

# Grassmannians and Stiefel Manifolds.

We begin with matrices. As we know,

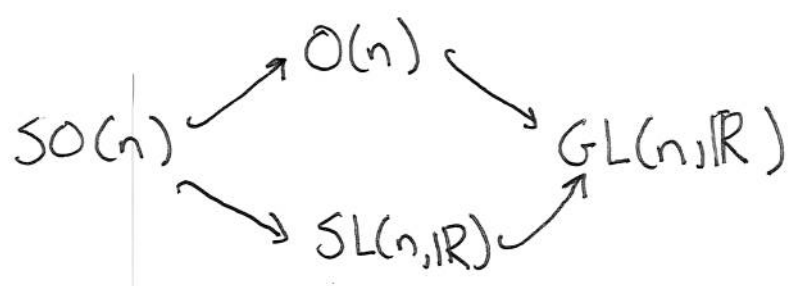
$GL(n, \mathbb{R})$  - invertible matrices ( $n \times n$ )

$SL(n, \mathbb{R})$  - matrices with determinant 1 ( $n \times n$ )

$O(n)$  - matrices with  $AA^T = I$  ( $n \times n$ )

$SO(n)$  - matrices with  $AA^T = I$  and  $\det +1$  ( $n \times n$ )

form a nested collection of groups, and of smooth manifolds:



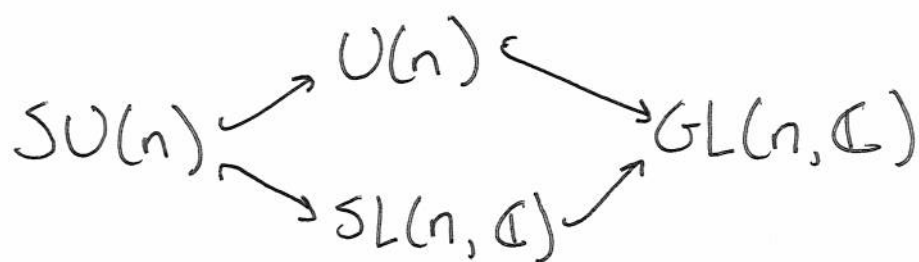
The orthogonal matrices have rows and columns forming orthonormal bases for  $\mathbb{R}^n$ , under the usual dot product.

(2)

If we recall the Hermitian dot product for complex vectors  $z$  in  $\mathbb{C}^n$ ,

$$\langle \vec{u}, \vec{v} \rangle = \sum u_i \bar{v}_i$$

we can define corresponding complex matrix groups



where  $SU(n)$  and  $U(n)$  are called unitary groups. Again, unitary matrices correspond to (Hermitian) orthonormal bases, this time for  $\mathbb{C}^n$ .

The situation is essentially the same over the quaternions, but the ~~special~~ groups of bases are symplectic groups.

3

We can think of  $O(n)$  as a sort of compact core of the (open) and much "floppier" space  $GL(n; \mathbb{R})$ .

Proposition.  $GL(n, \mathbb{R})$  deformation retracts onto  $O(n)$ .

Proof. The Gram-Schmidt process for orthonormalizing a basis is clearly continuous, preserves orthogonal matrices and can be done continuously.  $\square$

In fact, more is true, and we'll want to prove it. We'll need a cool idea from linear algebra. Recall that complex numbers can be written in the form

$$a + bi = re^{i\theta} \leftarrow \text{"polar form"}$$

where  $e^{i\theta}$  is a unit complex #  
corresponding to a rotation in the  
plane and  $r$  is a scaling. In fact,  
this generalizes to matrices.

(4)

Theorem (Polar decomposition, Gallier p.36 2.1.3,  
Halmos sec 83)

For any real  $n \times n$  invertible matrix  $A$ ,  
there is a unique orthogonal matrix  $R$   
and positive-definite symmetric matrix  $S$ , so that

$$A = RS$$

The same is true for complex invertible  
matrices, but  $R$  is unitary instead  
of orthogonal.

We also need one more idea.

