Coordinate Free Calculus

Def: A normed vector space (over $\mathbb{R}$) is a real vector space E together with a function $\|\cdot\| : E \longrightarrow \mathbb{R}$, satisfying

(1) $\|\cdot\|$ is positive definite: $\|e\| \geq 0 \quad \forall e \in E$ and

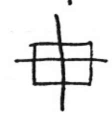
$$\|e\| = 0 \iff e = 0.$$

(2) $\|re\| = |r| \, \|e\|$ for all $r \in \mathbb{R}, e \in E$

(3) the triangle inequality: $\|e+f\| \leq \|e\| + \|f\| \quad \forall e, f \in E$.

Examples: (1) $\mathbb{R}^n$ and $\|x\| = \sqrt{\sum_i x_i^2}$ if $x = (x_1, \dots, x_n)$ 

(2) $\mathbb{R}^n$ and $\|x\| = \sum |x_i|$ if $x = (x_1, \dots, x_n)$ 

(3) $\mathbb{R}^n$ and $\|x\| = \sup_i |x_i|$ if $x = (x_1, \dots, x_n)$ 

(4) $\mathrm{Hom}(E,F)$, the vector space of linear transformations from E to F and $\|T\| = \sup_{e \neq 0} \dfrac{\|Te\|}{\|e\|} = \sup_{\|e\|=1} \|Te\|$

Def: Norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exist real numbers A and B, $A > 0$, $B > 0$, such that

$$A \|v\|_1 \leq \|v\|_2 \leq B \|v\|_1$$

for every $v \in E$.

Prop: $d(x,y) = \|x-y\|$ makes $(E, \|\cdot\|)$ into metric space, and hence a topological space. The topologies induced by equivalent norms are the same.

<u>Prop</u>: Any two norms on a finite dimensional normed vector space are equivalent.

<u>Cor</u> (Kolmogorov's Thm) There is a unique norm topology on a finite dimensional vector space.

<u>Prop</u>: (1) If $V_1, \dots, V_n$ are finite dimensional vector spaces with the norm topology, the product topology on $V_1 \times \cdots \times V_n$ is the same as the norm topology

(2) If $V \subset W$ is a subspace of a finite dimensional vector space the subspace topology on V is the same as the norm topology.

(3) If $f: V \to W$ is a linear map between finite dimensional vector spaces, then f is continuous.

Unless stated otherwise, all finite dimensional vector spaces will have their norm topology. With topological preliminaries complete, we are ready for some calculus.

<u>Def</u>: If $f: U \to F$, where U is an open subset of E, and E and F are finite dimensional vector spaces, the <u>derivative of f</u>,

$$Df: U \longrightarrow \mathrm{Hom}(E, F)$$

is

$$Df(x)(v) = \lim_{t \to 0} \frac{1}{t}\left(f(x+tv) - f(x)\right)$$

if this limit exists at each $x \in U$.

We say f is <u>differentiable</u> if Df exists, and we say f is <u>smooth</u> if $D^n f = D(D \cdots (Df) \cdots)$ exists for every $n \geq 0$.

<u>Notes</u>: (1) If $f : \mathbb{R}^m \to \mathbb{R}^n$, with coordinate functions $f_1, \ldots, f_n$ (so that $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$), then

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_m}(x) \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1}(x) & \cdots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix}$$

$$\in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$$

is the familiar <u>Jacobian</u> of f at x.

(2) Boris Mordukhovich informs me that the definition above is called the <u>Gateaux</u> derivative. We will note shortly that it is also the <u>Frechet derivative</u>, or best linear approximation.

(3) Since $Df : \mathcal{U} \to \mathrm{Hom}(E, F)$, we see that

$$D^2 f : \mathcal{U} \to \mathrm{Hom}(E, \mathrm{Hom}(E, F)) \cong \mathrm{Hom}(E \otimes E, F)$$

$$D^3 f : \mathcal{U} \to \mathrm{Hom}(E, \mathrm{Hom}(E, \mathrm{Hom}(E, F))) \cong \mathrm{Hom}(E \otimes E \otimes E, F)$$

etcetera. Here, the vector space $E \otimes E \otimes \cdots \otimes E$ (n factors) has basis the set $\{e_{i_1} \otimes \cdots \otimes e_{i_n} \mid e_{i_j} \in \mathcal{E}\}$ where $\mathcal{E}$ is a basis for E. Thus, if $d = \dim(E)$ then $d^n = \dim(E \otimes \cdots \otimes E)$. This is a special case of the <u>tensor product</u> of modules. The isomorphism

$$\mathrm{Hom}(E, \mathrm{Hom}(E, F)) \cong \mathrm{Hom}(E \otimes E, F)$$

is the correspondence $(e_i \mapsto (e_j \mapsto f_{ij})) \leftrightarrow (e_i \otimes e_j \mapsto f_{ij})$

(4) If E is n-dimensional, then $E \otimes E$ is n^2-dimensional. We may thus think of an element of $\mathrm{Hom}(E \otimes E, F)$ as an $n \times n$ matrix with entries in F. Equality of mixed partials for smooth functions implies that $D^2 f$ lies in the symmetric matrices.

