

BIOMATH 208 Week 6

Aidan Jan

February 25, 2025

Review

Groups are a practical example of manifold data we encounter a lot

- e.g., rotation matrices, and other types of invertible matrix functions.
- Axioms are a relaxation of $+$ in a vector space. We need a set G and a binary operation $\circ : G \times G \rightarrow G$ that satisfy three axioms:
 1. Associative
 2. Neutral Element
 3. Inverse Element

There are a few subsets of groups:

- Discrete groups: \circ can be written with a Cayley table (always can write a Cayley table if the group is finite, which is most of the time)
- Lie group: Elements can be represented as a smooth manifold, and $\circ, {}^{-1}$ functions need to be smooth functions of the parameters
- Matrix groups are an important example of Lie groups.
 - GL group (invertible $N \times N$ matrices)
 - Affine transformation group (homogeneous coordinates), e.g., linear transformations and translations

Use group actions in imaging:

- affine matrix acting on points Ax
 - $(d+1) \times N$ matrix of N points in d dimensions in homogeneous coordinates
- on normal vectors
 - $A^{-T}n$
- on images
 - $(A \cdot I)(x) = I(A^{-1}x)$
 - If images are discrete, we will discretize them again after transforming, this is not a group action.

Tangent Spaces

We want to build up a tool to understand differentiation on manifolds. This will enable us to do gradient based optimization (e.g., gradient descent), and understand rates of change of various quantities.

We will approach this by considering parameterized curves ($\mathbb{R} \rightarrow \mathcal{M}$), and consider how functions ($\mathcal{M} \rightarrow \mathcal{R}$) change with respect to the curve's velocity.

What is Velocity?

We typically define velocity as "change in position over time". which fundamentally depends on choice of a chart. Here, we will use a different description, based only on curves.

Parameterized Curves

Consider $(a, b) \in \mathbb{R}$ and a curve $\gamma : (a, b) \rightarrow \mathcal{M}$. Assume $0 \in (a, b)$ and let $\gamma(0) = p$ for some $p \in \mathcal{M}$.

- γ is a function with domain \mathbb{R} and range \mathcal{M} .

Definition: Velocity of a Curve

Let f be some smooth function from $\mathcal{M} \rightarrow \mathbb{R}$ ($f \in \mathbb{C}^\infty(\mathcal{M})$). Then the velocity of the curve is a map

$$v_{\gamma,p} : \mathbb{C}^\infty \rightarrow \mathbb{R}$$

defined by

$$v_{\gamma,p}(f) = \left. \frac{d}{dt} f \circ \gamma(t) \right|_{t=0}$$

Note the right side is just a map from $\mathbb{R} \rightarrow \mathbb{R}$, so we can use standard calculus here.

Velocities are linear maps

Based on the above definition, these maps are linear. To prove this, we need to show that they are compatible with $+$ and \cdot .

