II Escuela de Verano en Matemáticas

INSTITUTO DE MATEMÁTICAS, UNAM
UNIDAD CUERNAVACA

(20-25 Agosto 2001)

Grupos de Lie de Matrices

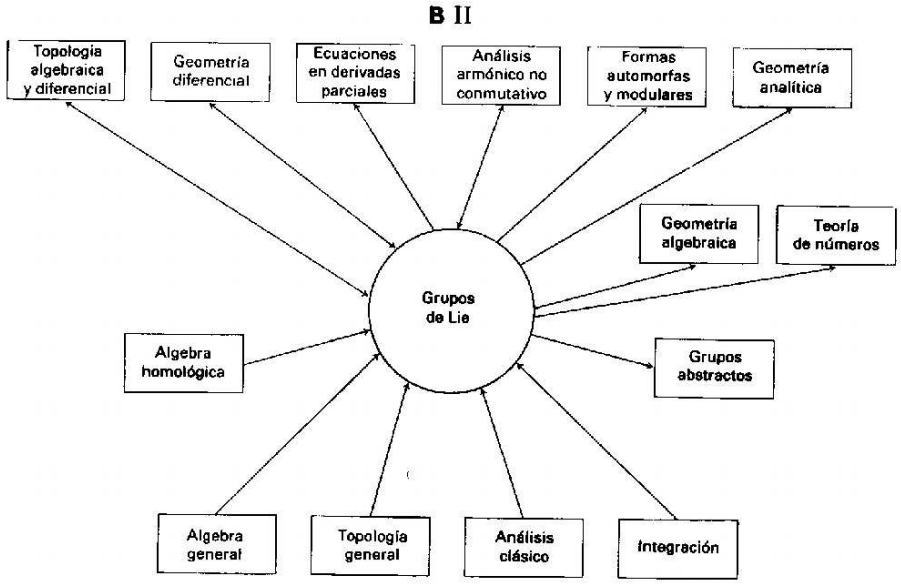
José Luis Cisneros Molina

Índice general

Capítulo 1

Introducción

La teoría de Grupos de Lie fué creada a partir de 1873 por el matemático noruego Sophus Lie (1842-1899). Los grupos de Lie son muy importantes pues están relacionados con muchas ramas de las matemáticas como lo muestra el siguiente diagrama tomado de [3, **B**II]:



Como su nombre lo indica, este curso es una breve introducción a dicha teoría. Se estudiarán los grupos de matrices, que son los ejemplos más sencillos de grupos de Lie. Las presentes notas siguen muy de cerca los primeros cuatro capítulos de [2]. Otras buenas referencias son [6, 5, 4, 1].

Las presentes notas se encuentran disponibles en la siguiente dirección:

`http://www.matcuer.unam.mx/~jlcm`.

Cualquier sugerencia para mejorarlas será bienvenido (`jlcm@matcuer.unam.mx`).

Capítulo 2

Preliminares

2.1. Grupos

Un **grupo** G es un conjunto no vacío, junto con una **operación binaria**, es decir, una función

$$\cdot : G \times G \to G,$$

en donde se satisfacen las siguientes propiedades:

1. La operación es **asociativa**, es decir, para todo $a, b, c \in G$ tenemos (denotando al elemento $\cdot(a, b) \in G$ por $a \cdot b$)

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

2. Existe un **elemento identidad** $e \in G$, es decir, tal que para todo $a \in G$ tenemos $e \cdot a = a \cdot e = e$.

3. Existen los **inversos**, esto es, para todo $a \in G$ existe un elemento que denotaremos por $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$.

De las propiedades anteriores se puede demostrar que un grupo posee únicamente un elemento identidad y que cada $a \in G$ tiene un único inverso.
Ejemplos:

1. El conjunto $\mathbb{Z}$ de los números enteros es un grupo bajo la adición, el elemento identidad es el 0 y el inverso de a es $-a$.

   Sin embargo, $\mathbb{Z}$ no es un grupo bajo la multiplicación, pues a pesar de que dicha operación es asociativa y tiene al 1 como elemento identidad, no existen los inversos.

