A QUICK AND PRATICAL INTRO TO MATRIX LIE GROUPS

ADAM WILLIAMS

Contents

1. Matrix Lie Groups and Lie Algebras

1.1. **Matrix Groups as Transformation Groups.** A **matrix Lie group** is a differentiable manifold whose elements form a matrix group. Key examples include:

- GL(n), the group of invertible linear transformations on $\mathbb{R}^n$.
- SO(n), the group of $n \times n$ rotation matrices.
- SE(n), the group of $(n+1) \times (n+1)$ matrices of the form
$$x = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix},$$
where $R \in \mathrm{SO}(n)$ and $t \in \mathbb{R}^n$, viewed as the set of rigid transformations on $\mathbb{R}^n$ using homogeneous coordinates.

A matrix Lie group G of $n \times n$ matrices has a natural **group action** on $\mathbb{R}^n$, by matrix multiplication: $p \mapsto xp$ for $x \in G$. We view the action of a matrix x as a **change-of-coordinate-frame**, taking a point p_A expressed with respect to some frame A, to *the same point*, but expressed with respect to a second frame, B. When it helps with clarity, we augment the notation to represent this viewpoint: We write a matrix $x \in G$ as x_{BA}, and write
$$p_B = x_{BA} p_A.$$
for $p_A \in \mathbb{R}^n$.

Transforms chain and invert in the natural way:
$$x_{CA} = x_{CB} x_{BA}$$
and
$$x_{BA}^{-1} = x_{AB}.$$

1.2. **Derivatives of Time-Parameterized Transforms (The Wrong Way).** Consider a **time-parameterized transform**, or, a curve $x : \mathbb{R} \to G$. Naively, one might think that the most "natural" way to consider the derivative of x is to simply take the coordinate-wise derivative of the matrix, $\dot{x}$. This is undesirable for (at least) one simple reason: we want the derivatives of curves to be vectors, and the naive matrix curve derivatives do not form a vector space! These quantities are closed under scaling, as
$$\frac{d}{dt} x(at) = a\dot{x},$$
but since $x + y \notin G$ in general, they are not closed under addition. That is, the quantity $\dot{x} + \dot{y}$ is generall not the derivative of any curve in G.

Now, without motivation, say we consider the "true" derivative to of a curve to be the derivative of the matrix, right-mulitplied by the inverse of the matrix: $\dot{x}x^{-1}$. Clearly, these quantities are still closed under scaling, but now they are also closed under addition: A simple calculation shows that, if $x, y : \mathbb{R} \to G$, then the sum of their derivatives
$$\dot{x}x^{-1} + \dot{y}y^{-1}$$

at $t = t_0$ is actually the derivative of the curve

$$t \mapsto x(t)x(t_0)^{-1}y(t)$$

at $t = t_0$. Hence we have a notion of derivative where the "tangent vectors" to curves actually form a vector space, which provides motivation from a mathematical perspective. In the next section, we will derive this notion of derivative $\dot{x}x^{-1}$ (and its symmetric counterpart, $-x^{-1}\dot{x}$) from physical intuition.

1.3. **Derivatives of Time-Parameterized Transforms (The Right Way).** Consider a **time-parameterized transform**, or, a curve $x_{BA} : \mathbb{R} \to G$. We can motivate the notion of derivative for such a curve by asking, "what instantaneous velocity does this transform induce on points via the group action?" That is, we can view the "velocity" of a transform implicitly as a a **velocity field on the points it transforms**, assigning to each point in space the vector rooted at that point indicating the instantaneous direction in which the transform sends that point.

We can derive such a velocity field two different ways, by taking two different viewpoints of the transformation. For any $p \in \mathbb{R}^n$, we have the relation

$$p_B = x_{BA}p_A$$

for all $t \in \mathbb{R}$.

Now we can either take a **right-invariant view**, where we view the *left* (output) frame as moving with time, and the right (input) frame as static; or the **left-invariant view**, where we view the right frame as static and the left frame as moving. In the right-invariant view, the quantity p_A is fixed, while p_B is a function of time, so differentiating yields

$$\begin{aligned}\dot{p}_B &= \dot{x}_{BA}p_A \\ &= \dot{x}_{BA}x_{BA}^{-1}x_{BA}p_A \\ &= \left[\dot{x}_{BA}x_{BA}^{-1}\right]p_B.\end{aligned}$$

Hence we have a differential equation

$$\dot{p}_B = \left[\dot{x}_{BA}x_{BA}^{-1}\right]p_B$$

describing the velocity field expressed in B coordinates, at any given time.

