GEOMETRÍA EN GRUPOS DE MATRICES

SILVIO REGGIANI

RESUMEN. En estas charlas presentamos algunos grupos de matrices como espacios en donde se puede hacer geometría (medir distancias, ángulos, longitudes de curvas, etc.). Trabajaremos principalmente con el grupo ortogonal (es decir, el grupo formado por las matrices cuya inversa es la matriz transpuesta), pero las ideas que presentamos se generalizan a otros grupos. En la segunda parte del curso, siguiendo a J.-H. Eschenburg, aplicaremos métodos geométricos para probar un conocido resultado de álgebra lineal. Más precisamente, el teorema que dice que toda matriz simétrica diagonaliza en una base ortonormal. La demostración geométrica se generaliza a matrices con coeficientes complejos, e incluso con coeficientes cuaterniónicos u octoniónicos (las pruebas clásicas no hacen eso). Se asumen conocimientos elementales de álgebra lineal y análisis de funciones de varias variables.

1. SUBGRUPOS DE MATRICES

Primeramente recordemos la definición abstracta de grupo. Un *grupo* es un conjunto G junto con una operación, llamada *multiplicación*, que a cada par de elementos $g, h \in G$ le asigna un nuevo elemento $gh \in G$ con las siguientes propiedades:

(G1) la multiplicación es *asociativa*: $g(hk) = (gh)k$ para todos $g, h, k \in G$;
(G2) existe un *elemento neutro* $e \in G$ tal que $eg = ge = g$ para todo $g \in G$;
(G3) todo elemento $g \in G$ posee un *inverso* $g^{-1} \in G$ tal que $gg^{-1} = g^{-1}g = e$.

Ejercicio 1.1. Probar que el elemento neutro y los inversos son únicos.

Un *subgrupo* H de un grupo G es un subconjunto de G que también es un grupo, con la multiplicación heredada de G.

Ejemplo 1.2. (a) El conjunto de los números reales $\mathbb{R}$, junto con la suma, forma un grupo. El elemento neutro es $0 \in \mathbb{R}$ y el inverso de $x \in \mathbb{R}$ es $-x$. Notar que este grupo además es *abeliano* o *conmutativo*: es decir, vale $x + y = y + x$ para todos $x, y \in \mathbb{R}$. Más generalmente, $\mathbb{R}^n$ con la suma de vectores es un grupo abeliano.

(b) Los conjuntos $\mathbb{R}^\times = \{x \in \mathbb{R} : x \neq 0\}$ y $\mathbb{R}^{>0} = \{x \in \mathbb{R} : x > 0\}$ son grupos con la multiplicación de números reales. Más aún, $\mathbb{R}^{>0}$ es un subgrupo de $\mathbb{R}^\times$. Notar que el elemento neutro es 1 y el inverso de x es $1/x$.

(c) El conjunto $M(n, \mathbb{R})$ que consiste de las matrices de tamaño $n \times n$ con coeficientes reales, forma un grupo con la suma de matrices, el cual se puede identificar con $\mathbb{R}^{n^2}$. Sin embargo, $M(n, \mathbb{R})$ no es un grupo con la multiplicación de matrices (pues no toda matriz posee una matriz inversa).

Date: 4 de agosto de 2012.
Notas del minicurso dicatado en el Encuentro de Geometría Diferencial, Rosario 2012.

Definición 1.3. Sea $n \geq 1$. El *grupo lineal general* es

$$\mathrm{GL}(n,\mathbb{R}) := \{A \in M(n,\mathbb{R}) : A \text{ es inversible}\} = \{A \in M(n,\mathbb{R}) : \det(A) \neq 0\}.$$

Observemos que $\mathrm{GL}(n,\mathbb{R})$ es efectivamente un grupo con la multiplicación de matrices, pues el producto de matrices es asociativo. El elemento neutro es la matriz identidad $I \in \mathrm{GL}(n,\mathbb{R})$ y el inverso de una matriz A es la matriz inversa de A, $A^{-1} \in \mathrm{GL}(n,\mathbb{R})$. Recordar que si $\det(A) \neq 0$, entonces

$$\det(A^{-1}) = \frac{1}{\det(A)} \neq 0$$

y esto implica que $A^{-1} \in \mathrm{GL}(n,\mathbb{R})$.

Definición 1.4. Sea $n \geq 1$. El *grupo ortogonal* es

$$\mathrm{O}(n) = \{A \in \mathrm{GL}(n,\mathbb{R}) : AA^T = A^T A = I\},$$

es decir el subgrupo de $\mathrm{GL}(n,\mathbb{R})$ formado por aquellas matrices cuya matriz inversa es la matriz traspuesta. Recordar que la matriz traspuesta de A es la matriz A^T definida por $(A^T)_{ij} = A_{ji}$.

Observemos que una matriz $A \in \mathrm{O}(n)$ si y sólo si las filas (o columnas) de A forman una base ortonormal de $\mathbb{R}^n$.

Por otro lado, si $A \in \mathrm{O}(n)$, entonces $1 = \det(AA^T) = (\det A)^2$. Luego, se tiene que $\det A = \pm 1$. Así, tiene sentido definir el siguiente grupo.

Definición 1.5. El *grupo ortogonal especial* es

$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) : \det A = 1\}.$$

Claramente $\mathrm{SO}(n)$ es un subgrupo de $\mathrm{O}(n)$.

Ejercicio 1.6. Probar que

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}.$$

Ejercicio 1.7. Probar que $A \in \mathrm{SO}(3)$ si y sólo si las filas (o columnas) de A forman una base ortonormal positivamente orientada de $\mathbb{R}^3$. Recordar que tres vectores $u, v, w \in \mathbb{R}^3$ forman una base *positivamente orientada* si $\langle u \times v, w\rangle > 0$, en donde $\langle \cdot, \cdot \rangle$ denota el producto escalar en $\mathbb{R}^3$ y $\times$ el producto vectorial en $\mathbb{R}^3$. ¿Se puede generalizar esto para $n > 3$?

Supongamos que $A \in \mathrm{O}(n)$ y $\det A = -1$. Si $B = \mathrm{diag}(-1, 1, \ldots, 1) = B^{-1}$. Entonces $B \in \mathrm{O}(n)$ y $A = B(BA)$ con $BA \in \mathrm{SO}(n)$, pues $\det(BA) = -\det A = 1$. Luego, podemos pensar a $\mathrm{O}(n)$ como dos copias de $\mathrm{SO}(n)$,

$$\mathrm{O}(n) = \mathrm{SO}(n) \cup B \cdot \mathrm{SO}(n) \qquad \text{(unión disjunta)}. \tag{1.1}$$

(Notar que $B \cdot \mathrm{SO}(n)$ no es un grupo, pues no contiene a la matriz identidad.) Más aún, la función $L_B : M(n,\mathbb{R}) \to M(n,\mathbb{R})$ definida por $L_B(A) = BA$ es diferenciable, pensada como función de $\mathbb{R}^{n^2}$ en $\mathbb{R}^{n^2}$, con inversa $(L_B)^{-1} = L_B$ y manda $\mathrm{O}(n)$ en sí mismo (aunque intercambia las componentes de la descomposición 1.1).

