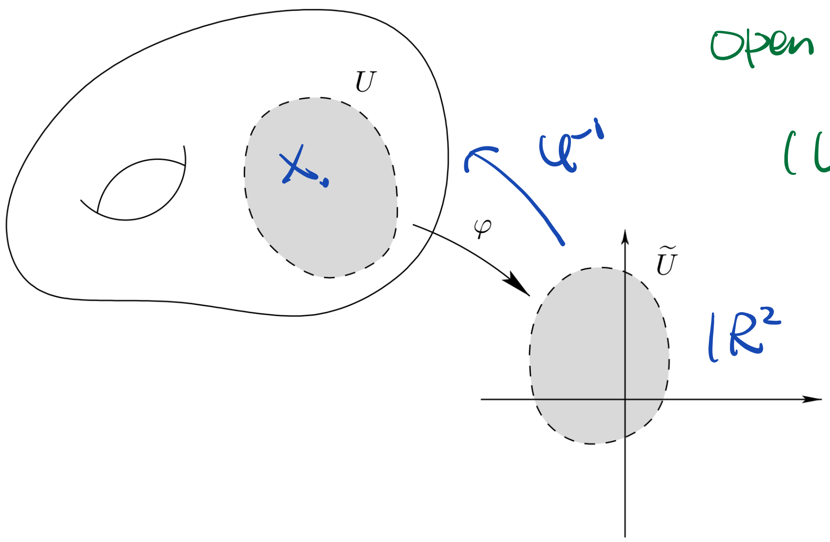
Plan First Order Geometry of Manifolds

Definition & Examples
- Tangent Space
- Riemannian Metric
- Riemannian gradient

Review of a smooth manifold:

- A topological manifold M of dimension n is a second countable Hausdauff topological space and any $x \in M$ has a neighborhood U that is homeomorphic to an open set $\tilde{U}$ in $\mathbb{R}^n$

φ and φ^{-1} are continuous

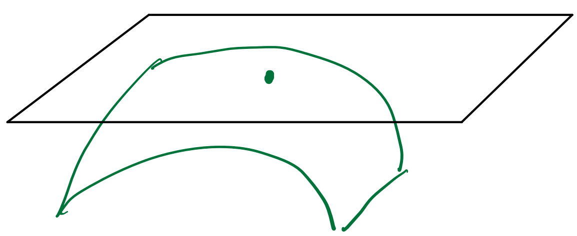


Open set U is called neighborhood

(U, φ) is called a (coordinate) chart

Examples: ① $\mathbb{R}^n$ is a manifold of dimension n

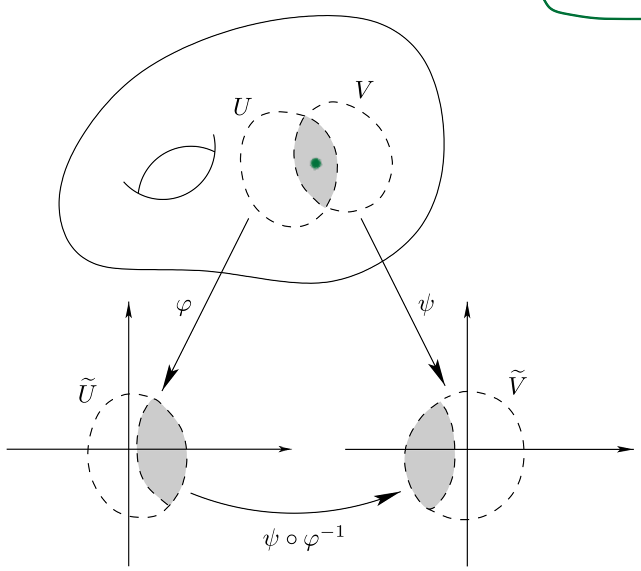
② Any surface is a manifold of dim 2



(U, φ) can be obtained by mapping U to tangent plane at x.