2. El conjunto $\mathbb{Q}$ de los números racionales es un grupo bajo la adición.

3. El conjunto $\mathbb{Q} \setminus \{0\}$ es un grupo bajo la multiplicación.

4. El conjunto $\mathbb{R}$ de los números reales con la adición.

5. El conjunto $\mathbb{R}^+ = \{ x \in \mathbb{R} \mid x > 0 \}$ de los números reales positivos forma un grupo bajo la multiplicación.
6. El conjunto $\mathbb{R} \setminus \{0\}$ bajo la multiplicación.
7. El conjunto $\mathbb{R}^n$ forma un grupo bajo la operación de adición de vectores.
8. El conjunto de todas las biyecciones de un conjunto de n elementos bajo la composición forma un grupo.

Un grupo es **abeliano** si todos sus elementos conmutan entre si. De los ejemplos anteriores, todos son abelianos excepto el último.

De las funciones entre dos grupos, nos interesan las que preservan las operaciones.

Sean G y H dos grupos. Una función $\sigma\colon G \to H$ es un **homomorfismo** si para toda $a, b \in G$ tenemos

$$\sigma(ab) = \sigma(a)\sigma(b).$$

Es fácil demostrar que un homomorfismo manda a la identidad en G a la identidad en H y que también manda inversos en inversos.

Un homomorfismo que es biyectivo se llama **isomorfismo** y decimos que dos grupos son **isomorfos** si existe un isomorfismo entre ellos. Dos grupos isomorfos son esencialmente iguales.

2.1 Ejemplo. Como ejemplo encontraremos un homomorfismo entre los ejemplos 4 y 6 descritos anteriormente que nos relaciones las operaciones de adición y multiplicación en los reales, es decir, una aplicación $\phi\colon \mathbb{R} \to \mathbb{R} \setminus \{0\}$ tal que

$$\phi(x + y) = \phi(x)\phi(y), \qquad \phi(0) = 1.$$

Queremos encontrar a ϕ explícitamente, para lo cual, consideremos su derivada

$$\begin{aligned}
\phi'(x) &= \lim_{h \to 0} \frac{\phi(x + h) - \phi(x)}{h} \\
&= \lim_{h \to 0} \frac{\phi(x)\phi(h) - \phi(x)}{h} \\
&= \lim_{h \to 0} \phi(x) \frac{\phi(h) - 1}{h} \\
&= \phi(x) \lim_{h \to 0} \frac{\phi(h) - 1}{h}.
\end{aligned}$$

Nótese que

$$\lim_{h \to 0} \frac{\phi(h) - 1}{h} = \phi'(0),$$

entonces

$$\phi'(x) = \phi'(0)\phi(x).$$

Resolviendo esta ecuación diferencial obtenemos que

$$\phi(x) = e^{cx},$$

donde $c = \phi'(0)$. Este ejemplo nos muestra que para estudiar los homomorfismos es muy útil contar con el concepto de derivada.

2.2. Cuaternios

Recordemos que un **campo** k es un conjunto que tiene operaciones de adición y multiplicación que satisfacen las siguientes propiedades:

1. Propiedad distributiva:
$$a(b+c) = ab + ac.$$
2. k es un grupo abeliano bajo la adición, donde la identidad es denotada por 0.
3. $k \setminus \{0\}$ es un grupo abeliano bajo la multiplicación.

Los ejemplos mas comunes de campos son los números racionales $\mathbb{Q}$, los números reales $\mathbb{R}$ y los números complejos $\mathbb{C}$. Recordemos que podemos interpretar a los números complejos como al conjunto $\mathbb{R}^2$ dotado de una multiplicación que junto con la adición de vectores extiende las operaciones de $\mathbb{R}$ para dar a $\mathbb{R}^2$ estructura de campo.