Proof

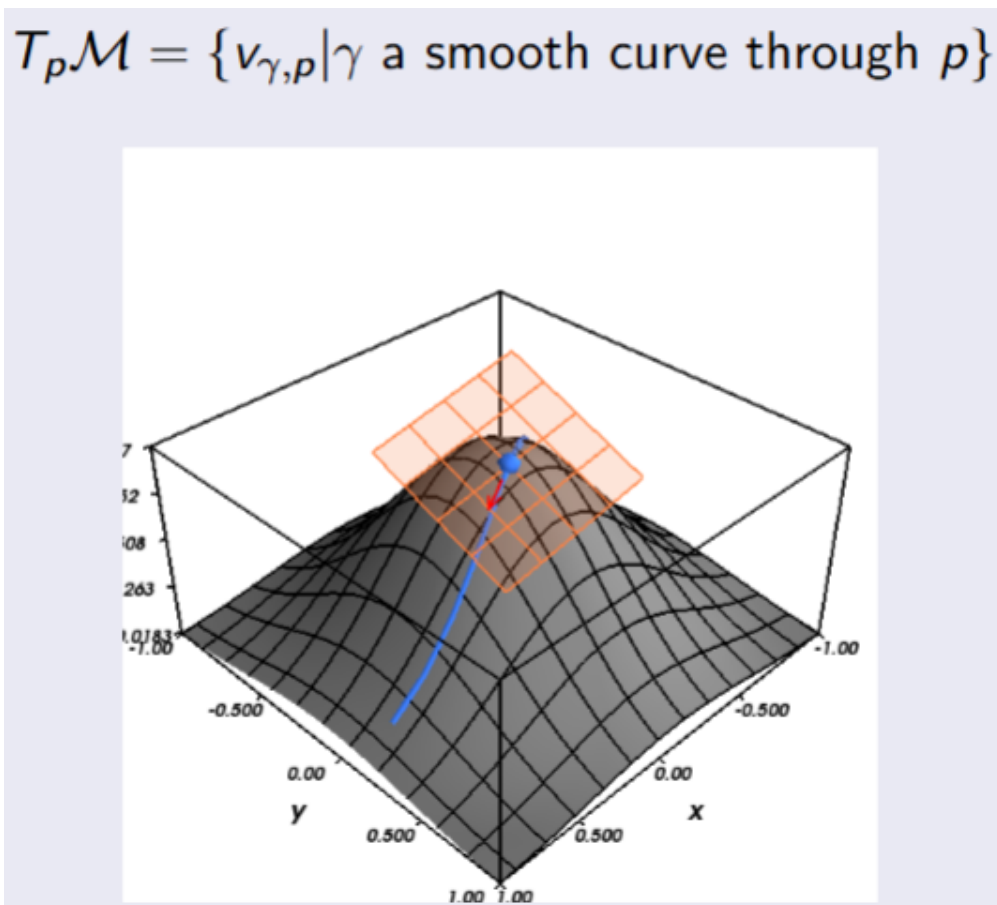
Let $f, g \in \mathbb{C}^\infty$ and $c \in \mathbb{R}$. Then,

$$\begin{aligned} v_{\gamma,p}(f+g) &= \left. \frac{d}{dt} (f+g)(\gamma(t)) \right|_{t=0} \\ &= \left. \frac{d}{dt} f(\gamma(t)) + g(\gamma(t)) \right|_{t=0} \\ &= \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} + \left. \frac{d}{dt} g(\gamma(t)) \right|_{t=0} \\ &= v_{\gamma,p}(f) + v_{\gamma,p}(g) \end{aligned}$$

and,

$$\begin{aligned} v_{\gamma,p}(cf) &= \left. \frac{d}{dt} (c \cdot f)(\gamma(t)) \right|_{t=0} \\ &= \left. \frac{d}{dt} c \cdot f(\gamma(t)) \right|_{t=0} \\ &= c \cdot v_{\gamma,p}(f) \end{aligned}$$

The Tangent Space to \mathcal{M} at the point p



- Note that there are an infinite number of curves that pass through p .
- If we want to check if a linear map is a velocity, we need to find which curve it corresponds to.
- On a smooth manifold, every line passing through p has the same velocity!
- \mathcal{M} is a curved surface. The orange plane, where all the tangents passing through p , is referred to as the **tangent plane**.
- The tangent plane is orthogonal to the normal vector at p .

The Tangent Space is a Vector Space

Definition: Addition of Velocities

We define $(v_{\alpha,p} + v_{\beta,p})(f) = v_{\alpha,p}(f) + v_{\beta,p}(f)$. We must show there is a curve in $T_p\mathcal{M}$ corresponding to the left side.

Proof: We chose some chart U, x , with $p \in U$ and assume (without loss of generality) $x(p) = 0$. Then consider

$$\gamma(t) = x^{-1}(x(\alpha(t)) + x(\beta(t)))$$

- p is in the coordinate neighborhood, U .