• A smooth manifold is a topological manifold with a smooth structure, i.e., any two charts must be smoothly compatible

transition map $\psi \circ \varphi^{-1}$ is smooth



Examples:

① Sphere in $\mathbb{R}^n$

$$S^{n-1} = \{ x \in \mathbb{R}^n : x_1^2 + x_2^2 + \dots + x_n^2 = 1 \} = \{ x \in \mathbb{R}^n : x^T x = 1 \}$$

② (Compact) Stiefel manifold

$$St(n,p) = \{ U \in \mathbb{R}^{n \times p} : \text{columns of } U \text{ are orthonormal} \}$$
$$= \{ U \in \mathbb{R}^{n \times p} : U^T U = I_p \}$$

③ Fixed Rank positive semi-definite matrices form a manifold

PSD

$$S_+^{n,p} = \{ X \in \mathbb{R}^{n \times n} : X^T = X ,\ X \succeq 0 ,\ \text{rank}(X) = p \}$$

④ PSD matrices of rank at most p do NOT form a manifold

$\{X \in \mathbb{R}^{n\times n} : X^T = X, X \succcurlyeq 0, \text{rank}(X) \le p\}$ is not a manifold but an Algebraic Variety

---

Heuristics : derivatives along a surface

Consider a surface described by

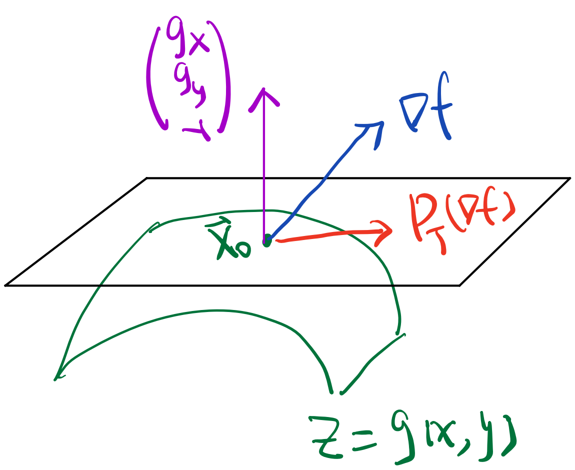
$$S = \{\vec{x} = (x,y,z) : z = g(x,y)\}$$

Consider $\min_{\vec{x}\in S} f(\vec{x})$

Lagrangian $L(\vec{x},\lambda) = f(\vec{x}) - \lambda(g(x,y) - z)$

KKT system
$$\begin{cases} \dfrac{\partial L}{\partial \vec{x}} = \begin{pmatrix} f_x - \lambda g_x \\ f_y - \lambda g_y \\ f_z + \lambda \end{pmatrix} = \nabla f - \lambda \begin{pmatrix} g_x \\ g_y \\ -1 \end{pmatrix} = 0 & ① \\ \dfrac{\partial L}{\partial \lambda} = g(x,y) - z = 0 & ② \end{cases}$$

$$① \Rightarrow \nabla f \parallel \begin{pmatrix} g_x \\ g_y \\ -1 \end{pmatrix}$$



Tangent Plane Equation at $\vec{x}_0$

$$g_x(x-x_0) + g_y(y-y_0) - (z-z_0) = 0$$

$$\Updownarrow$$

$$z = g(x_0,y_0) + g_x(x-x_0) + g_y(y-y_0)$$

So ∇f is parallel to normal vector

$\Rightarrow$ Projection of ∇f to tangent space should be 0

$P_T(\nabla f)$ is the gradient along the surface

$\Rightarrow$ At minimizer, $P_T(\nabla f) = 0$ (∇f can be nonzero)

---

① Next, want to extend it to manifolds

② Need to define "tangent vectors" for manifolds, which however are abstract sets.

