An Introduction to Grassmann Manifolds and their Matrix Representation

Daniel Karrasch*†
Technische Universität München
Center for Mathematics, M3
Boltzmannstr. 3
85748 Garching bei München, Germany

September 26, 2017

In the following we give a self-contained introduction to Grassmann manifolds and their representation by matrix manifolds. The presentation is based on [6, 1] and all results are taken from these references if not stated otherwise. Familiarity with calculus and elementary notions of differential geometry are assumed, for a textbook reference see, for instance, [12].

Throughout, let $n \in \mathbb{N}_{>0}$ and $k \in \{1, \dots, n\}$. The object of investigation is the set of all k-dimensional linear subspaces $V \subseteq \mathbb{R}^n$, i.e.

$$\operatorname{Gr}(k, \mathbb{R}^n) := \{V \subseteq \mathbb{R}^n;\ V \text{ is a linear subspace},\ \dim V = k\}.$$

The objective is to establish $\operatorname{Gr}(k, \mathbb{R}^n)$ as a Riemannian (differentiable) manifold.

1 Topological & Metric Structure

In this section, we follow [6]. We consider $A \in \mathbb{R}^{n \times k}$ as

$$A = (a_{ij})_{i \in \{1,\dots,n\}, j \in \{1,\dots,k\}} = \begin{pmatrix} a_1 & \cdots & a_k \end{pmatrix},$$

where $a_i = (a_{ji})_{j \in \{1,\dots,n\}} \in \mathbb{R}^n$, $i \in \{1, \dots, k\}$. We endow $\mathbb{R}^{n \times k}$ with the norm induced by the Frobenius inner product, i.e.

$$\|A\| = \sqrt{\operatorname{tr}(A^\top A)} = \left|\begin{pmatrix} |a_1|_2 & \cdots & |a_k|_2 \end{pmatrix}^\top\right|_2 = \sqrt{\sum_{i,j} a_{ij}^2}. \tag{1}$$

*These notes were written while I was doing my PhD at Fachrichtung Mathematik, Technische Universität Dresden, 01062 Dresden, Germany. Funding by ESF grant “Landesinnovationspromotion” No. 080942988 is gratefully acknowledged.
†E-mail: karrasch@ma.tum.de

We define the *(non-compact) Stiefel manifold* as

$$\begin{aligned}\mathrm{St}(k,\mathbb{R}^n) &:= \left\{A \in \mathbb{R}^{n\times k};\ \ \mathrm{rk}(A) = k\right\}\\ &= \left\{A = \begin{pmatrix}a_1 & \cdots & a_k\end{pmatrix} \in \mathbb{R}^{n\times k};\ a_1,\dots,a_k \text{ are linearly independent}\right\}.\end{aligned}$$

$\mathrm{St}(k,\mathbb{R}^n)$ is often referred to as the set of all *k-frames* (of linearly independent vectors) in $\mathbb{R}^n$. Note that $\mathrm{St}(k,\mathbb{R}^n)$ is an open subset of $\mathbb{R}^{n\times k}$ (as the preimage of the open set $\mathbb{R}\setminus\{0\}$ under the continuous map $A \mapsto \det(A^\top A)$). We define the *compact Stiefel manifold* as

$$\mathrm{St}^*(k,\mathbb{R}^n) := \left\{A \in \mathrm{St}(k,\mathbb{R}^n);\ A^\top A = I_k\right\}.$$

Clearly, $\mathrm{St}^*(k,\mathbb{R}^n)$ is a bounded (by $\sqrt{k}$) and closed (as the preimage of the closed set $\{I_k\}$ under the continuous map $A \mapsto A^\top A$) and hence compact subset of $\mathrm{St}(k,\mathbb{R}^n)$ and $\mathrm{St}^*(k,\mathbb{R}^n) \hookrightarrow \mathrm{St}(k,\mathbb{R}^n)$. It is often referred to as the set of all *orthonormal k-frames* of linearly independent vectors in $\mathbb{R}^n$.

We consider the map

$$\pi\colon\ \mathrm{St}(k,\mathbb{R}^n) \to \mathrm{Gr}(k,\mathbb{R}^n), \qquad A = \begin{pmatrix}a_1 & \cdots & a_k\end{pmatrix} \mapsto \mathrm{span}\left\{a_1,\dots,a_k\right\}, \tag{2}$$

and endow $\mathrm{Gr}(k,\mathbb{R}^n)$ with the final topology with respect to π, which we call the *Grassmann topology*, i.e. $U \subseteq \mathrm{Gr}(k,\mathbb{R}^n)$ is open if and only if $\pi^{-1}[U]$ is open in $\mathrm{St}(k,\mathbb{R}^n)$. Let $f \in \mathcal{L}(\mathbb{R}^k,\mathbb{R}^n)$ be injective and $A \in \mathrm{St}(k,\mathbb{R}^n)$ be the matrix representation with respect to the canonical bases in $\mathbb{R}^k$ and $\mathbb{R}^n$. Then $\pi(A)$ can be equivalently considered as $f[\mathbb{R}^k]$, i.e. the image of f.

1.1 Lemma. *Consider* $\mathrm{Gr}(k,\mathbb{R}^n)$ *with the Grassmann topology.*

(i) The map π *is surjective and for any* $A \in \mathrm{St}(k,\mathbb{R}^n)$ *we have*

$$\pi^{-1}[\pi(A)] = \{AP;\ P \in GL(k,\mathbb{R})\}. \tag{3}$$

(ii) The map π *is continuous and open.*

Proof. Obviously, π is surjective and continuous. Furthermore, $A, B \in \mathrm{St}(k,\mathbb{R}^n)$ span the same subspace, i.e. $\pi(A) = \pi(B)$ if and only if there exists a $P \in GL(k,\mathbb{R})$ such that $B = AP$. To see the sufficiency note that by assumption we have $A = BP$ and $B = AQ = BPQ$ for some $P, Q \in \mathbb{R}^{k\times k}$, from which we directly read off that $P = Q^{-1} \in GL(k,\mathbb{R})$. On the other hand, the necessity is obvious. For an open subset $V \subseteq \mathrm{St}(k,\mathbb{R}^n)$ we have $\pi^{-1}[\pi[V]] = \bigcup_{P\in GL(k,\mathbb{R})}[V]P$, i.e. $\pi^{-1}[\pi[V]]$ is a union of open subsets. Thus, $\pi[V] \subseteq \mathrm{Gr}(k,\mathbb{R}^n)$ is open by the definition of the Grassmann topology. □

Eq. (3) identifies $\mathrm{Gr}(k,\mathbb{R}^n)$ with the quotient space

$$\mathrm{St}(k,\mathbb{R}^n)/GL(k,\mathbb{R}) := \{A[GL(k,\mathbb{R})];\ A \in \mathrm{St}(k,\mathbb{R}^n)\}, \tag{4}$$

which can be endowed with the structure of a quotient manifold; see [2, Section 3.4]. However, we define a topological and differentiable structure on $\mathrm{Gr}(k,\mathbb{R}^n)$ directly.

We consider the restriction of π to $\mathrm{St}^*(k,\mathbb{R}^n)$, i.e. $\overline{\pi} := \pi|_{\mathrm{St}^*(k,\mathbb{R}^n)}$, and the "Gram-Schmidt orthonormalization map", i.e. $GS(A)$ is the matrix obtained by the application of the Gram-Schmidt orthonormalization method to the columns $a_1, \dots, a_n$ of A. It is well-known that the Gram-Schmidt method defines a continuous map. Thus, we have $\overline{\pi} \circ GS = \pi$. Clearly, $\overline{\pi}$ is surjective and continuous.

Next, we show a characterization of the continuity of functions f defined on a topological space endowed with the final topology with respect to some other function and mapping to another topological space.

1.2 Lemma. *Consider* $\mathrm{Gr}(k,\mathbb{R}^n)$ *with the Grassmann topology. Let* Ω *be a topological space and*

$$f\colon\ \mathrm{Gr}(k,\mathbb{R}^n) \to \Omega.$$

Then the following statements are equivalent:

(a) f *is continuous;* *(b)* $f \circ \pi$ *is continuous;* *(c)* $f \circ \overline{\pi}$ *is continuous.*

Proof. (a) $\Leftrightarrow$ (b) follows from the definition of final topology with respect to π.

(b) $\Rightarrow$ (c) is obvious since $\overline{\pi}$ is the restriction of π.

(c) $\Rightarrow$ (b) follows from the representation $f \circ \pi = f \circ \overline{\pi} \circ GS$ and the continuity of GS. □

One can show that the set

$$\Delta := \left\{(A,B) \in \mathrm{St}(k,\mathbb{R}^n)^2;\ \exists\, P \in GL(k,\mathbb{R}) \colon \quad AP = B\right\}$$

is closed in $\mathrm{St}(k,\mathbb{R}^n) \times \mathrm{St}(k,\mathbb{R}^n)$ by realizing that $(A,B) \in \Delta$ if and only if $\mathrm{rk}\begin{pmatrix} A & B \end{pmatrix} = k$. This condition can be characterized by the requirement that all (continuous) $k+1$ minors of $\begin{pmatrix} A & B \end{pmatrix}$ vanish. The set of matrices satisfying this condition is hence closed. Since $\mathrm{Gr}(k,\mathbb{R}^n) = \overline{\pi}[\mathrm{St}^*(k,\mathbb{R}^n)]$, $\overline{\pi}$ is continuous and $\mathrm{St}^*(k,\mathbb{R}^n)$ compact, we have that $\mathrm{Gr}(k,\mathbb{R}^n)$ is compact, too, and it remains to prove that the Grassmann topology is Hausdorff. To this end, let $L, M \in \mathrm{Gr}(k,\mathbb{R}^n)$ with $L \neq M$, $A \in \pi^{-1}[\{L\}]$, $B \in \pi^{-1}[\{M\}]$. By Lemma 1.1 we have $(A,B) \notin \Delta$ and since Δ is closed, there exist neighborhoods U_A and U_B of A and B, respectively, such that $(U_A \times U_B) \cap \Delta = \varnothing$. By Lemma 1.1 $\pi[U_A]$ and $\pi[U_B]$ are disjoint open sets containing $\pi(A) = L$ and $\pi(B) = M$. Summarizing, we obtain the following result.

