Matrix Groups

Vyas Krishnamurthy

1 Prelims

The text works in three generalized spaces, $\mathbb{R}^n$, $\mathbb{C}^n$ and $\mathbb{H}^n$. As usual $\mathbb{R}$ is the set of real numbers and $\mathbb{R}^n$ is the set of n-tuples with real-valued entries and $\mathbb{C}^n$ is similarly the set of n-tuples with complex valued entries. The other space is pretty new; $\mathbb{H} = \{a + bi + cj + dk : a, b, c, d \in \mathbb{R}\}$ is an extension of the complex numbers where $i^2 = j^2 = k^2 = -1$ and $\mathbb{H}^n$ is the usual n-tuple. I'll follow the incredible confusing but sometimes useful notation that Morton Curtis used and shorthand $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ with k^n as the set of n-tuples.

2 Inner Products

In order to understand a little bit more about Orthogonal Groups, and to get through some topology on Matrix Groups later on, we'll need the concept of an inner product. The particular inner product that we're interested in is defined as follows:

Definition 1. For vectors $x, y \in k^n$ we define an inner product $\langle x, y\rangle := \sum_{i=1}^{n} x_i\overline{y_i}$.

This inner product may seem intimidating, but if we're looking at vectors in only $\mathbb{R}^n$ then it's just the usual dot product. The really nice property of the dot product is that for $x \in \mathbb{R}^n$ the length of the vector, or the norm, is simply $\|x\| = \sqrt{x \cdot x}$. This particular inner product generalizes the dot product to both $\mathbb{C}$ and $\mathbb{H}$ while keeping the above definition of length to be a real number. Due to the way that this inner product is defined, we can deduce a property of the inner product that comes in handy when defining Orthogonal Groups.

Proposition 1. For all $x, y \in k^n,$ and $A \in \mathrm{M_n}(k)$ we have $\langle xA, y\rangle = \langle x, y\overline{A}^T\rangle$.

This fact is used very briefly later on, but we can actually turn our groups into metric spaces by defining the metric $\mathrm{dist}(x, y) = \|x - y\|$ where $x, y \in k^n$.

3 Orthogonal Groups

Definition 2. The **Orthogonal Group**, $\mathcal{O}(n, k)$, is the set of matrices $A \in \mathrm{M_n}(k)$ such that for all $x, y \in k^n$ we have $\langle xA, yA\rangle = \langle x, y\rangle$.

Applying Proposition 1 to the definition, $\langle xA, yA\rangle = \langle x, yA\overline{A}^T\rangle = \langle x, y\rangle$, so we see that $yA\overline{A}^T = y$ or $A\overline{A}^T = I$. This is a definition that you may have seen before in a linear algebra or otherwise for orthogonal matrices, but the inner product definition gives us some nice intuition on how matrices in the Orthogonal Group act on vectors. Take $x \in k^n$ with $A \in \mathcal{O}(n, k)$, then $\|xA\| = \sqrt{\langle xA, xA\rangle} = \sqrt{\langle x, x\rangle} = \|x\|$ which is just the length of x. This means that the orthogonal group is the set of matrices that preserves the length of a vector under transformation. In fact consider for $x, y \in \mathbb{R}^n$ we have that $\cos(\theta) = \frac{x \cdot y}{\|x\|\|y\|}$ where θ is the angle between the vectors. If we take $A \in \mathcal{O}(n, \mathbb{R})$, then $\cos(\theta) = \frac{\langle xA, yA\rangle}{\|xA\|\|yA\|} = \frac{\langle x, y\rangle}{\|x\|\|y\|}$ so that even the angle between the two vectors is preserved. We will work with a few orthogonal groups in particular:

Definition 3. The **Orthogonal Group**, $\mathrm{O}(n)$, is the set of matrices $A \in \mathrm{M_n}(\mathbb{R})$ such that $AA^T = A^TA = I$. The subgroup of matrices $A \in \mathrm{O}(n)$ such that $\det(A) = 1$ is called the **Special Orthogonal Group**, $\mathrm{SO}(n)$.
The **Unitary Group**, $\mathrm{U}(n)$, is the set of matrices $A \in \mathrm{M_n}(\mathbb{C})$ such that $A\overline{A}^T = \overline{A}^T A = I$. The subgroup of matrices $A \in \mathrm{U}(n)$ such that $\det(A) = 1$ is called the **Special Unitary Group**, $\mathrm{SU}(n)$.
The **Symplectic Group**, $\mathrm{Sp}(n)$, is the set of matrices $A \in \mathrm{M_n}(\mathbb{H})$ such that $A\overline{A}^T = \overline{A}^T A = I$.