1.1. Otros subgrupos de matrices. En esta sección mencionamos algunos subgrupos importantes de matrices, tanto a coeficientes reales como a coeficientes complejos. Queda como ejercicio verificar que estos subconjuntos son efectivamente grupos.

El *grupo lineal especial* se define como

$$\mathrm{SL}(n,\mathbb{R}) = \{A \in \mathrm{GL}(n,\mathbb{R}) : \det A = 1\}.$$

En particular, se tiene

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}).$$

Observemos que el grupo ortogonal $\mathrm{O}(n)$ puede definirse geométricamente de la siguiente manera. Si $\langle \cdot, \cdot \rangle$ es el producto interno usual en $\mathbb{R}^n$, es decir

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n$$

para todos $x, y \in \mathbb{R}^n$, entonces $\mathrm{O}(n)$ es el subgrupo de todas las matrices inversibles que preservan $\langle \cdot, \cdot \rangle$, o sea,

$$A \in \mathrm{O}(n) \qquad \text{si y sólo si} \qquad \langle Ax, Ay \rangle = \langle x, y \rangle$$

para todos $x, y \in \mathbb{R}^n$.

Análogamente puede definirse el *grupo simpléctico* $\mathrm{Sp}(n) \subset \mathrm{GL}(2n, \mathbb{R})$ como el grupo formado por todas las matrices inversibles de tamaño $2n \times 2n$ que preservan la 2-forma simpléctica[1] canónica ω en $\mathbb{R}^{2n}$, la cual se define como

$$\omega(x, y) = x_1 y_2 - x_2 y_1 + \cdots + x_{2n-1} y_{2n} - x_{2n} y_{2n-1}.$$

para todos $x, y \in \mathbb{R}^{2n}$. Observar que la dimensión del espacio debe ser par para que podamos definir ω. Así,

$$A \in \mathrm{Sp}(n) \qquad \text{si y sólo si} \qquad \omega(Ax, Ay) = \omega(x, y)$$

para todos $x, y \in \mathbb{R}^{2n}$.

Ejercicio 1.8. Probar que $\mathrm{Sp}(n) \subset \mathrm{SL}(2n, \mathbb{R})$.

A continuación definimos algunos subgrupos de matrices a coeficientes complejos. El *grupo lineal general complejo* es

$$\mathrm{GL}(n, \mathbb{C}) = \{A \in M(n, \mathbb{C}) : \det A \neq 0\}.$$

Observemos que toda matriz a coeficientes complejos $A \in M(n, \mathbb{C})$ puede escribirse como $A = B + iC$, con $B, C \in M(n, \mathbb{R})$. Es decir, una matriz compleja está determinada por $2n^2$ coordenadas reales. Uno puede identificar $\mathrm{GL}(n, \mathbb{C})$ con un subgrupo $G \subset \mathrm{GL}(2n, \mathbb{R})$ de la siguiente manera.

Veamos un caso particular y dejemos el caso general como ejercicio. Si $n = 1$, entonces $\mathrm{GL}(1, \mathbb{C}) = \mathbb{C}^\times = \{z \in \mathbb{C} : z \neq 0\}$. Sean $z, w \in \mathbb{C}$, digamos $z = a + ib$, $w = c + id$. Observemos que

$$zw = (a + ib)(c + id) = ac - bd + i(ad + bc).$$

Por otro lado, dado un número complejo $z = a + ib$, tenemos asociada una matriz

$$A_z = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Como $\det A_z = a^2 + b^2 = |z|^2$, se tiene que $A_z \in \mathrm{GL}(2, \mathbb{R})$ si y sólo si $z \in \mathrm{GL}(1, \mathbb{C})$. Un cálculo directo nos da

$$A_z A_w = \begin{pmatrix} ac - bd & ad + bc \\ -ad - bc & ac - bd \end{pmatrix} = A_{zw}.$$

Luego, $\mathrm{GL}(1, \mathbb{C})$ se identifica con el subgrupo G de $\mathrm{GL}(2, \mathbb{R})$ definido por

$$G = \{A_z \in \mathrm{GL}(2, \mathbb{R}) : z \in \mathrm{GL}(1, \mathbb{C})\}.$$

[1] Una *forma simpléctica* es una forma bilineal antisimétrica y no-degenerada ω en $\mathbb{R}^{2n}$. Es decir, $\omega(x, y)$ es lineal en x fijado y y viceversa; $\omega(x, y) = -\omega(y, x)$ para todos $x, y \in \mathbb{R}^{2n}$; y $\omega(x, y) = 0$ para todo y implica $x = 0$.

Ejercicio 1.9. Si $Z \in M(n,\mathbb{C})$ y $Z = B + iC$ con $B, C \in M(n,\mathbb{R})$ definimos

$$A_Z = \begin{pmatrix} B & C \\ -C & B \end{pmatrix}.$$

Probar que $A_Z \in \mathrm{GL}(2n,\mathbb{R})$ si y sólo si $Z \in \mathrm{GL}(n,\mathbb{C})$ y que $A_Z A_W = A_{ZW}$ para todas $Z, W \in \mathrm{GL}(n,\mathbb{C})$. De este modo, se identifica $\mathrm{GL}(n,\mathbb{C})$ con el subgrupo

$$G = \{A_Z \in \mathrm{GL}(2n,\mathbb{R}) : Z \in \mathrm{GL}(n,\mathbb{C})\} \subset \mathrm{GL}(2n,\mathbb{R}).$$

Observación 1.10. Observemos que $\mathrm{GL}(n,\mathbb{R}) \subset \mathrm{GL}(n,\mathbb{C}) \subset \mathrm{GL}(2n,\mathbb{R})$ (con la identificación anterior) y todas las inclusiones son propias. Más aún, las dimensiones de los espacios vectoriales reales $M(n,\mathbb{R})$, $M(n,\mathbb{C})$ y $M(2n,\mathbb{R})$ son respectivamente $n^2 < 2n^2 < 4n^2$. Como veremos (y definiremos) más adelante, estas son las dimensiones de los correspondientes grupos de matrices.

