Exercise 2

Due date:

Name : Vivi
Student ID : 24S153073
Grade : ________

Problem 1 Score: ________. Small Stiefel Manifold
The Stiefel manifold is the set of matrices of size $n \times p$ whose columns are orthonormal. In this introductory exercise, we suggest that you work out the fact that pairs of orthonormal vectors indeed form a manifold.

(1) Show that the set
$$\mathcal{M} = \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x^\top x = 1, y^\top y = 1, x^\top y = 0\}$$
is an embedded submanifold of $\mathcal{E} = \mathbb{R}^d \times \mathbb{R}^d$.

(2) What are the tangent spaces of $\mathcal{M}$? What is the dimension of $\mathcal{M}$?

Solution: (1) Let $f : \mathcal{E} \to \mathbb{R}^3$ be defined by $f(x, y) = \begin{pmatrix} x^\top x - 1 \\ y^\top y - 1 \\ x^\top y \end{pmatrix} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^3$. Then $\mathcal{M} = f^{-1}(0)$.

(a) f is smooth.
Since f is a polynomial, it is smooth.

(b) f is a submersion.
The Jacobian matrix of f is
$$Df(x, y) = J_f(x, y) = \begin{pmatrix} 2x^\top & 0 \\ 0 & 2y^\top \\ y^\top & x^\top \end{pmatrix} \in \mathbb{R}^{3\times 2d}$$
$\because x^\top x = 1$ and $y^\top y = 1$
$\therefore x \neq 0$ and $y \neq 0$
$\therefore Df(x, y)$ has full rank.

$\therefore \mathcal{M}$ is an embedded submanifold of $\mathcal{E}$.

(2) the tangent spaces of $\mathcal{M}$ at $x \in \mathcal{M}$ is
$$T_x\mathcal{M} = \ker Df(x, y) = \{(v, w) \in \mathbb{R}^d \times \mathbb{R}^d : x^\top v = 0, y^\top w = 0, x^\top w + y^\top v = 0\}$$
$\therefore \dim \mathcal{M} = \dim \mathbb{R}^d \times \mathbb{R}^d - \operatorname{rank} Df(x, y) = 2d - 3$

□

Problem 2 Score: ________. Rank-1 matrices as a manifold
Let $\mathcal{M} = \mathbb{R}_r^{m\times n}$ be the set of real $m \times n$ matrices of rank r. You will show that this set is an embedded submanifold of $\mathbb{R}^{m\times n}$ when $r = 1$.

(1) Show that $\mathcal{M} = \mathbb{R}_1^{m\times n}$ is an embedded submanifold of $\mathcal{E} = \mathbb{R}^{m\times n}$.

(2) What is the dimension of $\mathcal{M} = \mathbb{R}_1^{m\times n}$?

(3) For $X \in \mathcal{M} = \mathbb{R}_1^{m\times n}$, write $X = \sigma uv^\top$, where $\sigma > 0$ and $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ with $\|u\| = \|v\| = 1$. Show that
$$T_X\mathcal{M} = \{auv^\top + wv^\top + uz^\top : a \in \mathbb{R}, w \in \mathbb{R}^m, z \in \mathbb{R}^n, u^\top w = 0, v^\top z = 0\}.$$

Solution: (1) We consider the set
$$\mathcal{M} = \mathbb{R}_1^{m\times n} = \left\{X \in \mathbb{R}^{m\times n} : \operatorname{rank} X = 1\right\}$$
Since $\operatorname{rank} X = 1$, X must have a nonzero entry, sayX_{ij}.
First,we define the neighborhoods of X
$$U_{ij} = \left\{X \in \mathbb{R}^{m\times n} : X_{ij} \neq 0\right\}$$
where $i = 1, \cdots, m$ and $j = 1, \cdots, n$.
For $\forall X \in U_{ij}$, we have a neighborhood V_{ij} of X such that
$$\left\{V_{ij} \in \mathbb{R}^{m\times n} : \|X - V_{ij}\| \lneq |X_{ij}|/2\right\} \subseteq U_{ij}.$$

So, U_{ij} is open, therefore can be the neighborhood of X.
Then, we need to build the local defining functions $h_{ij} : U_{ij} \to \mathbb{R}^{(m-1)\times(n-1)}$.
First, we try to build h_{11}. Let $X \in U_{11}$, i.e., $X_{11} \neq 0$.X then can be written in block form as

$$X = \begin{pmatrix} X_{11} \in \mathbb{R} & X_{12} \in \mathbb{R}^{1\times(n-1)} \\ X_{21} \in \mathbb{R}^{(m-1)} \times 1 & X_{22} \in \mathbb{R}^{(m-1)\times(n-1)} \end{pmatrix}.$$

