a)}$ maps $T_{f(a)}\mathbb{R}^n$ to $\mathbb{R}$, so this definition makes sense. We write the pullback as $f^*\omega$. Note that f^* will be a map $\Omega^1(\mathbb{R}^n) \to \Omega^1(\mathbb{R}^m)$–the opposite direction of f. Intuitively, this is because we dualize when defining 1-forms.

Example. Let ω be the 1-form $x\ dx + dy$. Let $g(x, y) = x^2 + y$. Then $g\omega = (x^2+y)x\ dx + (x^2+y)\ dy$, as one might expect.

Now we show some nice properties of the pullback.

Date: 4/5/17.

Lemma. Let $u : \mathbb{R}^m \to \mathbb{R}^n$ be a function, $s : \mathbb{R}^n \to \mathbb{R}$ be another function, and $\omega, \omega' \in \Omega^1(\mathbb{R}^n)$ be 1-forms. Then:

(1) $u^*(dx_i) = \displaystyle\sum_{j=1}^{m} \frac{\partial u_i}{\partial x_j} dx_j$. (We can think of this as du_i.)
(2) $u^*(\omega + \omega') = u^*(\omega) + u^*(\omega')$.
(3) $u^*(s\omega) = (s \circ u)u^*\omega$. These last two conditions may be thought of as 'linearity' of pullback, as the previous is additivity and this one is (kind of) scalar multiplication, as functions evaluated at any given point are just scalars and 1-forms are really just elements of the cotangent space of every point.

Proof. For (1), as dx_j is dual to e_j in the tangent space, the dx_j component is just the 1-form evaluated at e_j. Thus it suffices to show that $u^*(dx_i)(e_j) = \frac{\partial u_i}{\partial x_j}$. But by definition, $u^*(dx_i)(e_j) = dx_i(Du(e_j))$, and again by duality this will be the e_i component of $Du(e_j)$. This will simply give us the (i, j) component of the total derivative, which is simply $\frac{\partial u_i}{\partial x_j}$, as desired.

For (2) and (3), all we need observe is that at each point a, u^* acts as a linear operator $T^*_{f(a)}\mathbb{R}^n \to T^*_a\mathbb{R}^m$, and the linearity of this operator gives us that it preserves addition and scalar multiplication at each point. $\square$

Example. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $f(a, b, c) = (a^2 + b^2, c^2)$. Then $f^*\omega$ will be

$$(a^2 + b^2)d(a^2 + b^2) + d(c^2) = (a^2 + b^2)(2a\ da + 2b\ db) + 2c\ dc.$$

So in some sense you can think of pullback as substitution.

Similarly, we can have k-forms, which will be assignments of elements of $\Omega^k(T^*_a\mathbb{R}^n)$ for each point a, so a k-form ω will have $\omega|_a \in T^*_a(\mathbb{R}^n)$. We will refer to the space of k-forms as Ω^k (or $\Omega^k(\mathbb{R}^n)$).

Now, given the existence of the wedge product, we have a map $\Omega^k \times \Omega^\ell \to \Omega^{k+\ell}$ given by $(\omega_1, \omega_2) \to \omega_1 \wedge \omega_2$. We can define a pullback similarly on k-forms, as a map of vector spaces additionally induces a map on their exterior powers, and by definition we have that $f^*(\omega_1 \wedge \omega_2) = f^*\omega_1 \wedge f^*\omega_2$.

On $\mathbb{R}^n$, a special class of forms are the n-forms. Why is this? Well, we know that at each point a, Ω^n looks like $\Omega^n(T^*_a\mathbb{R}^n)$. We know that $T^*_a\mathbb{R}^n$ has dimension n, so our resulting space will have dimension $\binom{n}{n} = 1$. Also, we know that a (1-element) basis is given by $dx_1 \wedge \cdots \wedge dx_n$ by our discussion in the last class, as the dx_i generate the cotangent space at each point. So, if we have a function $f : \mathbb{R}^n \to \mathbb{R}^n$, at each point f^* will be a linear operator on the 1-dimensional cotangent space. But a linear operator on a 1-dimensional space will be multiplication by a scalar! Therefore, $f^*(dx_1 \wedge \cdots \wedge dx_n) = g\ dx_1 \wedge \cdots \wedge dx_n$ for some function $g : \mathbb{R}^n \to \mathbb{R}$. Which function? Let's compute.

By our above observation that taking pullbacks commutes with taking wedge products, $f *$ $(dx_1 \wedge \cdots \wedge dx_n) = f^*dx_1 \wedge \cdots \wedge f^*dx_n$. We know by the earlier lemma that $f^*dx_i = \displaystyle\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} dx_j$.

