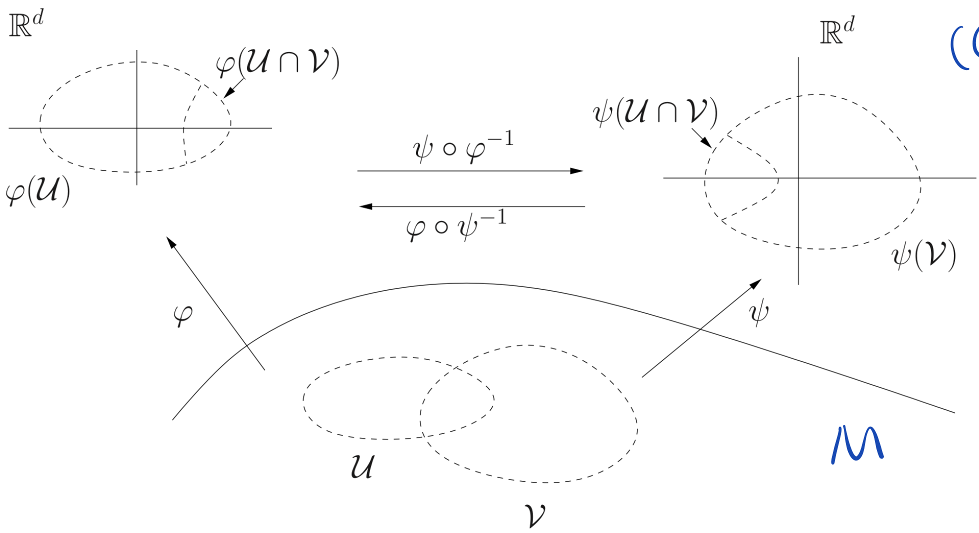
- Recall that the generic definition of a smooth manifold M can be very abstract



(φ, U) and (ψ, V) are coordinate charts

- For simplicity, we have consider a manifold that can be described as a level set of some $h(x)$

1) $\mathcal{E}$ is d-dim linear space

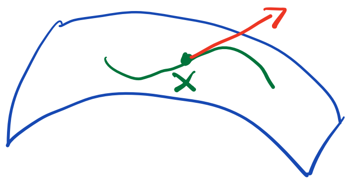
2) $M \subseteq \mathcal{E}$ and $\forall x \in \mathcal{E}, h(x) = 0 \iff x \in M$

3) $\text{Rank}[Dh(x)] = k$, $\forall x \in M$

Then $\begin{cases} M \text{ is an embedded submanifold of dim } d-k \\ T_xM = \text{Ker}[Dh(x)] \text{ is the tangent space at } x. \\ T_xM \text{ is a linear space of dim } d-k. \end{cases}$

- The general definition of tangent space of a manifold $M \subseteq \mathcal{E}$

$$T_xM = \{c'(0) \mid c: (-\varepsilon, \varepsilon) \to M \text{ is smooth } \& \ c(0) = x\}$$



$c(t)$ is a curve on M
$c'(t)$ is tangent to the curve

Theorem If $M \subseteq \mathcal{E}$ is describle by $h(x)=0$, then
$$\ker[Dh(x)] = T_x M$$

Proof: ① $\forall\, v \in T_x M$, $\exists$ a curve $c: (-\varepsilon, \varepsilon) \to M$

s.t. $\begin{cases} c(0) = x \\ c'(0) = v \end{cases}$

Since $c(t) \in M$, $\forall t$, we have $h(c(t)) = 0$

$$\Rightarrow 0 = \frac{d}{dt} h(c(t)) = \langle \operatorname{grad} h(c(t)), c'(t) \rangle = Dh(c(t))[c'(t)]$$

$$\left(Dh(c(t))[c'(t)] = \lim_{s \to 0} \frac{h(c(t) + s c'(t)) - h(c(t))}{s} \right)$$

$$\Rightarrow \quad v = c'(0) \in \ker[Dh(x)] \Rightarrow T_x M \subseteq \ker[Dh(x)]$$

② We can show $T_x M$ contains a linear subspace of dimension $n-k$ (skipped, see book B6)

$$\Rightarrow T_x M = \ker[Dh(x)]$$

- Definition/Notation

M is a manifold
$T_x M$ is the tangent space of M at x

Any element in T_xM is a tangent vector at x

Tangent Bundle TM

$$TM = \{ \underbrace{V_x}_{} \in T_xM : \forall x \in M \} = \bigcup_{x \in M} T_xM$$

subscript x deotes it's a tangent vector at x

For convenience, we sometimes ignore the subscript.

