Optimization on smooth manifolds
MATH-512
Fall, 2024

Exercise 3

Due date:

Name : Vivi
Student ID : 24S153073
Grade : ________

Problem 1 Score: ________. Smooth maps and differentials

(1) let $\mathcal{M}, \mathcal{M}', \mathcal{M}''$ be embedded submanifolds of the linear spaces $\mathcal{E}, \mathcal{E}', \mathcal{E}''$, respectively.
For smooth maps $F : \mathcal{M} \to \mathcal{M}'$ and $G : \mathcal{M}' \to \mathcal{M}''$, show that $G \circ F : \mathcal{M} \to \mathcal{M}''$ is smooth, and the chain rule is satisfied:

$$D(G \circ F)(x) = DG(F(x)) \circ DF(x).$$

(2) Give an example of an embedded submanifold $\mathcal{M}$ in a linear space $\mathcal{E}$ and a smooth function $f : \mathcal{M} \to \mathbb{R}$ for which there does not exist a smooth extension $\tilde{f} : \mathcal{E} \to \mathbb{R}$ smooth on all of $\mathcal{E}$. Aim for an example where f is bounded on $\mathcal{M}$.

Solution: (1) (a) smoothness
Since F and G are smooth, we have two smooth extensions $\bar{F} : U \to \mathcal{E}'$ and $\bar{G} : U' \to \mathcal{E}''$.
Let $\tilde{F} : U \cap \bar{F}^{-1}(U') \to U'$ be the restriction of $\bar{F}$.
Since $\bar{F}$ is smooth, $\bar{F}^{-1}(U')$ is open in U, then $U \cap \bar{F}^{-1}(U')$ is open in U.
Moreover, for $x \in \mathcal{M}$, we have $\bar{F}(x) = F(x) \in \mathcal{M}' \subseteq U'$, i.e., $x \in \bar{F}^{-1}(U')$.
Then $x \in U \cap \bar{F}^{-1}(U')$, i.e., $\mathcal{M} \subseteq U \cap \bar{F}^{-1}(U')$.
So $U \cap \bar{F}^{-1}(U')$ is a neighborhood of $\mathcal{M}$ in $\mathcal{E}$, and $\tilde{F}$ is a smooth extension of F.
Then, $\bar{G} \circ \tilde{F} : U \cap \bar{F}^{-1}(U') \to \mathcal{E}''$ is a smooth extension of $G \circ F$, therefore $G \circ F$ is smooth.

(b) chain rule
Let $(x, v) \in T\mathcal{M}$, then $x \in \mathcal{M}$ and $v \in T_x\mathcal{M}$, and a smooth curve $c : \mathbb{R} \to \mathcal{M}$ with $c(0) = x$ and $c'(0) = v$.
Then $F \circ c : \mathbb{R} \to \mathcal{M}'$ is a smooth curve with $(F \circ c)(0) = F(x)$ and $(F \circ c)'(0) = DF(x)[v]$.
Therefore, we have

$$\begin{aligned}
D(G \circ F)(x)[v] &= D(G \circ F)(c(0))[c'(0)] \\
&= \frac{d}{dt}(G \circ F)(c(t))|_{t=0} \\
&= DG((F \circ c)(0))[(F \circ c)'(0)] \\
&= DG(F(x))[DF(x)[v]] \\
&= DG(F(x)) \circ DF(x)[v].
\end{aligned}$$

(2) $\mathcal{E} = \mathbb{R}, \mathcal{M} = \mathbb{R} \setminus 0, f(x) = \frac{1}{x}$.

□

Problem 2 Score: ________. Submanifolds of submanifolds

Let $\mathcal{M}$ be an embedded submanifold of a linear space $\mathcal{E}$, and $\mathcal{N}$ a subset of $\mathcal{M}$ defined by $\mathcal{N} = g^{-1}(0)$, where $g : \mathcal{M} \to \mathbb{R}^l$ is smooth and $\operatorname{rank}(Dg(x)) = l \geq 1$ for all $x \in \mathcal{N}$.
Show that $\mathcal{N}$ is itself an embedded submanifold of $\mathcal{E}$, of dimension $\dim(\mathcal{M}) - l$, with tangent spaces $T_x\mathcal{N} = \ker(Dg(x)) \subset T_x\mathcal{M}$.

