Exercise 4

Due date:

Name : Vivi
Student ID : 24S153073
Grade : ________

Problem 1 Score: ________. Fréchet mean on hemisphere
Consider the $(d-1)$-dimensional hemisphere

$$\mathcal{M} = \{x = (x^{(1)}, \cdots, x^{(d)}) \in \mathbb{R}^d : \|x\| = 1, x^{(d)} > 0\},$$

viewed as a Riemannian submanifold of $\mathbb{R}^d$ (with standard Euclidean inner product). The Riemannian distance between two points $x, y \in \mathcal{M}$ is given by

$$d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}, \quad d(x,y) = \arccos\left(x^\top y\right).$$

Let $x_1, \cdots, x_n \in \mathcal{M}$, and define

$$f : \mathcal{M} \to \mathbb{R}, \quad f(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2} d(x, x_i)^2.$$

A minimizer of f can be interpreted as an intrinsic Fréchet mean of the points $x_1, \cdots, x_n$ on the hemisphere.

(1) Show that $f : \mathcal{M} \to \mathbb{R}$ is smooth.
using the flowing fact. Define the function

$$g : (-1, 1] \to \mathbb{R}, \quad g(t) = \frac{1}{2}\arccos(t)^2.$$

There is a smooth function $\bar{g} : (-1, 2) \to \mathbb{R}$ such that $\bar{g} = g$ for all $t \in (-1, 1]$, and $\bar{g}'(t) = -\frac{\arccos(t)}{\sqrt{1-t^2}}$ for all $t \in (-1, 1)$ and $\bar{g}'(1) = -1$.

(2) Given $x \in \mathcal{M}$, give an expression for the Riemannian gradient of f at x.

Solution: (1) To show that $f : \mathcal{M} \to \mathbb{R}$ which is a sum of functions is smooth, it suffices to show that the function

$$h : \mathcal{M} \times \mathcal{M} \to \mathbb{R}, \quad h(x,y) = \frac{1}{2} d(x,y)^2 = \frac{1}{2}\arccos(x^\top y)^2 = g(x^\top y)$$

is smooth. To do so, we build a smooth extension $\bar{h}$ of h.
As $\|x\| = \|y\| = 1$ and $x^{(d)}, y^{(d)} > 0$, we have $x^\top y = \sum_{i=1}^{d} x^{(i)} y^{(i)} = \sum_{i=1}^{d-1} x^{(i)} y^{(i)} + x^{(d)} y^{(d)} \in (-1, 1]$.
Thus we can have $\bar{h}$ on $d^{-1}(-1, 2)$ by defining $\bar{h}(x, y) = \bar{g}(x^\top y)$ for all $x, y \in$ some neighborhood of $\mathcal{M}$. Then $\bar{h}$ is smooth and therefore h is smooth, which implies that f is smooth.

(2) Consider the following function:

$$h : \mathcal{M} \to \mathbb{R}, \quad h(x) = g(x^\top y)$$

Given $x \in \mathcal{M}$, we compute the Riemannian gradient of h at first:

$$\begin{aligned}
gradh(x) &= Proj(grad\bar{h}(x)) \\
&= (I - xx^\top) grad\bar{h}(x) \\
&= (I - xx^\top) y \bar{g}'(x^\top y) \\
&= -(I - xx^\top) y \frac{\arccos(x^\top y)}{\sqrt{1 - (x^\top y)^2}} \\
&= (x\cos(d(x,y)) - y)\frac{d(x,y)}{\sin(d(x,y))}.
\end{aligned}$$

Therefore, the Riemannian gradient of f at x is

$$\begin{aligned}
gradf(x) &= \frac{1}{n}\sum_{i=1}^{n} gradh(x_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} (x\cos(d(x,x_i)) - x_i)\frac{d(x,x_i)}{\sin(d(x,x_i))}.
\end{aligned}$$

□

Problem 2 Score: ________. Product of spheres
Let $\mathcal{M} = \mathbb{S}^{m-1} \times \mathbb{S}^{n-1}$ which is an embedded submanifold of $\mathcal{E} = \mathbb{R}^m \times \mathbb{R}^n$. Endow $\mathbb{R}^m \times \mathbb{R}^n \cong \mathbb{R}^{m+n}$ with its usual Euclidean structure:

$$\langle (u,v),(u,v)\rangle = \begin{pmatrix} u^\top & v^\top \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = u^\top u + v^\top v$$

for $(u,v) \in \mathbb{R}^m \times \mathbb{R}^n$. We can turn $\mathcal{M}$ into a Riemannian manifold by using the Euclidean structure of the ambient space $\mathcal{E} = \mathbb{R}^m \times \mathbb{R}^n$. Let $M \in \mathbb{R}^{m\times n}$. Maximizers of the following function are closely related to the singular value decomposition:

$$f : \mathcal{M} \to \mathbb{R}, \quad f(x,y) = x^\top M y.$$

(1) Show that $f : \mathcal{M} \to \mathbb{R}$ is smooth.