Dually, if we take the left-invariant view, the quantity p_B is fixed and p_A varies with time, so differentiating yields

$$0 = \dot{x}_{BA}p_A + x_{BA}\dot{p}_A,$$

and we can solve for

$$\dot{p}_A = -x_{BA}^{-1}\dot{x}_{BA}p_A,$$

giving us a velocity field in the right frame.

The term *left-* (resp. *right-*) *invariant* is justified by noting that the value of $\dot{x}x^{-1}$ (resp. $x^{-1}\dot{x}$) is not changed when x is right- (resp. left-) multiplied by a constant group element c.

The **Lie algebra** $\mathfrak{g}$ is a vector space consisting of all such **infinitesimal transformation** matrices, or, derivatives of smooth curves in G. The left- and right-invariant views are equivalent and dual to each other:

$$\begin{aligned}\mathfrak{g} &= \left\{\dot{x}x^{-1} : x : \mathbb{R} \to G\right\} \\ &= \left\{-x^{-1}\dot{x} : x : \mathbb{R} \to G\right\}\end{aligned}$$

When we restrict to derivatives at the identity element $e \in G$, the left- and right-invariant views become identical, as $x^{-1} = e$. Hence to avoid making the arbitrary choice of left versus right, is it standard to define the Lie algebra as the **tangent space at the identity**:

$$\mathfrak{g} = \{\dot{x}(0) : x : \mathbb{R} \to G, \;\; x(0) = e\}$$

Hence the Lie algebra consists of matrices of the same size as elements of G, and has the same (vector space) dimension as the (manifold dimension of the) group. The Lie algebra structure can be computed by differentiating the group. $R \in \mathrm{SO}(n)$ satisfies $RR^\top = R^\top R = I$, differentiating this yields

$$\dot{R}R^\top + R\dot{R}^\top = 0,$$

meaning that the elements $\dot{R}R^\top$ of the Lie algebra $\mathfrak{so}(n)$ are skew-symmetric matrices.

1.4. **Example:** SO(3)**.** Consider the 3D rotation group SO(3). Let R be a time-parameterized rotation. Then the right-invariant velocity field induced on $\mathbb{R}^3$ at an instant is given by a skew-symmetric matrix: $p \mapsto \Omega p$ for $\Omega = \dot{R}R^\top \in \mathfrak{so}(3)$. For $n = 3$, set of $n \times n$ skew-symmetric matrices is isomorphic to $\mathbb{R}^3$ by the mapping

$$\begin{aligned}\mathbb{R}^3 &\cong \mathfrak{so}(3) \\ \omega &\mapsto \omega_\times,\end{aligned}$$

where $\omega_\times$ is the matrix for the linear mapping given by the cross product with ω:

$$\omega_\times p = \omega \times p.$$

Letting $\omega_\times = \Omega$, and noting that $\Omega\omega = \omega \times \omega = 0$, we can see that the instantaneous velocity of $\omega \in \mathbb{R}^3$ induced by R is zero. Hence ω is the **instantaneous axis of rotation** of R (expressed in the left frame), and the velocity field induced by R is given by the cross product with this axis:

$$\dot{p} = \omega \times p.$$

1.5. **Example:** SE(3)**.** Now consider the rigid transformation group SE(3) on $\mathbb{R}^3$, consisting of 4×4 matrices

$$x = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix},$$

This group acts on homogeneous vectors $\left(\begin{smallmatrix} p^\top & 1 \end{smallmatrix}\right)^\top$:

$$p \mapsto x \begin{pmatrix} p \\ 1 \end{pmatrix} = Rp + t.$$

The inverse of x is

$$x^{-1} = \begin{pmatrix} R^\top & -R^\top t \\ 0 & 1 \end{pmatrix},$$

Hence elements of the Lie algebra $\mathfrak{se}(3)$ have the form

$$\begin{aligned} \dot{x}x^{-1} &= \begin{pmatrix} \dot{R} & \dot{t} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R^\top & -R^\top t \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \dot{R}R^\top & -\dot{R}R^\top t + \dot{t} \\ 0 & 0 \end{pmatrix}. \\ &= \begin{pmatrix} \omega_\times & -\omega \times t + \dot{t} \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence the instantaneous velocity field (in the left frame) is given by a rigid transformation with rotational component R and translation t is

$$\dot{p} = \omega \times (p - t) + \dot{t}.$$

Note that the general form of an $\mathfrak{se}(3)$ element is

$$\begin{pmatrix} \omega_\times & v \\ 0 & 0 \end{pmatrix},$$

so we have an ismorphism with $\mathbb{R}^6$ by concatenating ω and v. We will use ξ to denote such a 6-vector representing an $\mathfrak{se}(3)$ element.