1.3 Proposition. $\mathrm{Gr}(k,\mathbb{R}^n)$ *is a compact space with respect to the Grassmann topology.*

Often, $\mathrm{Gr}(k,\mathbb{R}^n)$ is endowed and investigated with the so-called *gap metric*, which we introduce next.

1.4 Definition (Gap metric)**.** We define

$$\Theta\colon\ \mathrm{Gr}(k,\mathbb{R}^n) \times \mathrm{Gr}(k,\mathbb{R}^n) \to \mathbb{R}_{\geq 0}, \qquad (L,M) \mapsto \|\Pi_L - \Pi_M\|, \tag{5}$$

where Π_X denotes the orthogonal projection on a subspace X and $\|\cdot\|$ denotes the operator norm. Θ is called the *gap metric* on $\mathrm{Gr}(k,\mathbb{R}^n)$.

In the literature, $\Theta(L, M)$ is also referred to as the *opening* between the subspaces L and M and was originally introduced in [11]. By the Projection Theorem for Hilbert spaces, there is a one-to-one correspondence between (closed) subspaces U and orthogonal projections Π_U with $\operatorname{im}\Pi_U = U$ and $\ker\Pi_U = U^\perp$. Therefore, it is easy to see that Θ is indeed a metric on $\mathrm{Gr}(k, \mathbb{R}^n)$. Next we show some further properties of the gap metric.

1.5 Proposition ([7, Theorems 13.1.1 & 13.1.2])**.** *For $L, M \in \mathrm{Gr}(k, \mathbb{R}^n)$*

(i) one has

$$\Theta(L, M) = \max\left(\sup_{x\in\mathcal{S}\cap L} d(x, M), \sup_{x\in\mathcal{S}\cap M} d(x, L)\right), \tag{6}$$

where $d(x, M) := \inf\{|x - y|\,;\ y \in M\}$;

(ii) $\Theta(L, M) = \Theta(L^\perp, M^\perp) \le 1$*;*

(iii) $\Theta(L, M) < 1 \quad\Leftrightarrow\quad L\cap M^\perp = L^\perp \cap M = \{0\} \quad\Leftrightarrow\quad \mathbb{R}^n = L \oplus M^\perp = L^\perp \oplus M$*.*

Proof. (i) Let $x \in L \cap \mathcal{S}$. Then $\Pi_M x \in M$ and it follows

$$|x - \Pi_M x| = |(\Pi_L - \Pi_M)\, x| \le \|\Pi_L - \Pi_M\|\,.$$

Therefore,

$$\sup_{x\in\mathcal{S}\cap L} d(x, M) \le \|\Pi_L - \Pi_M\|\,.$$

Analogously we have

$$\sup_{x\in\mathcal{S}\cap M} d(x, L) \le \|\Pi_L - \Pi_M\|$$

and we obtain

$$\max\left(\sup_{x\in\mathcal{S}\cap L} d(x, M), \sup_{x\in\mathcal{S}\cap M} d(x, L)\right) \le \|\Pi_L - \Pi_M\| = \Theta(L, M).$$

To derive the complementary inequality, observe that by the Projection Theorem in Hilbert spaces we have

$$\begin{aligned} \rho_L &:= \sup_{x\in\mathcal{S}\cap L} d(x, M) = \sup_{x\in\mathcal{S}\cap L} |x - \Pi_M x| = \sup_{x\in\mathcal{S}\cap L} |(I - \Pi_M)x|\,, \\ \rho_M &:= \sup_{x\in\mathcal{S}\cap M} d(x, L) = \sup_{x\in\mathcal{S}\cap M} |x - \Pi_L x| = \sup_{x\in\mathcal{S}\cap M} |(I - \Pi_L)x|\,. \end{aligned}$$

Consequently, we have for every $x \in \mathbb{R}^n$ that

$$|(I - \Pi_L)\Pi_M x| \le \rho_M\,|\Pi_M x|\,, \qquad \text{and} \qquad |(I - \Pi_M)\Pi_L x| \le \rho_L\,|\Pi_L x|\,. \tag{7}$$

We estimate

$$\begin{aligned} |\Pi_M(I - \Pi_L)x|^2 &= \langle \Pi_M(I - \Pi_L)x, \Pi_M(I - \Pi_L)x\rangle \\ &= \langle \Pi_M(I - \Pi_L)x, (I - \Pi_L)x\rangle \\ &= \langle \Pi_M(I - \Pi_L)x, (I - \Pi_L)(I - \Pi_L)x\rangle \\ &= \langle (I - \Pi_L)\Pi_M(I - \Pi_L)x, (I - \Pi_L)x\rangle \\ &\le |((I - \Pi_L)\Pi_M(I - \Pi_L)x)|\,|(I - \Pi_L)x|\,, \end{aligned}$$

thus by Eq. (7)

$$|\Pi_M(I-\Pi_L)x|^2 \le \rho_M\, |\Pi_M(I-\Pi_L)x|\, |(I-\Pi_L)x|$$

and hence

$$|\Pi_M(I-\Pi_L)x| \le \rho_M\, |(I-\Pi_L)x|\,. \tag{8}$$

On the other hand, using

$$\Pi_M - \Pi_L = \Pi_M(I-\Pi_L) - (I-\Pi_M)\Pi_L$$

and the orthogonality of Π_M, we obtain by Eqs. (8) and (7)

$$\begin{aligned}
|(\Pi_M-\Pi_L)x|^2 &= \langle (\Pi_M(I-\Pi_L)-(I-\Pi_M)\Pi_L)\,x, (\Pi_M(I-\Pi_L)-(I-\Pi_M)\Pi_L)\,x\rangle \\
&= |\Pi_M(I-\Pi_L)x|^2 + |(I-\Pi_M)\Pi_L x|^2 \\
&\le \rho_M^2\, |(I-\Pi_L)x|^2 + \rho_L^2\, |\Pi_L x|^2 \\
&\le \max\left\{\rho_M^2, \rho_L^2\right\} |x|^2\,.
\end{aligned}$$

Thus, we obtain

$$\|\Pi_L - \Pi_M\| \le \max\left\{\rho_M, \rho_L\right\},$$

which finishes the proof of (a).

(ii) The equality $\Theta(L,M) = \Theta(L^\perp, M^\perp)$ follows directly from the definition, since $\Pi_{L^\perp} = I - \Pi_L$. The estimate $\Theta(L,M) \le 1$ holds by the observation that for $x \in \mathcal{S} \cap L$ we have $d(x,M) = |\Pi_{M^\perp} x| = |x| - |\Pi_M x| \le |x| = 1$.

(iii) First note that the second equivalence is clear by the definition of decompositions. The first one follows from the following observation.

1.6 Claim. Let $M \in \mathrm{Gr}(k,\mathbb{R}^n)$ and $x \in \mathcal{S}$. Then $x \in M^\perp$ if and only if $d(x,M) = 1$.

Proof of claim. The proof relies on the relation $|x|^2 = |\Pi_M x|^2 + |\Pi_{M^\perp} x|^2$. From the relation we directly read off that

$$1 = |x|^2 = |\Pi_{M^\perp} x|^2 = |(I-\Pi_M)x|^2 = |x - \Pi_M x|^2 = d(x,M).$$

Conversely, from $d(x,M) = |\Pi_{M^\perp} x|^2 = |x|^2 = 1$ follows $|\Pi_M x| = 0$ and hence $x \in M^\perp$. □

Due to the continuity of the norm and the compactness of $\mathcal{S} \cap L$ the value of $\Theta(L,M)$ is attained, so from $\Theta(L,M) < 1$ follows $L \cap M^\perp = L^\perp \cap M = \{0\}$. On the other hand, from $\Theta(L,M) = 1$, which we (w.l.o.g.) assume to be attained by $\sup_{x\in\mathcal{S}\cap L} d(x,M) = 1$, follows that there exists some $z \in \mathcal{S} \cap L$ such that $d(z,M) = 1$ and by the claim $z \in M^\perp$ holds. □

As a consequence, we get that for any $L \in \mathrm{Gr}(k,\mathbb{R}^n)$

$$B[L,1] = \mathrm{Gr}(k,\mathbb{R}^n), \qquad \text{and} \qquad B(L,1) = \left\{M \in \mathrm{Gr}(k,\mathbb{R}^n);\ \mathbb{R}^n = L \oplus M^\perp\right\}.$$

Our next aims are to prove that the gap metric on $\mathrm{Gr}(k,\mathbb{R}^n)$ is equivalent to the Hausdorff metric on the sections of two subspaces with the unit sphere (called the spherical gap), and second to show the equivalence of the Grassmann topology to the topology induced by the gap metric/Hausdorff metric.

1.7 Definition (Spherical gap metric, cf. also [10])**.** We define

$$\widetilde{\Theta}\colon \mathrm{Gr}(k,\mathbb{R}^n)\times\mathrm{Gr}(k,\mathbb{R}^n)\to\mathbb{R}_{\geq 0},$$
$$(L,M)\mapsto d_{\mathrm{H}}(\mathcal{S}\cap L,\mathcal{S}\cap M)=\max\left(\sup_{x\in\mathcal{S}\cap L} d(x,\mathcal{S}\cap M),\ \sup_{x\in\mathcal{S}\cap M} d(x,\mathcal{S}\cap L)\right),$$

where d_{H} denotes the Hausdorff metric defined on closed subsets of $\mathbb{R}^n$. $\widetilde{\Theta}$ is called the *spherical gap metric* on $\mathrm{Gr}(k,\mathbb{R}^n)$.