Since x has rank 1, each collumn of $\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}$ is a scalar multiple of the first column $\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}$.
Thus, $\exists w \in \mathbb{R}^{n-1}$ such that

$$\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix} = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} w^\top = \begin{pmatrix} X_{11} w^\top \\ X_{21} w^\top \end{pmatrix}.$$

Considering that $X_{11} \neq 0$, we can get from the first row of the above equation that $w^\top = X_{11}^{-1} X_{12}$, and from the second row that $X_{22} = X_{21} w^\top = X_{21} X_{11}^{-1} X_{12}$.
This shows that for $X \in \mathcal{M} \cap U_{11}$, $X_{22} - X_{21} X_{11}^{-1} X_{12} = 0$, which gives us a local defining function

$$h_{11} : U_{11} \to \mathbb{R}^{(m-1)\times(n-1)}, \quad h_{11}(X) = X_{22} - X_{21} X_{11}^{-1} X_{12}.$$

Apparently, h_{11} is smooth because it is a componentwise polynomial.
Moreover, the equation above shows that $\mathcal{M} \cap U_{11} \subseteq h_{11}^{-1}(0)$. To get the reverse inclusion, we need to find $h_{11}^{-1}(0)$.For $h(X) = 0$, we have

$$\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix} = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} X_{11}^{-1} X_{12}.$$

Therefore

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} \begin{pmatrix} 1 & X_{11}^{-1} X_{12} \end{pmatrix}.$$

So, $\operatorname{rank} X \leq 1$ and $\operatorname{rank} X \geq 1$ due to the fact that $X_{11} \neq 0$, which implies that $\operatorname{rank} X = 1$, i.e., $X \in \mathcal{M} \cap U_{11}$.
Therefore, $\mathcal{M} \cap U_{11} = h_{11}^{-1}(0)$.
Then, we need to show that $\operatorname{rank} Dh_{11} = (m-1)(n-1)$, by definition we have

$$Dh_{11}(X)[V] = \frac{d}{dt} \left[h_{11}(X + tV) \right] |_{t=0} = V_{22} - X_{11}^{-1} V_{21} X_{12} - X_{11}^{-1} X_{21} V_{12} + X_{11}^{-2} V_{11} X_{21} X_{12}.$$

We can set $V_{11} = 0$ and $V_{12} = 0$ and $V_{21} = 0$, $Dh_{11}(X)[V] = V_{22}$, which can be any $(m-1) \times (n-1)$ matrix.
So $Dh_{11}(X)$ is a surjection for $\forall X$, i.e., $\operatorname{rank} Dh_{11} = (m-1)(n-1)$.
Therefore h_{11} is indeed a local defining function on U_{11}.
Next, for U_{ij}, we can define h_{ij} in a similar way as h_{11}.
For $\forall X \in U_{ij}$, $P_i X Q_j \in U_{11}$, where P_i is $m \times m$ permutation matrix exchanging the i-th and the first row of X, and Q_j is $n \times n$ permutation matrix exchanging the j-th and the first column of X.
Then we can define h_{ij} as

$$h_{ij}(X) = h_{11}(P_i X Q_j).$$

By chain rule, we have

$$Dh_{ij}(X)[V] = Dh_{11}(P_i X Q_j)[P_i V Q_j].$$

As the map $\mathcal{M} \cap U_{ij} \to \mathcal{M} \cap U_{11}, X \to P_i X Q_j$ is

(a) smooth, so h_{ij} is smooth as a composition of two smooth maps;
(b) bijective, so $h_{ij}^{-1}(0) = \mathcal{M} \cap U_{ij}$ and $\operatorname{rank} Dh_{ij} = (m-1)(n-1)$.

Thus, h_{ij} is a local defining function on U_{ij}.
Finally, we can show that $\mathcal{M}$ is an embedded submanifold of $\mathcal{E} = \mathbb{R}^{m\times n}$ under the atlas $\{(U_{ij}, h_{ij})\}$.