Veamos que no es posible extender las operaciones de $\mathbb{C}$ para dar a $\mathbb{R}^3$ estructura de campo:

2.2 Proposición. *A $\mathbb{R}^3$ no se le puede dar estructura de campo.*

Demostración. Sea $\{1, i, j\}$ una base para $\mathbb{R}^3$, entonces para todo $\mathbf{a} \in \mathbb{R}^3$

$$\mathbf{a} = a + bi + cj \quad \text{con} \quad a, b, c \in \mathbb{R}.$$

Si extendemos la multiplicación de $\mathbb{C}$ a $\mathbb{R}^3$ deberemos tener

$$ij = a + bi + cj \quad \text{con} \quad a, b, c \in \mathbb{R},$$

y así

$$\begin{aligned} i(ij) &= ai - b + cij \\ -j &= ai - b + c(a + bi + cj) \\ -j &= (ac - b) + (a + bc)i + c^2 j \end{aligned}$$

lo cual implicaría que $c^2 = -1$, contradiciendo el hecho de que los coeficientes son reales. □

Sin embargo, es posible definir una multiplicación en $\mathbb{R}^4$ que extiende las operaciones de $\mathbb{C}$, pero debilitando la condición (3), pidiendo únicamente que $\mathbb{R}^4 \setminus \{0\}$ sea un grupo bajo la multiplicación, sin la condición de que sea abeliano. A continuación describiremos dicha multiplicación.

Sea 1, i, j, k una base para $\mathbb{R}^4$. Definimos el producto de los elementos de la base mediante las siguientes relaciones

$$\begin{gathered} \mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = -1 \\ \mathsf{ij} = -\mathsf{ji} = \mathsf{k} \quad \mathsf{jk} = -\mathsf{kj} = \mathsf{k} \quad \mathsf{ki} = -\mathsf{ik} = \mathsf{j}. \end{gathered} \tag{2.1}$$

Con lo anterior podemos multiplicar cualquier cuádrupla de números reales:

$$\begin{aligned}(a+\mathsf{i}b+\mathsf{j}c+\mathsf{k}d)(x+\mathsf{i}y+\mathsf{j}z+\mathsf{k}w) = \\ (ax-by-cz-dw)+\mathsf{i}(ay+bx+cw-dz) \\ +\mathsf{j}(az+cx+dy-bw)+\mathsf{k}(aw+dx+bz-cy).\end{aligned}$$

A $\mathbb{R}^4$ con esta multiplicación lo denotaremos por $\mathbb{H}$ y es llamado los **cuaternios**. Es fácil verificar que esta multiplicación extiende la multiplicación de los complejos $\mathbb{C}$ tomando $c = 0 = d$ y $z = 0 = w$ en la fórmula anterior. También es fácil ver que dicha multiplicación se distribuye con respecto a la adición. Para ver que el conjunto de los elementos distintos de cero forman un grupo bajo la multiplicación solo basta mostrar que todo elemento $p = x+\mathsf{i}y+\mathsf{j}z+\mathsf{k}w$ distinto de cero, tiene un inverso multiplicativo. Un cuaternio $p = x+\mathsf{i}y+\mathsf{j}z+\mathsf{k}w$ tiene un **conjugado** $\bar{p} = x-\mathsf{i}y-\mathsf{j}z-\mathsf{k}w$. Entonces $p\bar{p} = \|p\|^2$, donde la norma, es la norma usual de $\mathbb{R}^4$. Por lo tanto, para todo cuaternio p distinto de cero, su inverso puede escribirse como $p^{-1} = \frac{\bar{p}}{\|p\|^2}$.

Por lo tanto los cuaternios satisfacen las propiedades de campo, salvo por el hecho de que la multiplicación no es conmutativa.

Podemos hacernos la pregunta en general. ¿Para cuales valores de n podemos dotar a $\mathbb{R}^n$ con un producto

$$\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

sin divisores de cero? El siguente teorema nos da la respuesta.