Tangent Spaces and the Exponential Map

Another set of preliminaries here, we need to understand the concept of tangent spaces to matrix groups, then we can use this to investigate the exponential and logarithm of a matrix, as well as some interesting properties of the tangent space itself.

Definition 4. A **curve** on a vector space V is a continuous function γ from an open interval $(a, b) \in \mathbb{R}$ to V.

Definition 5. We say that a curve γ on a vector space V is **differentiable** at a point $c \in (a, b)$ if

$$\lim_{x \to c} \frac{\gamma(x) - \gamma(c)}{x - c}$$

exists.

The idea of a curve on a matrix group may seem strange, but to make it a bit more concrete we can look at an example of one. For $t \in (-\pi, \pi)$ define

$$\gamma(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix}.$$

The first thing to note is for all $t \in (-\pi, \pi)$, $\det(\gamma(t)) = 1$ so that $\gamma(t)$ is a curve on $\mathrm{SO}(2)$. The derivative of $\gamma(t)$ evaluated at 0 is

$$\gamma'(0) = \begin{bmatrix} -\sin(0) & \cos(0) \\ -\cos(0) & -\sin(0) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

What we're really trying to accomplish here is extend familiar results from calculus and vector calculus to spaces these spaces of matrix groups, with curves on the matrix group playing the role of a function on x.

Definition 6. The **Tangent Space**, T, to a matrix group, G, is the set of all tangent vectors $\gamma'(0)$ to curves $\gamma : (a, b) \to G$ where $\gamma(0) = I$.

Proposition 2. T is a **linear subspace** of $\mathrm{M_n}(k)$.

Definition 7. **(i)** Matrix $A \in \mathrm{M_n}(\mathbb{R})$ is **skew-symmetric** if $A + A^T = 0$.

(ii) Matrix $A \in \mathrm{M_n}(\mathbb{C})$ is **skew-hermitian** if $A + A^T = 0$.

(iii) Matrix $A \in \mathrm{M_n}(\mathbb{H})$ is **skew-symplectic** if $A + A^T = 0$.

Proposition 3. Let $\mathrm{so}(n)$ be the set of skew-symmetric matrices. Then $\mathrm{so}(n)$ is a subspace of $\mathrm{M_n}(\mathbb{R})$.

This result follows similarly for su(n) and sp(n), the set of skew-hermitian and skew-symplectic matrices respectively. We'll now prove a fairly straightforward result that is leads us part-way to the goal of this presentation.

Theorem 1. Let $\gamma(t)$ be a curve through $\mathrm{O}(n)$, then $\gamma'(0) \in \mathrm{so}(n)$.

Proof. We simply want to show here that if $\gamma(t)$ is a curve through $\mathrm{O}(n)$ then it is implied that $\gamma'(0) + \gamma'(0)^T = 0$ so that $\gamma'(0) \in \mathrm{so}(n)$. By definition of an orthogonal matrix $\gamma(t)\gamma(t)^T = I$. Taking the derivative of both sides with respect to t using the usual product rule then evaluating at 0 we have

$$\gamma'(0)\gamma(0)^T + \gamma(0)\gamma'(0)^T = 0.$$

Again by definition $\gamma(0)^T = \gamma(0) = I$ thus we have $\gamma'(0) + \gamma'(0)^T = 0$ so that $\gamma'(0) \in \mathrm{so}(n)$. In particular if $\gamma(t)$ is a curve through $\mathrm{SO}(n)$ then $\gamma'(t)$ is a curve through $\mathrm{so}(n)$ as well. □

Definition 8. Let $A \in \mathrm{M_n}(\mathbb{R})$ and set

$$e^A := I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$$

We say that this series **converges** if the sequence of partial sums converges for each of the n^2 matrix entries.