Proof: Let's show it has the correct action:

$$\begin{aligned} v_{\gamma,p}(f) &= \left. \frac{d}{dt} f(\alpha(t)) \right|_{t=0} \\ &= \left. \frac{d}{dt} f(x^{-1}(x(\alpha(t)) +_{\mathbb{R}^+} x(\beta(t))) \right|_{t=0} \end{aligned}$$

Look at $f \circ x^{-1} : \mathbb{R}^3 \rightarrow \mathbb{R}$, $x \cdot \alpha$ or $x \cdot \beta : \mathbb{R} \rightarrow \mathbb{R}^d$

$$\begin{aligned} &= \left. \partial_i(f \circ x^{-1}) \right|_{x(\alpha(t)) + x(\beta(t_1))} \cdot \left. \frac{d}{dt} (x^i(\alpha(t)) + x^i(\beta(t_1))) \right|_{t=0} \\ &= \left. \partial_i(f \circ x^{-1}) \right|_{0 \in \mathbb{R}^d} \cdot \left. \frac{d}{dt} (x^i(\alpha(t)) + x^i(\beta(t))) \right|_{t=0} \\ &= \left. \partial_i(f \circ x^{-1}) \right|_0 \cdot \left. \frac{d}{dt} x^i(\alpha(t)) \right|_{t=0} + \left. \partial_i(f \circ x^{-1}) \right|_0 \cdot \left. \frac{d}{dt} x^i(\beta(t)) \right|_{t=0} \end{aligned}$$

Now, undo the derivative

$$= \nu_{\alpha_1} \rho(f) + \nu_{\beta_1} \rho(f)$$

This leads to the pointwise definition of what we expect for linear maps.

Definition: Scalar Multiplication of Velocities

We define scalar multiplication by changing the speed we move along a curve. For $c \in \mathbb{R}$, let $\beta(t) = \gamma(ct)$. Then,

$$\begin{aligned} (c \cdot v_{\gamma,p})(f) &= c \cdot v_{\gamma,p}(f) \\ &= \left. \frac{d}{dt} f(\gamma(ct)) \right|_{t=0} = c \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0} \end{aligned}$$

Action of a velocity in a chart

Let's evaluate the action of a velocity v on a function f in a coordinate neighborhood (U, x) . With $p \in U$, $\gamma(0) = p$, and $x(p) = 0$.

- T, \mathcal{M} should have the same dimension and some basis (assuming dimension is finite).

$$\begin{aligned} v_{\gamma,p}(f) &= \left. \frac{d}{dt} f \cdot \gamma(t) \right|_{t=0} \\ &= \left. \frac{d}{dt} f \circ x^{-1} \circ x \circ \delta(t) \right|_{t=0} \end{aligned}$$

Note that $\gamma : \mathbb{R} \rightarrow \mathcal{M}$, $x : \mathcal{M} \rightarrow \mathbb{R}^d$, $x^{-1} : \mathbb{R}^d \rightarrow \mathcal{M}$, $f : \mathcal{M} \rightarrow \mathbb{R}$

$$\begin{aligned} &= \left. \frac{d}{dt} (f \circ x^{-1}) \circ (x \circ \gamma)(0) \right|_{t=0} \\ &= [(f \circ x^{-1}) \circ (x \circ \gamma)]'(0) \\ &= \partial_i(f \circ x^{-1})(x(p)) \cdot [(x \cdot \gamma)^i]'(0) \end{aligned}$$

A More Familiar Notation

We rewrite the first term as:

$$\partial_i(f \circ x^{-1})(x(p)) \doteq \left(\frac{df}{dx^i} \right)_p = \left(\frac{\partial}{\partial x^i} \right)_p (f)$$

We rewrite the second term as

$$[(x \circ \gamma)^i]'(0) \doteq \dot{\gamma}_x^i(0)$$

- The dot above the $\dot{\gamma}$ means "derivative with respect to time"

[FILL]

