VOLUME OF THE STIEFEL $C$-MANIFOLD

ADRIAN CHUN PONG CHU

Abstract. We introduce the Stiefel $C$-manifolds, which can be seen as generalized spheres to the case of matrices. Given two $n$-spheres, the ratio of their volumes is well known to be the ratio of their radius to the $n$-th power. We are going to prove an analogous formula for the Stiefel $C$-manifolds.

1. Introduction

In this article we assume every number and matrix to be real, and denote by $\mathbb{R}^{n \times k}$ the set of $n \times k$ matrices.

We are all familiar with the $n$-sphere $S^n$ of radius $c$, it is the set
\[ \{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = c^2 \} \].

We can generalize this to the case of matrices as follows. Observe that the equation defining $S^n$ is $x^T x = c^2$, if we write $x = (x_1, \ldots, x_n)^T$.

So we naturally want to consider the set
\[ V_{n,C} := \{ X \in \mathbb{R}^{n \times k} : X^T X = C \} \],

where $n$ and $k$ are fixed positive integers and $C$ is a fixed $k \times k$ matrix. But there are three things to be aware of.

First, any matrix of the form $X^T X$ must be positive semi-definite. To see this, let $y \in \mathbb{R}^n$, then we have
\[ y^T X^T X y = (Xy)^T (Xy) = \langle Xy, Xy \rangle \geq 0. \]

So the given matrix $C$ should better be positive semi-definite. Second, if $C$ is merely positive semi-definite instead of positive definite, then $V_{n,C}$ might have singularities and thus fail to be a manifold. So we assume $C$ to be positive definite. Third, if $n < k$, the equation $X^T X = C$ might be over-determined and has no solution. So we will also assume $n \geq k$.

In conclusion, we have the following definition. Given positive integers $n \geq k$ and an $k \times k$ positive definite matrix $C$, we define $V_{n,C}$ to be the submanifold
\[ \{ X \in \mathbb{R}^{n \times k} : X^T X = C \} \]
of $\mathbb{R}^{n \times k}$. This manifold, called the Stiefel $C$-manifold, is our generalized sphere.

Notice that if $C$ is the $n \times n$ identity matrix $I_n$ then $V_{n,C} = SO(n)$, and if $C = I_1$ then $V_{n,C} = S^{n-1}$. Now what is the dimension of $V_C$? Each entry of $C$ gives one relation between the $nk$ entries of $X$. Recall that $C$ is symmetric, so it has $k(k+1)/2$ independent entries. Thus $\dim V_C$ should be $nk - k(k+1)/2$.

We may also ask, for which $C$ is the Stiefel $C$-manifold analogous to the unit sphere? The obvious choice would be the $I_k$, as we note $V_{n,I_k} = S^n$. For simplicity we write $V_{n,k} := V_{n,I_k}$ and just call it the Stiefel manifold.

For two $n$-spheres of radius 1 and $c$ respectively, their $n$-dimensional volumes differ by a factor, namely $c^n$. This is not the case for $V_{n,k}$ and $V_{n,C}$. The main goal of this article is to find out the ratio between their $(nk - k(k+1)/2)$-dimensional volumes, which we denote by $\text{vol} V_{n,k}$ and $\text{vol} V_{n,C}$ respectively.

1
Before stating the theorem, we must clarify what we mean by “volume”. We treat \( V_{n,k} \) as a submanifold of \( \mathbb{R}^{n \times k} \cong \mathbb{R}^{nk} \), and borrow the Riemannian metric from it. More precisely, given \( X = (X_{ij}) \) and \( Y = (Y_{ij}) \) in \( \mathbb{R}^{n \times k} \), their usual inner product \( (X, Y) \) is just \( \sum_{i,j} x_{ij} y_{ij} \), or \( \text{tr}(XY) \) equivalently. The induced Riemannian metric on \( V_{n,k} \) is thus \( (X, Y) = \text{tr}(X^TY) \), for any \( X_0 \in V_{n,k} \), and \( X \) and \( Y \) in the tangent space \( T_{X_0}V_{n,k} \). This metric will then give rise to a volume form.

**Theorem 1.1.** Given an \( k \times k \) positive definite matrix \( C \), the volume ratio \( \frac{\text{vol}V_{n,C}}{\text{vol}V_{n,k}} \) is equal to

\[
\frac{1}{2^{k(k-1)/4}} \sqrt{\det(C)^{n-k} \det(C \otimes I_k + I_k \otimes C)},
\]

where \( \otimes \) denotes the Kronecker product.