Análogamente al caso real, el *grupo lineal especial complejo* se define como

$$\mathrm{SL}(n,\mathbb{C}) = \{A \in \mathrm{GL}(n,\mathbb{C}) : \det A = 1\}.$$

El *grupo unitario* $\mathrm{U}(n)$ se define como el subgrupo de $\mathrm{GL}(n,\mathbb{C})$ que consiste de las matrices que preservan el producto interno complejo o hermitiano canónico de $\mathbb{C}^n$, es decir, el definido por

$$\langle z, w\rangle = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n \tag{1.2}$$

para todos $z, w \in \mathbb{C}^n$. O sea,

$$A \in \mathrm{U}(n) \qquad \text{si y sólo si} \qquad \langle Az, Aw\rangle = \langle z, w\rangle$$

para todos $z, w \in \mathbb{C}^n$. El *grupo unitario especial* se define como

$$\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C}).$$

No es difícil probar que el grupo unitario puede verse como

$$\mathrm{U}(n) = \{A \in \mathrm{GL}(n,\mathbb{C}) : AA^* = A^*A = I\},$$

en donde A^* denota la matriz traspuesta y conjugada de A.

Existen algunas identificaciones[2], a veces para dimensiones bajas, entre los distintos subgrupos que acabamos de definir. Como estos resultados no nos interesan particularmente en este curso, los dejamos como ejercicio para el lector interesado.

Ejercicio 1.11. Probar que S^1 se identifica con $\mathrm{SO}(2)$, en donde

$$S^1 = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\theta} : \theta \in \mathbb{R}\}$$

es la circunferencia unidad en el plano, mirada como subgrupo de $\mathrm{GL}(1,\mathbb{C})$.

Ejercicio 1.12. Probar que existe una biyección entre $\mathrm{U}(n)$ y $S^1 \times \mathrm{SU}(n)$. ¿Se puede identificar $\mathrm{U}(n)$ con $S^1 \times \mathrm{SU}(n)$ como grupo (en donde el producto en el segundo caso se define coordenada a coordenada)?

Ejercicio 1.13. Probar que $\mathrm{SU}(2)$ se identifica con $\mathrm{Sp}(1)$.

[2]En estas notas, cuando hablamos de identificación entre dos grupos H y G queremos decir que existe una función biyectiva $\varphi : H \to G$ tal que $\varphi(gh) = \varphi(g)\varphi(h)$ para todos $g, h \in H$. Una tal φ es llamada un *isomorfismo de grupos*.

1.2. Funciones diferenciables y espacios tangentes. En este apartado nos restringimos, por simplicidad, a los subgrupos $G = \mathrm{GL}(n, \mathbb{R})$, $G = \mathrm{O}(n)$ o $G = \mathrm{SO}(n)$, pero las ideas que trataremos se generalizan a otros subgrupos de matrices (en particular, a los subgrupos definidos en la subsección anterior). Más aún, estas nociones se generalizan a ciertos subconjuntos de $\mathbb{R}^k$ llamados subvariedades, que no son necesariamente subgrupos de matrices.

Tengamos presente la identificación $M(n, \mathbb{R}) \simeq \mathbb{R}^{n^2}$.

Definición 1.14. Sea $G \subset \mathrm{GL}(n, \mathbb{R})$ un subgrupo de matrices.

(a) Una función $f : U \subset \mathbb{R}^k \to G$ se dice *diferenciable* si la función $f : U \to \mathbb{R}^{n^2}$ es diferenciable. Es decir, si cada una de sus funciones coordenadas es una función diferenciable.
(b) Una *curva diferenciable* en G es una función diferenciable $c : I \to G$, en donde I es un intervalo en $\mathbb{R}$.
(c) Una función $f : G \to \mathbb{R}$ se dice *diferenciable*, si para cada $A \in G$ existe un entorno abierto de A, $U \subset M(n, \mathbb{R})$ y una función diferenciable $F : U \to \mathbb{R}$ tal que $F|_{G\cap U} = f|_{G\cap U}$.

Ejercicio 1.15. ¿Cómo se definiría una función diferenciable $f : G \to H$ entre dos subgrupos de matrices?

Definición 1.16. El *espacio tangente* a G en $A \in G$ se define como

$$T_A G = \{c'(0) : c(t) \text{ es una curva diferenciable en } G \text{ con } c(0) = A\},$$

en donde las entradas de $c'(0)$ se obtienen de derivar las entradas de la matriz $c(t)$ en $t = 0$.

Definimos la *dimensión* de G como $\dim G = \dim T_A G$. Más adelante probaremos que $T_A G$ es un espacio vectorial real de dimensión finita (la misma dimensión para cada $A \in G$).

Ejemplo 1.17 (La aplicación exponencial)**.** Si $A \in M(n, \mathbb{R})$ definimos la aplicación exponencial $A \mapsto e^A$ por

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k. \tag{1.3}$$

Para probar la buena definición de la aplicación exponencial debemos analizar la convergencia de la serie 1.3. En efecto, si $|A_{ij}| \leq N$ para todos i, j, entonces $|(A^k)_{ij}| \leq n^{k-1}N^k$ (se ve por inducción). Luego, por el M-test de Weierstrass, cada coordenada de e^A converge uniformemente en la región $|A_{ij}| \leq N$, $1 \leq i, j \leq n$. Por consiguiente, la aplicación exponencial $e : M(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ está bien definida y es diferenciable, como sigue del siguiente lema.

Lema 1.18. *Sean $A, B \in M(n, \mathbb{R})$. Entonces,*

(a) $e^0 = I$;
(b) $e^{sA}e^{tA} = e^{(s+t)A}$ *para todos* $s, t \in \mathbb{R}$;
(c) $e^A e^{-A} = e^{-A}e^A = I$, *por lo tanto,* e^A *toma valores en* $\mathrm{GL}(n, \mathbb{R})$;
(d) *si* $AB - BA = 0$, *entonces* $e^A e^B = e^{A+B}$;
(e) $(e^A)^T = e^{A^T}$;
(f) *si* $C \in \mathrm{GL}(n, \mathbb{R})$, *entonces* $e^{CAC^{-1}} = Ce^AC^{-1}$.