1.8 Lemma ([8])**.** *For any $L,M\in\mathrm{Gr}(k,\mathbb{R}^n)$ holds*

$$\Theta(L,M)\leq\widetilde{\Theta}(L,M)\leq 2\Theta(L,M).$$

Proof. We follow [10, p. 199]. The first inequality follows trivially from $d(x,M)\leq d(x,M\cap\mathcal{S})$, $x\in\mathbb{R}^n$. To show the second inequality it suffices to prove $d(x,M\cap\mathcal{S})\leq 2d(x,M)$ for $x\in\mathcal{S}$. By the definition of infimum, for any $\varepsilon\in\mathbb{R}_{>0}$ there exists $y\in M\setminus\{0\}$ such that $|x-y|<d(x,M)+\varepsilon$. Define $y_0:=\frac{y}{|y|}\in\mathcal{S}\cap M$. Then the estimate $d(x,M\cap\mathcal{S})\leq|x-y_0|\leq|x-y|+|y-y_0|$ holds. Furthermore, since y and y_0 are parallel, we have

$$|y-y_0|=\left|y-\frac{y}{|y|}\right|=||y|-1|=||y|-|x||\leq|y-x|\,.$$

In summary, we have $d(x,M\cap\mathcal{S})<2d(x,M)+2\varepsilon$, and since ε is arbitrary, the prove is finished. ∎

1.9 Corollary. *The topologies on $\mathrm{Gr}(k,\mathbb{R}^n)$ induced by Θ and $\widetilde{\Theta}$, respectively, coincide.*

1.10 Remark. For a sequence $(X_n)_{n\in\mathbb{N}}\in(\mathrm{Gr}(k,\mathbb{R}^n))^{\mathbb{N}}$ converging to $X\in\mathrm{Gr}(k,\mathbb{R}^n)$ with respect to Θ, i.e. $\|\Pi_{X_n}-\Pi_X\|\xrightarrow{n\to\infty}0$, we have by the definition of Hausdorff metric that for any $x\in X$ there exists a sequence $(x_n)_{n\in\mathbb{N}}\in(\mathcal{S})^{\mathbb{N}}$ with $x_n\in X_n$ for each $n\in\mathbb{N}$ such that $x_n\xrightarrow{n\to\infty}x$, and hence with respect to any norm.

1.11 Remark. When Θ is introduced by Eq. (6) in the set of subspaces of general Banach spaces, Θ does not necessarily satisfy the triangle inequality and can therefore not be used to define a metric/topology. However, in Hilbert space one can show equality to the expression given in Eq. (5), which we have used for the definition and which obviously satisfies the conditions for a metric. On the other hand, $\widetilde{\Theta}$ defines a proper metric even in the general Banach space case; see, for instance, [10, IV.2.1] and the references therein.

Next we address our second aim, to show that the Grassmann topology coincides with the topology induced by the gap metric Θ (or equivalently $\widetilde{\Theta}$). To this end, we need to prove that the identity map from the topological space $\mathrm{Gr}(k,\mathbb{R}^n)$ with the Grassmann topology to the topological space $\mathrm{Gr}(k,\mathbb{R}^n)$ with the gap topology is continuous together with its inverse. Since $\mathrm{Gr}(k,\mathbb{R}^n)$ with the Grassmann topology is compact and the metric space $\mathrm{Gr}(k,\mathbb{R}^n)$ with the metric Θ is Hausdorff, the continuity of the inverse follows directly from the continuity of the identity map. Furthermore, by Lemma 1.2 it suffices to show that $\overline{\pi}\colon \mathrm{St}^*(k,\mathbb{R}^n) \to (\mathrm{Gr}(k,\mathbb{R}^n),\Theta)$ is continuous.

1.12 Proposition. *The Grassmann topology in* $\mathrm{Gr}(k,\mathbb{R}^n)$ *coincides with the topologies induced by* Θ *and* $\widetilde{\Theta}$*, respectively.*

Proof. Let $A,B \in \mathrm{St}^*(k,\mathbb{R}^n)$ and $\overline{\pi}(A) = L$, $\overline{\pi}(B) = M$, i.e. $L,M \in \mathrm{Gr}(k,\mathbb{R}^n)$. It is well-known that the orthogonal projections onto L and M have the form $\sum_{i=1}^k \langle a_i,\cdot\rangle a_i$ and $\sum_{i=1}^k \langle b_i,\cdot\rangle b_i$, respectively. Now we estimate

$$\begin{aligned}
\Theta(\overline{\pi}(A),\overline{\pi}(B)) = \Theta(L,M) &= \|\Pi_L - \Pi_M\| \\
&= \left\| \sum_{i=1}^k \langle a_i,\cdot\rangle a_i - \sum_{i=1}^k \langle b_i,\cdot\rangle b_i \right\| \\
&= \left\| \sum_{i=1}^k \langle a_i,\cdot\rangle (a_i - b_i) + \sum_{i=1}^k \langle a_i - b_i,\cdot\rangle b_i \right\| \\
&\le \sum_{i=1}^k |a_i|\,|a_i - b_i| + |a_i - b_i|\,|b_i| \\
&\le 2k \sum_{i=1}^k |a_i - b_i| = 2k \langle (1)_{i\in\{1,\dots,k\}}, (|a_i - b_i|)_{i\in\{1,\dots,k\}} \rangle \\
&\le 2k\sqrt{k}\,\|A - B\|,
\end{aligned}$$

which implies the Lipschitz-continuity of $\overline{\pi}$. □

1.13 Corollary. *The metric space* $(\mathrm{Gr}(k,\mathbb{R}^n),\Theta)$ *is complete.*

2 Differentiable Structure

This section merges ideas of [6, 1]. To introduce the differentiable structure, we make use of *local affine cross sections*. Let $A \in \mathrm{St}(k,\mathbb{R}^n)$ and define

$$\begin{aligned}
S_A &:= \left\{ B \in \mathrm{St}(k,\mathbb{R}^n);\ A^\top (B - A) = 0 \right\} = \left\{ B \in \mathrm{St}(k,\mathbb{R}^n);\ A^\top B = A^\top A \right\} \\
&\subseteq \left\{ B \in \mathrm{St}(k,\mathbb{R}^n);\ \det(A^\top B) \neq 0 \right\} =: T_A \subset \mathrm{St}(k,\mathbb{R}^n),
\end{aligned}$$

orthogonal to the fiber $A[GL(k,\mathbb{R})]$ crossing through A. For $B \in \mathrm{St}(k,\mathbb{R}^n)$ the equivalence class $B[GL(k,\mathbb{R})] = \pi^{-1}[\pi(B)]$ intersects the cross section S_A if and only if $B \in T_A$, and then in $B(A^\top B)^{-1}A^\top A$. This can be seen by plugging BP with $P \in GL(k,\mathbb{R})$ in the definition of S_A:

$$A^\top (BP - A) = 0 \quad \Leftrightarrow \quad A^\top BP = A^\top A \quad \Leftrightarrow \quad P = (A^\top B)^{-1} A^\top A.$$

Since for each $A \in \mathrm{St}(k, \mathbb{R}^n)$ we have $A \in T_A$ and T_A is open, $(T_A)_{A \in \mathrm{St}(k,\mathbb{R}^n)}$ is an open covering of $\mathrm{St}(k, \mathbb{R}^n)$. On the other hand, for $B \in \mathrm{St}(k, \mathbb{R}^n)$ the set

$$\{A \in \mathrm{St}(k, \mathbb{R}^n);\ B \in T_A\} = T_B$$

and is therefore open as well. Let

$$U_A := \pi[T_A] = \Big\{\pi(B);\ B \in \mathrm{St}(k, \mathbb{R}^n),\ \ \det(A^\top B) \neq 0\Big\} \subset \mathrm{Gr}(k, \mathbb{R}^n)$$

be the set of subspaces whose representing fibers $B[GL(k, \mathbb{R})]$ intersect the cross section S_A. We call the mapping

$$\sigma_A \colon U_A \to S_A, \qquad\qquad L = \pi(B) \mapsto B(A^\top B)^{-1} A^\top A \tag{9}$$

the *cross section mapping*.

2.1 Lemma. *The following statements hold:*

(i) $(U_A)_{A\in \mathrm{St}(k,\mathbb{R}^n)}$ *is an open covering of* $\mathrm{Gr}(k, \mathbb{R}^n)$*.*

(ii) For each $A \in \mathrm{St}(k, \mathbb{R}^n)$ *one has* $\pi^{-1}[U_A] = T_A$*.*

(iii) For each $A \in \mathrm{St}(k, \mathbb{R}^n)$ *one has that* $\pi\big|_{T_A} \colon T_A \subset \mathrm{St}(k, \mathbb{R}^n) \to U_A$ *is continuous and open.*

(iv) For each $A \in \mathrm{St}(k, \mathbb{R}^n)$ *one has that* σ_A *is continuous,* $\pi \circ \sigma_A = \mathrm{id}_{U_A}$ *and* $\sigma_{AP}(L) = \sigma_A(L)P$ *for* $P \in GL(k, \mathbb{R})$ *and* $L \in U_A$*.*

(v) For each $B \in \mathrm{St}(k, \mathbb{R}^n)$ *one has that* $T_B \ni A \mapsto \sigma_A(\pi(B))$ *is differentiable.*

Proof. (i) The covering property is clear by the surjectivity of π and the fact that $(T_A)_{A\in\mathrm{St}(k,\mathbb{R}^n)}$ is an open covering of $\mathrm{St}(k, \mathbb{R}^n)$. The fact that for each $A \in \mathrm{St}(k, \mathbb{R}^n)$ the set U_A is open follows from the fact that both T_A and π are open.

(ii) We have by definition that $\pi^{-1}[U_A] = \pi^{-1}[\pi[T_A]] = \bigcup_{P\in GL(k,\mathbb{R})}[T_A]P = T_A$, since T_A is invariant under right-multiplication with elements from $GL(k, \mathbb{R})$.

(iii) This is clear by the continuous embedding of T_A in $\mathrm{St}(k, \mathbb{R}^n)$ via the identity.