(2) Give a formula for orthogonal projection from $\mathcal{E}$ onto the tangent space $T_{(x,y)}\mathcal{M}$.

(3) Given $(x,y) \in \mathcal{M}$, give an expression for the Riemannian gradient of f at (x,y).

Solution: (1) Clearly, the function $\bar{f} : \mathbb{R}^{m+n} \to \mathbb{R}, \bar{f}(x,y) = x^\top M y$ is smooth, which can be a smooth extension of f. Therefore, f is smooth.

(2) Given $(x,y) \in \mathcal{M}$, we have the tangent space

$$\begin{aligned} T_{(x,y)}\mathcal{M} &= T_{(x,y)}(\mathbb{S}^{m-1} \times \mathbb{S}^{n-1}) \\ &= T_x\mathbb{S}^{m-1} \times T_y\mathbb{S}^{n-1} \end{aligned}$$

As the orthogonal projection from $\mathbb{R}^m$ onto $T_x\mathbb{S}^{m-1}$ is $Proj_x(u) = (I - xx^\top)u$ and from $\mathbb{R}^n$ onto $T_y\mathbb{S}^{n-1}$ is $Proj_y(v) = (I - yy^\top)v$, the orthogonal projection from $\mathbb{R}^m \times \mathbb{R}^n$ onto $T_{(x,y)}\mathcal{M}$ is

$$Proj_{(x,y)}(u,v) = ((I - xx^\top)u, (I - yy^\top)v).$$

(3) Given $(x,y) \in \mathcal{M}$, which is a Riemannian submanifold of $\mathbb{R}^m \times \mathbb{R}^n$,

$$\begin{aligned} gradf(x,y) &= Proj_{(x,y)}(grad\bar{f}(x,y)) \\ &= Proj_{(x,y)}(My, M^\top x) \\ &= ((I - xx^\top)My, (I - yy^\top)M^\top x) \end{aligned}$$

□

Problem 3 Score: ________. The product metric
Let $\mathcal{M}, \mathcal{M}'$ be embedded submanifolds of Euclidean spaces $\mathcal{E}, \mathcal{E}'$, respectively.
Turn $\mathcal{M}$ and $\mathcal{M}'$ into Riemannian manifolds by giving them the Riemannian metrics $\langle\cdot,\cdot\rangle^a$ and $\langle\cdot,\cdot\rangle^b$, respectively. We can turn $\mathcal{M} \times \mathcal{M}'$ into a Riemannian manifold by giving it the Riemannian product metric: for all $(u,u'),(v,v')$ in the tangent space $T_{(x,x')}\mathcal{M} \times \mathcal{M}'$,

$$\langle (u,u'),(v,v')\rangle_{(x,x')} := \langle u,v\rangle_x^a + \langle u',v'\rangle_{x'}^b .$$

(1) What are the tangent spaces of $\mathcal{M} \times \mathcal{M}'$? How do they relate to $T_x\mathcal{M}$ and $T_{x'}\mathcal{M}'$?

(2) For a smooth function $f : \mathcal{M} \times \mathcal{M}' \to \mathbb{R}$, show that

$$gradf(x,x') = (grad(x \mapsto f(x,x'))(x), grad(x' \mapsto f(x,x'))(x')),$$

where $x \mapsto f(x,x')$ and $x' \mapsto f(x,x')$ are functions from $\mathcal{M}$ to $\mathbb{R}$ obtained from f by fixing the other argument.

Solution: (1) $T_{(x,x')}\mathcal{M} \times \mathcal{M}' = T_x\mathcal{M} \times T_{x'}\mathcal{M}'$.

(2) Let $(x,u) \in T\mathcal{M}, (x',u') \in T\mathcal{M}'$ and smooth curves:

$$\begin{aligned} c &: \mathbb{R} \to \mathcal{M}, \quad c(0) = x, \quad c'(0) = u, \\ c' &: \mathbb{R} \to \mathcal{M}', \quad c'(0) = x', \quad c'(0) = u'. \end{aligned}$$

Then define two new smooth curves:

$$C : \mathbb{R} \to \mathcal{M} \times \mathcal{M}', \quad C(t) = (c(t), x'),$$

$$C' : \mathbb{R} \to \mathcal{M} \times \mathcal{M}', \quad C'(t) = (x, c'(t)).$$

Then,

$$\begin{aligned}
\langle (u, u'), gradf(x, x') \rangle_{(x,x')} &= Df(x, x')[u, u'] \\
&= Df(x, x')[u, 0] + Df(x, x')[0, u'] \\
&= D(x \mapsto f(x, x'))(x)[u] + D(x' \mapsto f(x, x'))(x')[u'] \\
&= \langle u, grad(x \mapsto f(x, x'))(x) \rangle_x^a + \langle u', grad(x' \mapsto f(x, x'))(x') \rangle_{x'}^b \\
&= \langle (u, u'), (grad(x \mapsto f(x, x')), grad(x' \mapsto f(x, x'))) \rangle_{(x,x')} .
\end{aligned}$$

As the quation above holds for all $(u, u') \in T_{(x,x')}\mathcal{M} \times \mathcal{M}'$, $gradf(x, x') = (grad(x \mapsto f(x, x'))(x), grad(x' \mapsto f(x, x'))(x'))$.