Demostración. La parte (a) es trivial. La parte (b) sigue de (d). Para probar (c), observemos que como la aplicación exponencial converge uniformemente en subconjuntos acotados, el producto de las series $e^A e^{-A}$ converge a la serie producto. Por una manipulación formal (la misma que hacemos para la función exponencial de una variable real), sabemos que los coeficientes de la serie producto son todos nulos, excepto por el primero, que es igual a 1. Luego, $e^A e^{-A} = e^{-A} e^A = I$. Las partes (e) y (f) siguen de que $(A^k)^T = (A^T)^k$ y $(CAC^{-1})^k = CA^k C^{-1}$. Sólo falta probar la parte (d), la cual sigue de cálculos estándares (ver, por ejemplo, [6]). $\square$

Ejemplo 1.19 (El espacio tangente a $\mathrm{GL}(n,\mathbb{R})$ en la identidad)**.** Observemos primeramente que si $c(t)$ es una curva en $\mathrm{GL}(n,\mathbb{R})$ con $c(0) = I$, entonces su derivada $c'(0) \in M(n,\mathbb{R})$. Es decir, $T_I \mathrm{GL}(n,\mathbb{R}) \subset M(n,\mathbb{R})$. Recíprocamente, dada $A \in M(n,\mathbb{R})$, se tiene que $c(t) = e^{tA}$ es una curva en $\mathrm{GL}(n,\mathbb{R})$ con $c(0) = I$. Además

$$c'(0) = \frac{d}{dt}\bigg|_0 \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = A.$$

Luego $M(n,\mathbb{R}) \subset T_I \mathrm{GL}(n,\mathbb{R})$. Y por consiguiente $T_I \mathrm{GL}(n,\mathbb{R}) = M(n,\mathbb{R})$. En particular $\dim \mathrm{GL}(n,\mathbb{R}) = \dim M(n,\mathbb{R}) = n^2$.

Ejemplo 1.20 (El espacio tangente a $\mathrm{SO}(n)$ en la identidad)**.** Antes que nada observemos que $T_I \mathrm{SO}(n) = T_I \mathrm{O}(n)$. En efecto, como $\mathrm{SO}(n) \subset \mathrm{O}(n)$, claramente se tiene que $T_I \mathrm{SO}(n) \subset T_I \mathrm{O}(n)$. Por otro lado, si $c(t)$ es una curva en $\mathrm{O}(n)$ con $c(0) = I$, entonces $\det c(0) = 1$. Como el determinante es una función continua, no cambia de signo para valores de t próximos a 0. Luego $\det c(t) = 1$ para t cerca de 0, lo cual implica que $c(t) \in \mathrm{SO}(n)$. Así, $T_I \mathrm{SO}(n) = T_I \mathrm{O}(n)$.

Ahora, si $c(t)$ es una curva en $\mathrm{SO}(n)$ con $c(0) = I$, entonces $c(t)c(t)^T = I$. Derivando en $t = 0$ se obtiene

$$c'(0) + c'(0)^T = 0.$$

Por lo tanto, $T_I \mathrm{SO}(n) \subset \mathfrak{so}(n) := \{A \in M(n,\mathbb{R}) : A + A^T = 0\}$, es decir, el subespacio de matrices antisimétricas de tamaño $n \times n$.

Recíprocamente, si $A \in \mathfrak{so}(n)$ entonces

$$(e^{tA})^{-1} = e^{-tA} = e^{tA^T} = (e^{tA})^T$$

por el Lema 1.18. Luego $c(t) = e^{tA}$ es una curva en $\mathrm{SO}(n)$ con $c'(0) = A$ (con las mismas cuentas de antes). Así, $T_I \mathrm{SO}(n) = \mathfrak{so}(n)$.

Para calcular la dimensión de $\mathfrak{so}(n)$ observemos que una matriz antisimétrica queda completamente determinada por sus entradas arriba de la diagonal (en la diagonal hay ceros). Luego,

$$\dim \mathrm{SO}(n) = \dim \mathfrak{so}(n) = 1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}.$$

Ejercicio 1.21. Encontrar los espacios tangentes en la identidad para los grupos $\mathrm{SL}(n,\mathbb{R})$, $\mathrm{Sp}(n)$, $\mathrm{GL}(n,\mathbb{C})$, $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{U}(n)$ y $\mathrm{SU}(n)$. Concluir que $\dim \mathrm{SL}(n,\mathbb{R}) = n^2 - 1$, $\dim \mathrm{Sp}(n) = 2n^2 + n$, $\dim \mathrm{GL}(n,\mathbb{C}) = 2n^2$, $\dim \mathrm{SL}(n,\mathbb{C}) = 2(n^2 - 1)$, $\dim \mathrm{U}(n) = n^2$ y $\dim \mathrm{SU}(n) = n^2 - 1$.

Sea $G \subset \mathrm{GL}(n,\mathbb{R})$ un subgrupo de matrices. Si uno conoce el espacio tangente en la identidad $T_I G$, es fácil obtener el espacio tangente $T_A G$ en $A \in G$. En efecto, como ya mencionamos antes para un caso particular, la función $L_A : M(n,\mathbb{R}) \to M(n,\mathbb{R})$, definida por $L_A(B) = AB$ es una función diferenciable, con inversa diferenciable $(L_A)^{-1} = L_{A^{-1}}$, que manda G en sí mismo. Luego, si $c'(0) \in T_I G$, para cierta curva $c(t)$ en G con $c(0) = I$. Entonces $\alpha(t) = L_A(c(t))$ es una curva en G con $\alpha(0) = L_A(I) = A$. Luego,

$$\alpha'(0) = \left.\frac{d}{dt}\right|_0 L_A(c(t)) = \left.\frac{d}{dt}\right|_0 Ac(t) = Ac'(0).$$

Más aún, toda curva $\alpha(t)$ con $\alpha(0) = A$ es de la forma $\alpha(t) = L_A(c(t))$, con $c(0) = I$ (simplemente tomando $c(t) = L_{A^{-1}}(\alpha(t))$. Luego,

$$T_A G = \{AX : X \in T_I G\} = A \cdot T_I G.$$

En los ejemplos anteriores se obtiene

$$T_A \operatorname{GL}(n, \mathbb{R}) = A \cdot M(n, \mathbb{R}) = M(n, \mathbb{R}),$$

pues A es inversible, y

$$T_A \operatorname{SO}(n) = A \cdot \mathfrak{so}(n) = \{AX : X + X^T = 0\}.$$

1.3. Un poco de geometría. La idea básica para hacer geometría en un espacio es tener definida una manera de medir distancias y ángulos. En nuestro caso, quisiéramos medir longitudes de curvas y ángulos entre velocidades de curvas que se crucen en un punto. Sea G un subgrupo de $\operatorname{GL}(n, \mathbb{R})$. Pedimos que en cada espacio tangente $T_A G$, $A \in G$, exista un producto interno $\langle \cdot, \cdot \rangle_A$. Además se pide que la asignación $A \mapsto \langle \cdot, \cdot \rangle_A$ sea "diferenciable" en el siguiente sentido. Si $c(t)$, $t \in [a, b]$, es una curva diferenciable (a trozos) en G definimos la *longitud* de $c(t)$ como

$$L(c) = \int_a^b \sqrt{\langle c'(t), c'(t) \rangle_{c(t)}}\, dt.$$

El hecho de que $A \mapsto \langle \cdot, \cdot \rangle_A$ se diferenciable lo interpretaremos como que la función $t \mapsto \langle c'(t), c'(t) \rangle_{c(t)}$ sea diferenciable, cualquiera sea la curva diferenciable $c(t)$. Llamamos a la asignación $A \mapsto \langle \cdot, \cdot \rangle_A$ una *métrica riemanniana* en G.