(2) Since h_{ij} maps U_{ij} to $\mathbb{R}^{(m-1)\times(n-1)}$

$$\dim \mathcal{M} = \dim \mathbb{R}^{m\times n} - \dim \mathbb{R}^{(m-1)\times(n-1)} = mn - (m-1)(n-1) = m + n - 1.$$

(3) Since X has rank 1, using the SVD decomposition, X can be written as $X = \sigma u v^\top$, where $\sigma > 0$ and $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ with $\|u\| = \|v\| = 1$.
We define 3 smooth curves:

$$\sigma : \mathbb{R} \to \mathbb{R}, \sigma(0) = \sigma, \sigma'(0) = a$$

$$u : \mathbb{R} \to \mathbb{S}^{m-1}, u(0) = u, u'(0) = w/\sigma$$
$$v : \mathbb{R} \to \mathbb{S}^{n-1}, v(0) = v, v'(0) = z/\sigma$$

Then we have a smooth curve on $\mathcal{M}$:
$$c = \sigma u v^\top, c(0) = X.$$

Differentiating c at 0, we have

$$\begin{aligned} c'(0) &= \sigma'(0)u(0)v^\top(0) + \sigma(0)u'(0)v^\top(0) + \sigma(0)u(0)v'^\top(0) \\ &= auv^\top + wv^\top + uz^\top. \end{aligned}$$

As $u \in \mathbb{S}^{m-1}$ and $v \in \mathbb{S}^{n-1}$, we have $u^\top w = 0$ and $v^\top z = 0$.
Therefore,
$$\{auv^\top + wv^\top + uz^\top : a \in \mathbb{R}, w \in \mathbb{R}^m, z \in \mathbb{R}^n, u^\top w = 0, v^\top z = 0\} \subseteq T_X\mathcal{M}.$$

On the other hand,
$$\{auv^\top + wv^\top + uz^\top : a \in \mathbb{R}, w \in \mathbb{R}^m, z \in \mathbb{R}^n, u^\top w = 0, v^\top z = 0\}$$

has dimension $m + n - 1 = \dim \mathcal{M}$.
In conclusion, $T_X\mathcal{M} = \{auv^\top + wv^\top + uz^\top : a \in \mathbb{R}, w \in \mathbb{R}^m, z \in \mathbb{R}^n, u^\top w = 0, v^\top z = 0\}$.

□

Problem 3 Score: ________. Product manifolds
Let $\mathcal{M}$ be an embedded submanifold of a linear space $\mathcal{E}$. Likewise, let $\mathcal{M}'$ be an embedded submanifold of a (possibly different) linear space $\mathcal{E}'$.

(1) Show that $\mathcal{M} \times \mathcal{M}'$ is an embedded submanifold of $\mathcal{E} \times \mathcal{E}'$.

(2) How are the tangent spaces of the product manifold related to the tangent spaces of the base manifolds?

Solution: (1) Suppose that $\dim \mathcal{M} = m, \dim \mathcal{E} = p, m \le p$ and $\dim \mathcal{M}' = n, \dim \mathcal{E}' = q, n \le q$.

(a) $m = p$ and $n = q$, i.e., $\mathcal{M}$ and $\mathcal{M}'$ are open in $\mathcal{E}$ and $\mathcal{E}'$.
Clearly, $\mathcal{M} \times \mathcal{M}'$ is open in $\mathcal{E} \times \mathcal{E}'$.

(b) $m < p$ and $n = q$, i.e., $\mathcal{M}'$ is open in $\mathcal{E}'$ and there exists a local defining function $h : U \to \mathbb{R}^{p-m}$ on $\mathcal{M}$.
Thus we can define a local defining function

$$H : U \times \mathcal{M}' \to \mathbb{R}^{p-m}, H(x, x') = h(x)$$

which is smooth and satisfies $H^{-1}(0) = \mathcal{M} \times \mathcal{M}'$, $\operatorname{rank} DH(x, x') = \operatorname{rank} Dh(x) = p - m = p + q - (m + n)$.