Proposition 4. For any matrix $A \in \mathrm{M_n}(\mathbb{R})$, the series

$$e^A = I + \sum_{i=1}^{\infty} \frac{A^i}{i!}$$

converges.

Proof. Let m be the largest entry in the matrix A. Take matrix M defined as the n by n matrix with all entries as m, then each entry in M^2 will be nm^2 and clearly $a_{ij} \leq nm^2$ where a_{ij} is an arbitrary entry of A. Continue on in this way to see that the entries in A^k are bounded above by $n^{k-1}m^k$ in M^k, so it suffices to show that the exponential of M converges. Take the expansion deduced above

$$1 + \sum_{i=0}^{\infty} \frac{n^{i-1}m^i}{i!}.$$

We can apply the ratio test to this series as follows:

$$\limsup_{i\to\infty} \left|\frac{a_{i+1}}{a_i}\right| = \limsup_{i\to\infty} \left|\frac{n^i m^{i+1}(i-1)!}{n^{i-1}m^i i!}\right| = \limsup_{i\to\infty} \frac{nm}{i} = 0.$$

Thus we have that the radius of converge of the exponential function is infinite and e^A converges entry-wise for all $A \in \mathrm{M_n}(\mathbb{R})$. □

Proposition 5. If the matrices A and B commute, then $e^{A+B} = e^A e^B$.

Corollary 1. For any real n×n matrix A, e^A is invertible.

Proposition 6. If $A \in \mathrm{so}(n)$ then e^A is orthogonal.

Proof. By definition, $A \in \mathrm{so}(n) \Rightarrow A + A^T = A^T + A = 0$, then by Proposition 5 $e^{A+A^T} = e^A e^{A^T} = e^0 = I$. The exponential is defined entrywise, so $e^{A^T} = (e^A)^T$ thus $e^A(e^A)^T = (e^A)^T e^A = I$ so we have our result, $e^A \in \mathcal{O}(n, \mathbb{R})$. □

Definition 9. A set $D \in \mathbb{R}^n$ is **path connected** if: Given arbitrary $x, y \in D$ there exists a continuous function $\gamma : [0, 1] \to D$ with $\gamma(0) = x$ and $\gamma(1) = y$.

The concept of path connectedness gives us one really interesting result: $\mathrm{SO}(n)$ is path connected, but $\mathrm{O}(n)$ is not!. In order to reach a contradiction assume that $\mathrm{O}(n)$ is in fact path connected. Choose $A \in \mathrm{SO}(n), B \in \mathrm{O}(n) - \mathrm{SO}(n)$. By assumption $\mathrm{O}(n)$ is path connected, so there exists a continuous function $\gamma : [0, 1] \to \mathrm{O}(n)$ such that $\gamma(0) = A$ and $\gamma(1) = B$. Recalling a fact from analysis, the composition of continuous functions is similarly continuous, so we are going to now look at the continuous function $h(t) = (det \circ \gamma)(t)$, or the composition of the determinant function with our path. Then the continuous image of a path connected space is path connected, so there must be some $A \in \mathrm{O}(n)$ with determinant 0, which is clearly a contradiction.
Note: Won't do this proof in full, just in case someone asks I wanted to have the full one available.

Proposition 7. The tangent space of $\mathrm{SO}(n)$ is $\mathrm{so}(n)$.

Proof. We've actually already proved the bulk of what is necessary here. We've shown that the derivative of a curve through $\mathrm{SO}(n)$ is in fact an element of $\mathrm{so}(n)$, so all that is left to show is that any curve in $\mathrm{so}(n)$ is tangent to some curve through $\mathrm{SO}(n)$ which will give us that the tangent space of $\mathrm{SO}(n)$ is ALL of $\mathrm{so}(n)$. Take $B \in \mathrm{so}(n)$ to be arbitrary. Consider the map $\rho : [0, 1] \to \mathrm{SO}(n)$ where $\rho(t) = e^{tB}$. This map is a path between I and e^B in $\mathrm{SO}(n)$. By path connectedness, for any matrix e^{tB} we must have $\det(e^{tB}) = 1$ thus we have our result. □