Si uno sabe medir longitudes de curvas, se puede definir la *distancia* entre dos puntos $A, B \in G$ por

$$d(A, B) = \operatorname{ínf}\{L(c) : c \text{ es una curva que une } A \text{ con } B\}.$$

Una función biyectiva $f : G \to G$ que preserva distancias se llama una *isometría*. Otra forma de ver a una isometría es la siguiente: f es una función diferenciable tal que

$$\langle X, Y \rangle_A = \langle df(X), df(Y) \rangle_{f(A)}$$

para todo $A \in G$, $X, Y \in T_A G$, en donde

$$df(X) = \left.\frac{d}{dt}\right|_0 f(c(t))$$

para una curva $c(t)$ con $c(0) = A$ y $c'(0) = X$. Observemos que df es la llamada diferencial de f (en el punto A). En otras palabras, f es una isometría de G si y sólo si su diferencial $df : T_A G \to T_{f(A)} G$ es una isometría lineal entre los respectivos espacios tangentes.

Una familia muy importante de curvas en G es la de las llamadas curvas *geodésicas*, que son las curvas que minimizan localmente la distancia. Se puede probar que siempre existe una geodésica por cualquier punto y con una velocidad inicial dada, resolviendo una ecuación diferencial de segundo orden (ver [2]).

1.4. Métricas invariantes a izquierda. En un grupo de matrices G cualquier producto interno $\langle\cdot,\cdot\rangle$ en T_IG induce una métrica riemanniana, simplemente definiendo

$$\langle X,Y\rangle_A = \langle A^{-1}X, A^{-1}Y\rangle$$

para todos $X,Y \in T_AG$. Es decir, $\langle\cdot,\cdot\rangle_I = \langle\cdot,\cdot\rangle$ y las translaciones a izquierda L_A son isometrías para todo $A \in G$. Notar que si $X \in T_AG$, entonces $L_{A^{-1}}(X) = A^{-1}X \in T_IG$. Estas métricas se llaman *métricas invariantes a izquierda* (pues las multiplicaciones a izquierda resultan isometrías). Una métrica invariante a izquierda queda completamente determinada por su valor $\langle\cdot,\cdot\rangle$ en T_IG. Por eso, para la simplificar la notación, cuando el punto base se sobreentienda denotaremos con el mismo símbolo $\langle\cdot,\cdot\rangle$ a la métrica riemanniana invariante a izquierda en G.

Ejemplo 1.22 (La métrica bi-invariante en $\mathrm{SO}(n)$)**.** Observar que dados $X,Y \in \mathfrak{so}(n) = T_I\,\mathrm{SO}(n)$, la asignación

$$\langle X,Y\rangle = -\operatorname{traza}(XY) = \sum_{i,j} X_{ij}Y_{ij}$$

define un producto interno en $\mathfrak{so}(n)$ (probarlo). La métrica invariante a izquierda inducida, es llamada la métrica bi-invariante en $\mathrm{SO}(n)$. Esta terminología está motivada por el hecho de que las traslaciones a derecha $R_A : \mathrm{SO}(n) \to \mathrm{SO}(n)$, $A \in \mathrm{SO}(n)$, definidas por $R_A(B) = BA$, también son isometrías. En efecto, si $X,Y \in \mathfrak{so}(n)$, entonces

$$\begin{aligned}
\langle dR_A(X), dR_A(Y)\rangle_A &= \langle A^{-1}dR_A(X), A^{-1}dR_A(Y)\rangle_I \\
&= \langle A^{-1}XA, A^{-1}YA\rangle_I \\
&= -\operatorname{traza}(A^{-1}XAA^{-1}YA) \\
&= -\operatorname{traza}(A^{-1}XYA) = -\operatorname{traza}(XYAA^{-1}) \\
&= -\operatorname{traza}(XY) = \langle X,Y\rangle_I.
\end{aligned}$$

Luego $dR_A : T_I\,\mathrm{SO}(n) \to T_A\,\mathrm{SO}(n)$ es una isometría lineal. Análogamente se ve que $dR_A : T_B\,\mathrm{SO}(n) \to T_{BA}\,\mathrm{SO}(n)$ es una isometría lineal.

2. Un teorema de álgebra lineal

Esta sección está dedicada a probar el Teorema 2.1, el cual es un resultado bien conocido de álgebra lineal. Esta sección está basada en las notas [3] de J.-H. Eschenburg.

Teorema 2.1. *Sea $S \in M(n,\mathbb{R})$ una matriz simétrica, es decir, $S^T = S$. Entonces S tiene una base ortonormal de autovectores. En otras palabras, $S = ADA^{-1}$ para alguna matriz diagonal $D = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$ y para alguna $A \in \mathrm{O}(n)$.*

Observemos que puede asumirse $A \in \mathrm{SO}(n)$, reemplazando la primera columna A_1 de A por $-A_1$ (pues el negativo de un autovector también es un autovector).

Ejemplo 2.2. Consideremos la matriz

$$S = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

El polinomio característico de S es x^2-x-1 de donde sigue que los autovalores de S son $\lambda_1 = \varphi$ y $\lambda_2 = -2/\varphi$, en donde $\varphi = (1+\sqrt{5})/2 \approx 1{,}618...$ es el llamado número de oro. Podemos obtener una base de autovectores $v_1 = (\varphi, 1)$ y $v_2 = (-1, \varphi)$. Como $\langle v_1, v_2\rangle = 0$ y $\|v_1\|^2 = \|v_2\|^2 = 1 + \varphi^2$, normalizando esta base, obtenemos que S diagonaliza en una base ortonormal.