2.3 Teorema. *Las únicas dimensiones para las cuales se puede definir una multiplicación $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ de manera que no se tengan divisores de ceros son $n = 1, 2, 4$ y 8.*

Dichos productos corresponden a $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ y el caso $n = 8$ es conocido como **números de Cayley** u **octonios** en el cual se ha pagado un precio más alto, la multiplicación ya no es asociativa, aunado a la no conmutatividad.

2.3. Vectores y Matrices

En esta sección definiremos vectores y matrices con entradas en $\mathbb{R}$, $\mathbb{C}$ o $\mathbb{H}$ y para hacerlo de manera unificada escribiremos $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$.

Sea $\mathbb{F}^n$ el conjunto de todas las n-adas ordenadas de elementos de $\mathbb{F}$, y definimos la adición de dos elementos de $\mathbb{F}^n$ de la manera usual, es decir, mediante la suma de las entradas respectivas de dichos elementos. Como la adición de los elementos de $\mathbb{F}$ es conmutativa, tenemos que con esta adición $\mathbb{F}^n$ es un grupo abeliano con identidad $(0, 0, \ldots, 0)$.

Además, si $c \in \mathbb{F}$ y $x = (x_1, \ldots, x_n) \in \mathbb{F}^n$ definimos

$$cx = (cx_1, \ldots, cx_n) \tag{2.2}$$

$$xc = (x_1c, \ldots, x_nc). \tag{2.3}$$

En el caso de que $\mathbb{F}$ sea $\mathbb{R}$ o $\mathbb{C}$, como la multiplicación es conmutativa, estas definiciones coinciden, y junto con la operación de adición, le dan a $\mathbb{F}^n$ la estructura de espacio vectorial usual.

Relajando la definición usual de espacio vectorial que insiste en que $\mathbb{F}$ sea un campo, podemos hacer lo mismo con $\mathbb{F} = \mathbb{H}$, pero tenemos que tener en cuenta que en este caso (2.2) y (2.3) no coinciden, pues la multiplicación en $\mathbb{H}$ no es conmutativa, por lo tanto tenemos que escoger una de ellas. Aunque no sea lo usual, escogeremos (2.3), pues esto nos permitirá mas adelante, multiplicar matrices con vectores de la manera usual.

Por la misma razón, también tenemos que modificar nuestra definición usual de aplicación lineal y pedir que la multiplicación por escalares se haga por la derecha.

Decimos que una aplicación $\phi\colon \mathbb{F}^n \to \mathbb{F}^n$ es lineal si para todo $c, d \in \mathbb{F}$ y $x, y \in \mathbb{F}^n$ tenemos que

$$\phi(xc + yd) = \phi(x)c + \phi(y)d.$$

En particular podemos ver (haciendo $c = d = 1$) que una aplicación lineal es un homomorfismo del grupo aditivo de $\mathbb{F}^n$.

Es fácil ver que la composición de dos aplicaciones lineales es nuevamente una aplicación lineal.

A continuación, como es usual en álgebra lineal, relacionaremos las aplicaciones lineales con matrices, pero tenemos que tener cuidado con algunos detalles para considerar que en el caso de los cuaternios la multiplicación no es conmutativa.

Sea $M_n(\mathbb{F})$ el conjunto de matrices $n \times n$ con entradas en $\mathbb{F}$. Si $M \in M_n(\mathbb{F})$ y $M = (m_{ij})$ $(m_{ij} \in \mathbb{F})$, podemos definir una aplicación lineal $\phi(M)$ de la siguiente manera:

$$\phi(M)(x) = Mx,$$

donde $x \in \mathbb{F}^n$ es visto como un vector columna y el producto es la multiplicación usual por la izquierda de una matriz por un vector columna. Es fácil ver que la aplicación definida de esta manera es lineal en el sentido que hemos definido anteriormente, es decir, con la multiplicación escalar por la derecha.

Nótese además, que de haber definido la multiplicación escalar de la manera usual por la izquierda, la aplicación $\phi(M)$ **no** sería lineal en el sentido usual, si la multiplicación no es conmutativa, como en el caso de $\mathbb{H}$.