(iv) In order to prove continuity of σ_A, by Lemma 1.2 it is equivalent to show the continuity of $\sigma_A \circ \pi \colon T_A \subset \mathrm{St}(k, \mathbb{R}^n) \to S_A$, which is obvious by the representation in Eq. (9). To see that $\pi \circ \sigma_A = \mathrm{id}$ observe that for $B \in T_A$ we have $(A^\top B)^{-1}A^\top A \in GL(k, \mathbb{R})$. Let $B \in T_A$ and $L := \pi(B) \in U_A$, then $\pi(\sigma_A(L)) = \pi(B(A^\top B)^{-1}A^\top A) = \pi(B) = L$. We show the homogeneity property by calculating

$$\begin{aligned}\sigma_{AP}(\pi(B)) &= B((AP)^\top B)^{-1}(AP)^\top AP = B(P^\top A^\top B)^{-1}(AP)^\top AP \\ &= B(A^\top B)^{-1}(P^\top)^{-1}P^\top A^\top AP = B(A^\top B)^{-1}A^\top AP \\ &= \sigma_A(\pi(B))P.\end{aligned}$$

(v) This is clear with the fact that W_B is open and with the representation in Eq. (9). □

Next, choose for each $A \in \mathrm{St}(k, \mathbb{R}^n)$ an $A_\perp \in \mathrm{St}^*(n-k, n)$ such that $A^\top A_\perp = 0 \in \mathbb{R}^{k\times(n-k)}$ and $A_\perp^\top A = 0 \in \mathbb{R}^{(n-k)\times k}$ in the following way. Let $A = U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} V^\top$ be a singular value decomposition (SVD) with $U \in O(n)$, $V \in O(k)$ and $\Sigma \in \mathbb{R}^{k\times k}$ a diagonal matrix with positive diagonal entries. Define $A_\perp := U \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix}$ for some $W \in \mathbb{R}^{(n-k)\times(n-k)}$ satisfying $WW^\top = I_{n-k}$. We then verify

$$\begin{aligned} A^\top A_\perp &= V \begin{pmatrix} \Sigma^\top & 0_{k\times(n-k)} \end{pmatrix} U^\top U \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} \\ &= V \begin{pmatrix} \Sigma^\top & 0_{k\times(n-k)} \end{pmatrix} \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} = 0_{k\times(n-k)}, \end{aligned} \tag{10}$$

and consequently $A_\perp^\top A = 0_{(n-k)\times k}$. This justifies the notation $A_\perp$ since we have $\pi(A)^\perp = \pi(A_\perp)$ due to the full rank of both matrices and their orthogonality from Eq. (10). For preparatory reasons, we first consider

$$A_\perp A_\perp^\top = U \begin{pmatrix} 0_{k\times(n-k)} \\ W \end{pmatrix} \begin{pmatrix} 0_{k\times(n-k)} & W^\top \end{pmatrix} U^\top = U \begin{pmatrix} 0_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & I_{n-k} \end{pmatrix} U^\top$$

and secondly

$$\begin{aligned} A(A^\top A)^{-1}A^\top &= U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} V^\top \left(V\Sigma\Sigma V^\top \right)^{-1} V \begin{pmatrix} \Sigma & 0_{(n-k)\times k} \end{pmatrix} U^\top \\ &= U \begin{pmatrix} \Sigma \\ 0_{(n-k)\times k} \end{pmatrix} \Sigma^{-1}\Sigma^{-1} \begin{pmatrix} \Sigma & 0_{(n-k)\times k} \end{pmatrix} U^\top \\ &= U \begin{pmatrix} I_k \\ 0_{(n-k)\times k} \end{pmatrix} \begin{pmatrix} I_k & 0_{(n-k)\times k} \end{pmatrix} U^\top \\ &= U \begin{pmatrix} I_k & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & 0_{(n-k)\times(n-k)} \end{pmatrix} U^\top \\ &= U \left(I_n - \begin{pmatrix} 0_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & I_{n-k} \end{pmatrix} \right) U^\top \\ &= I_n - A_\perp A_\perp^\top. \end{aligned} \tag{11}$$

Note that $(A^\top A)^{-1}A^\top$ is the Moore-Penrose inverse to A since A has full rank. It is well-known that for $B \in \mathrm{St}^*(k, \mathbb{R}^n)$ the matrix $BB^\top$ represents the orthogonal projection onto $\pi(B)$. Thus, we derived in Eq. (11) that $A(A^\top A)^{-1}A^\top$ represents the orthogonal projection onto $\pi(A)$ and equivalently $I_n - A(A^\top A)^{-1}A^\top$ represents the orthogonal projection onto $\pi(A)^\perp$.

With respect to $A \in \mathrm{St}(k, \mathbb{R}^n)$ we define the following family of functions

$$\begin{aligned} \varphi_A \colon U_A &\to \mathbb{R}^{(n-k)\times k}, \\ L = \pi(B) &\mapsto A_\perp^\top \sigma_A(L) = A_\perp^\top B(A^\top B)^{-1}A^\top A. \end{aligned} \tag{12}$$

2.2 Theorem. $\{\varphi_A;\ A \in \mathrm{St}(k,\mathbb{R}^n)\}$ *define a structure of differentiable manifold of dimension* $k(n-k)$ *in* $\mathrm{Gr}(k,\mathbb{R}^n)$ *with parametrization*

$$p_A\colon \mathbb{R}^{(n-k)\times k} \to U_A, \qquad K \mapsto \pi(A + A_\perp K)$$

for $A \in \mathrm{St}(k,\mathbb{R}^n)$.

Proof. We need to prove that (a) the domains of the charts cover $\mathrm{Gr}(k,\mathbb{R}^n)$, (b) all charts are injective, (c) the image of any chart is mapped by any other chart to an open subset of $\mathbb{R}^{(n-k)\times k}$ and (d) coordinate changes are smooth.

(a) By Lemma 2.1(a) the domains of the charts φ_A, $A \in \mathrm{St}(k,\mathbb{R}^n)$, are a covering for $\mathrm{Gr}(k,\mathbb{R}^n)$.

(b) To prove the invertibility of φ_A for an arbitrary $A \in \mathrm{St}(k,\mathbb{R}^n)$ we need to check if the parametrization p_A is its inverse to obtain surjectivity. To this end, let $K \in \mathbb{R}^{(n-k)\times k}$ and we calculate

$$\begin{aligned}
\varphi_A(p_A(K)) &= A_\perp^\top (A + A_\perp K)(A^\top(A + A_\perp K))^{-1} A^\top A \\
&= \left(A_\perp^\top A + A_\perp^\top A_\perp K\right)\left(A^\top A + A^\top A_\perp K\right)^{-1} A^\top A \\
&= K\left(A^\top A\right)^{-1} A^\top A = K.
\end{aligned}$$

To prove injectivity, let $L \in U_A$, $B \in \pi^{-1}[\{L\}]$ and by making use of Eq. (11) we calculate

$$\begin{aligned}
p_A(\varphi_A(L)) &= \pi(A + A_\perp A_\perp^\top B(A^\top B)^{-1} A^\top A) \\
&= \pi(A + B(A^\top B)^{-1} A^\top A - A(A^\top A)^{-1} A^\top B(A^\top B)^{-1} A^\top A) \\
&= \pi(B(A^\top B)^{-1} A^\top A) = \pi(B) = L.
\end{aligned}$$

(c) Clearly, φ_A and p_A are continuous for every $A \in \mathrm{St}(k,\mathbb{R}^n)$, i.e. φ_A is a homeomorphism. For $A \in \mathrm{St}(k,\mathbb{R}^n)$ the domain of φ_A is open, thus $U_A \cap U_B$ is open and consequently $\varphi_B[U_A \cap U_B] = p_B^{-1}[U_A \cap U_B] \subseteq \mathbb{R}^{(n-k)\times k}$ is open due to continuity of p_B for any $B \in \mathrm{St}(k,\mathbb{R}^n)$.

(d) Let $A, B \in \mathrm{St}(k,\mathbb{R}^n)$ be such that $U_A \cap U_B \neq \varnothing$. Then for $K \in \varphi_A[U_B]$ we have

$$\begin{aligned}
\varphi_B \circ p_A(K) &= B_\perp^\top (A + A_\perp K)(B^\top(A + A_\perp K))^{-1} B^\top B \\
&= \left(B_\perp^\top A + B_\perp^\top A_\perp K\right)\left(B^\top A + B^\top A_\perp K\right)^{-1} B^\top B,
\end{aligned}$$

which is smooth as the composition of smooth functions (multiplication and inversion). □

From Eq. (12) we directly read off the differentiability of π.

2.3 Corollary. *Let* $\mathrm{Gr}(k,\mathbb{R}^n)$ *be endowed with the differentiable structure from Theorem 2.2. Then* $\pi\colon \mathrm{St}(k,\mathbb{R}^n) \to \mathrm{Gr}(k,\mathbb{R}^n)$ *as defined in Eq.* (2) *is differentiable.*

Furthermore, with the charts at hand we obtain that σ_A is an immersion for any $A \in \mathrm{St}(k,\mathbb{R}^n)$.

2.4 Lemma. *For any $A \in \mathrm{St}(k, \mathbb{R}^n)$ the function $\sigma_A \colon U_A \to S_A$ is an immersion.*

Proof. First, differentiability with respect to the charts is obvious. We need to show that $\partial(\mathrm{id} \circ \sigma_A \circ \varphi_A^{-1})(\varphi_A(L)) \in L(\mathbb{R}^{(n-k)\times k}, \mathbb{R}^{n\times k})$ is injective for any $L \in U_A$. By Theorem 2.2 we have that $\varphi_A^{-1} = p_A$ and that $\sigma_A \circ p_A = (K \mapsto A + A_\perp K)$, which has the constant derivative $A_\perp$. By construction, $A_\perp$ has full rank and the injectivity is proved. □

3 (Fiber) Bundle Structure

In this section, we follow [6]. It is known that $GL(k, \mathbb{R})$ forms a real Lie-group, see for instance [5, Example 2.4]. Next we establish $\pi \colon \mathrm{St}(k, \mathbb{R}^n) \to \mathrm{Gr}(k, \mathbb{R}^n)$ as a principal $GL(k, \mathbb{R})$-bundle, i.e. a differentiable bundle with a differentiable right $GL(k, \mathbb{R})$-action. This differentiable right $GL(k, \mathbb{R})$-action $\cdot$ on $\mathrm{St}(k, \mathbb{R}^n)$ is the right-multiplication with matrices $P \in GL(k, \mathbb{R})$

$$\mathrm{St}(k, \mathbb{R}^n) \times GL(k, \mathbb{R}) \to \mathrm{St}(k, \mathbb{R}^n), \qquad (A, P) \mapsto AP. \tag{13}$$

Also, if $U \subset \mathrm{Gr}(k, \mathbb{R}^n)$ is open, we consider the $GL(k, \mathbb{R})$-action on $U \times GL(k, \mathbb{R})$ defined by

$$(U \times GL(k, \mathbb{R})) \times GL(k, \mathbb{R}) \to U \times GL(k, \mathbb{R}), \qquad ((L, P), Q) \mapsto (L, PQ). \tag{14}$$

It is immediate that both are differentiable and effective, i.e. $AP = A$ implies $P = I$ and $(L, PQ) = (L, P)$ implies $Q = I$ with $A \in \mathrm{St}(k, \mathbb{R}^n)$, $P, Q \in GL(k, \mathbb{R})$ and $L \in \mathrm{Gr}(k, \mathbb{R}^n)$.