1.6. **The Lie Bracket.** We haven't yet discussed what makes the Lie algebra an **algebra**. The answer is that the Lie algebra is endowed with an antisymmetric vector product, called the **Lie bracket**, defined as

$$[X, Y] = XY - YX$$

for $X, Y \in \mathfrak{g}$.

The Lie bracket on $\mathfrak{so}(3)$ is given by the cross product:

$$[\omega_\times, \eta_\times] = (\omega \times \eta)_\times,$$

whereas the bracket of $\mathfrak{se}(3)$ elements is given by

$$[\xi_1, \xi_2] = \begin{pmatrix} \omega_1 \times \omega_2 \\ \omega_1 \times v_2 - \omega_2 \times v_1 \end{pmatrix}.$$

2. The Adjoint Representation

In general, if we have a transform T_A on a vector space, expressed in a coordinate system A, we change the coordinate representation of T_A to a coordinate system B by **conjugation** by the change-of-coordinates x_{BA} from A to B:

$$T_B = x_{BA} T_A x_{BA}^{-1},$$

so that if

$$q_A = T_A p_A,$$

then

$$q_B = \left[x_{BA} T_A x_{BA}^{-1}\right] p_B.$$

Changing the coordinates of a Lie algebra element is of particular signficance; the map $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ taking $x \in G$ to the linear isomorphism "conjugate by x" is called the **adjoint representation** of the group G:

$$\mathrm{Ad}_x X = x X x^{-1}.$$

For instance, the (negative) adjoint maps between the left- and right-invariant derivatives of a time-parameterized transform:

$$\begin{aligned} -x^{-1}\dot{x} &= -\mathrm{Ad}_{x^{-1}}\left[\dot{x}x^{-1}\right] \\ \dot{x}x^{-1} &= -\mathrm{Ad}_x\left[-x^{-1}\dot{x}\right]. \end{aligned}$$

The map Ad is a *group homomorphism*, so that

$$\mathrm{Ad}_{xy} = \mathrm{Ad}_x \mathrm{Ad}_y$$

and

$$\mathrm{Ad}_{x^{-1}} = \mathrm{Ad}_x^{-1}.$$

The adjoint representation as a function on the 3-vector realization of $\mathfrak{so}(3)$ (given by $\omega \mapsto \omega_\times$) is simply multiplication by the respective matrix:

$$\mathrm{Ad}_R = R.$$

We say that SO(3) is **self-adjoint**. For the 6-vector realization of SE(3), it is

$$\mathrm{Ad}_x = \begin{pmatrix} R & 0 \\ t_\times R & R \end{pmatrix},$$

so that

$$\mathrm{Ad}_x \xi = \begin{pmatrix} R\omega \\ t \times (R\omega) + Rv \end{pmatrix},$$

3. Differential Calculus

3.1. **Left- and Right-Invariant Pushforwards.** The **pushforward** of a function $f : G \to H$ between Lie groups is a linear mapping from $\mathfrak{g} \to \mathfrak{h}$ which takes derivatives of a curve x in G to derivatives of the curve $f \circ x$ in H. For the right-invariant view, this means

$$f_*\left[\dot{x}x^{-1}\right] = [f(\dot{x})] f(x)^{-1}.$$

That is, the pushforward of a velocity field $\dot{x}x^{-1}$ is the velocity field given by the same curve x mapped through f. Note that here there is an implicit "base point" x for the pushforward (i.e. we could write $(f_*)_x$, but the base point x is already expressed by the $\dot{x}x^{-1}$ notation).

We can check basic properties: if $f : G \to H$ is constant, then clearly, $f_* = 0$, and if $f : G \to G$ is the identity, then the derivative is the identity as well. These properties, together with the **product rule for Lie groups**, allow us to compute derivatives of more complicated functions between Lie groups. The product rule can be derived as follows:

$$\begin{aligned}(fg)_* \left[\dot{x}x^{-1}\right] &= [f(x)\dot{g}(x)] \left[f(x)g(x)\right]^{-1} \\ &= \left([\dot{f(x)}]g(x) + f(x)[\dot{g(x)}]\right) g(x)^{-1}f(x)^{-1} \\ &= [\dot{f(x)}]f(x)^{-1} + f(x)[\dot{g(x)}]g(x)^{-1}f(x)^{-1} \\ &= f_* \left[\dot{x}x^{-1}\right] + \mathrm{Ad}_{f(x)}g_* \left[\dot{x}x^{-1}\right].\end{aligned}$$