Recíprocamente, dada una aplicación lineal $\phi\colon \mathbb{F}^n \to \mathbb{F}^n$ podemos encontrar una matriz M $n \times n$ tal que $\phi = \phi(M)$. La primer columna de M es la n-ada $\phi(1, 0, 0, \ldots, 0)$, la segunda columna de M es $\phi((0, 1, 0, \ldots, 0)$ y así sucesivamente.

Nótese que si la matriz A da la aplicación lineal ϕ y la matriz B da la aplicación lineal ψ, entonces la matriz BA da la aplicación $\psi \circ \phi$.

Una aplicación lineal es un **isomorfismo** si es inyectiva y suprayectiva. Entonces ϕ^{-1} es también un isomorfismo lineal y $\phi \circ \phi^{-1} = \text{identidad} = \phi^{-1} \circ \phi$. Para las matrices correspondientes esto significa que $M(\phi^{-1})M(\phi) = I = MM(\phi)M(\phi^{-1})$, de tal manera que $M(\phi^{-1})$ es un inverso bilateral para $M(\phi)$.

Daremos al conjunto $M_n(\mathbb{F})$ estructura de espacio vectorial de la siguiente manera:

1. Si $a = (a_{ij})$ y $B = (b_{ij})$, entonces
$$A + B = (a_{ij} + b_{ij});$$
2. Si $A = (a_{ij})$ y $c \in \mathbb{F}$, entonces
$$cA = (ca_{ij}).$$

Nótese que $M_n(\mathbb{F})$ no solamente es un espacio vectorial, pues posee una multiplicación que se distribuye respecto a la suma por ambos lados y además tiene un identidad multiplicativa (la matriz identidad). Un sistema como este es llamado un **álgebra**.

Si $\mathcal{A}$ es un álgebra, un elemento $x \in \mathcal{A}$ es una **unidad** si existe alguna $y \in \mathcal{A}$, tal que $xy = 1 = yx$, es decir, si tiene un inverso multiplicativo.

De la definición anterior, es inmediato ver que si $\mathcal{A}$ es un álgebra con multiplicación asociativa y U es el conjunto de sus unidades, entonces U es un grupo bajo la multiplicación.

Un **Grupo de Lie de Matrices** es un subgrupo del grupo de unidades del álgebra de matrices $M_n(\mathbb{F})$.

2.4. Grupos Generales Lineales

Nuestros primeros ejemplos de grupos de Lie de matrices son los **grupos generales lineales** que son los siguientes:

- El grupo de unidades en el álgebra $M_n(\mathbb{R})$ es denotado por $GL_n(\mathbb{R})$.
- El grupo de unidades en el álgebra $M_n(\mathbb{C})$ es denotado por $GL_n(\mathbb{C})$.
- El grupo de unidades en el álgebra $M_n(\mathbb{H})$ es denotado por $GL_n(\mathbb{H})$.

Nótese que $A \in M_n(\mathbb{F})$ es una unidad si y sólo si A representa un isomorfismo de $\mathbb{F}^n$.

Una matriz 1×1 sobre $\mathbb{F}$ es solamente un elemento de $\mathbb{F}$ y la multiplicación de dichas matrices es la multiplicación como elementos de $\mathbb{F}$. Como la tercera propiedad que define a $\mathbb{F}$ nos dice que $\mathbb{F} \setminus \{0\}$ es un grupo, tenemos que

$$\begin{aligned} GL_1(\mathbb{R}) &= \mathbb{R} \setminus \{0\} \\ GL_1(\mathbb{C}) &= \mathbb{C} \setminus \{0\} \\ GL_1(\mathbb{H}) &= \mathbb{H} \setminus \{0\}. \end{aligned}$$

Para los casos de $\mathbb{R}$ y $\mathbb{C}$, tenemos la función determinante definida en $M_n(\mathbb{R})$ y $M_n(\mathbb{C})$ y sabemos que una matriz es invertible, es decir tiene un inverso, si y sólo si su determinante es distinto de cero, es decir

$$\begin{aligned} GL_n(\mathbb{R}) &= \{\, a \in M_n(\mathbb{R}) \mid \det A \neq 0 \,\} \\ GL_n(\mathbb{C}) &= \{\, a \in M_n(\mathbb{C}) \mid \det A \neq 0 \,\}. \end{aligned}$$