Solution: Assume that $\dim(\mathcal{E}) = d, \dim(\mathcal{M}) = m \leq d$.
For $m = d$, (N) is apparently an embedded submanifold of $\mathcal{E}$, and $\dim(\mathcal{N}) = d - l = m - l$.
For $m < d$, let the local defining function of $\mathcal{M}$ be $f : U \to \mathbb{R}^{d-m}$, where U is a neighborhood of $\mathcal{M}$ in $\mathcal{E}$.
We can build a smooth extension of $\bar{g} : V \to \mathbb{R}^l$ of g where V is another neighborhood of $\mathcal{M}$ in $\mathcal{E}$.
Then we have a local defining funcion $F : U \cap V \to \mathbb{R}^{d-m+l}, F(x) = (f(x), \bar{g}(x))$.
Apparently F is smooth.
Assume $F(x) = 0$, then $f(x) = 0$ and $\bar{g}(x) = 0$, i.e., $x \in \mathcal{M}$ and $x \in \mathcal{N}$; conversely assume $x \in \mathcal{M}$ and $x \in \mathcal{N}$, then $f(x) = 0$ and $\bar{g}(x) = g(x) = 0$, i.e., $F(x) = 0$.
Then focus on the differential of F at x, $DF(x) : \mathcal{E} \to \mathbb{R}^{d-m+l}, DF(x)[v] = (Df(x)[v], D\bar{g}(x)[v])$.

$$\begin{aligned}
\ker Dg(x) &= \{v \in T_x\mathcal{M} : Dg(x)[v] = 0\} \\
&= \{v \in T_x\mathcal{M} : D\bar{g}(x)[v] = 0\} \\
&= \{v \in \mathcal{E} : Df(x)[v] = 0, D\bar{g}(x)[v] = 0\} \\
&= \ker DF(x).
\end{aligned}$$

Since, $\operatorname{rank}(Dg(x)) = l$, we have $\operatorname{rank}(\ker DF(x)) = \operatorname{rank}(\ker Dg(x)) = \dim(\mathcal{M}) - \operatorname{rank}(Dg(x)) = m - l$.
Then, $\operatorname{rank}(DF(x)) = \dim(\mathcal{E}) - \operatorname{rank}(\ker DF(x)) = d - (m - l) = d - m + l$.
Therefore, F is a local defining function of $\mathcal{N}$, and $\mathcal{N}$ is an embedded submanifold of $\mathcal{E}$, of dimension $m - l = \dim(\mathcal{M}) - l$, with tangent spaces $T_x\mathcal{N} = \ker DF(x) = \ker Dg(x) \subset T_x\mathcal{M}$. □

Problem 3 Score: _________. Stereographic projection
For $(x, v) \in T\mathbb{S}^{d-1}$, let $R_x(v)$ denote the point which lies on $\mathbb{S}^{d-1}$ and on the line connecting $x+v$ and $-x$, and which is not $-x$. Show that $(x, v) \mapsto R_x(v)$ is well defined on the whole tangent bundle, and that it is retraction.

Solution: The line connecting $x+v$ and $-x$ can be written as

$$R_x(v) = t(x+v) + (1-t)(-x) = tx + tv + tx - x = (2t-1)x + tv$$

where $t \in \mathbb{R} \setminus 0$.
Since $x \in \mathbb{S}^{d-1}$, we have $x \cdot x = 1$ and $x \cdot v = 0$.Then, we have

$$\begin{aligned} R_x(v) \cdot R_x(v) &= (2t-1)^2 x \cdot x + t^2 v \cdot v + 2t(2t-1) x \cdot v \\ &= (2t-1)^2 + t^2 \|v\|^2 \\ &= (4 + \|v\|^2)t^2 - 4t + 1 = 1. \end{aligned}$$

We get $t = \frac{4}{4+\|v\|^2}$, then $R_x(v) = \frac{4(2x+v)}{4+\|v\|^2} - x$, which is smooth.
For $c(t) = R_x(tv)$, we have

$$\begin{aligned} c(0) &= \frac{4(2x)}{4} - x = x. \\ c'(0) &= \frac{4v}{4} = v. \end{aligned}$$

Therefore, $(x, v) \mapsto R_x(v)$ is well defined on the whole tangent bundle, and it is retraction. □

Problem 4 Score: _________. QR retraction for small Stiefel
We've showed that

$$\mathcal{M} = \{X = (x, y) \in \mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{d \times 2} : x^\top x = 1, y^\top y = 1, x^\top y = 0\}$$

is an embedded submanifold of $\mathcal{E} = \mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{d \times 2}$.