A vector field on M is a map

$$V : M \longrightarrow TM$$

s.t. $V(x) \in T_xM$ for any $x \in M$

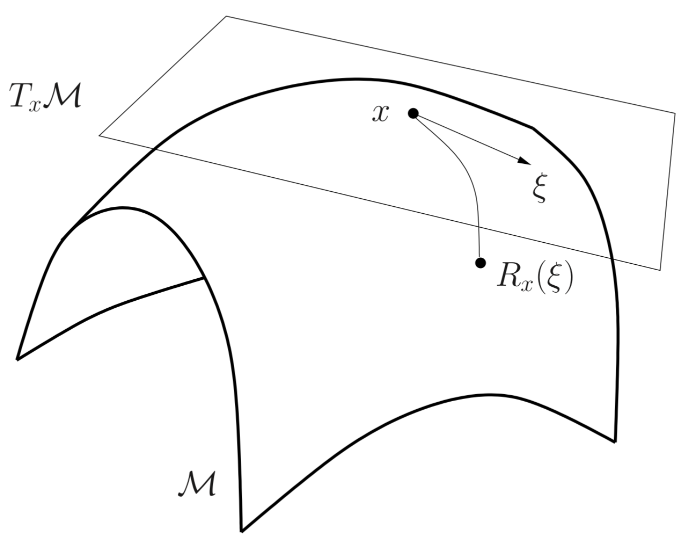
A Retraction on M is a smooth map

$$R : TM \longrightarrow M$$
$$V_x \longmapsto R_x(V)$$

Such that each curve $c(t) = R_x(tV)$ satisfies

$$\begin{cases} c(0) = x \\ c'(0) = V \end{cases}$$

$$R : TM \to M$$
$$V_x \mapsto R_x(V_x)$$

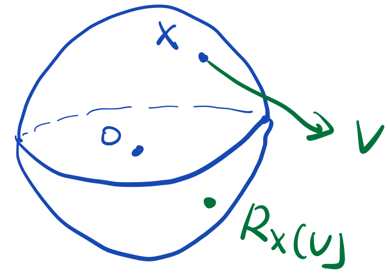


$$c(t) = R_x(t\xi)$$

$$c(0) = x$$

$$c'(0) = \xi$$

Examples of Retraction on S^{n-1}



① $R_x(v) = \dfrac{x+v}{\|x+v\|} = \dfrac{x+v}{\sqrt{\|x+v\|^2}}$

$= \dfrac{x+v}{\sqrt{1+\|v\|^2}}$

② $R_x(v) = \cos(\|v\|)\,x + \dfrac{\sin(\|v\|)}{\|v\|}\,v$

$c(t) = R_x(tv)$ is the great circle passing x with $c'(0)=v$.

<u>Def</u> A metric is a choice of inner product $\langle \cdot\, , \cdot \rangle_x$ for T_xM at each $x \in M$.

<u>Def</u> A metric is Riemannian if $\langle \cdot\, , \cdot \rangle_x$ is smooth w.r.t. x

i.e. $\langle V(x), W(x) \rangle_x$ is smooth for smooth vector fields V, W.

<u>Def</u> Let $g_x(\cdot\, , \cdot) = \langle \cdot\, , \cdot \rangle_x$ be a Riemannian metric

(M, g) is a Riemannian manifold.

Example: $\langle\, , \rangle$ is the inner product in $\mathcal{E}$

M is an embedded submanifold in $\mathcal{E}$

Then $\langle\, , \rangle$ is a Riemannian metric of M.