Identifying Components

We can rewrite the above action as

$$v_{\gamma,p}(f) = \dot{\gamma}_x(0) \left(\frac{\partial}{\partial x^i} \right) (f)$$

or just the linear operator as

$$v_{\gamma,p} = \dot{\gamma}_x(0) \left(\frac{\partial}{\partial x^i} \right)$$

Unlike ∂_i , the notation $\frac{\partial}{\partial x^i}$ reminds us we are using the chart x . Previously in vector calculus we would say "the directional derivative is the gradient dot the direction"

Chart Induced Basis

The operators $\frac{\partial}{\partial x^i}$ form a basis of $T_p\mathcal{M}$ for $p \in U$. If the dimension of \mathcal{M} is d , there are d basis elements in this d dimensional vector space.

Proof

We already showed that these vectors span the vector space, so we must show these vectors are linearly independent, i.e., show that

$$0 = \lambda^i \left(\frac{\partial}{\partial x^i} \right)_p \implies \lambda^i = 0 \forall i$$

Let's do that:

$$0 = \lambda^i \left(\frac{\partial}{\partial x^i} \right)_p x^j$$

By definition of the partial derivative symbol, we get the following:

$$= \lambda \partial_i (x^j \circ x^{-1})(x(p))$$

This essentially means, "what is the derivative of the j -th component with respect to the i -th component?"

$$\begin{aligned} &= \lambda^i \delta_i^j \\ &= \lambda^j \end{aligned}$$

If we compare the left hand side with the right hand side, it implies that $\lambda^i = 0$. Therefore, we have a d dimensional vector space.

Chart Induced Change of Components

Let $X \in T_p\mathcal{M}$ and let (U, x) and (V, y) be two charts with $p \in U \cap V$. We can express this vector in either chart

$$X_{(y)}^i \left(\frac{\partial}{\partial y^i} \right) = X = X_{(x)}^i \left(\frac{\partial}{\partial x^i} \right)_p$$

These are related by

$$X_{(y)}^i = \left(\frac{\partial y^i}{\partial x^j} \right)_p x_{(x)}^j$$

- ∂x^j is the "down index"
- The entire fraction (coefficient of $X_{(x)}^j$) is a Jacobian matrix of chart transition map.
- A Jacobian is a non-linear map of linear maps.

Proof

Consider acting on a smooth function f .

$$\left(\frac{\partial}{\partial x^i} \right)_p f = \partial_i (f \circ x^{-1})(x(p))$$

Now we will insert the identity $y^{-1} \circ y$

$$= \partial_i [(f \circ y^{-1}) \circ (y \circ x^{-1})](x(p))$$

Notice that $(y \circ x^{-1})$ is a chart transition map. Applying the chain rule...

$$\begin{aligned} &= \partial_j (f \circ y^{-1})(y(p)) \cdot \partial_i (y \circ x^{-1})^j(x(p)) \\ &= \left(\frac{\partial y^j}{\partial x^i} \right)_p \left(\frac{\partial}{\partial y^j} \right)_p \end{aligned}$$

Now, we plug this result into the original expression:

$$= X_{(y)}^i \cdot \left(\frac{\partial}{\partial y^i} \right)_p = X = X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i} \right)_p \left(\frac{\partial}{\partial y^j} \right)_p$$

This quantity is multiplying the basis vector. Therefore, these are the components in the new basis

- Essentially, components in chart (y) (left hand side) is equal to the components in chart (x) times the Jacobian matrix (right hand side)

Example: Vectors in Cartesian and Polar Coordinates

- Consider the manifold $\mathcal{M} = \mathbb{R}^2$, with a standard topology and atlas smoothly compatible with the identity chart. Let (\mathbb{R}^2, x) be the identity chart and (V, y) the polar chart where V has the nonpositive x axis cut out.
- Given a vector with components $X_{(x)}^i$ in chart x , we will find its components in chart y and vice versa.
- To do so, we need to evaluate the Jacobian matrix of the chart transition map:

$$\left(\frac{\partial y^i}{\partial x^j}\right)_p = \partial_j(y^i \circ x^{-1})(x(p))$$