Substituting and distributing the wedge product gives us

$$\begin{aligned} f^*(dx_1 \wedge \cdots \wedge dx_n) &= f^*dx_1 \wedge \cdots \wedge f^*dx_n \\ &= \left(\sum_{j=1}^{n} \frac{\partial f_1}{\partial x_j} dx_j\right) \wedge \cdots \wedge \left(\sum_{j=1}^{n} \frac{\partial f_n}{\partial x_j} dx_j\right) \\ &= \sum_{1 \le j_1, \ldots, j_n \le n} \left(\prod_{i=1}^{n} \frac{\partial f_i}{\partial x_{j_i}}\right) dx_{j_1} \wedge \cdots \wedge dx_{j_n}. \end{aligned}$$

Now, when any two of the j_i are the same, we know that the wedge product vanishes. Thus in any nonvanishing term, the j_i are a permutation of $\{1, \ldots, n\}$. But since the wedge product is anticommutative, we can switch any two adjacent terms in the wedge product while gaining a factor of -1, so we can sort the terms into $dx_1 \wedge \cdots \wedge dx_n$ and the factor we gain is simply the sign of the permutation given by the j_i. Thus we can rewrite our sum as

$$\begin{aligned} f^*(dx_1 \wedge \cdots \wedge dx_n) &= \sum_{1 \le j_1, \ldots, j_n \le n} \left(\prod_{i=1}^{n} \frac{\partial f_i}{\partial x_{j_i}} \right) dx_{j_1} \wedge \cdots \wedge dx_{j_n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{i=1}^{n} \frac{\partial f_i}{\partial x_{\sigma(i)}} \right) dx_1 \wedge \cdots \wedge dx_n \\ &= \left(\det \left[\frac{\partial f_i}{\partial x_j} \right] \right) dx_1 \wedge \cdots dx_n \\ &= (\det Df)\, dx_1 \wedge \cdots \wedge dx_n. \end{aligned}$$

Thus our factor is just the determinant of Df! (Philosophically, in some sense this is where the factor of the Jacobian in the change of variables formula comes from, as we can write it as $\displaystyle\int h\omega = \int (h \circ f)(f^*\omega)$, where ω is an n-form.)

If ω is a differentiable k-form, we can actually take the derivative of it, which we will write $d\omega$. Similarly to how a function f is just a 0-form and its derivative df is a 1-form, the derivative $d\omega$ of a k-form will be a $k+1$-form. (Note that this means that the derivative of an n-form on $\mathbb{R}^n$ will be 0, as there will be no nonzero $n+1$-forms.) We will now define this derivative.

Definition. Let $\omega = \displaystyle\sum_J g_J dx_{J_1} \wedge \cdots \wedge dx_{J_k}$ be a k-form. (We know from last time that $dx_{J_1} \wedge \cdots \wedge dx_{J_k}$ with $J_1 < \cdots < J_k$ indeed form a basis for k-forms.) Then

$$d\omega = \sum_J dg_J dx_{J_1} \wedge \cdots \wedge dx_{J_k} = \sum_J \sum_{i=1}^{n} \frac{\partial g_J}{\partial x_i} dx_i \wedge dx_{J_1} \wedge \cdots \wedge dx_{J_k}.$$

Theorem. The following properties are true.

(1) d is additive. That is, if $\omega_1, \omega_2 \in \Omega^k$ are k-forms, then $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.
(2) d satisfies a version of the product rule, up to a sign. In particular, if $\omega_1 \in \Omega^k$ and $\omega_2 \in \Omega^\ell$, then $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$
(3) $d^2 = 0$. That is, $d(d\omega) = 0$ for all k-forms ω.
(4) d commutes with pullback. That is, if f is differentiable, then $f^*(d\omega) = d(f^*\omega)$.

Proof. The first three proofs simply follow from computation and the definitions. For (4) you can proceed by induction, showing that if this holds for ω it holds for $\omega \wedge dx_i$, by using the other three properties. □

Note that (3) is a generalization (and indeed just follows from) the fact that partials commute. If $f : \mathbb{R}^2 \to \mathbb{R}$ is a function, then

$$\begin{aligned}
d^2 f &= d(df) \\
&= d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) \\
&= \frac{\partial^2 f}{\partial x^2}dx \wedge dx + \frac{\partial^2 f}{\partial y \partial x}dy \wedge dx + \frac{\partial^2 f}{\partial x \partial y}dx \wedge dy + \frac{\partial^2 f}{\partial y^2}dy \wedge dy \\
&= \frac{\partial^2 f}{\partial y \partial x}dy \wedge dx + \frac{\partial^2 f}{\partial x \partial y}dx \wedge dy \\
&= \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}\right) dx \wedge dy.
\end{aligned}$$