Dado que el producto de cuaternios no es conmutativo, no es posible definir un determinante convencional en este caso, sin embargo, es posible definir un determinante con valores complejos para la matrices con entradas cuaterniónicas, de tal modo que éstas sean invertibles si y sólo si dicho determinante es no nulo. Para ello, definiremos un homomorfismo inyectivo

$$\Psi\colon GL_n(\mathbb{H}) \to GL_{2n}(\mathbb{C}),$$

por lo tanto, $GL_n(\mathbb{H})$ será isomorfo a su imagen en $GL_{2n}(\mathbb{C})$, bajo Ψ y para $A \in GL_n(\mathbb{H})$ asignaremos como determinante de A el determinante de $\Psi(A)$.

Primeramente, consideremos la siguiente aplicación:

$$\psi\colon \mathbb{H} \to M_2(\mathbb{C}),$$

definida por

$$\psi(x + \mathsf{i}y + \mathsf{j}z + \mathsf{k}w) = \begin{pmatrix} x + iy & -z - iw \\ z - iw & x - iy \end{pmatrix}.$$

Es rutina verificar que ψ es un homomorfismo de álgebras, es decir, que preserva la adición y la multiplicación y además que es inyectiva.

Pasemos ahora a definir $\Psi\colon M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$. Sea $A \in M_n(\mathbb{H})$, definimos

$$\Psi(A) = (\psi(a_{ij})), \tag{2.4}$$

es decir, $\psi(A)$, es la matriz compleja $2n \times 2n$ cuyo bloque 2×2 en la posición ij está dado por $\psi(a_{ij})$.

2.4 Lema. $\Psi(AB) = \Psi(A)\Psi(B)$.

Demostración. Sea $A = (\alpha_{\mu\nu})$ y $B = (\beta_{\mu\nu})$. Entonces

$$\begin{aligned} (AB)_{ij} &= \alpha_{i1}\beta_{1j} + \cdots + \alpha_{in}\beta_{nj} \\ (\Psi(AB)_{ij}) &= \psi(\alpha_{i1})\psi(\beta_{1j}) + \cdots + \psi(\alpha_{in})\psi(\beta_{nj}), \end{aligned}$$

ya que ψ es un homomorfismo de álgebras, y tenemos que esta es precisamente la entrada ij de $\Psi(A)\Psi(B)$. □

2.5 Proposición. *Si $A \in M_n(\mathbb{H})$, entonces $A \in GL_n(\mathbb{H})$ si y sólo si* $\det \Psi(A) \neq 0$.

Demostración. Si $A \in GL_n(\mathbb{H})$ implica que existe $A^{-1} \in GL_n(\mathbb{H})$ con $AA^{-1} = I = A^{-1}A$. Entonces $\Psi(A)$ tiene como inverso a $\Psi(A)^{-1}$ en $GL_{2n}(\mathbb{C})$ y entonces $\det \Psi(A)$ es no nulo. Por otro lado, si $\det \Psi(A) \neq 0$, tenemos que demostrar que $\Psi(A)^{-1} \in \Psi(GL_n(\mathbb{H}))$. Esto es consecuencia del siguiente resultado general sobre álgebras. □

2.6 Proposición ([2, Prop. 2.10]). *Sea $\mathcal{B}$ un álgebra asociativa de dimensión finita con 1. Sea $\mathcal{A}$ una subálgebra de $\mathcal{B}$ (con $1 \in \mathcal{A}$). Si $U(\mathcal{A})$ y $U(\mathcal{B})$ denotan al grupo de unidades respectivo, entonces*

$$U(\mathcal{A}) = \mathcal{A} \cap U(\mathcal{B}).$$

Uno de los objetivos del presente curso es estudiar algunos subgrupos de los grupos generales lineales los cuales definiremos en la siguiente capítulo.