Now, the definition of Df implicitly assumed the following Lemma.

* <u>Lemma</u>: $Df(x)$ is linear: $Df(x)(av+bw) = a\,Df(x)(v) + b\,Df(x)(w)$.

If f is linear, then Df is constant:

<u>Lemma</u>: If f is linear then $Df(x) = f$ for all x.

Our derivative can also be characterized as the Frechet derivative. Assume f is smooth, $f: U \to F$, U open in E.

<u>Lemma</u>: For each $x \in U$ there is a neighborhood N and a positive constant C such that if $x+v \in N$ then

$$\| f(x+v) - f(x) - Df(x)(v) \| < C \|v\|^2 .$$

For each $x \in U$, $Df(x)$ is the only linear transformation with this property.

As noted above, $D^2 f$ is symmetric (still assuming f smooth).

<u>Lemma</u> $D^2 f(x)(v)(w) = D^2 f(x)(w)(v)$

When we construct the tangent bundle to a smooth manifold, we will need to know the derivative of the adjoint

$$\widetilde{Df} : U \times E \longrightarrow F$$

of the derivative Df, defined by $\widetilde{Df}(x,v) = Df(x)(v)$.

<u>Prop</u>: $D(\widetilde{Df})(x,v) : E \times E \longrightarrow F$ is $\begin{pmatrix} D^2f(x)(v) & Df(x) \end{pmatrix}$

That is,
$$D(\widetilde{Df})(x,v)(e_1,e_2) = D^2f(x)(v)(e_1) + Df(x)(e_2)$$

<u>Proof</u>: We simply compute

$$D(\widetilde{Df})(x,v)(e_1,e_2) = \lim_{t \to 0} \frac{1}{t}\left(\widetilde{Df}(x+te_1, v+te_2) - \widetilde{Df}(x,v)\right)$$

$$= \lim_{t \to 0} \frac{1}{t}\left(\widetilde{Df}(x+te_1, v) - \widetilde{Df}(x,v) + t\,\widetilde{Df}(x+te_1, e_2)\right)$$

$$= \lim_{t \to 0} \left[\frac{1}{t}\left(Df(x+te_1)(v) - Df(x)(v)\right) + Df(x+te_1)(e_2)\right]$$

$$= D^2f(x)(e_1)(v) + Df(x)(e_2)$$

$$= D^2f(x)(v)(e_1) + Df(x)(e_2). \quad //$$

Derivatives of linear maps are quite simple. Bilinear maps are also quite useful, and their derivatives are only a little more complicated. Recall that $B : E \times E \longrightarrow F$ is <u>bilinear</u> if it is linear in each variable separately:

$$B(a_1 v_1 + a_2 v_2, w) = a_1 B(v_1, w) + a_2 B(v_2, w), \text{ and}$$

$$B(v, a_1 w_1 + a_2 w_2) = a_1 B(v, w_1) + a_2 B(v, w_2) \quad \text{for}$$

$a_1, a_2 \in \mathbb{R}$, $v, v_1, v_2, w, w_1, w_2 \in E$. (In fact a bilinear map

$$B : E \times E \longrightarrow F$$

is essentially the same thing as a linear map $E \otimes E \to F$, though we shall not need this.)

<u>Lemma</u>: $DB : E \times E \longrightarrow Hom(E \times E, F)$ is

$$DB(x, y)(v, w) = B(x, w) + B(v, y)$$

<u>Proof</u>: B is linear in each variable separately, so it is simply a matter of saying this precisely: $DB = (D_1 B \quad D_2 B) = (D(Bi_1), D(Bi_2))$, or in other words,

$$DB(x, Y)(v, w) = DB(x, Y)(v, 0) + DB(x, Y)(0, w)$$

$$= D(Bi_1(Y))(x)(v) + D(Bi_2(x))(Y)(w)$$

where

$$i_1(Y) : E \longrightarrow E \times E \quad \text{by} \quad i_1(Y)(x) = (x, Y)$$

and

$$i_2(x) : E \longrightarrow E \times E \quad \text{by} \quad i_2(x)(Y) = (x, Y).$$

Since $Bi_1(Y)$ and $Bi_2(x)$ are linear we get

$$DB(x, Y)(v, w) = Bi_1(Y)(v) + Bi_2(x)(w) = B(v, Y) + B(x, w). //$$

The derivative of a composite

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & F \\ \cap & & \cap & & \\ E_1 & & E_2 & & \end{array}$$

obeys the <u>chain rule</u>

<u>Prop</u>: $D(gf)(x) = Dg(f(x)) \circ Df(x)$

$$\begin{tikzcd} E_1 \arrow[r, "Df(x)"] \arrow[dr, "D(gf)(x)"'] & E_2 \arrow[d, "Dg(f(x))"] \\ & F \end{tikzcd}$$