Un hecho un tanto curioso es que iterando k veces la matriz S (observemos que iterar una matriz simétrica nos da una matriz simétrica), obtenemos que

$$S^k = \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix}$$

en donde F_k es la sucesión de Fibbonacci $F_0 = 0$, $F_1 = 1$, $F_{k+1} = F_k + F_{k-1}$. En efecto, para $k = 1$ esto es cierto, si suponemos que vale para k, entonces

$$\begin{aligned} S^{k+1} = SS^k &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} F_{k+1} + F_k & F_k + F_{k-1} \\ F_{k+1} & F_k \end{pmatrix} = \begin{pmatrix} F_{k+2} & F_{k+1} \\ F_{k+1} & F_k \end{pmatrix}. \end{aligned}$$

2.1. Demostraciones clásicas. En [5], L. A. Steen menciona que el Teorema 2.1, a veces conocido como teorema espectral, estaba ya implícito en los trabajos de Fermat (1679, póstumo) y Descartes (1637). Más precisamente, toda forma cuadrática $ax^2 + 2bxy + cy^2$ puede ser transformada por una rotación en el plano a una forma normal $\alpha x^2 + \beta y^2$. Esto se conoce también como el teorema de los ejes principales. El termino “ejes principales” fue introducido por Euler (1748, 1765) quien también estudió la reducción de formas cuadráticas en dos y tres dimensiones. La forma general del teorema de los ejes principales dice que una forma cuadrática simétrica $\sum A_{ij}x_ix_j$ (simétrica significa que $A_{ij} = A_{ji}$) en $\mathbb{R}^n$ puede ser escrita, mediante una transformación ortogonal, en forma normal $\sum \lambda_i x_i^2$ y aparece en los trabajos de Lagrange (1759) sobre máximos y mínimos de funciones de varias variables. Un tiempo después, Cauchy probó en 1829 y 1830 que los coeficientes λ_i deben ser números reales.

La presentación moderna del teorema de los ejes principales, en términos de matrices, data de la segunda mitad del siglo XIX. En 1852, Sylvester probó que los coeficientes λ_i son las raíces del polinomio característico $\det(xI - A) = 0$. Y en 1858, Cayley muestra que la reducción a la forma normal corresponde al proceso de diagonalización de la matriz A. Todas las referencias pueden consultarse en [5].

La clave para probar el Teorema 2.1 es la observación de que el complemento ortogonal de un autoespacio de S es también invariante por S, y de este modo se puede razonar inductivamente. Más precisamente, consideremos el producto interno usual en $\mathbb{R}^n$,

$$\langle v, w\rangle = v_1w_1 + \cdots + v_nw_n$$

para todos $v, w \in \mathbb{R}^n$. Que una matriz S sea simétrica, significa que es autoadjunta con respecto a este producto interno, es decir $\langle Sv, w\rangle = \langle v, Sw\rangle$. Recordemos que si pensamos a v y w como vectores columna, entonces $\langle v, w\rangle = v^T w$. Por ende

$$\langle Sv, w\rangle = (Sv)^T w = v^T S^T w = v^T Sw = \langle v, Sw\rangle.$$

Lema 2.3. *Sea $S \in M(n, \mathbb{R})$ una matriz simétrica y sea $V \subset \mathbb{R}^n$ un subespacio S-invariante, es decir, $SV \subset V$. Entonces el complemento ortogonal*

$$V^\perp = \{w \in \mathbb{R}^n : \langle v, w\rangle = 0 \text{ para todo } v \in V\}$$

de V también es S-invariante.

Demostración. Debemos probar que si $w \in V^\perp$ entonces $Sw \in V^\perp$. Pero esto es cierto pues $\langle Sw, v\rangle = \langle w, Sv\rangle = 0$ pues $Sv \in V$ para todo $v \in V$. □

La demostración del Teorema 2.1 sigue por inducción de la siguiente manera. Si λ es un autovalor de S y V es el autoespacio de autovalor λ, es decir $V = E_\lambda = \{v \in \mathbb{R}^n : Sv = \lambda v\}$, entonces $W = V^\perp$ es un subespacio de $\mathbb{R}^n$ de dimensión menor en el cual S también es autoadjunto (si fuera $\dim W = 0$, entonces $S = \lambda I$ y la conclusión vale trivialmente). Por hipótesis inductiva, $S|_W$ diagonaliza en una base ortonormal. Uniendo dicha base a una base ortonormal de $V = E_\lambda$, se tiene que S diagonaliza en una base ortonormal de $\mathbb{R}^n$.

En el párrafo anterior hemos dado por sentado un hecho no trivial: ¿por qué una transformación autoadjunta S tiene un autovalor real? En la literatura se encuentran diversas forma de probar esto.

2.1.1. Teorema fundamental del álgebra. Consideramos a S como una matriz compleja (usando que $\mathbb{R} \subset \mathbb{C}$). Usando el teorema fundamental del álgebra (que dice que todo polinomio sobre los números complejos tiene una raíz) podemos encontrar un autovalor $\lambda_o \in \mathbb{C}$ de S. Más precisamente, λ_o es una raíz del polinomio característico $\det(\lambda I - S)$ de S. Como S es simétrica, se tiene que $\lambda_o \in \mathbb{R}$. En efecto, como S es autoadjunta con respecto al producto escalar en $\mathbb{R}^n$, se tiene que S es también autoadjunta con respecto al producto interno hermitiano en $\mathbb{C}^n$

$$\langle v, w\rangle = v^* w$$

definido en 1.2. Recordar que $*$ significa trasponer y conjugar. Como S tiene todas sus entradas reales, se tiene que $S^* = S^T = S$. En particular, para todo $w \in E_{\lambda_o}$, $w \neq 0$, se tiene

$$\lambda_o\langle w, w\rangle = \langle Sw, w\rangle = \langle w, Sw\rangle = \bar{\lambda}_o\langle w, w\rangle,$$

con lo cual $\lambda_o = \bar{\lambda}_o$ y así $\lambda_o \in \mathbb{R}$.

Cabe aclarar que el teorema fundamental del álgebra no es sencillo de probar.

2.1.2. Teorema de los valores intermedios. Consideremos la matriz $S_\lambda = S - \lambda I$ definida para $\lambda \in \mathbb{R}$. Para valores grandes de $|\lambda|$, el término dominante de S_λ es $-\lambda I$, luego $\langle S_\lambda v, v\rangle$ tiene el signo opuesto a λ para todo $v \neq 0$. En particular, S_λ es definida positiva para $\lambda < 0$ y suficientemente grande en módulo. Sea λ_o el valor más pequeño con la propiedad que S_{λ_o} no es más definida positiva. En este caso tenemos que $\langle S_{\lambda_o} v, v\rangle \geq 0$ para todo v, y además existe $v_o \neq 0$ tal que $\langle S_{\lambda_o} v_o, v_o\rangle = 0$. Esto implica que $S_{\lambda_o} v_o = 0$ y por ende $Sv_o = \lambda_o v_o$.