3.1 Proposition. *The function π together with the $GL(k, \mathbb{R})$-actions defined in Eqs. (13) and (14) is a principal $GL(k, \mathbb{R})$-bundle.*

Proof. We need to show that

(i) π is a fiber bundle, i.e. the "local triviality" is satisfied: for every $L \in \mathrm{Gr}(k, \mathbb{R}^n)$ there is an open neighborhood U of L in $\mathrm{Gr}(k, \mathbb{R}^n)$ and a diffeomorphism

$$\psi \colon \pi^{-1}[U] \to U \times GL(k, \mathbb{R}),$$

such that the diagram

$$\begin{array}{ccc} \pi^{-1}[U] & \xrightarrow{\psi} & U \times GL(k, \mathbb{R}) \\ \pi \downarrow & \swarrow_{\mathrm{proj}_0} & \\ U & & \end{array}$$

commutes; here, proj_0 denotes the projection onto the first component;

(ii) for every $L \in \mathrm{Gr}(k, \mathbb{R}^n)$ its fiber $\pi^{-1}[\{L\}]$ is diffeomorphic to $GL(k, n)$, and

(iii) the function ψ in a) is equivariant.

Concerning (i) let $L \in \mathrm{Gr}(k, \mathbb{R}^n)$ and $A \in \pi^{-1}[\{L\}]$. Then U_A is an open neighborhood of L and with

$$\begin{aligned} \psi_A &: T_A \to U_A \times GL(k, \mathbb{R}), & B &\mapsto (\pi(B), (A^\top A)^{-1} A^\top B), \\ \theta_A &: U_A \times GL(k, \mathbb{R}) \to T_A, & (L, P) &\mapsto \sigma_A(L) P, \end{aligned}$$

the diagram commutes, as is easily checked. In particular, ψ_A and θ_A are differentiable (by the differentiability of π and σ_A), inverse functions of each other and consequently ψ_A is a diffeomorphism. Requirement (ii) is satisfied by Lemma 1.1(i). Finally, ψ_A is equivariant by the following calculation: let $B \in T_A$ and $P \in GL(k, \mathbb{R})$, then

$$\psi_A(BP) = (\pi(BP), (A^\top A)^{-1} A^\top BP) = (\pi(B), (A^\top A)^{-1} A^\top BP) = \psi_A(B)P. \qquad \square$$

As a consequence, we have that π is a surjective submersion.

4 Riemannian Structure

The Stiefel manifold can be endowed with the structure of a Riemannian manifold by the Riemannian metric

$$\langle X, Y \rangle_A := \mathrm{tr}((A^\top A)^{-1} X^\top Y)^{\frac{1}{2}}, \tag{15}$$

where $A \in \mathrm{St}(k, \mathbb{R}^n)$ and $X, Y \in T_A \mathrm{St}(k, \mathbb{R}^n)$. Note that this is different from the Euclidean Riemannian structure we will also consider later. Since the Grassmann manifold can be equivalently considered as a quotient manifold with respect to the Stiefel manifold modulo $GL(k, \mathbb{R})$ (see Eq. (4)), it can inherit the Riemannian structure from the Stiefel manifold by turning it into a Riemannian quotient manifold. Since subspaces are represented by matrices that span the subspaces, the aim of this section is to find representations for notions connected with the tangent bundle of the Grassmann manifold by objects connected with the tangent bundle of the Stiefel manifold. To this end, the two notions of *vertical bundle* and *horizontal bundle* are introduced as follows.

Let $A \in \mathrm{St}(k, \mathbb{R}^n)$. Since $\mathrm{St}(k, \mathbb{R}^n)$ is open in $\mathbb{R}^{n\times k}$ one has that $T_A \mathrm{St}(k, \mathbb{R}^n) = T_A \mathbb{R}^{n\times k} = \mathbb{R}^{n\times k}$. Next we decompose the tangent bundle $T\,\mathrm{St}(k, \mathbb{R}^n)$ into two subbundles, the aforementioned vertical and horizontal bundle. First, observe that by Lemma 1.1 π is surjective and as a consequence of the fiber bundle property π is a submersion. Hence, by the Submersion Theorem each fiber is a submanifold of dimension k^2. The well-defined tangent space to the fiber $\pi^{-1}[\{\pi(A)\}]$ is a subspace of $T_A \mathrm{St}(k, \mathbb{R}^n)$ and is called *vertical space* at A, denoted by V_A, i.e.

$$V_A := T_A(\pi^{-1}[\{\pi(A)\}]) = T_A(A[GL(k, \mathbb{R})]) \cong A[\mathbb{R}^{k\times k}],$$

since $GL(k, \mathbb{R})$ is open in $\mathbb{R}^{k\times k}$. Equivalently, the vertical space can be represented as

$$V_A = \ker D\pi\big|_A,$$

since π is constant on fibers. The *horizontal space* H_A at A is, in the case of Riemannian manifolds considered here, the orthogonal complement in $T\,\mathrm{St}(k,\mathbb{R}^n)$ to V_A. This yields the tangent space to the local cross section through A, i.e.

$$H_A := V_A^\perp = \left\{ Y \in T\,\mathrm{St}(k,\mathbb{R}^n);\ A^\top Y = 0 \right\} \cong A_\perp[\mathbb{R}^{(n-k)\times k}] \cong T_A S_A,$$

i.e. all matrices "orthogonal" to A. We denote by $V\,\mathrm{St}(k,\mathbb{R}^n)$ and $H\,\mathrm{St}(kn)$ the vertical and horizontal bundle associated to $\mathrm{St}(k,\mathbb{R}^n)$, respectively.

4.1 Lemma. *For any* $A \in \mathrm{St}(k,\mathbb{R}^n)$ *one has* $T_A\,\mathrm{St}(k,\mathbb{R}^n) = V_A \oplus H_A$.

Proof. This is easily verified making use of the full rank of A and $A_\perp$ and their orthogonality. □

For short, we have the decomposition of the tangent bundle

$$T\,\mathrm{St}(k,\mathbb{R}^n) = V\,\mathrm{St}(k,\mathbb{R}^n) \oplus H\,\mathrm{St}(k,\mathbb{R}^n),$$

which is meant fiberwise and abbreviates the statement of Lemma 4.1. By Theorem 2.2 we have that $\mathrm{Gr}(k,\mathbb{R}^n)$ is a differentiable manifold of dimension $k(n-k)$. Hence, the tangent space at an arbitrary point $L \in \mathrm{Gr}(k,\mathbb{R}^n)$ is a vector space of the same dimension. By Lemma 2.4 σ_A is an immersion and, hence, the differential $D\sigma_A\colon TU_A \subset T\,\mathrm{Gr}(k,\mathbb{R}^n) \to TS_A \subset H_A \subset T\,\mathrm{St}(k,\mathbb{R}^n)$ is injective. Furthermore, the tangent spaces have equal dimension so that $D\sigma_A$ is a vector space isomorphism between the tangent space $T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$ and the horizontal space H_A of the Stiefel manifold.

We denote by π_{Gr} and π_{St} the tangent bundle projections of $\mathrm{Gr}(k,\mathbb{R}^n)$ and $\mathrm{St}(k,\mathbb{R}^n)$, respectively.

4.2 Definition (Horizontal lift)**.** We define

$$\{(A,X) \in \mathrm{St}(k,\mathbb{R}^n) \times T\,\mathrm{Gr}(k,\mathbb{R}^n);\ \pi(A) = \pi_{\mathrm{Gr}}(X)\} \to H\,\mathrm{St}(k,\mathbb{R}^n),$$
$$(A,X) \mapsto \overline{X}_A := D\sigma_A\big|_{\pi(A)} X,$$

and call $\overline{X}_A$ the *horizontal lift of* $X \in T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$ *at* A.

By what has been said before, $\bar{\cdot}_A\colon T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n) \to H_A$ is an isomorphism, i.e. $\overline{Y}_A$ is the *only* horizontal vector that represents $Y \in T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$ in a sense specified in Proposition 4.3(ii). In the next proposition we collect some simple properties of the horizontal lift. We denote by $\partial_A f(x)$ the directional derivative of f in the direction A at x.

4.3 Proposition. *Let* $A \in \mathrm{St}(k,\mathbb{R}^n)$, $f \in \mathcal{F}_{\pi(A)}(\mathrm{Gr}(k,\mathbb{R}^n))$, $X \in \mathcal{X}(\mathrm{Gr}(k,\mathbb{R}^n))$, $Y \in T_{\pi(A)}\,\mathrm{Gr}(k,\mathbb{R}^n)$, $Z \in U_A$ *and* $P \in GL(k,\mathbb{R})$. *Then the following statements hold:*

(i) $\overline{X \circ \pi} := (A \mapsto \overline{X(\pi(A))}_A) \in \mathcal{X}(\mathrm{St}(k,\mathbb{R}^n))$, *i.e. the horizontal lift of a vector field on* $\mathrm{Gr}(k,\mathbb{R}^n)$ *is a vector field on* $\mathrm{St}(k,\mathbb{R}^n)$.

(ii) $D\pi\big|_A \overline{Y}_A = Y$*, or, equivalently,* $D\pi\big|_A$ *restricted to* H_A *is the inverse to* $\bar{\cdot}_A$*. By definition, that means that* $\overline{Y}_A$ *is* π*-related to* Y*.*

(iii) $\overline{X}_{AP} = \overline{X}_A P$*.*

(iv) $A^\top \overline{Y}_{AP} = 0$*.*

(v) $Yf \cong \partial_{\overline{Y}_A}(f \circ \pi)(A)$*.*

Proof. (i) This holds by the definition of the push-forward and by Lemma 2.1(v).

(ii) By Lemma 2.1(iv) we have that

$$\mathrm{id}_{U_A} = \pi \circ \sigma_A.$$

Differentiating both sides yields in particular

$$Y = D(\pi \circ \sigma_A)Y = D\pi\big|_A D\sigma_A\big|_{\pi(A)} Y = D\pi\big|_A \overline{Y}_A.$$

(iii) This follows from the homogeneity property of σ_A proved in Lemma 2.1(iv).

(iv) This is a paraphrase of the orthogonality of A and H_A.