Example: $\mathcal{E}$ is also a manifold

Differential of a map

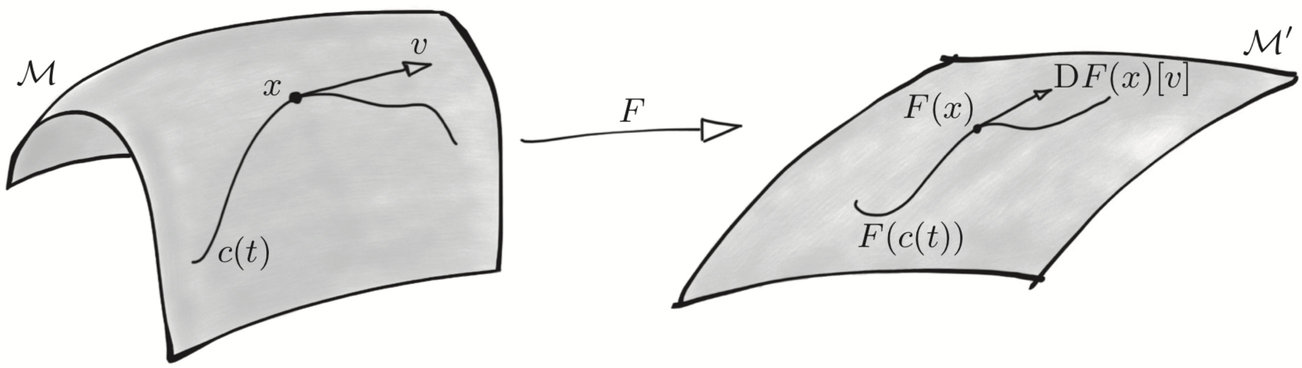
- $F: M \to M'$ is smooth function where M, M' are manifolds

  The differential is the linear map

  $DF(x): T_x M \to T_{F(x)} M'$ defined by

  $$DF(x)[v] = \frac{d}{dt} F(c(t)) \Big|_{t=0} = (F \circ c)'(0)$$

  where $c(t)$ is a smooth curve on M $\begin{cases} c(0) = x \\ c'(0) = v \end{cases}$



Facts: ① $DF(x)[v]$ is unique, regardless of choice of $c(t)$

② $DF(x)$ is a linear map

- Let $M \subset \mathcal{E}$ and $M' \subset \mathcal{E}'$ be embedded submanifolds in $\mathcal{E}$ and $\mathcal{E}'$

  Then a smooth map $F: M \to M'$ has a smooth extension $\bar{F}: \mathcal{E} \to \mathcal{E}'$

Example: $\min\limits_{x \in S^{n-1}} f(x) = x^T A x$, $f: S^{n-1} \to \mathbb{R}$ can be regarded as

$\bar{f}: \mathbb{R}^n \to \mathbb{R}$

Notice that $F(c(t)) = \bar{F}(c(t))$ for a curve $c(t)$ in M.

$$\frac{d}{dt}\bar{F}(c(t))\Big|_{t=0} = D\bar{F}(c(0))[c'(0)] = D\bar{F}(x)[v]$$

$$\parallel$$

$$\frac{d}{dt}F(c(t))\Big|_{t=0} = DF(x)[v]$$

For a smooth map $\bar{F}: \mathcal{E} \to \tilde{\mathcal{F}}$

its differential at $x_0 \in U$ is $D\bar{F}(x_0): \mathcal{E} \to \tilde{\mathcal{F}}$ defined by

$$\forall u \in \mathcal{E}, \quad D\bar{F}(x_0)[u] = \frac{d}{dt}\bar{F}[x_0 + tu]\Big|_{t=0} = \lim_{t \to 0} \frac{\bar{F}(x_0+tu) - \bar{F}(x_0)}{t}$$

$$\bar{f}(x) = x^T A x \Rightarrow D\bar{f}(x)[v] = 2x^T A v \quad \begin{cases} f: S^{n-1} \to \mathbb{R} \\ \bar{f}: \mathbb{R}^n \to \mathbb{R} \end{cases}$$

$$\Rightarrow Df(x)[v] = 2x^T A v$$

$$\forall v \in T_x S^{n-1} = \{v \in \mathbb{R}^n : x^T v = 0\}$$

$D\bar{f}(x): \mathbb{R}^n \to \mathbb{R}$ but $Df(x): T_x S^{n-1} \to \mathbb{R}$