□

Problem 4 Score: ________. Distorted $\mathbb{R}^d$

Let U be an open subset of $\mathcal{E} = \mathbb{R}^d$, and denote the Euclidean inner product on $\mathcal{E}$ by $\langle u, v \rangle_{\mathcal{E}} = u^\top v$. Let $G : U \to \mathbb{R}^{d \times d}$ be a smooth map such that $G(x)$ is symmetric and positive definite for all $x \in U$. Let $\mathcal{M}$ be U equipped with the Riemannian metric $\langle u, v \rangle_{\mathcal{M}} = u^\top G(x)v$.

(1) Show that $\langle \cdot, \cdot \rangle_{\mathcal{M}}$ is a Riemannian metric on U.

(2) Let $f : U \to \mathbb{R}$ be a smooth function. Derive an expression for the Riemannian gradient of $f, grad_{\mathcal{M}} f$, in terms of the Euclidean gradient of $f, grad_{\mathcal{E}} f$.

(3) If $U = \mathbb{R}^d$, argue that $R_x(u) = x + u$ is a retraction on $\mathcal{M}$. Write down Riemannian gradient descent on $\mathcal{M}$ with retractions R. Compare this algorithm to

  (a) gradient descent on $\mathcal{E}$ with a preconditioner

  (b) Newton's method on $\mathcal{E}$

(4) Consider a particular case known as the Poincaré ball model of hyperbolic space. Let $r > 0$ and $U = \{x \in \mathbb{R}^d := \|x\|_{\mathcal{E}} < r\}$, and

$$G(x) = \frac{4r^4}{(r^2 - \|x\|_{\mathcal{E}}^2)^2} I, \quad \forall x \in U.$$

With $f : \mathcal{M} \to \mathbb{R}$ smooth, give an expression for the Riemannian gradient of f, $grad_{\mathcal{M}} f$ in terms of the Euclidean gradient of f, $grad_{\mathcal{E}} f$.

Solution: (1) As U is an open subset of $\mathbb{R}^d$, for $x \in U, T_xU = \mathbb{R}^d$.
Then for $u, v \in T_xU = \mathbb{R}^d$, $\langle u, v \rangle_x = u^\top G(x)v$ is a inner product on T_xU as $G(x)$ is symmetric and positive definite. Furthermore, for smooth vector fields $V_1, V_2 : U \to \mathbb{R}^d$, the map $x \mapsto \langle V_1(x), V_2(x) \rangle_x = V_1(x)^\top G(x)V_2(x)$ is smooth. Therefore, $\langle \cdot, \cdot \rangle_{\mathcal{M}}$ is a Riemannian metric on U.

(2) For $v \in T_xU = \mathbb{R}^d$, we have

$$\begin{aligned}
\langle grad_{\mathcal{M}} f(x), v \rangle_{\mathcal{M}} &= Df(x)[v] \\
&= v^\top G(x) grad_{\mathcal{M}} f(x) \\
&= \langle v, G(x) grad_{\mathcal{M}} f(x) \rangle_{\mathcal{E}} \\
&= \langle v, grad_{\mathcal{E}} f(x) \rangle_{\mathcal{E}} .
\end{aligned}$$

Therefore, $grad_{\mathcal{E}} f(x) = G(x) grad_{\mathcal{M}} f(x)$, i.e., $grad_{\mathcal{M}} f(x) = G(x)^{-1} grad_{\mathcal{E}} f(x)$.

(3) For $x \in \mathbb{R}^d, u \in T_x\mathbb{R}^d = \mathbb{R}^d, R_x(u) = x + u$ is clearly smooth and $R_x(0) = x$.

$$\frac{d}{dt} R_x(tu) \bigg|_{t=0} = \frac{d}{dt}(x + tu) \bigg|_{t=0} = u.$$

Therefore, $R_x(u) = x + u$ is a retraction on $\mathcal{M}$.
The Riemannian gradient descent on $\mathcal{M}$ with retractions R is given by

$$\begin{aligned}
x_{k+1} &= R_{x_k}(-\eta_k grad_{\mathcal{M}} f(x_k)) \\
&= x_k - \eta_k G(x_k)^{-1} grad_{\mathcal{E}} f(x_k),
\end{aligned}$$

which can be interpreted as the gradient descent on $\mathcal{E}$ with a preconditioner $G(x_k)^{-1}$; and by setting $\eta_k = 1$, it is equivalent to Newton's method on $\mathcal{E}$.

(4) From (2), we have

$$\begin{aligned} grad_{\mathcal{M}} f(x) &= G(x)^{-1} grad_{\mathcal{E}} f(x) \\ &= \frac{(r^2 - \|x\|_{\mathcal{E}}^2)^2}{4r^4} grad_{\mathcal{E}} f(x). \end{aligned}$$

□