(v) First observe that $f \circ \pi \in \mathcal{F}(\mathrm{St}(k, \mathbb{R}^n))$. In submanifolds of $\mathbb{R}^l$ derivations (tangent vectors) and directional derivatives can be identified:

$$\partial_{\overline{Y}_A}(f \circ \pi)(A) \cong \overline{Y}_A(f \circ \pi) = D\pi\big|_A \overline{Y}_A(f) \overset{(ii)}{=} Yf.$$ □

Next we define a Riemannian metric on $\mathrm{Gr}(k, \mathbb{R}^n)$, which is in fact induced by the horizontal lift and the Riemannian metric on $\mathrm{St}(k, \mathbb{R}^n)$ given in Eq. (15). It is natural in the following sense: it is the only Riemannian metric that turns π into a *Riemannian submersion*, i.e. a submersion whose restriction of the differential to the horizontal bundle is an isometry; see Eq. (16) in connection with Proposition 4.3(ii).

4.4 Proposition (Riemannian metric)**.** *For* $L \in \mathrm{Gr}(k, \mathbb{R}^n)$*,* $A \in \pi^{-1}[\{L\}] \subset \mathrm{St}(k, \mathbb{R}^n)$ *and* $X, Y \in T_L \mathrm{Gr}(k, \mathbb{R}^n)$ *we define*

$$\langle X, Y \rangle_L = \langle X, Y \rangle_{\pi(A)} := \mathrm{tr}((A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A)^{\frac{1}{2}} = \left\langle \overline{X}_A, \overline{Y}_A \right\rangle_A. \tag{16}$$

Then the family $(\langle \cdot, \cdot \rangle_L)_{L \in \mathrm{Gr}(k, \mathbb{R}^n)}$ *is a Riemannian metric on* $\mathrm{Gr}(k, \mathbb{R}^n)$*.*

Proof. The inner product nature of $\langle \cdot, \cdot \rangle_L$ for every $L \in \mathrm{Gr}(k, \mathbb{R}^n)$ is obvious as well as the differentiability, so it remains to check whether the definition is independent of the choice of the representing matrix. Let therefore $A, B \in \pi^{-1}[\{L\}]$ and $P \in GL(k, \mathbb{R})$ such that $B = AP$. Then

$$\begin{aligned}
\left\langle \overline{X}_{AP}, \overline{Y}_{AP} \right\rangle_{AP} &= \mathrm{tr}(((AP)^\top AP)^{-1} \overline{X}_{AP}^\top \overline{Y}_{AP}) \\
&= \mathrm{tr}((P^\top A^\top AP)^{-1} P^\top \overline{X}_A^\top \overline{Y}_A P) \\
&= \mathrm{tr}(P^{-1}(A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A P) \\
&= \mathrm{tr}((A^\top A)^{-1} \overline{X}_A^\top \overline{Y}_A) \\
&= \left\langle \overline{X}_A, \overline{Y}_A \right\rangle_A
\end{aligned}$$

by the similarity-invariance of the trace. □

4.5 Proposition (Lie bracket)**.** *Let* $X, Y \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$ *and* $A \in \mathrm{St}(k, \mathbb{R}^n)$*. Then*

$$\overline{[X, Y]_{\mathrm{Gr}}(\pi(A))}_A = \Pi_{\pi(A_\perp)}[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A), \tag{17}$$

where $\Pi_{\pi(A_\perp)} = A_\perp A_\perp^\top = I_n - A(A^\top A)^{-1}A^\top$ *(recall Eq. (11)) and*

$$[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) = \partial_{\overline{X(\pi(A))}_A}(\overline{Y \circ \pi})(A) - \partial_{\overline{Y(\pi(A))}_A}(\overline{X \circ \pi})(A)$$

is the Lie bracket for vector fields on $\mathrm{St}(k, \mathbb{R}^n)$*.*

Proof. By [13, Lemma 1.22] we have that $[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) \in T_A \mathrm{St}(k, \mathbb{R}^n)$ is π-related to $[X, Y]_{\mathrm{Gr}}(\pi(A))$. The projection $\Pi_{\pi(A_\perp)}$ makes it horizontal, such that

$$\Pi_{\pi(A_\perp)}[\overline{X \circ \pi}, \overline{Y \circ \pi}]_{\mathrm{St}}(A) \in H_A$$

and is π-related to $[X, Y]_{\mathrm{Gr}}(\pi(A))$. These are the characterizing properties of $\overline{[X, Y]_{\mathrm{Gr}}(\pi(A))}_A$ and Eq. (17) is proved. □

In the following we endow $\mathrm{Gr}(k, \mathbb{R}^n)$ with the Riemannian structure induced by the Riemannian metric given in Eq. (16).

The *gradient* of a function $f \in \mathcal{F}(\mathrm{Gr}(k, \mathbb{R}^n))$ with respect to the Riemannian metric, denoted by $\mathrm{grad}_{\mathrm{Gr}} f$, is the vector field satisfying $\langle \mathrm{grad}_{\mathrm{Gr}} f, X \rangle = Xf$ for any $X \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$. Recall that in $\mathrm{St}(k, \mathbb{R}^n)$ the Euclidean, i.e. induced by $\mathbb{R}^{n \times k}$ as in Eq. (1), Riemannian metric is given by $\langle A, B \rangle_{\mathrm{St}} = \sqrt{\mathrm{tr}(A^\top B)}$. Thus, the Euclidean gradient for $g \in \mathcal{F}(\mathrm{St}(k, \mathbb{R}^n))$ with

$$\mathrm{grad}_{\mathrm{St}}\, g(A) = (\partial_{ij} g(A))_{i,j} =: \partial \otimes g(A) \in \mathbb{R}^{n \times k} \tag{18}$$

is characterized by

$$X(f \circ \pi) \cong \partial_X (f \circ \pi)(A) = \mathrm{tr}(X^\top \mathrm{grad}_{\mathrm{St}}(f \circ \pi)(A))$$

for $A \in \mathrm{St}(k, \mathbb{R}^n)$, $X \in T_A \mathrm{St}(k, \mathbb{R}^n)$ and $f \in \mathcal{F}(\mathrm{Gr}(k, \mathbb{R}^n))$. Alternatively, we can define another gradient with respect to the metric defined in Eq. (15) and denoted by $\mathrm{grad}_A f$ accordingly. Note that for $f \in \mathcal{F}(\mathrm{Gr}(k, \mathbb{R}^n))$ we have that $f \circ \pi$ is constant on fibers and hence it follows that

$$\langle \mathrm{grad}_A(f \circ \pi)(A), X \rangle_A = \partial_X (f \circ \pi)(A) = 0$$

for all $X \in V_A$. Consequently, $\mathrm{grad}_A(f \circ \pi)(A) \in H_A$.

4.6 Lemma (Gradient)**.** *Let* $A \in \mathrm{St}(k, \mathbb{R}^n)$ *and* $f \in \mathcal{F}(\mathrm{Gr}(k, \mathbb{R}^n))$*. Then*

$$\overline{\mathrm{grad}_{\mathrm{Gr}} f(\pi(A))}_A = \mathrm{grad}_A(f \circ \pi)(A) = \mathrm{grad}_{\mathrm{St}}(f \circ \pi)(A) A^\top A. \tag{19}$$

Proof. Since $\mathrm{grad}_A(f \circ \pi)(A) \in H_A$ it suffices to consider horizontal tangent vectors, that can be uniquely represented as the horizontal lift $\overline{X}_A$ of a tangent vector $X \in T_{\pi(A)} \mathrm{Gr}(k, \mathbb{R}^n)$. Consider the following equalities:

$$\overline{X}_A([f \circ \pi]_{\sim_A}) = \left\langle \overline{X}_A, \mathrm{grad}_A(f \circ \pi)(A) \right\rangle_A \tag{20}$$

$$= \mathrm{tr}\big((A^\top A)^{-1} \overline{X}_A^\top \, \mathrm{grad}_A(f \circ \pi)(A)\big), \tag{21}$$

$$\overline{X}_A([f \circ \pi]_{\sim_A}) = \left\langle \overline{X}_A, \mathrm{grad}_{\mathrm{St}}(f \circ \pi)(A) \right\rangle_{\mathrm{St}}$$

$$= \mathrm{tr}\big(\overline{X}_A^\top \, \mathrm{grad}_{\mathrm{St}}(f \circ \pi)(A)\big), \tag{22}$$

$$\overline{X}_A([f \circ \pi]_{\sim_A}) = X([f]_{\sim_{\pi(A)}}) = \langle X, \mathrm{grad}_{\mathrm{Gr}} \, f(\pi(A)) \rangle_{\pi(A)}$$

$$= \left\langle \overline{X}_A, \overline{\mathrm{grad}_{\mathrm{Gr}} \, f(\pi(A))}_A \right\rangle_A . \tag{23}$$

Now, from Eqs. (21) and (22) we read off the second claimed equality, and from Eqs. (20) and (23) the first one. ☐

If M is a submanifold of a Euclidean space, then the Riemannian connection $\nabla(X, Y)$, $X \in T_pM$ and $Y \in \mathcal{X}(M)$, consists in taking the directional derivative of Y in the direction of X in the ambient Euclidean space and projecting the result into T_pM. The projection step can be omitted in the case of $\mathrm{St}(k, \mathbb{R}^n)$ since it is an open subset of $\mathbb{R}^{n\times k}$. In other words, in $\mathrm{St}(k, \mathbb{R}^n)$ the Riemannian connection ∇_{St} is the directional derivative. On the other hand, in $\mathrm{Gr}(k, \mathbb{R}^n)$ the Riemannian metric gives rise to a unique Riemannian connection ∇_{Gr}. The next result shows that both Riemannian connections are related to each other through a horizontal projection.