So in general tangent vector of a manifold has an abstract definition

③ To make it easier, for now, we only consider a linear space $\mathcal{E}$ and its subset M

Examples: 1) $\mathcal{E} = \mathbb{R}^n$, $M = S^{n-1} = \{x \in \mathbb{R}^n : x^T x = 1\}$

$p \le n$ 2) $\mathcal{E} = \mathbb{R}^{n \times p}$, $M = St(n,p) = \{X \in \mathbb{R}^{n \times p} : X^T X = I_p\}$

Let $\mathcal{E}$ be a real Linear Space, $M \subset \mathcal{E}$ be a manifold

① An inner product on $\mathcal{E}$ is defined by requiring

$\forall u, v, w \in \mathcal{E}$, $\forall a, b \in \mathbb{R}$

1) symmetric $\langle u, v \rangle = \langle v, u \rangle$

2) linearity $\langle au+bv, w\rangle = a\langle u,w\rangle + b\langle v,w\rangle$

3) positive $\langle u,u\rangle \geq 0$ & $\langle u,u\rangle = 0 \Leftrightarrow u=0$

② Euclidean distance $\|u\| = \sqrt{\langle u,u\rangle}$

Example: $\mathcal{E} = \mathbb{R}^{n\times p}$, Frobenius inner product is

$$\langle U,V\rangle = \sum_{i=1}^{n}\sum_{j=1}^{p} U_{ij}V_{ij} = tr(V^TU) = tr(U^TV)$$

$$\|U\|_F = \sqrt{\langle U,U\rangle} = \sqrt{\sigma_1^2+\cdots+\sigma_p^2}$$

③ u, v are orthogonal if $\langle u,v\rangle = 0$

④ open ball centered at x_0 with radius R

$$B(x_0,R) = \{x\in\mathcal{E} : \|x-x_0\| < R\}$$

open sets can be defined as:

$S\subset\mathcal{E}$ is open if $\forall x_0\in S$, $\exists R>0$ s.t. $B(x_0,R)\subseteq S$

⑤ Smooth map & differential

Consider two vector spaces $\mathcal{E}$ and $\mathcal{F}$

U is an open set in $\mathcal{E}$

V is an open set in $\mathcal{F}$

Consider a smooth map $F: U\to V$,

differential of F at $x_0\in U$ is the linear map

$DF(x_0): \mathcal{E}\to\mathcal{F}$ defined by

$\forall u\in\mathcal{E}$

$$DF(x_0)[u] = \frac{d}{dt}F(x_0+tu)\Big|_{t=0} = \lim_{t\to 0}\frac{F(x_0+tu)-F(x_0)}{t}$$

For a curve $c : \mathbb{R} \to \mathcal{E}$,

$c'(t) = \frac{d}{dt} c(t)$ is the velocity

⑥ Gradient

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

the Euclidean gradient $\operatorname{grad} f : \mathcal{E} \to \mathcal{E}$ is defined by

$$\langle \operatorname{grad} f(x), v \rangle = Df(x)[v] \quad \forall x, v \in \mathcal{E}$$

Example: $\mathcal{E} = \mathbb{R}^{n \times p}$,

$$f(X) = \langle X, X \rangle = \|X\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{p} X_{ij}^2$$

$$Df(X)[U] = \lim_{t \to 0} \frac{f(X + tU) - f(X)}{t}$$

$$= \lim_{t \to 0} \sum_{i=1}^{n} \sum_{j=1}^{p} \frac{|X_{ij} + tU_{ij}|^2 - |X_{ij}|^2}{t}$$

$$= 2 \sum_{i=1}^{n} \sum_{j=1}^{p} X_{ij} \cdot U_{ij} = \langle 2X, U \rangle$$

$$\Rightarrow \operatorname{grad} f(X) = 2X$$

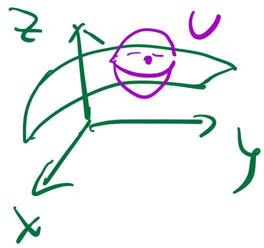
⑦ Embedded submanifold of $\mathcal{E}$

Definition Let $\mathcal{E}$ be a vector space of dim d and $\mathcal{M}$ be an non-empty subset.