Hence we have the **right-invariant product rule**:

$$(fg)_* = f_* + \mathrm{Ad}_f g_*.$$

Dually, we can take the left-invariant view and get an analogous definition:

$${}_*f\left[-x^{-1}\dot{x}\right] = -f(x)^{-1}[\dot{f(x)}].$$

Note that we moved the $_*$ to the other side of f to disambiguate the left vs the right pushforward. This yields a **left-invariant product rule**:

$${}_*(fg) = \left[\mathrm{Ad}_g^{-1}\right] {}_*f + {}_*g.$$

Naturally, the left- and right-invariant pushforwards are invariant under left- and right-translation of the function, respectively. That is, if $c \in G$ is constant, then

$$(fc)_* = f_*$$

and

$${}_*(cf) = f.$$

3.2. **Example: Differentiating the Group Inverse.** The pushforwards of the group inverse follows from the product rule:

$$\imath_* = -\mathrm{Ad}_\imath$$

and

$${}_*\imath = -\mathrm{Ad}_{\mathrm{id}_G}.$$

where $\imath(x) = x^{-1}$.

3.3. **Example: Differentiating the Group Action.** Let $A_p(x) = xp$ be the group action on a point $p \in \mathbb{R}^n$. Then

$$(A_p)_*\dot{x}x^{-1} = [\dot{xp}] = \dot{x}x^{-1}xp = \dot{x}x^{-1}A_p(x).$$

(note that we can drop the $-A_px$ on the right-hand side since the output space is a vector space). Hence the (right-invariant) pushforward of the group action with respect to the group is the mapping "right-multiply by the transformed point".

For example, consider SO(3). The pushforward of the action A_p with respect to a roation R is

$$\omega_\times \mapsto \omega_\times Rp = -(Rp)_\times \omega.$$

So in the 3-vector representation, the matrix for the pushforward of A_p evaluated at group element R is simply $-(Rp)_\times$.

4. The Exponential Map

4.1. Lie Algebra Elements as Infinitesimal Generators. Consider a time-parameterized transform M_{BA} whose instantaneous velocity field is constant:

$$\dot{p}_B(t) = Xp_B(t),$$

where $X \in \mathfrak{g}$ does not vary with time. Then the velocity field is given by a time-independent homogeneous linear ODE, for which the solution is well-known: we can write x_B over time as

$$p_B(t) = \exp(tX)p_B(0),$$

where exp is the **matrix exponential**, given by

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

When viewed as a mapping $\exp : \mathfrak{g} \to G$, the matrix exponential is called the **exponential map**. For a Lie algebra element X, $\exp(tX)$ transforms points by moving according to the **flow** determined by the velocity field $x \mapsto Xx$ for time t.

4.2. The Baker-Campbell-Hausdorff Formula.

4.3. Example: SO(3)**.** For certain Lie groups, the exponential map has a closed form. Consider SO(3). The exponential map is

$$\exp(\omega) = \sum_{k=0}^{\infty} \frac{\omega_\times^k}{k!}.$$

Writing $\theta = \|\omega\|$ and $\hat{\omega} = \omega/\theta$, we can apply properties of the cross product to see that $\hat{\omega}_\times^3 = -\hat{\omega}_\times$, and so higher powers of $\hat{\omega}_\times$ in the power series collapse:

$$\begin{aligned}\hat{\omega}_\times^3 &= -\hat{\omega}_\times \\ \hat{\omega}_\times^4 &= -\hat{\omega}_\times^2 \\ \hat{\omega}_\times^5 &= \hat{\omega}_\times.\end{aligned}$$

Hence we can group all of the power series terms into a $\hat{\omega}_\times$ term and a $\hat{\omega}_\times^2$ term:

$$\begin{aligned}\exp(\omega) &= I + \left(1 - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots\right)\hat{\omega}_\times + + \left(\frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \frac{\theta^6}{6!}\right)\hat{\omega}_\times^2 \\ &= I + \left(\frac{\sin\theta}{\theta}\right)\omega_\times + \left(\frac{1-\cos\theta}{\theta^2}\right)\omega_\times^2.\end{aligned}$$

Hence we have a practical formula for the exponential map on SO(3). The geometric interpretation of the exponential map on SO(3) is that $\exp(\omega)$ is the (right-handed) rotation of θ radians about the axis $\hat{\omega}$.

4.4. Example: SE(3)**.** Now consider SE(3).