We derived $y \circ x^{-1} = (r, \text{sign}(x^1) \arccos(x^0/r))$, where $r = \sqrt{(x^0)^2 + (x^1)^2}$. We can use some identities,

$$\begin{aligned}\partial_i r(x^0, x^1) &= \frac{x^i}{r(x^0, x^1)} \\ \partial_i \frac{x^j}{r} &= \frac{\delta_i^j r^2 - x^j x^i}{r^3} \\ \arccos'(t) &= -\frac{1}{\sqrt{1-t^2}}\end{aligned}$$

and compute

$$\partial_j y^i \circ x^{-1}(x(p)) = \begin{pmatrix} \frac{x^0}{r} & \frac{x^1}{r} \\ -\frac{x^1}{r^2} & \frac{x^0}{r^2} \end{pmatrix}$$

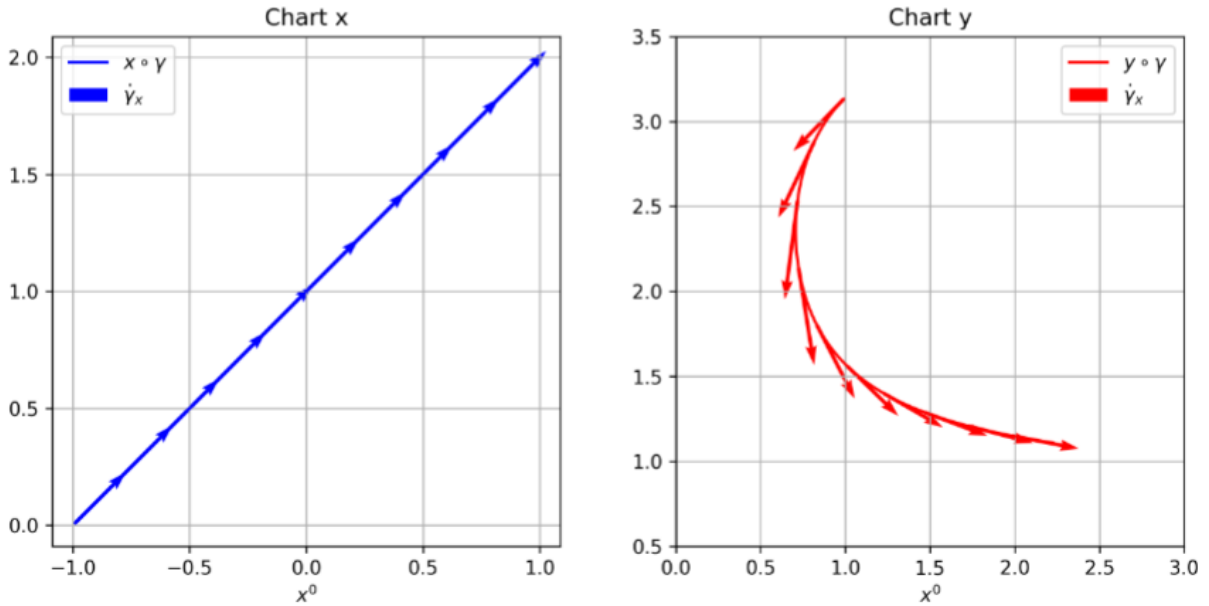
Consider a straight line in chart x : $\gamma(t) = (t, t+1)$, for $t \in (-1, 1)$. In chart y , the curve takes the form

$$y \circ \gamma(t) = \left(\sqrt{t^2 + (t+1)^2}, \text{sign}(t+1) \arccos\left(\frac{t}{\sqrt{t^2 + (t+1)^2}}\right) \right)$$

and the components of its tangent vectors take the form

$$\dot{\gamma}_y = \begin{pmatrix} \frac{x^0}{r} & \frac{x^1}{r} \\ -\frac{x^1}{r^2} & \frac{x^0}{r^2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{x^0+x^1}{r} \\ \frac{-x^1+x^0}{r^2} \end{pmatrix}$$

If we graph the two charts, we get:



This shows that whether a line is straight or not depends on the choice of chart.