Riemannian Gradient

For a smooth $f: M \to \mathbb{R}$ on a Riemannian manifold (M, g)

the Riemannian gradient of f is the vector field $\operatorname{grad} f$

defined by $g_x(\operatorname{grad} f(x), v_x) = Df(x)[v_x]$, $\forall x \in M$, $\forall v_x \in T_x M$

Remark: $\operatorname{grad} f(x)$ is a tangent vector

For a smooth function $\bar{f} : \mathcal{E} \to \mathbb{R}$,

the Euclidean gradient $\nabla \bar{f} : \mathcal{E} \to \mathcal{E}$ is defined by

$$\langle \nabla \bar{f}(x), v \rangle = D\bar{f}(x)[v] \quad \forall x, v \in \mathcal{E}$$

If $f : M \to \mathbb{R}$ can be extended to

$\bar{f} : \mathcal{E} \to \mathbb{R}$, then

$\nabla \bar{f}$ is the Euclidean gradient of $\bar{f}$

$\operatorname{grad} f$ is the Riemannian gradient of f

and $\forall v \in T_x M$, $g_x(\operatorname{grad} f(x), v) = Df(x)[v]$

$$= D\bar{f}(x)[v]$$

$$= \langle \nabla \bar{f}(x), v \rangle$$

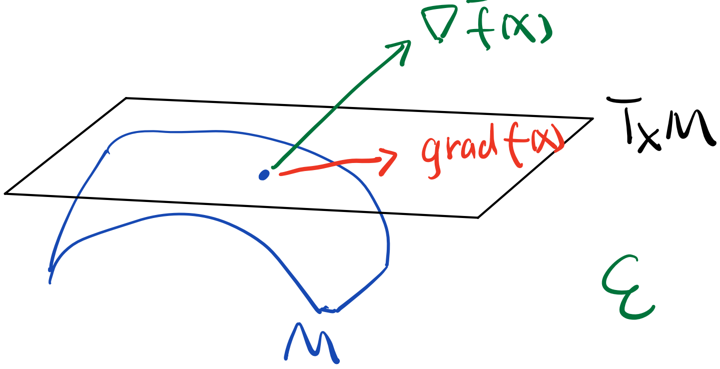
Notice that the metric g_x does not need to be the same as $\langle\, , \rangle$.

Riemannian Gradient of a Riemannian submanifold of $\mathcal{E}$

meaning metric is $\langle\, , \rangle$ in $\mathcal{E}$

$$\langle \operatorname{grad} f(x), v \rangle = g_x(\operatorname{grad} f(x), v) = \langle \nabla \bar{f}(x), v \rangle$$

$\Rightarrow \operatorname{grad} f(x)$ is the projection of $\nabla \bar{f}(x)$ onto $T_x M$



$$\langle \nabla \bar{f}(x) - \operatorname{grad} f(x), v \rangle = 0$$

$$\forall v \in T_xM$$

$$\Rightarrow [\nabla \bar{f}(x) - \operatorname{grad} f(x)] \perp T_xM$$

$$\Updownarrow$$

$$[\nabla \bar{f}(x) - w] \perp T_xM$$

$$\Updownarrow \quad \text{critical point}$$

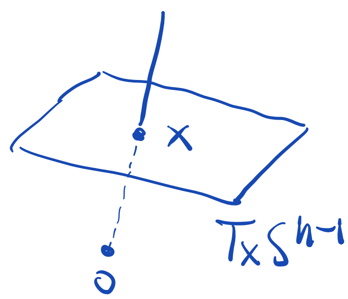
$$w = \underset{v \in T_xM}{\operatorname{argmin}} \|\nabla \bar{f}(x) - v\|^2 = \underset{v \in T_xM}{\operatorname{argmin}} \langle \nabla \bar{f}(x) - v, \nabla \bar{f}(x) - v \rangle$$