(1) Show that for all $(X, V) \in T\mathcal{M}$, there is a unique way to write $X + V = QR$ where $Q \in \mathcal{M}$ and R is upper triangular with positive diagonal entries. Then define $\mathcal{R} : T\mathcal{M} \to \mathcal{M}$ by $\mathcal{R}_X(V) = Q$. Hint: When is the QR decomposition unique for a matrix $A \in \mathbb{R}^{d \times m}$?

(2) Derive an explicit formula for $R_X(V)$, and use it to show that $\mathcal{R}$ is a retraction for $\mathcal{M}$.

(3) For $X \in \mathcal{M}$, is $\mathcal{R}_X : T_X\mathcal{M} \to \mathcal{M}$ surjective?

Solution: (1) For $(X, V) \in T\mathcal{M}$, $(X+V)^\top(X+V) = I + V^\top V \succ 0$, i.e., $X + V$ has full collumn rank.
Then, we have the QR decomposition $X + V = QR$ where $Q \in \mathcal{M}$ and R is upper triangular with positive diagonal entries is unique.

(2) For $X = (x_1, x_2), V = (v_1, v_2)$, we have $X + V = (x_1 + v_1, x_2 + v_2) = QR$.
We can apply the Gram-Schmidt process to $x_1 + v_1, x_2 + v_2$ to get $Q = (q_1, q_2)$:

$$\begin{aligned} q_1 &= \frac{x_1 + v_1}{\|x_1 + v_1\|}, \\ q_2 &= \frac{x_2 + v_2 - (x_2 + v_2) \cdot q_1}{\|x_2 + v_2 - (x_2 + v_2) \cdot q_1\|}. \end{aligned}$$

To show that $\mathcal{R}$ is a retraction for $\mathcal{M}$, we need to show, for the curve $c(t) = R_X(tV)$, that $c(0) = X$ and $c'(0) = V$. That is to say, for two curves $q_1(t) = \frac{x_1 + tv_1}{\|x_1 + tv_1\|}$ and $q_2(t) = \frac{x_2 + tv_2 - (x_2 + tv_2) \cdot q_1(t)}{\|x_2 + tv_2 - (x_2 + tv_2) \cdot q_1(t)\|}$, $q_1(0) = x_1, q_2(0) = x_2$ and $q_1'(0) = v_1, q_2'(0) = v_2$.

(3) For $X \in \mathcal{M}$, $\mathcal{R}_X : T_X\mathcal{M} \to \mathcal{M}$ is not surjective.

□

Problem 5 Score: _________. Metric projection retraction for Stiefel
For $p \leq n$, consider the Stiefel

$$\mathcal{M} = St(n, p) = \{X \in \mathbb{R}^{n \times p} : X^\top X = I_p\}.$$

(1) Show that $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^{n\times p}$. As usual, we endow $\mathbb{R}^{n\times p}$ with the inner product $\langle X,Y\rangle = \mathrm{Tr}(X^\top Y)$. What is the dimension of $\mathcal{M}$? What are the tangent spaces $T_X\mathcal{M}$?

(2) For $(X,V)\in T\mathcal{M}$, let $U\Sigma W^\top$ be a thin SVD of $X+V$ (i.e., $U\in\mathcal{M}$, $W\in O(p)$ and $\Sigma\in\mathbb{R}^{p\times p}$ is diagonal with positive entries). Show that $UW^\top$ is the unique metric projection of $X+V$ to $\mathcal{M}$, i.e., $Y=UW^\top$ is the unique solution of

$$\min_{Y\in\mathcal{M}} \|X+V-Y\|^2.$$

For $(X,V)\in T\mathcal{M}$, define $\mathcal{R}_X(V)=UW^\top$.