We remark that this ratio is indeed 1 when \( C = I_k \), and is \( C^{n-1} \) when \( k = 1 \).

**Acknowledgment.** The author wish to thank Alex Tung for the helpful discussions.

2. **Sketch of Proof**

We know that any two spheres with the same dimension and center, but different radius, can be identified via a scaling. But this is not the case for Stiefel \( C \)-manifolds. In fact, \( V_{n,k} \) and \( V_{n,C} \) is identified via the linear map \( F: \mathbb{R}^{n \times k} \to \mathbb{R}^{n \times k} \) defined by \( F(X) := X \sqrt{C} \), as explained in the following.

We first say a few words about \( \sqrt{C} \). Since \( C \) is positive definite, it can be diagonalized by some orthogonal matrix \( Q \) into the diagonal matrix \( \text{diag}(\lambda_1, \ldots, \lambda_k) \) consisting of the positive eigenvalues of \( C \). Hence we can define

\[
\sqrt{C} := Q^T \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_k}) Q \in \mathbb{R}^{k \times k},
\]

giving \( \sqrt{C}^2 = C \). Now for \( X \in V_{n,k} \),

\[
F(X)^T F(X) = \sqrt{C}^T X^T X \sqrt{C} = \sqrt{C} \sqrt{C} = C,
\]

verifying that \( F \) maps \( V_{n,k} \) into \( V_{n,C} \). Since \( F \) is a linear map in \( \mathbb{R}^{n \times k} \), it is a diffeomorphism between \( V_{n,k} \) and \( V_{n,C} \).

Then according to the area formula,

\[
\text{vol}V_{n,C} = \int_{V_{n,k}} \sqrt{\det(|D(F)|_{V_{n,k}}) \det(D(F)|_{V_{n,k}}))},
\]

Here \( D(F)|_{V_{n,k}} \) is the Jacobian of the restriction of \( F \) on \( V_{n,k} \).

Fix \( X_0 \in V_{n,k} \). We will construct an orthonormal basis \( B \) for the tangent space \( T_{X_0}V_{n,k} \). Then the above integrand evaluated at \( X_0 \) is equal to

\[
\frac{\text{volume of the parallelepiped spanned by } \{DF(W): W \in B\}}{\text{volume of the parallelepiped spanned by } B}
\]

Here \( DF \) is the Jacobian of \( F: \mathbb{R}^{n \times k} \to \mathbb{R}^{n \times k} \). It suffices compute this fraction, which, as we later show, is independent of the choice of \( X_0 \in V_{n,k} \). So \( \text{vol}V_{n,C} \) is equal to the above constant fraction times \( \text{vol}V_{n,k} \).

3. **Proof**

3.1. **Assuming \( C \) to be diagonal.** Without loss of generality, we can in fact assume \( C \) to be diagonal. To see this, we again let \( C = Q^T DQ \) where \( D \) is diagonal and \( Q \) is orthogonal. We can check as before that the linear map \( G: \mathbb{R}^{n \times k} \to \mathbb{R}^{n \times k} \)
defined by \( G(X) = XQ \) is a diffeomorphism between \( V_{n,D} \) and \( V_{n,C} \). Moreover, \( G \) preserves the inner product on \( \mathbb{R}^{n \times k} \). Namely,

\[
\langle G(X), G(Y) \rangle = \text{tr}(Q^T X^T Y Q) = \text{tr}(QQ^T X^T Y) = \text{tr}(X^T Y) = \langle X, Y \rangle.
\]

Therefore \( \text{vol} V_{n,D} = \text{vol} V_{n,C} \), and we may assume \( C \) to be a diagonal matrix \( \text{diag}(\lambda_1, ..., \lambda_k) \).