Example ① Regard S^{n-1} as a Riemannian submanifold of $\mathbb{R}^n$

$$\left.\begin{array}{l} f(x) = \frac{1}{2} x^T A x \text{ defined on } S^{n-1} \\ \bar{f}(x) = \frac{1}{2} x^T A x \text{ defined on } \mathbb{R}^n \end{array}\right\} A^T = A$$

$$\nabla \bar{f}(x) = Ax$$

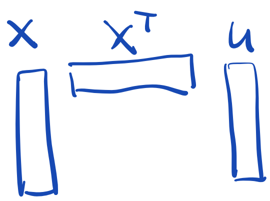
Projection of u onto T_xS^{n-1}



$$P_{T_xS^{n-1}}(u) = x - \frac{\langle x, u \rangle}{\langle x, x \rangle} x \quad , x \in S^{n-1}$$

$$= u - \langle x, u \rangle x$$

$$= (I - xx^T) u$$



$$\operatorname{grad} f(x) = P_{T_xS^{n-1}}(\nabla \bar{f}(x))$$

$$= (I - xx^T)Ax = Ax - (x^TAx)x$$

Remark: $\operatorname{grad} f(x) = 0 \iff Ax = \underbrace{(x^TAx)}_{\text{scalar}} x$

So critical point on $S^{n-1} \iff$ eigenvectors

② $M = \Delta_+^{n-1} = \{x \in \mathbb{R}^n : x_i > 0, \sum_{i=1}^n x_i = 1\}$

relative interior of the simplex or probability manifold

M is an embedded manifold of $\mathbb{R}^n$

$T_xM = \{v \in \mathbb{R}^n : \sum_{i=1}^n v_i = 0\}$

The Fisher-Rao metric

$$g_x(u,v) = \sum_{i=1}^n \frac{u_i v_i}{x_i}$$

$f : M \to \mathbb{R}$

$\bar{f} : \mathbb{R}^n \to \mathbb{R}$ $\quad\quad \nabla \bar{f}(x)$

$$g_x(\operatorname{grad} f(x), v) = \langle \nabla \bar{f}(x), v \rangle = v^T \cdot \nabla \bar{f}(x)$$

To find an expression

Let $\operatorname{grad} f(x) = u \in T_xM$

$\frac{u}{x}$ denotes the vector $\left[\frac{u_1}{x_1} \cdots \frac{u_n}{x_n}\right]^T$

$$v^T \cdot \frac{\operatorname{grad} f(x)}{x} = v^T \cdot \nabla \bar{f}(x) \quad\quad \forall v \in T_xM$$

Given $w \in \mathbb{R}^n$, solve $v^T u = v^T w$, $\forall v \in T_x\mathcal{M}$

$$\Rightarrow v^T(u-w) = 0$$

$$\Rightarrow P_{T_x\mathcal{M}}(u-w) = 0$$

$$\Rightarrow u-w \parallel \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\Rightarrow u = w + \lambda \mathbb{1}, \quad \lambda \in \mathbb{R}$$

$$\Rightarrow \frac{\operatorname{grad} f(x)}{x} = \nabla \bar{f}(x) + \lambda \mathbb{1}$$

$$\Rightarrow \operatorname{grad} f(x) = x \circ \nabla \bar{f}(x) + \lambda x$$

entry wise product

$$\operatorname{grad} f(x) \in T_x\mathcal{M} \Rightarrow \sum_{i=1}^{n} \left(x_i [\nabla \bar{f}(x)]_i + \lambda x_i \right) = 0$$

$$\Rightarrow \lambda = -\sum_{i=1}^{n} x_i [\nabla \bar{f}(x)]_i$$

$$\Rightarrow \operatorname{grad} f(x) = x \circ \nabla \bar{f}(x) - \langle x, \nabla \bar{f}(x) \rangle x$$