4.7 Proposition (Riemannian connection)**.** *Let* $A \in \mathrm{St}(k, \mathbb{R}^n)$ *and* $X, Y \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n))$*. Then*

$$\overline{\nabla_{\mathrm{Gr}}(X(\pi(A)), Y)}_A = \Pi_{\pi(A_\perp)} \nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y \circ \pi}\big) \tag{24}$$

with

$$\nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y \circ \pi}\big) = \partial_{\overline{X(\pi(A))}_A} \overline{Y \circ \pi}(A). \tag{25}$$

Proof. Obviously, both sides of Eq. (24) are elements of H_A. Eq. (24) holds if both sides have the same scalar product with every horizontal tangent vector, which can be represented by $\overline{Z}_A$ for $Z \in T_{\pi(A)} \mathrm{Gr}(k, \mathbb{R}^n)$. Since the vertical component of $\nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y \circ \pi}\big)$ is orthogonal to the horizontal space and hence there is no contribution to the scalar product, it suffices to prove for each $Z \in T_{\pi(A)} \mathrm{Gr}(k, \mathbb{R}^n)$ that

$$\left\langle \nabla_{\mathrm{St}}\big(\overline{X(\pi(A))}_A, \overline{Y \circ \pi}\big), \overline{Z}_A \right\rangle_A = \langle \nabla_{\mathrm{Gr}}(X(\pi(A)), Y), Z \rangle_{\pi(A)} . \tag{26}$$

This follows by expanding both sides in the Koszul formula, where it suffices to consider terms of the following two types. Let $X, Y, Z \in \mathcal{X}(\mathrm{Gr}(k, \mathbb{R}^n)$ and denote by $\overline{X \circ \pi} :=$

$(A \mapsto \overline{X \circ \pi}_A) \in \mathcal{X}(\mathrm{St}(k, \mathbb{R}^n))$ the horizontal vector field and consider the following as functions on $\mathrm{St}(k, \mathbb{R}^n)$:

$$\begin{aligned}
\overline{X \circ \pi} \left\langle \overline{Y \circ \pi}, \overline{Z \circ \pi} \right\rangle &= \left(A \mapsto \overline{X \circ \pi}_A \left([\langle \overline{Y \circ \pi}, \overline{Z \circ \pi} \rangle]_{\sim_A} \right)\right) \\
&= \left(A \mapsto \overline{X \circ \pi}_A \left(\left[\langle Y \circ \pi, Z \circ \pi \rangle_{\pi(\cdot)} \right]_{\sim_A} \right)\right) \\
&= \left(A \mapsto \overline{X \circ \pi}_A \left([\langle Y, Z \rangle_{\cdot} \circ \pi]_{\sim_A} \right)\right) \\
&= \left(A \mapsto D\pi\big|_A \overline{X \circ \pi}_A \left([\langle Y, Z \rangle]_{\sim_{\pi(A)}} \right)\right) \\
&= \left(A \mapsto X(\pi(A)) \left([\langle Y, Z \rangle]_{\sim_{\pi(A)}} \right)\right) \\
&= (X \langle Y, Z \rangle) \circ \pi, \\
\left\langle \overline{X \circ \pi}, \left[\overline{Y \circ \pi}, \overline{Z \circ \pi}\right] \right\rangle &= \left(A \mapsto \left\langle \overline{X \circ \pi}_A, \left[\overline{Y \circ \pi}, \overline{Z \circ \pi}\right]_{\mathrm{St}} (A) \right\rangle_A \right) \\
&= \left(A \mapsto \left\langle \overline{X \circ \pi}_A, \overline{[Y, Z]_{\mathrm{Gr}} (\pi(A))}_A \right\rangle_A \right) \\
&= \left(A \mapsto \langle X(\pi(A)), [Y, Z]_{\mathrm{Gr}}(\pi(A)) \rangle_{\pi(A)} \right) \\
&= (\langle X, [Y, Z] \rangle) \circ \pi.
\end{aligned}$$

In summary, corresponding terms in the expansion in the Koszul formula of both sides of Eq. (26) equal each other and the proof is finished. □

In the following we consider regular curves on $\mathrm{Gr}(k, \mathbb{R}^n)$. We do not state the domains explicitly, but assume implicitly that they contain 0.

Let $t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ be a regular curve on $\mathrm{Gr}(k, \mathbb{R}^n)$. Recall that a vector field $X \in \mathcal{X}_C(\mathrm{Gr}(k, \mathbb{R}^n))$ along C is said to be parallel transported along C if $\underset{C}{\nabla} X = 0$. A (regular) curve $A\colon t \mapsto c(t) \in \mathrm{St}(k, \mathbb{R}^n)$ on $\mathrm{St}(k, \mathbb{R}^n)$ is called *horizontal* if $\dot{A}(t) \in H_{A(t)}$ for any $t \in D(A)$.

Let $t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ be a regular curve on $\mathrm{Gr}(k, \mathbb{R}^n)$, $\pi(A_0) = C(0) \in \mathrm{Gr}(k, \mathbb{R}^n)$, $A_0 \in \mathrm{St}(k, \mathbb{R}^n)$. Then there exists a unique horizontal curve $A\colon t \mapsto A(t)$ such that $A(0) = A_0$ and $\pi(A(t)) = C(t)$ for all $t \in D(C)$. The reason for that is that locally the image of C is a submanifold in $\mathrm{Gr}(k, \mathbb{R}^n)$. For simplicity assume that this holds globally. The preimage of C under π is by the Submersion Theorem a submanifold of $\mathrm{St}(k, \mathbb{R}^n)$, on which we can define a horizontal vector field that is constant on fibers, i.e. for $B \in \pi^{-1}[C[D(C)]] \subseteq \mathrm{St}(k, \mathbb{R}^n)$ define

$$X(B) \coloneqq \overline{\dot{C}(\pi(B))}_B.$$

For each A_0 there exists a unique integral curve A through A_0, i.e. a solution to $A(0) = A_0$ and $\dot{A}(t) = X(A(t))$, satisfying $\pi(A(t)) = C(t)$ for all $t \in D(C)$. By definition of X we have that A is horizontal and the projection property follows from uniqueness. The curve A is called the *horizontal lift of C through A_0*.

4.8 Proposition (Parallel transport)**.** *Let C be a regular curve on* $\mathrm{Gr}(k, \mathbb{R}^n)$*,* $X \in \mathcal{X}_C(\mathrm{Gr}(k, \mathbb{R}^n))$*, A be a horizontal lift of C. Then*

$$\overline{X}_A := \left(t \mapsto \overline{X(t)}_{A(t)}\right) \in \mathcal{X}_C(\mathrm{St}(k, \mathbb{R}^n))$$

is the horizontal lift of X along A and X is parallel transported along C if and only if

$$\dot{\overline{X}}_A(t) + A(t)\left(A(t)^\top A(t)\right)^{-1} \dot{A}(t)^\top \overline{X}_A(t) = 0, \tag{27}$$

where

$$\dot{\overline{X}}_A(t) = (s \mapsto \overline{X(s)}_{A(s)})'(t).$$

Proof. Let $t \in D(C)$. We have that $\dot{A}(t) = \overline{\dot{C}(t)}_{A(t)}$ and

$$\begin{aligned}
\left(\nabla_C X\right)(t) &= \Pi_{\pi(A(t)_\perp)} \nabla_{\mathrm{St}}\left(\overline{\dot{C}(t)}_{A(t)}, \overline{X}_A\right) \\
&= \Pi_{\pi(A(t)_\perp)} \nabla_{\mathrm{St}}\left(\overline{\dot{C}(t)}_{A(t)}, \overline{X}_A\right) \\
&= \Pi_{\pi(A(t)_\perp)} \partial_{\overline{\dot{C}(t)}_{A(t)}} \overline{X}_A(A(t)) \\
&= \Pi_{\pi(A(t)_\perp)} \partial_{\dot{A}(t)} \overline{X}_A(A(t)) \\
&= \Pi_{\pi(A(t)_\perp)} \dot{\overline{X}}_A(t)
\end{aligned}$$

by Proposition 4.7. Thus, X is parallel transported along C if and only if

$$0 = \left(\nabla_C X\right)(t) = \Pi_{\pi(A(t)_\perp)} \dot{\overline{X}}_A(t),$$

i.e. $\dot{\overline{X}}_A(t) \in V_{A(t)}$. Consequently it is of the form

$$\dot{\overline{X}}_A(t) = A(t)M(t) \tag{28}$$

for some $M\colon t \mapsto M(t) \in \mathbb{R}^{k\times k}$. By Proposition 4.3(iv) we have

$$A(t)^\top \overline{X}_A(t) = 0.$$

Differentiation of the last equation yields

$$\dot{A}(t)^\top \overline{X}_A(t) + A(t)^\top \dot{\overline{X}}_A(t) = 0.$$

Plugging in Eq. (28) and solving for $M(t)$ we obtain

$$M(t) = -\left(A(t)^\top A(t)\right)^{-1} \dot{A}(t)^\top \overline{X}_A(t),$$

and the proof is finished. □

Next, we consider geodesics $C\colon t \mapsto C(t) \in \mathrm{Gr}(k, \mathbb{R}^n)$ with some reference point $C(0) = C_0$ and initial direction $\dot{C}_0 \in T_{C(0)} \mathrm{Gr}(k, \mathbb{R}^n)$. The unique geodesic C is characterized by $\underset{C}{\nabla}\dot{C} = 0$ and $\dot{C}(0) = \dot{C}_0$. With this notation the following holds.

4.9 Proposition (Geodesics)**.** *Let $A_0 \in \pi^{-1}[\{C_0\}] \subset \mathrm{St}(k, \mathbb{R}^n)$, $\overline{\dot{C}_0}_{A_0} =: \dot{A}_0$ and*

$$\dot{A}_0(A_0^\top A_0)^{-1/2} = U\Sigma V^\top$$

be a thin singular value decomposition (SVD), i.e. $U \in \mathrm{St}^*(k, \mathbb{R}^n)$, $V \in \mathrm{St}^*(k, k)$ and $\Sigma \in \mathbb{R}^{k\times k}$ is diagonal with nonnegative entries; see, for instance, [9, Section 2.5.4]. Then

$$C(t) = \pi(A_0(A_0^\top A_0)^{-1/2} V \cos(t\Sigma) + U \sin(t\Sigma)).$$

We define $\exp(\dot{C}_0) := C(1)$.