Push Forward of Vectors

Consider two manifolds \mathcal{M} and \mathcal{N} , and a smooth map $\phi : \mathcal{M} \rightarrow \mathcal{N}$. Let f be a smooth function on \mathcal{N} .

If $X \in T_p\mathcal{M}$, we denote its push forward to $T_{\phi(p)}\mathcal{N}$ by $\phi_*(p)X \doteq d\phi(p)X$

It is defined by its action on a function $(\phi_*(p)X)f = X(f \circ \phi)$.

- $d\phi$ is the derivative of ϕ , or the Jacobian of the function.

Push Forward in Local Coordinates

Consider \mathcal{M} with a chart (U, x) and \mathcal{Y} with a chart (V, y) . Let $p \in U$ and $d\phi(p) \in V$. Consider $X \in T_p\mathcal{M}$ given by $X = X^i \frac{\partial}{\partial x^i}$. The coordinates of $\phi_*(p)X$ in the chart y are given by

$$(\phi_*(p)X)_{(y)}^j = \left(\frac{\partial y^j \circ \phi}{\partial x^i} \right)_p X_{(x)}^i$$

That is, we multiply by the Jacobian matrix of the map ϕ in local coordinates.

Proof.

Act on a function f and start with the definition

$$(\phi_*(p)X)(f) = X(f \circ \phi)$$

Solving,

$$= X_{(x)}^i \left(\frac{\partial}{\partial x} \right)_p f \cdot \phi$$

By the definition of the partial derivative,

$$= X_{(x)}^i \partial_i (f \circ \phi \circ x^{-1})(x(p))$$

Now, insert the identity map $y^{-1} \circ y$ and use the chain rule.

$$\begin{aligned} &= X_{(x)}^i \partial_i ((f \circ y^{-1}) \circ (y \circ \phi \circ x^{-1}))(x(p)) \\ &= X_{(x)}^i \partial_j (f \circ y^{-1}) \Big|_{y \circ \phi \circ x^{-1} \circ x(p)} \cdot \partial_i (y^j \circ \phi \circ x^{-1}) \Big|_{x(p)} \\ &= X_{(x)}^i \left(\frac{\partial f}{\partial y^j} \right)_{\phi(p)} \left(\frac{\partial y^j \circ \phi}{\partial x^i} \right)_p \\ &= [FILL] \end{aligned}$$

Example: Extrinsic coordinates on the sphere

Here we will push forward a vector from the sphere

$$\mathcal{M} = \{(p^0, p^1, p^2) \in \mathbb{R}^3 | (p^0)^2 + (p^1)^2 + (p^2)^2 = 1\}$$

with the spherical coordinates chart

$$x^{-1}(u, v) = (\cos(u) \cos(v), \cos(u) \sin(v), \sin(u))$$

into \mathbb{R}^3 with the identity chart, using the map $\phi(p) = p$.