(3) Show that

$$R_X(V)=(X+V)(I_p+V^\top V)^{-1/2}.$$

(4) Show that R is a retraction for $\mathcal{M}$, which is known as the polar retraction.

(5) Is $\mathcal{R}_X : T_X\mathcal{M}\to\mathcal{M}$ surjective?

Solution: (1) Define a map

$$h:\mathbb{R}^{n\times p}\to Sym(p), h(X)=X^\top X-I_p,$$

where $Sym(p):=\{A\in\mathbb{R}^{p\times p}: A=A^\top\}$.
As h is clearly smooth and $\mathcal{M}=h^{-1}(0)$, we just need to show that $Dh(X)$ has full rank for all $X\in\mathcal{M}$.

$$Dh(X)(V)=\frac{d}{dt}h(X+tV)|_{t=0}=V^\top X+X^\top V.$$

For $W\in Sym(p), V=\frac{1}{2}XW$, we have

$$Dh(X)(V)=\frac{1}{2}W^\top X^\top X+\frac{1}{2}X^\top XW=W,$$

i.e., $Dh(X)$ has full rank for all $X\in\mathcal{M}$.
Thus, $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^{n\times p}$, and $\dim(\mathcal{M})=\dim(\mathbb{R}^{n\times p})-\dim(Sym(p))=np-\frac{p(p+1)}{2}=p(n-\frac{p+1}{2})$.
Lastly, the tangent space $T_X\mathcal{M}$ is the kernel of $Dh(X)$, i.e., $T_X\mathcal{M}=\{V\in\mathbb{R}^{n\times p}: V^\top X+X^\top V=0\}$.

(2) As the map $Y\to YW, W\in O(p)$from $St(n,p)$ to $St(n,p)$ is bijective,

$$\begin{aligned}
\min_{Y\in\mathcal{M}}\|X+V-Y\|^2 &= \min_{Y\in\mathcal{M}}\left\|U\Sigma W^\top-Y\right\|^2\\
&=\min_{Y\in\mathcal{M}}\|U\Sigma-YW\|^2\\
&=\min_{Z\in\mathcal{M}}\|U\Sigma-Z\|^2\\
&=\min_{Z\in\mathcal{M}}\left(\sum_{i=1}^p\|\sigma_i u_i-z_i\|^2\right)\\
&=\min_{Z\in\mathcal{M}}\left(\sum_{i=1}^p\sigma_i^2-2\sigma_i\langle u_i,z_i\rangle+1\right)\\
&\geq\sum_{i=1}^p(\sigma_i^2-2\sigma_i+1)
\end{aligned}$$

where the equality holds when $Z=YW=U$, i.e., $Y=UW^\top$

(3) For $V\in T_X\mathcal{M}$,

$$\begin{aligned}
(I_p+V^\top V)^{-1/2}&=((X+V)^\top(X+V))^{-1/2}\\
&=(W\Sigma U^\top U\Sigma W^\top)^{-1/2}\\
&=(W\Sigma^2W^\top)^{-1/2}\\
&=W\Sigma^{-1}W^\top
\end{aligned}$$

Then, $(X+V)(I_p+V^\top V)^{-1/2}=U\Sigma W^\top W\Sigma^{-1}W^\top=UW^\top=R_X(V)$.

(4) Define a curve $c(t) = R_X(tV)$, then obviously, $c(0) = X$. We then show that $c'(0) = V$.

$$c'(0) = \frac{d}{dt} R_X(tV)|_{t=0} = \frac{d}{dt}(X + tV)(I_p + tV^\top V)^{-1/2}|_{t=0} = V.$$

Therefore, R is a retraction for $\mathcal{M}$.

(5) For $X \in \mathcal{M}$, $\mathcal{R}_X : T_X\mathcal{M} \to \mathcal{M}$ is not surjective.

□

Problem 6 Score: _________. Exponential map on rotations
Let $\mathcal{M} = SO(n) = \{X \in \mathbb{R}^{n\times n} : X^\top X = I, \det(X) = 1\}$ be the special orthogonal group. The matrix exponential map $exp : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is the smooth function defined by

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \text{where} A^0 = I.$$

(1) Show that $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^{n\times n}$. What is the dimension of $\mathcal{M}$? What are the tangent spaces $T_X\mathcal{M}$?