$\mathcal{M}$ is an embedded submanifold of $\mathcal{E}$ of dim n if

$\forall x \in \mathcal{M}$, $\exists$ an open set U in $\mathcal{E}$, $x \in U$

and a smooth $h: U \to \mathbb{R}^k$



such that ① $n = d - k$

② $\forall y \in U$, $h(y) = 0 \Leftrightarrow y \in M$

③ $\text{Rank}[\underbrace{Dh(x)}_{\text{regard it as a matrix}}] = k$, $\forall x \in M$

differential of F at x_0 is the linear map

$DF(x_0): \mathcal{E} \to \mathcal{F}$ defined by

Example: 1) $\mathcal{E} = \mathbb{R}^d$, $M = S^{d-1} = \{x^T x = 1\}$

$$\left.\begin{aligned} h &: \mathcal{E} \to \mathbb{R}^1 \\ h(y) &= y^T y - 1 \end{aligned}\right\} \Rightarrow \begin{cases} k = 1 \\ n = d-1 \end{cases}$$

$$Dh(y): \mathcal{E} \to \mathbb{R}$$

$$u \mapsto \lim_{t \to 0} \frac{h(y+tu) - h(y)}{t} = \langle 2y, u \rangle$$

The matrix representation of $L: \underset{\text{dim } n}{\mathcal{E}} \to \underset{\text{dim } m}{\mathcal{F}}$ is $m \times n$ matrix

The matrix representation is $[2y]^T$ of size $1 \times n$

2) $\mathcal{E} = \mathbb{R}^3$, $S = \left\{ \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} : F(x,y,z) = c \right\}$

$$h(\vec{x}) = F(x,y,z) - c$$

The matrix for $Dh(\vec{x})$ has size

3) $\mathcal{E} = \mathbb{R}^{n \times p}$, $St(n,p) = \{X \in \mathbb{R}^{n \times p} : X^T X = I_p\}$

$h(X) = X^T X - I$ 　　 dim $= n \times p - p(p+1)/2$

$$Dh(X)[U] = \lim_{t \to 0} \frac{[X+tU]^T [X+tU] - X^T X}{t}$$

$$= \lim_{t \to 0} \frac{tU^T X + tX^T U + t^2 U^T U}{t}$$

$$= U^T X + X^T U$$

$$Dh(X) : \mathbb{R}^{n \times p} \longrightarrow \mathbb{R}^{p \times p}$$

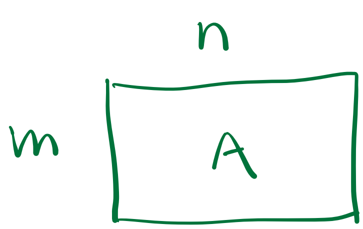
To find the rank:

$$Dh(X)[\tfrac{1}{2} X U] = (\tfrac{1}{2} X U)^T X + X^T (\tfrac{1}{2} X U)$$

$$= \tfrac{1}{2} U^T X^T X + \tfrac{1}{2} X^T X U = U$$

$\Rightarrow$ $Dh(X)$ can map to any $\underline{U \in \mathbb{R}^{p \times p}, U^T = U}$

$\Rightarrow$ $Dh(X)$ has rank $(p+1)p/2$ $(p \le n)$



$L_A : \mathbb{R}^n \to \mathbb{R}^m$

$x \longmapsto Ax$

If $\forall y \in \mathbb{R}^m$, $\exists x \in \mathbb{R}^n$, $Ax = y$, then $\{Ax : x \in \mathbb{R}^n\} = \mathbb{R}^m$

$\qquad\qquad \shortparallel$

$\qquad\qquad$ Column Space of A

dim of Col Space is rank

Read Chapter 3 in Reference Book 6.