A positive-definite matrix, like a positive number, has a kind of "square root."

Theorem (Cholesky decomposition, Demmel)

Any symmetric positive definite matrix  $S$  can be written uniquely as  $S = LL^*$  where  $L$  is a lower-triangular matrix with strictly positive diagonal entries.

~~Lemma~~ This is true for both real and complex valued  $S$ , but  $L$  is real valued  $\Leftrightarrow S$  is. Combining these we see

Theorem. Up to homeomorphism,

$$GL(n, \mathbb{R}) = O(n) \times \mathbb{R}^{\frac{n(n+1)}{2}}$$

$$GL(n, \mathbb{C}) = U(n) \times \mathbb{C}^{\frac{n(n+1)}{2}}$$

The proof is kind of obvious at this

6

point, with the observations that

~~the number of nonzero~~

$\frac{n(n+1)}{2}$  = # of nonzero elements in an  $n \times n$  lower triangular matrix and  $(0, \infty) \cong (-\infty, \infty)$  up to homeomorphism.

This tells us that  $O(n)$  and  $U(n)$  really are rigid, in the sense that we can define a projection from  $GL(n, \mathbb{R})$  or  $GL(n, \mathbb{C})$  and understand the fiber.

~~\*~~

Now if  $n$ -frames in  $\mathbb{R}^n$  are special, what about  $K$ -frames? Here we meet

Definition. The Stiefel manifold  $V_k(\mathbb{R}^n)$  or  $V_k(\mathbb{C}^n)$  is the manifold of orthonormal (or Hermitian orthonormal)  $K$ -frames in  $\mathbb{R}^n$  or  $\mathbb{C}^n$ .

⑦

~~We can generalize the polar decomposition to this case, by writing.~~

Clearly,  $V_n(\mathbb{R}^n) = O(n)$ ,  $V_n(\mathbb{C}^n) = U(n)$ , so we understand these in depth. But what about  $V_k(\mathbb{R}^n)$  and  $V_k(\mathbb{C}^n)$ ?

Lemma.  $V_k(\mathbb{R}^n) \cong \{ A \in n \times k \text{ matrices} \mid A^T A = I_k \}$

Proof.  $A^T A$  is the Gram matrix of the columns: the matrix of dot products.

It is  $= I_k \iff$  the columns are an orthonormal  $k$ -frame.  $\square$

A similar statement holds for  $V_k(\mathbb{C}^n)$ .

We then have:

Theorem. For every ~~rank  $k$~~   $n \times k$  matrix  $A$  of full rank (that is, rank  $k$ ), there are ~~is a~~ unique matrices  $R \in V_k(\mathbb{R}^n)$ ,  $S$  in symmetric positive definite  $k \times k$  matrices so  $A = RS$ .

Using the Cholesky decomposition as before, we get that

Theorem. Up to homeomorphism,

$$\text{rank } K, n \times K \text{ matrices} = \begin{matrix} V_K(\mathbb{R}^n) \times \mathbb{R}^{\frac{K(K+1)}{2}} \\ V_K(\mathbb{C}^n) \times \mathbb{C}^{\frac{K(K+1)}{2}} \end{matrix}$$

depending on whether they are real or complex.

This gives us some immediate insight into the structure of the Stiefel manifolds: they are essentially determined by linear independence of their vectors, not by orthonormality, at least in a very strong topological sense.



Further, this gives us our first numerical insight into the Stiefel manifolds.

Definition. If  $A$  is an  $n \times m$  matrix, let the Frobenius norm of  $A$  be defined by  $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\text{trace}(A^*A)}$ .

Theorem. Given an ~~n~~  $n \times k$  matrix  $A$  of rank  $k$ , the nearest matrix (in the Frobenius norm) to  $A$  is the matrix  $R$  in  $V_k(\mathbb{R}^n)$  given by the polar decomposition.

The same statement holds for  $V_k(\mathbb{C}^n)$ . This means that we have a natural numerical method for correcting any errors which creep into our orthonormal frames during a computation: compute the polar to project back!