En efecto, esto es un hecho general. Si T es una matriz simétrica semi-definida positiva, es decir $\langle Tv, v\rangle \geq 0$ para todo $v \neq 0$, y $\langle Tv_o, v_o\rangle = 0$ entonces $Tv_o = 0$. Para probar esto notemos que para todo w y para todo $t \in \mathbb{R}$ se tiene

$$0 \leq \langle T(tv_o + w), tv_o + w\rangle = 2t\langle Tv_o, w\rangle + \langle Tw, w\rangle,$$

lo cual es posible solamente si $\langle Tv_o, w\rangle = 0$ y por tanto $Tv_o = 0$.

2.1.3. Multiplicadores de Lagrange. Sea $M \subset \mathbb{R}^n$ una subvariedad y $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable tal que $f|_M$ alcanza un valor máximo en $x_o \in M$. Sea $(\nabla f)_{x_o}$ el gradiente de f en x_o y sea $T_{x_o}M \subset \mathbb{R}^n$ el espacio tangente a M en x_o. Entonces $(\nabla f)_{x_o} \perp T_{x_o}M$.

Aplicamos este argumento a la función $f(x) = \langle x, x\rangle/2$ y a la subvariedad $M = S^{n-1} = \{x \in \mathbb{R}^n : \langle x, x\rangle = 1\}$. Como M es compacta, $f|_M$ alcanza un máximo en algún $x_o \in M$. Luego $(\nabla f)_{x_o} \perp T_{x_o}M = \{x_o\}^\perp$. Como $(\nabla f)_x = Sx$, se tiene que $Sx_o = \lambda x_o$, luego λ es un autovalor real.

2.2. La prueba de Eschenburg. En las demostraciones anteriores empezamos con una matriz simétrica S y buscamos una matriz ortogonal A y una matriz diagonal D tales que $S = ADA^{-1}$. Esto es equivalente a probar que las clases de conjugación (por matrices ortogonales) de las matrices diagonales nos dan todo el subespacio de matrices simétricas. Recordemos que la clase de conjugación de una matriz D es

$$\mathscr{C}_D = \{ADA^{-1} : A \in \mathrm{SO}(n)\} = \{L_A(R_{A^{-1}}(D)) : A \in \mathrm{SO}(n)\}.$$

Para hacer esto podemos usar geometría en el espacio de matrices simétricas

$$\Sigma = \{S \in M(n, \mathbb{R}) : S^T = S\}.$$

Observar que Σ no es un subgrupo de matrices, pero sí es un subespacio vectorial de $M(n, \mathbb{R})$. En Σ también podemos definir un producto interno usando la función traza (análogamente a lo que hicimos en $\mathfrak{so}(n)$). Para $S, T \in \Sigma$ definimos

$$\langle S, T\rangle = \mathrm{traza}(ST) = \sum_{i,j} S_{ij}T_{ij}.$$

Lema 2.4. *Conjugar por una matriz ortogonal preserva el producto interno en Σ. En otras palabras, para toda $A \in \mathrm{O}(n)$ la aplicación lineal*

$$\mathrm{Ad}_A = L_A \circ R_{A^{-1}} : \Sigma \to \Sigma,$$

$\mathrm{Ad}_A(S) = ASA^{-1}$*, es una isometría lineal.*

Demostración. Se tiene que, dadas $S, T \in \Sigma$,

$$\begin{aligned}\langle ASA^{-1}, AYA^{-1}\rangle &= \mathrm{traza}(ASA^{-1}AYA^{-1})\\ &= \mathrm{traza}(ASTA^{-1})\\ &= \mathrm{traza}(ST) = \langle S, T\rangle\end{aligned}$$

pues la traza es invariante por conjugación. □

La idea de la prueba del Teorema 2.1 es la siguiente. Consideremos el subespacio $\Delta \subset \Sigma$ de todas las matrices diagonales

$$\Delta = \{\mathrm{diag}(\lambda_1, \dots, \lambda_n) : \lambda_1, \dots, \lambda_n \in \mathbb{R}\}.$$

Sea $D \in \Delta$, $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ tal que los λ_i son todos distintos, es decir, $\lambda_i \neq \lambda_j$ si $i \neq j$. La clase de conjugación $\mathscr{C}_D \subset \Sigma$ no es un subespacio lineal, pero es una subvariedad de Σ (se pueden definir curvas diferenciables y espacio tangente en cada punto de la misma forma que lo hicimos en las Definiciones 1.14 y 1.16) y veremos en el Lema 2.5 que el espacio tangente a $\mathscr{C}_D$ en D, es precisamente el complemento ortogonal de Δ,

$$T_D\mathscr{C}_D = \Delta^\perp.$$

Sea ahora $S \in \Sigma$ una matriz simétrica arbitraria. Como $\mathscr{C}_D$ es compacto (es cerrado y acotado en un espacio vectorial) existe un elemento $X_o \in \mathscr{C}_D$ que es el que más se aproxima a S, i.e., X_o es una matriz en donde la función continua $X \mapsto \|X - S\|^2 = \langle X - S, X - S\rangle$, con $X \in \mathscr{C}_D$, alcanza su valor mínimo. Sigue que $T = S - X_o$ es perpendicular a $T_{X_o}\mathscr{C}_D$. Si así no fuera, podríamos encontrar otro $X \in \mathscr{C}_D$ más cercano a S. Como $X_o \in \mathscr{C}_D$, existe $A \in \mathrm{SO}(n)$ tal que $\mathrm{Ad}_A(X_o) = D$. Como Ad_A es una isometría que preserva la clase $\mathscr{C}_D$, manda el vector $T \perp T_{X_o}\mathscr{C}_D$ en un vector $T' \perp T_D\mathscr{C}_D = \Delta^\perp$, lo cual implica que $T' \in \Delta$.

Ahora,

$$\mathrm{Ad}_A(S) = \mathrm{Ad}_A(X_o + T) = \mathrm{Ad}_A(X_o) + \mathrm{Ad}_A(T) = D + T' = S' \in \Delta,$$

y por ende $S = A^{-1}S'A$ es conjugada por una matriz ortogonal a la matriz diagonal S'.

Sólo falta calcular el espacio tangente a $\mathscr{C}_D$ en D.