(c) $m < p$ and $n < q$, i.e., there exists two local defining functions $h : U \to \mathbb{R}^{p-m}$ on $\mathcal{M}$ and $h' : U' \to \mathbb{R}^{q-n}$ on $\mathcal{M}'$.
Then we can define a local defining function

$$H : U \times U' \to \mathbb{R}^{p-m} \times \mathbb{R}^{q-n}, H(x, x') = (h(x), h'(x'))$$

which is clearly smooth, and

$$H^{-1}(0) = \{(x, x') \in \mathcal{M} \times \mathcal{M}' : h(x) = 0, h'(x') = 0\} = (U \cap \mathcal{M}) \times (U' \cap \mathcal{M}') = (U \times U') \cap (\mathcal{M} \times \mathcal{M}')$$

and $\operatorname{rank} DH(x, x') = \operatorname{rank} Dh(x) + \operatorname{rank} Dh'(x') = p - m + q - n = p + q - (m + n)$.

(2) From the definition of the tangent space, we have

$$T_{(x,x')}(\mathcal{M} \times \mathcal{M}') = \ker DH(x, x') = \ker Dh(x) \times \ker Dh'(x') = T_x\mathcal{M} \times T_{x'}\mathcal{M}'.$$

□

Problem 4 Score: ________. The cross is not a manifold
Show that the cross $\mathcal{X} = \{x \in \mathbb{R}^2 : x_1^2 = x_2^2\}$ is not an embedded submanifold of $\mathbb{R}^2$.

Solution: Clearly, $\mathcal{X}$ is open in $\mathbb{R}^2$, then $\dim \mathcal{X} \in \{0,1\}$ to ensure $\mathcal{X}$ is an embedded manifold.
Focusing on the point $x = (0,0) \in \mathcal{X}$, we can have two smooth curves

$$\gamma_1 : \mathbb{R} \to \mathcal{X}, \gamma_1(t) = (t,t)$$
$$\gamma_2 : \mathbb{R} \to \mathcal{X}, \gamma_2(t) = (t,-t)$$

which satisfy that $\gamma(0) = (0,0)$. Therefore $\gamma_1'(0) = (1,1)$ and $\gamma_2'(0) = (1,-1)$ are two linearly independent vectors in $T_{(0,0)}\mathcal{X}$.
Thus, $\dim T_{(0,0)}\mathcal{X} \geq 2$, which contradicts the fact that $\dim \mathcal{X} \in \{0,1\}$.
Therefore, $\mathcal{X}$ is not an embedded submanifold of $\mathbb{R}^2$. □

Problem 5 Score: _________. Differentiating the matrix inversion
We define $U \subseteq \mathbb{R}^{n\times n}$ by $U := \det^{-1}(\mathbb{R} \setminus \{0\})$, i.e., U is the set of invertible matrices. As $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous, we have that U is open.
Let $F : U \to U$ be given by $A \mapsto A^{-1}$, i.e., F is the matrix inversion. The map F is smooth, this can be seen by the cofactor formula, which shows that every entry of $F(A)$ is a rational function.
Compute the differential of F.

Solution: (1) For $A \in U, H \in \mathbb{R}^{n\times n}$, we have

$$\begin{aligned}
DF(A)[H] &= \frac{d}{dt}\left[F(A+tH)\right]|_{t=0} \\
&= \frac{d}{dt}\left[(A+tH)^{-1}\right]|_{t=0} \\
&= \lim_{t\to 0} \frac{(A+tH)^{-1} - A^{-1}}{t} \\
&= \lim_{t\to 0} \frac{A^{-1} + A^{-1}(t^2 HA^{-1}H - tH)A^{-1} - A^{-1}}{t} \\
&= \lim_{t\to 0} (tA^{-1}HA^{-1}HA^{-1} - A^{-1}HA^{-1}) \\
&= -A^{-1}HA^{-1}.
\end{aligned}$$

(2) Define $A(t) = A + tH, A(t) \in U$ for t small enough.
Then we have
$$F \circ A(t) \cdot A(t) = A^{-1}(t) \cdot A(t) = I.$$
Differentiating on both sides, we have
$$\begin{aligned}
0 &= \frac{d}{dt}\left[F \circ A(t) \cdot A(t)\right] \\
&= \frac{d}{dt}\left[F \circ A(t)\right] \cdot A(t) + F \circ A(t) \cdot \frac{d}{dt}\left[A(t)\right] \\
&= DF(A(t))[A'(t)] \cdot A(t) + F(A(t)) \cdot H \\
&= DF(A(t))[H] \cdot A(t) + F(A(t)) \cdot H
\end{aligned}$$
In particular, for $t = 0$, we have
$$DF(A)[H] \cdot A + A^{-1} \cdot H = 0.$$
Therefore, $DF(A)[H] = -A^{-1}HA^{-1}$.

□