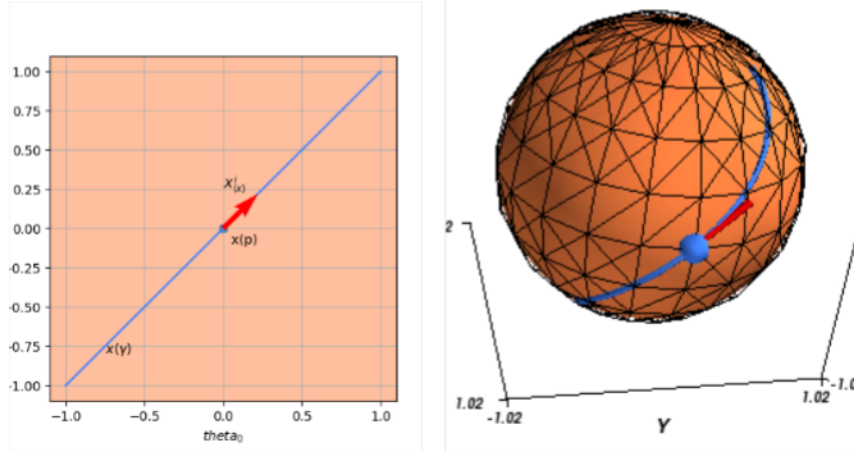
Example: The Jacobian

$y \circ \phi \circ x^{-1}(u, v) = (\cos(u) \cos(v), \cos(u) \sin(v), \sin(u))$, so

$$J = \begin{pmatrix} -\sin(x^0) \cos(x^1) & -\cos(x^0) \sin(x^1) \\ -\sin(x^0) \sin(x^1) & \cos(x^0) \cos(x^1) \\ [FILL] & [FILL] \end{pmatrix}$$

Example: The Push Forward

[FILL] Visualization:



Covectors and Function Gradients

We began with using curves to differentiate functions $v_{\gamma,p} = \dot{\gamma}^i \frac{\partial}{\partial x^i} f = \dot{\gamma}^i \frac{\partial f}{\partial x^i}$. We see the components of the derivatives act on the components of the velocities to give a real number. Therefore, gradients are covectors.

Definition: Dual Basis

If e_i are basis vectors for $T_p\mathcal{M}$, then we can define dual basis vectors ϵ^j for $T_p^*\mathcal{M}$ by $e_i\epsilon^j = \delta_i^j$. By convention, we use the familiar symbols

$$e_i = \left(\frac{\partial}{\partial x} \right)_p, \quad \epsilon^j = (dx^j)_p$$

The dx is not the same as the dx in the integral. They are, however, related.

Definition: The Gradient

We define the gradient of a function f , denoted $df(X)$, in a coordinate manner by its action on a vector X .

$$df(X) \doteq X(f)$$

Identifying Components

We can find the components of df in some chart by acting on the basis elements $\frac{\partial}{\partial x^i}$.

$$(df)_i = df \left(\frac{\partial}{\partial x^i} \right) = \frac{\partial}{\partial x^i} (f) \doteq \partial_i (f \circ x^{-1})$$

- We need a rchart to compute partial derivatives.

So in a chart induced bases, we can write

$$df = (df)_i \cdot dx^i$$

- $(df)_i$ is the component,
- x is the basis vector

Chart Induced Change of Components

Let $\chi \in T_p^* \mathcal{M}$ be expressed in two charts as

$$\chi_j^x dx^j = \chi = \chi_i^y dy^i$$

Then the components in chart y can be expressed in terms of those in chart x by

$$\chi_i^y = \frac{\partial x^i}{\partial y^j} \chi_j^x$$

Proof

Let them act on $X \in T_p \mathcal{M}$ with $X_x^j \frac{\partial}{\partial x^j} = X = X_y^i \frac{\partial}{\partial y^i}$. Then,

$$\begin{aligned} \chi(X) &= \chi_i^y X_y^i \\ &= \chi_i^x X_x^i \\ &= \chi_i^x \left(\frac{\partial x^i}{\partial y^j} \right) X_y^j \end{aligned}$$

$$\therefore \chi_j^y = \frac{\partial x^i}{\partial y^j} \chi_i^x [FILL]$$

Vector and Covector Comparison

Notice that we still have the "inverse transpose" property for covectors.

$$\begin{aligned} X_y^i &= \frac{\partial y^i}{\partial x^j} X_x^j \\ \chi_i^y &= \frac{\partial x^j}{\partial y^i} \chi_j^x \end{aligned}$$

Here, "inverse" is not just matrix inverse, but we must express x as a function of y instead of y as a function of x .