<u>Examples</u>: (1) Let $d : E \times E \to \mathbb{R}$ be an inner product, and hence bilinear. Let $\Delta : E \to E \times E$ be the diagonal $\Delta(e) = (e, e)$ a linear map. Let $f = d\Delta$, i.e.

$$f(e) = d\Delta(e) = d(e,e)$$

We get

$$\begin{aligned} Df(e)(v) &= D(d\Delta)(e)(v) && \text{Def.} \\ &= Dd(\Delta(e))(D\Delta(e)(v)) && \text{Chain Rule} \\ &= Dd(\Delta(e))(\Delta(v)) && \Delta \text{ is linear} \\ &= Dd(e,e)(v,v) && \text{Def. of } \Delta \\ &= d(e,v) + d(v,e) && d \text{ is bilinear.} \end{aligned}$$

(2) (Another bilinear map composed with Δ) Let $p : M_n(\mathbb{R})^2 \to \mathbb{R}$ be $p(A,B) = AB^t$, where B^t is the transpose of B. Let $f(A) = AA^t$, so $f = p\Delta$. Then

$$Df(A)(B) = Dp(A,A)(B,B) = AB^t + BA^t.$$

Examples (1) and (2) will give easy proofs that S^n is an n-manifold, and that $O(n) = \{A \in M_n(\mathbb{R}) \mid AA^t = I\}$ is an $\frac{n(n-1)}{2}$-manifold, shortly. The last example is left as an exercise:

(3) The derivative of the determinat $\det : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ at the identity $I \in M_n(\mathbb{R})$ is the trace $M_n(\mathbb{R}) \xrightarrow{tr} \mathbb{R}$.

$$D(\det)(I)(M) = tr(M).$$

Finally, we shall need some theorems related to the Inverse Function Theorem of advanced calculus.

<u>Inverse Function Theorem</u>: Let $f : \mathcal{U} \to F$ be a smooth map, $\mathcal{U}$ open in E, E and F real vector spaces. If $Df(x)$ is invertible, for some $x \in \mathcal{U}$, then there exist nhoods V of x and W of $f(x)$, such that $f|_V$ is a diffeomorphism from V to W.

$$\begin{array}{ccccc} & & E & & \\ & & \cup & & \\ x & \in & \mathcal{U} & \xrightarrow{f} & F \\ & & \cup & & \cup \\ x & \in & V & \xrightarrow[\cong]{f|_V} & W \end{array}$$

Implicit Function Theorem Let E_i and F be finite dimensional real vector spaces, $\mathcal{U} \subset E_1 \times E_2$ open, and $\mathcal{U} \xrightarrow{f} F$ smooth. Let $i_1 : E_1 \times \{y\} \longrightarrow E_1 \times E_2$ be the inclusion, where $y \in E_2$.

$$\begin{array}{ccccc} (x,y) \in \mathcal{U} \cap (E_1 \times \{y\}) & \xrightarrow{\quad i_1 \quad} & \mathcal{U} & \xrightarrow{\quad f \quad} & F \\ \cap & & \cap & & \\ E_1 \times \{y\} & \xrightarrow{\quad i_1 \quad} & E_1 \times E_2 & & \end{array}$$

If $D(fi_1)(x,y) : E_1 \longrightarrow F$ is invertible, then there exist nhoods V of y in E_2, and W of x in E_1, and a unique function $g : V \to W$ such that $g(y) = x$ and

$$f(g(y), y) \text{ is constant for } y \in V.$$

Further, g is smooth.

These are the usual inverse and implicit function theorems stated in a coordinate free fashion. Before stating our final version of these theorems, we make an observation.

Lemma: Let f be a smooth function $\mathcal{U} \xrightarrow{f} F$ and $x \in \mathcal{U}$. There is a nhood of x on which rank(Df) does not decrease.

<u>Proof</u>: If rank$(Df(x))$ is k, then there is a $k \times k$ minor in $Df(x)$ with non-zero determinant. This same minor will be non-zero in $Df(y)$ for y in some nhood of x, since Df and det are continuous. //

<u>Example</u>: rank $Df(x)$ may increase in every nhood. Consider $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Then $Df(x) = (2x)$ which has rank 0 when $x = 0$, but rank 1 elsewhere. Thus, rank $Df(x)$ increases in every nhood of $x = 0$.

Our final version of the inverse and implicit function theorems tells us that in many cases, a change of coordinates will transform a function into its derivative, locally. Let $E \supset U \xrightarrow{f} F$ be smooth.

<u>Theorem</u>: If $Df(x)$ has constant rank in some nhood of x_0, then there exist nhoods V of x_0 and W of $f(x_0)$ and diffeomorphisms $V \xrightarrow{h} V$ with $h(x_0) = x_0$ and $W \xrightarrow{k} W$ with $k(f(x_0)) = f(x_0)$ such that

$$kfh(x) = f(x_0) + Df(x_0)(x - x_0).$$