Lema 2.5. *Sea $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ una matriz diagonal con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces $T_C\mathscr{C}_D = \Delta^{\perp}$.*

Demostración. De acuerdo a nuestra definición, el espacio tangente a $\mathscr{C}_D$ consiste de las velocidades iniciales $X'(0)$ de curvas $X(t)$ en $\mathscr{C}_D$ con $X(0) = D$. Podemos asumir que $X(t) = \mathrm{Ad}_{A(t)} D$, en donde $A(t)$ es una curva en $\mathrm{SO}(n)$ con $A(0) = I$. Luego, recordando que $A(t)^{-1} = A(t)^T$, tenemos

$$0 = \frac{d}{dt}A(t)A(t)^T = A'(t)A(t)^T + A(t)\frac{d}{dt}A(t)^T$$

y por consiguiente

$$\frac{d}{dt}A(t)^T = -A(t)^T A'(t)A(t)^T.$$

Así,

$$\frac{d}{dt}X(t) = \frac{d}{dt}A(t)DA(t)^T = A'(t)DA(t)^T - A(t)DA(t)^T A'(t)A(t)^T.$$

Evaluando en $t = 0$ se obtiene

$$X'(0) = UD - DU =: [U, D]$$

en donde $U = A'(0) \in \mathfrak{so}(n)$ es una matriz antisimétrica. Luego,

$$T_D\mathscr{C}_D = \{[U, D] : U \in \mathfrak{so}(n)\}.$$

Las entradas de la matriz $[U, D]$ son

$$[U, D]_{ij} = U_{ij}(\lambda_i - \lambda_j).$$

En particular, $U_{ii} = 0$ y por consiguiente, $[U, D] \perp \Delta$. En efecto, si $M \in \Delta$, digamos $M = \mathrm{diag}(\mu_1, \dots, \mu_n)$, entonces

$$\langle [U, D], M\rangle = \sum_{i,j}[U, D]_{ij}M_{ij} = \sum_i [U, D]_{ii}\mu_i = 0.$$

Además, $[U, D] = 0$ sólo si $U = 0$, pues $\lambda_i - \lambda_j \neq 0$ si $i \neq j$. Luego $\dim T_D\mathscr{C}_D = \dim \mathfrak{so}(n) = n(n-1)/2$. Como $T_D\mathscr{C}_D$ contiene a todas las matrices simétricas con ceros en la diagonal (con la misma cuenta que hicimos antes), $T_D\mathscr{C}_D = \Delta^{\perp}$. □

3. Comentarios finales

3.1. Grupos de Lie. Los grupos con los que trabajamos en estas notas son ejemplos de los llamados grupos de Lie. En un grupo de Lie G, el espacio tangente en la identidad $T_eG = \mathfrak{g}$ admite un corchete de Lie $[\cdot, \cdot]$ que satisface para todos $X, Y, Z \in \mathfrak{g}$:

(a) $[\cdot, \cdot]$ es bilineal;
(b) $[X, Y] = -[Y, X]$, es decir el corchete es antisimétrico;
(c) se cumple la identidad de Jacobi

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

Informalmente, puede pensarse que el álgebra de Lie $\mathfrak{g}$ de un grupo de Lie G tiene codificada buena parte de la estructura de grupo de G. Aquí también se puede definir una función exponencial $\exp : \mathfrak{g} \to G$, métricas invariantes a izquierda, etc.

Las álgebras de Lie de algunos de los ejemplos que vimos aquí son:

- $\mathfrak{gl}(n, \mathbb{R}) = M(n, \mathbb{R})$ es el álgebra de Lie de $\mathrm{GL}(n, \mathbb{R})$;
- $\mathfrak{so}(n)$ es el álgebra de Lie de $\mathrm{O}(n)$ y $\mathrm{SO}(n)$;
- $\mathfrak{sl}(n, \mathbb{R}) = \{A \in \mathfrak{gl}(n, \mathbb{R}) : \mathrm{traza}\, A = 0\}$ es el álgebra de Lie de $\mathrm{SL}(n, \mathbb{R})$;
- $\mathfrak{gl}(n, \mathbb{C}) = M(n, \mathbb{C})$ es el álgebra de Lie de $\mathrm{GL}(n, \mathbb{C})$;
- $\mathfrak{sl}(n, \mathbb{C}) = \{A \in \mathfrak{gl}(n, \mathbb{C}) : \mathrm{traza}\, A = 0\}$ es el álgebra de Lie de $\mathrm{SL}(n, \mathbb{C})$;
- $\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n, \mathbb{C}) : A + A^* = 0\}$ es el álgebra de Lie de $\mathrm{U}(n)$;
- $\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$ es el álgebra de Lie de $\mathrm{SU}(n)$.

En todos los caso el corchete de Lie está definido por $[A, B] = AB - BA$ para todos $A, B \in \mathfrak{g}$.

Los grupos de Lie y las álgebras de Lie son un objeto de estudio en sí mismo y son la base de una teoría muy profunda. Pero también son un buen ejemplo de cómo la geometría se relaciona con otras ramas de la matemática y de cómo ideas geométricas pueden servir para obtener resultados en otras áreas.

3.2. Generalizaciones de la prueba del Teorema 2.1. Para terminar, quisiéramos observar, al igual que se hace en [3], que existe un teorema análogo al Teorema 2.1 para matrices hermitianas (i.e. autoadjuntas con respecto al producto hermitiano) de tamaño $n \times n$ sobre los números complejos $\mathbb{C} = \mathbb{R} + i\mathbb{R}$. El teorema anterior también es cierto sobre los cuaterniones $\mathbb{H} = \mathbb{C} + j\mathbb{C}$ (cfr. [4]) e incluso sobre los octoniones $\mathbb{O} = \mathbb{H} + \ell\mathbb{H}$, para matrices 3×3 (cfr. [1]). Notemos que la multiplicación en $\mathbb{H}$ y en $\mathbb{O}$ no es conmutativa (en $\mathbb{O}$ ni siquiera es asociativa) y por ende no es posible definir una función determinante de manera natural. Luego, las pruebas clásicas no se adaptan a estos casos. Sin embargo, la prueba geométrica de Eschenburg sí lo hace.

Referencias

[1] J. Baez, *The octonions*, Bull Amer. Math. Soc. **39** (2001), no. 2, 145–205.

[2] M. P. do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser, Boston, 1992.

[3] J.-H. Eschenburg, *How geometry may help understanding linear algebra*, http://www.math.uni-augsburg.de/~eschenbu/linalggeom.pdf (2010).

[4] D. Farenick and B. Pidkowich, *The spectral theorem in quaternions*, Linear Algebra Appl. **371** (2003), 75–102.

[5] L. A. Steen, *Highlights in the history of spectral theory*, Amer. Math. Monthly **80** (1973), 359–381.

[6] F. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983.

Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba

E-mail address: reggiani@famaf.unc.edu.ar