3.2. Setting up a basis. We now construct an orthonormal basis for the tangent space of \( V_{n,k} \) at a fixed point \( X_0 \). Note that the tangent space

\[
T_{X_0}V = \{ W \in \mathbb{R}^{n \times k} : W^T X_0 + WX_0^T = 0 \}.
\]

This can be seen by setting up a smooth curve \( X(t) \) on \( V_{n,k} \) with \( X(0) = X_0 \) and differentiating \( X(t)^T X(t) = I_k \) with respect to \( t \). We now let \( X_1, ..., X_k \) be the \( n \times 1 \) column vectors of \( X_0 \), and them to an orthonormal basis \( \{X_1, ..., X_k, X_{k+1}, ..., X_n\} \) of \( \mathbb{R}^n \). Also let \( e_i \) be the 1 \( \times \) \( k \) row vector \((0, ..., 1, ..., 0)\) with 1 at the \( i \)th place.

Let \( B_1 \) be the set consisting of the \((n \times k)\) matrices

\[
Y_{12} := X_2 e_1 - X_1 e_2, \quad Y_{13} := X_3 e_1 - X_1 e_3, ..., \quad Y_{1k} := X_k e_1 - X_1 e_k,
\]

\[
Y_{23} := X_3 e_2 - X_2 e_3, \quad Y_{24} := X_4 e_2 - X_2 e_4, ..., \quad Y_{2k} := X_k e_2 - X_2 e_k,
\]

\[
\ldots
\]

\[
Y_{(k-1)k} := X_k e_{k-1} - X_{k-1} e_k,
\]

and \( B_2 \) be the set consisting of the matrices

\[
Z_{(k+1)1} := X_{k+1} e_1, ..., \quad Z_{(k+1)k} := X_{k+1} e_k,
\]

\[
\ldots
\]

\[
Z_{nk} := X_n e_k.
\]

We claim that \( B := B_1 \cup B_2 \) is an orthogonal basis of \( T_{X_0}V_{n,k} \).

First note that the number of matrices in total is

\[
(k - 1) + (k - 2) + \cdots + 1 + k(n - k) = nk - k(k + 1)/2,
\]

the same as \( \dim T_{X_0}V \). The fact that they indeed lie in \( T_{X_0}V \) can be proven by substituting each of them into \( W^T X_0 + WX_0^T = 0 \). Orthogonality of \( B \) follows from the orthonormality of \( \{X_1, ..., X_n\} \). Hence \( B \) is a basis.

3.3. Volume of the parallelepiped spanned by \( B \). Since \( B \) is a orthogonal basis, the required volume is the product of norms of the basic elements. Each element of \( B_1 \), which has \( k(k-1)/2 \) elements in total, has norm \( \sqrt{2} \), and each element of \( B_2 \) has norm 1. The required volume is \( 2^{k(k-1)/4} \).

3.4. Volume of the parallelepiped spanned by \( \{DF(W) : W \in B\} \). Note that \( F \) is linear on \( \mathbb{R}^{n \times k} \), so \( DF(W) = F(W) = W \sqrt{C} \). Since each \( Y \in B_1 \) is orthogonal to \( Z \in B_2 \)

\[
\langle DF(Y), DF(Z) \rangle = \text{tr}(C^T Y^T Z C) = 0.
\]

So, the volume of the parallelepiped spanned by \( \{DF(W) : W \in B\} \) is equal to the product of that volume of the parallelepiped respectively spanned by \( \{DF(Y) : Y \in B_1\} \) and \( \{DF(Z) : Z \in B_2\} \).
3.5. Volume of the parallelepiped spanned by \( \{DF(Z) : Z \in B_2\} \). We can as before show that \( DF(Z_{pq}) \) and \( DF(Z_{rs}) \) is orthogonal if \( p \neq r \). Thus, the volume of the parallelepiped spanned by \( \{DF(Z) : Z \in B_2\} \) is equal to the product of parallelepipeds spanned by

\[
\{DF(Z_{(k+1)1}), ..., DF(Z_{(k+1)k})\}, ..., \{DF(Z_{n1}), ..., DF(Z_{nk})\}
\]

respectively.

Recall that given a set linear independent vectors \( v_1, ..., v_k \in \mathbb{R}^l \), the volume of the parallelepiped they span is equal to

\[
\sqrt{\det((v_i, v_j))},
\]

where \( \langle v_i, v_j \rangle \) is a \( k \times k \) matrix, called the Gram matrix of \( \{v_1, ..., v_k\} \).