(2) Let $\Omega \in \mathbb{R}^{n\times n}$ be skew-symmetric, i.e., $\Omega^\top = -\Omega$. Show that $\exp(\Omega) \in SO(n)$.

(3) Let $\Omega \in \mathbb{R}^{n\times n}$ be skew-symmetric. Show that $\frac{d}{dt}[exp(t\Omega)]|_{t=0} = \Omega$.

(4) Define $R_X(V) = X\exp(X^\top V)$. Show that $R_X(V) \in \mathcal{M}$ for all $(X, V) \in T\mathcal{M}$.

(5) Show that $R : T\mathcal{M} \to \mathcal{M}$ is a retraction.

(6) For, $X \in \mathcal{M}$, is $\mathcal{R}_X : T_X\mathcal{M} \to \mathcal{M}$ injective?

Solution: (1) Observe that $SO(n) = St(n, n) \cap (\det^{-1}(\{-1\}))^c$, where $St(n, n)$ is the Stiefel manifold.
Then, $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^{n\times n}$, because $(\det^{-1}(\{-1\}))^c$ is open.
Thus, $\dim(\mathcal{M}) = \dim(St(n, n)) = n(n - \frac{n+1}{2}) = \frac{n(n-1)}{2}$.
Moreover, $\mathcal{M}$ has the same tangent spaces as $St(n, n)$, i.e., $T_X\mathcal{M} = \{V \in \mathbb{R}^{n\times n} : V^\top X + X^\top V = 0\}$.

(2) For $\Omega \in \mathbb{R}^{n\times n}$, we have

$$\begin{aligned}
\exp(\Omega)^\top \exp(\Omega) &= \exp(\Omega^\top)\exp(\Omega) \\
&= \exp(-\Omega)\exp(\Omega) \\
&= \exp(\Omega - \Omega) \\
&= \exp(0) = I.
\end{aligned}$$

And

$$\det(\exp(\Omega)) = \exp(\mathrm{Tr}(\Omega)) = 1.$$

Therefore, $\exp(\Omega) \in SO(n)$.

(3)

$$\exp(t\Omega) = \sum_{k=0}^{\infty} \frac{(t\Omega)^k}{k!} = I + t\Omega + \frac{t^2\Omega^2}{2} + \cdots .$$

Then,

$$\begin{aligned}
\frac{d}{dt}[\exp(t\Omega)]|_{t=0} &= \sum_{k=1}^{\infty} \frac{d}{dt}[\frac{(t\Omega)^k}{k!}]|_{t=0} \\
&= \sum_{k=1}^{\infty} \frac{\Omega^k t^{k-1}}{(k-1)!}|_{t=0} \\
&= \Omega.
\end{aligned}$$

(4) For $(X, V) \in T\mathcal{M}$, we have $V^\top X + X^\top V = 0$, i.e., $X^\top V$ is skew-symmetric.
Then, $\exp(X^\top V) \in SO(n)$, and for $R_X(V) = X \exp(X^\top V)$:

$$\begin{aligned} R_X(V)^\top R_X(V) &= \exp(X^\top V)^\top X^\top X \exp(X^\top V) \\ &= \exp(-X^\top V) \exp(X^\top V) \\ &= \exp(X^\top V - X^\top V) \\ &= \exp(0) = I. \end{aligned}$$

And $\det(R_X(V)) = \det(X) \det(\exp(X^\top V)) = 1$.
Therefore, $R_X(V) \in \mathcal{M}$ for all $(X, V) \in T\mathcal{M}$.

(5) Define a curve $c(t) = R_X(tV)$, then obviously, $c(0) = X$. We then show that $c'(0) = V$.

$$\begin{aligned} c'(0) &= \frac{d}{dt} R_X(tV)|_{t=0} \\ &= \frac{d}{dt} X \exp(tX^\top V)|_{t=0} \\ &= X \frac{d}{dt} \exp(tX^\top V)|_{t=0} \\ &= XX^\top V = V. \end{aligned}$$

As $R_X(V)$ is smooth, R is a retraction for $\mathcal{M}$.

(6) For $X \in \mathcal{M}$, $\mathcal{R}_X : T_X\mathcal{M} \to \mathcal{M}$ is not injective.

□