Definition: Pull back of covectors

The opposite of a push forward is a pull back. We define the pull back in terms of the push forward by

$$(\phi^*(\chi))(X) = \chi(\phi_*(X))$$

In coordinates we recover a similar formula (transforming with inverse transpose of Jacobian).

Inner Products (yet again)

We can define higher order tensors as multilinear maps as before. An inner product g at the point p , can act on two vectors X and Y via

$$g(X, Y) = g_{kl} dx^k \otimes dx^l \left(u^i \frac{\partial}{\partial x^i}, v^j \frac{\partial}{\partial x^j} \right)$$

- The \otimes is notation saying that the first covector acts on the first vector, and the second covector acts on the second vector.

Expanding, we get:

$$\begin{aligned} &= g_{kl} u^i v^j \delta_i^\mu \delta_j^\mu \\ &= g_{ij} U^i V^j [FILL] \end{aligned}$$

Change of Coordinates for Inner Products

We can derive a change of basis formula for inner products by applying the change of basis formula for vectors:

- Note that x and y are charts. Since a transfer function can be used to transfer between charts, $g(X, Y)$ is independent of which chart is used since they cancel out.

$$\begin{aligned} g(X, Y) &= g_{ij}^y X_y^i Y_y^j = g_{ij}^x X_x^i Y_x^j \\ &= g_{ij}^x \left(\frac{\partial x^i}{\partial y^k} X_y^k \right) \left(\frac{\partial x^j}{\partial y^l} Y_y^l \right) \\ &= \left[\left(\frac{\partial x^i}{\partial y^k} \right) g_{ij}^x \left(\frac{\partial x^k}{\partial y^l} [FILL] \right) \right] \\ &= \end{aligned}$$

Example: "Standard" inner product in polar coordinates

- The "standard" inner product refers to the dot product.

Consider the space \mathbb{R}^2 with x the "identity chart" and y the "polar coordinates chart" $y(p) = (r, \text{sign}(p^1) \arccos(p^0/r))$.

First find the coordinates of g in the chart x .

$$g_{ij}^{(x)} = \delta_{ij}$$

Now, find the jacobian of the chart transition map

$$\begin{aligned} x \circ y^{-1} &= (y^0 \cos(y^1), y^0 \sin(y^1)) \\ \partial_j x^i \circ y^{-1} &= \begin{pmatrix} \cos(y^1) & -y^0 \sin(y^1) \\ \sin(y^1) & y^0 \cos(y^1) \end{pmatrix} \end{aligned}$$

- Each column in the matrix is a component, partial derivatives is a row.

Now, form the matrix product

$$\begin{aligned} &\begin{pmatrix} \cos(y^1) & \sin(y^1) \\ -y^0 \sin(y^1) & y^0 \cos(y^1) \end{pmatrix} \cdot I \cdot \begin{pmatrix} \cos(y^1) & -y^0 \sin(y^1) \\ \sin(y^1) & y^0 \cos(y^1) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(y^1) + \sin^2(y^1) & -y^0 \sin(y^1) \cos(y^1) + y^0 \sin(y^1) \cos(y^1) \\ 0 & (y^0)^2 \sin^2(y^1) + (y^0)^2 \cos^2(y^1) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & (y^0)^2 \end{pmatrix} \end{aligned}$$

Tensor Field

- We can define a tangent space $T_p \mathcal{M}$ for every point p on a manifold. The union of all these spaces is called the "tangent bundle".
- A vector field can be defined simply as a vector at every point. Same for covector or tensor fields.
- Because we cannot add or subtract vectors from different spaces, taking the derivative of a vector field is challenging to define.
- Mathematically, the tangent bundle is a type of object called a "fiber bundle" and a vector field is a "section" of this bundle.

$$\text{Tangent Bundle} = \bigcup_{p: p \in \mathcal{M}} T_p \circ \mathcal{M}$$