We need to compute the Gram matrix of \( \{DF(Z_{p1}), ..., DF(Z_{pk})\} \) for each \( p = k + 1, ..., n \). Now

\[
\langle DF(Z_{pq}), DF(Z_{ps}) \rangle = \text{tr}(\sqrt{C} Z_{pq}^T Z_{ps} \sqrt{C}) = \text{tr}(C e_p^T X_p X_s e_s)
\]

where \( \delta_{pr} \) is the Kronecker delta. The last equality follows from that \( C = \text{diag}(\lambda_1, ..., \lambda_k) \).

Hence, the Gram matrix of \( \{DF(Z_{p1}), ..., DF(Z_{pk})\} \) is the \( k \times k \) matrix \( (\lambda_q \delta_{qs}) \), whose determinant is \( \prod_i \lambda_i \). It follows that the volume of the parallelepiped spanned by \( \{DF(Z) : Z \in B_2\} \) is \( (\prod_i \lambda_i)^{(n-k)/2} = \sqrt{\det C^{n-k}} \).

3.6. Volume of the parallelepiped spanned by \( \{DF(Y) : Y \in B_1\} \). We need to compute the Gram matrix of \( \{DF(Y) : Y \in B_1\} \), which should be a \( k(k-1)/2 \times k(k-1)/2 \) matrix. Now

\[
\langle DF(Y_{pq}), DF(Y_{rs}) \rangle
\]

\[
= \text{tr}(\sqrt{C} (e_p^T X_p^T - e_q^T X_p^T)(X_r e_r - X_s e_s)) \sqrt{C})
\]

\[
= \text{tr}(C(e_p^T e_r \delta_{qs} - e_q^T e_r \delta_{ps} + e_q^T e_s \delta_{pr}))
\]

\[
= \lambda_p \delta_{pr} \delta_{qs} - \lambda_p \delta_{qs} \delta_{pr} - \lambda_q \delta_{qr} \delta_{ps} + \lambda_q \delta_{qs} \delta_{pr}
\]

\[
= (\lambda_p + \lambda_q)(\delta_{pr} \delta_{qs} - \delta_{ps} \delta_{qr})
\]

\[
= (\lambda_p + \lambda_q)\lambda_p \delta_{pr} \delta_{qs}.
\]

The last equality follows from that \( Y_{pq} \) and \( Y_{rs} \) are never both defined. The Gram matrix required is therefore a diagonal matrix, with diagonal entries \( \lambda_p + \lambda_q \), where \( 1 \leq p < q \leq k \). Hence, the volume of the parallelepiped spanned by \( \{DF(Y) : Y \in B_1\} \) is

\[
\prod_{1 \leq p < q \leq k} \sqrt{\lambda_p + \lambda_q}.
\]

3.7. Final step. We now piece everything back together.

\[
\sqrt{\det((DF|1,\ldots,k)^T (DF|1,\ldots,k))} \text{ evaluated at } X_0
\]

\[
= \frac{\text{volume of the parallelepiped spanned by } \{DF(W) : W \in B\}}{\text{volume of the parallelepiped spanned by } B}
\]

\[
= \frac{\text{volume of the parallelepiped spanned by } \{DF(Y) : Y \in B_1\}}{\text{volume of the parallelepiped spanned by } \{DF(Z) : Z \in B_2\}}
\]

\[
= \left( \prod_{1 \leq i < j \leq k} \sqrt{\lambda_i + \lambda_j} \right) \sqrt{\det C^{n-k}/2^{k(k-1)/4}},
\]
which is independent of $X_0$. Finally, applying the area formula, we have

$$\frac{\text{vol} V_{n,C}}{\text{vol} V_{n,k}} = \frac{\sqrt{\det C^{n-k}}}{2^{k(k-1)/4}} \prod_{1 \leq i < j \leq k} \sqrt{\lambda_i + \lambda_j}.$$ 

And using the well known fact from linear algebra that

$$\det(C \otimes I_k + I_k \otimes C) = \prod_{1 \leq i < j \leq k} (\lambda_i + \lambda_j),$$

the proof is completed.

4. A Remark

One may wonder what the volume of $V_{n,k}$ actually is. The answer, as given in [1], is

$$2^k \pi^{k(2n-k+1)/4} \frac{\prod_{i=1}^{k} \Gamma((n - i + 1)/2)}{\prod_{i=1}^{k} \Gamma((n - i + 1)/2)},'$$

where $\Gamma$ is the Gamma function.

REFERENCES


Department of Mathematics, The University of Chicago, Chicago, IL 60637, USA

E-mail address: acpc@uchicago.edu