Proof. Let $t \in D(C)$ and A be the unique horizontal lift of C through A_0, so that $\dot{A}(t) = \overline{\dot{C}(t)}_{A(t)}$. Then by (27) we have

$$\ddot{A}(t) + A(t)\left(A(t)^\top A(t)\right)^{-1} \dot{A}(t)^\top \dot{A}(t) = 0. \tag{29}$$

Since A is horizontal, one has

$$A(t)^\top \dot{A}(t) = 0 = \dot{A}(t)^\top A(t). \tag{30}$$

Thus, we have

$$\dot{\widehat{(A^\top A)}}(t) = \dot{A}(t)^\top A(t) + A(t)^\top \dot{A}(t) = 0,$$

saying that $t \mapsto A(t)^\top A(t)$ is a constant function. Differentiation of Eq. (30) yields

$$\dot{A}(t)^\top \dot{A}(t) + A(t)^\top \ddot{A}(t) = 0,$$

or equivalently

$$\dot{A}(t)^\top \dot{A}(t) = -A(t)^\top \ddot{A}(t). \tag{31}$$

By plugging Eq. (31) into Eq. (29) we get

$$\ddot{A}(t) - A(t)\left(A(t)^\top A(t)\right)^{-1} A(t)^\top \ddot{A}(t) = \Pi_{\pi(A(t))^\perp} \ddot{A}(t) = 0,$$

saying that $\ddot{A}(t) \in V_{A(t)}$ and consequently it is of the form

$$\ddot{A}(t) = A(t)M(t) \tag{32}$$

for some $M\colon t \mapsto M(t) \in \mathbb{R}^{k\times k}$. With Eqs. (32) and (31) we obtain that

$$\begin{aligned}
\dot{\widehat{(\dot{A}^\top A)}}(t) &= \ddot{A}(t)^\top \dot{A}(t) + \dot{A}(t)^\top \ddot{A}(t) \\
&= (A(t)M(t))^\top \dot{A}(t) + \dot{A}(t)^\top A(t)M(t) \\
&= M(t)^\top A(t)^\top \dot{A}(t) + \dot{A}(t)^\top A(t)M(t) = 0,
\end{aligned}$$

i.e. $t \mapsto \dot{A}(t)^\top \dot{A}(t)$ is constant, too. Now consider the thin SVD

$$\dot{A}_0(A_0^\top A_0)^{-1/2} = U\Sigma V^\top \tag{33}$$

as in the statement of the proposition. Eq. (29) right-multiplied with $(A_0^\top A_0)^{-1/2}$ yields

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2} + A(t)(A_0^\top A_0)^{-1/2}((A_0^\top A_0)^{-1/2}\dot{A}^\top)(\dot{A}(A_0^\top A_0)^{-1/2}) = 0.$$

Plugging Eq. (33) into the last equation we obtain

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2} + A(t)(A_0^\top A_0)^{-1/2} V\Sigma^2 V^\top = 0,$$

and after right-multiplication with V

$$\ddot{A}(t)(A_0^\top A_0)^{-1/2}V + A(t)(A_0^\top A_0)^{-1/2} V\Sigma^2 = 0.$$

With the abbreviation $B(t) := A(t)(A_0^\top A_0)^{-1/2}V$, the last equation reads

$$\ddot{B}(t) + B(t)\Sigma^2 = 0,$$

from which the solution

$$B(t) = B(0)\cos(t\Sigma) + \dot{B}(0)\Sigma^{-1}\sin(t\Sigma),$$

or equivalently

$$A(t)(A_0^\top A_0)^{-1/2}V = A_0(A_0^\top A_0)^{-1/2}V\cos(t\Sigma) + \dot{A}_0(A_0^\top A_0)^{-1/2}V\Sigma^{-1}\sin(t\Sigma),$$

can be read off directly. The assertion now follows from

$$C(t) = \pi(A(t)) = \pi(A(t)(A_0^\top A_0)^{-1/2}V),$$

since $(A_0^\top A_0)^{-1/2}V \in GL(k, \mathbb{R})$. □

Proposition 4.9 shows that the Grassmann manifold with the Riemannian structure given by the Riemannian metric in (16) is complete, i.e. the geodesics are defined on $\mathbb{R}$. By [14] the Grassmann manifold is connected and hence, by the Hopf-Rinow-theorem (see, for instance, [3, Theorem 10.4.16]), a complete metric space with respect to the Riemannian distance function, also referred to as the geodesic distance. However, the completeness of $\mathrm{Gr}(k, \mathbb{R}^n)$ with respect to the Riemannian distance function can be obtained alternatively with the notion of principal angles; for an introduction see, for instance, [9, 12.4.3]. The distance function induced by the Riemannian metric (16) is the 2-norm of the vector of principal angles denoted by θ. On the other hand, the gap metric Θ corresponds to the sin of the largest principal angle, i.e. $\sin|\theta|_\infty$. Now, one easily verifies the (strong) equivalence of the gap metric and the geodesic distance, i.e.

$$\sin|\theta|_\infty \leq |\theta|_2 \leq \frac{2\sqrt{k}}{\pi}\sin|\theta|_\infty,$$

from which the completeness of $\mathrm{Gr}(k, \mathbb{R}^n)$ with respect to the geodesic distance function follows. For other definitions of distance functions in terms of principal angles see [4, 4.3].

Acknowledgments

I am grateful to Marcus Köhler and Sascha Trostorff for helpful discussions.

References

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. Riemannian Geometry of Grassmann Manifolds with a View on Algorithmic Computation. *Acta Applicandae Mathematicae*, 80(2):199–220, 2004. doi:10.1023/B:ACAP.0000013855.14971.91.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

[3] L. Conlon. *Differentiable Manifolds*. Birkhäuser Advanced Texts. Birkhäuser, 2. edition, 2001.

[4] A. Edelman, T. A. Arias, and S. T. Smith. The Geometry of Algorithms with Orthogonality Constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, 1998. doi:10.1137/S0895479895290954.

[5] H. D. Fegan. *Introduction to compact Lie groups*, volume 13 of *Series in Pure Mathematics*. World Scientific, Singapore, 1991.

[6] J. Ferrer, M. I. Garcia, and F. Puerta. Differentiable families of subspaces. *Linear Algebra and its Applications*, 199:229–252, 1994. doi:10.1016/0024-3795(94)90351-4.

[7] I. C. Gohberg, P. Lancaster, and L. Rodman. *Invariant subspaces of matrices with applications*. Classics in applied mathematics. SIAM, 2006. doi:10.1137/1.9780898719093.

[8] I. C. Gohberg and A. S. Markus. Two theorems on the opening between subspaces of a Banach space. *Uspekhi Mat. Nauk*, 89:135–140, 1959. in Russian.

[9] G.H. Golub and C.F.V. Loan. *Matrix computations*. Johns Hopkins studies in the mathematical sciences. Johns Hopkins University Press, 3. edition, 1996.

[10] T. Kato. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der mathematischen Wissenschaften*. Springer, reprint of the 2nd edition, 1995. doi:10.1007/978-3-642-66282-9.

[11] M. G. Krein and M. A. Krasnoselski. Fundamental theorems concerning the extension of Hermitian operators and some of their applications to the theory of orthogonal polynomials and the moment problem. *Uspekhi Mat. Nauk*, 2(3), 1947. in Russian.

[12] J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2012. doi:10.1007/978-1-4419-9982-5.

[13] B. O'Neill. *Semi-Riemannian geometry. With applications to relativity.*, volume 103 of *Pure and Applied Mathematics*. Academic Press, 1983.

[14] Y. Wong. Differential Geometry of Grassmann Manifolds. *Proceedings of the National Academy of Sciences*, 57(3):589–594, 1967.

Nomenclature

In contrast to the usual differential geometric nomenclature we consequently write arguments of functions in brackets instead of as lower indices, cp. evaluations of vector fields.

$\mathrm{Gr}(k,n)$	Grassmann manifold, set of all k-dimensional subspaces
A, B	$n \times k$-matrices with full rank, elements of $\mathrm{St}(k,n)$
$\lvert\cdot\rvert$	Euclidean vector norm on $\mathbb{R}^n$
$\lVert\cdot\rVert$	Frobenius norm on $\mathbb{R}^{n\times k}$
$\mathrm{St}(k,n)$	Stiefel manifold, set of all $n \times k$-matrices with full rank
$\mathrm{St}^*(k,n)$	compact Stiefel manifold, set of all orthonormal $n \times k$-matrices
π	span operation
$GL(k,\mathbb{R})$	general linear group of all $k \times k$ invertible real matrices
P, Q	invertible matrices, elements of $GL(k,\mathbb{R})$
$\overline{\pi}$	restriction of π to $\mathrm{St}^*(k,n)$
GS	Gram-Schmidt orthonormalization map
L, M	subspaces, elements of $\mathrm{Gr}(k,n)$
Θ	gap metric
Π_L	orthogonal projection onto L
$\mathcal{S}$	unit sphere in $\mathbb{R}^n$
$L^\perp$	orthogonal complement of L
$L \oplus M$	direct sum of L and M
$\widetilde{\Theta}$	spherical gap metric
S_A	local cross section through A
T_A	set of matrices which span a subspace without any orthogonal component to A
σ_A	cross section mapping
$A_\perp$	$n \times (n-k)$-matrix that is orthogonal to A
φ_A	chart on U_A
p_A	paramtrization for U_A
V_A	vertical space at A
H_A	horizontal space at A
$V\,\mathrm{St}(k,n)$	vertical bundle
$H\,\mathrm{St}(k,n)$	horizontal bundle
T_pM	tangent space of a manifold M in $p \in M$
TM	tangent bundle of a manifold M

$D\big\|_p f$	push-forward/differential of a differentiable function f in p
$\pi_{\mathrm{Gr}}, \pi_{\mathrm{St}}$	tangent bundle projections of $\mathrm{Gr}(k,n)$ and $\mathrm{St}(k,n)$, resp.
$\lozenge$	horizontal lift
$\mathcal{F}_p(M)$	set of functions germs in $p \in M$ on a manifold M
$\mathcal{F}(M)$	set of smooth real-valued functions defined on a manifold M
$\mathcal{X}(M)$	set of (smooth) vector fields on a manifold M
X, Y	derivations/vector fields
$[X, Y]$	Lie bracket of X and Y
$\langle \cdot, \cdot \rangle_L$	Riemannian metric on $T_L \,\mathrm{Gr}(k,n)$
tr	trace operator
$\mathrm{grad}_{\mathrm{Gr}}, \mathrm{grad}_{\mathrm{St}}$	gradients on $\mathrm{Gr}(k,n)$ and $\mathrm{St}(k,n)$, resp.
$\nabla(X, Y)$	(Riemannian) connection, covariant derivative of Y in the direction of X
∇_{St}	Riemannian connection on $\mathrm{St}(k,n)$, directional derivative
∇_{Gr}	Riemannian connection on $\mathrm{Gr}(k,n)$
C	regular curve on $\mathrm{Gr}(k,n)$
$\underset{C}{\nabla} X$	induced covariant derivative for vector fields along C, cp. $\frac{\nabla}{\mathrm{d}t}$ in [3, Definition 10.1.10]
c	(horizontal) regular curve on $\mathrm{St}(k,n)$
$D(C)$	domain of C