Capítulo 3

Grupos ortogonales

3.1. Productos interiores

Como mencionamos anteriormente, al igual que los complejos, para todo cuaternio existe su elemento conjugado. Dado que los cuaternios son una extensión de los complejos y estos a su vez, lo son de los reales, es decir $\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$, definiremos la noción de conjugación de una manera consistente.

$$\begin{aligned} \text{Para } x \in \mathbb{R}, \quad & \overline{x} = x, \\ \text{para } \alpha = x + iy \in \mathbb{C}, \quad & \overline{\alpha} = x - iy, \\ \text{para } q = x + \mathsf{i}y + \mathsf{j}z + \mathsf{k}w \in \mathbb{H}, \quad & \overline{q} = x - \mathsf{i}y - \mathsf{j}z - \mathsf{k}w. \end{aligned}$$

La conjugación tiene las siguientes propiedades:

- $\overline{\overline{\alpha}} = \alpha$,
- $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$,
- $\overline{\alpha\beta} = \overline{\beta}\overline{\alpha}$, que en el caso de $\mathbb{R}$ y $\mathbb{C}$ es lo mismo que $\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}$.

Sea $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$, definimos un **producto interior** $\langle\ ,\ \rangle$ en $\mathbb{F}^n$ por

$$\langle x, y \rangle = \overline{x_1} y_1 + \cdots + \overline{x_n} y_n.$$

Este producto interior satisface las siguientes propiedades:

1. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$,
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
3. $\overline{a}\langle xa, y \rangle$, $\langle x, ya \rangle = \langle x, y \rangle a$,
4. $\overline{\langle x, y \rangle} = \langle y, x \rangle$,
5. $0 \leq \langle x, x \rangle \in \mathbb{R}$,

6. $\langle x, x\rangle = 0 \Leftrightarrow x = (0, \ldots, 0)$,

7. Si $e_1, \ldots, e_n$ es la base estándar para $\mathbb{F}^n$, es decir $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, entonces
$$\langle e_i, e_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

8. Es no degenerado, es decir:
$$\text{Si } \langle x, y\rangle = 0 \text{ para toda } y, \text{ entonces } x = (0, \ldots, 0);$$
$$\text{Si } \langle x, y\rangle = 0 \text{ para toda } x, \text{ entonces } y = (0, \ldots, 0).$$

La norma $\|x\|$ de $x \in \mathbb{F}^n$ se define como
$$\|x\| = \sqrt{\langle x, x\rangle},$$
la cual corresponde al valor absoluto en $\mathbb{R}$, al módulo en $\mathbb{C}$ y a la norma estándar de $\mathbb{R}^4$ para el caso de $\mathbb{H}$.

Recordemos que si $A \in M_n(\mathbb{F})$, su **conjugado** $\overline{A}$ se obtiene reemplazando cada entrada a_{ij} por $\overline{a_{ij}}$ y su **traspuesta** A^t se obtiene reemplazando cada entrada a_{ij} por a_{ji}. Estas dos operaciones conmutan, por lo que no hay ambigüedad al definir la **conjugada traspuesta** y definirla por $\overline{A}^t$.

3.1 Proposición. *Para todo $x, y \in \mathbb{F}^n$ y $A \in M_n(\mathbb{F})$ se tiene que*
$$\langle Ax, y\rangle = \langle x, \overline{A}^t y\rangle.$$

Demostración. Sea $A = (a_{ij})$. Entonces
$$Ax = (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{n1}x_1 + \cdots + a_{nn}x_n)$$
$$\overline{A}^t y = (\overline{a}_{11}y_1 + \cdots + \overline{a}_{n1}y_n, \ldots, \overline{a}_{1n}y_1 + \cdots + \overline{a}_{nn}y_n).$$
Entonces el miembro izquierdo de la igualdad $\langle Ax, y\rangle$ está dado por
$$(\overline{x}_1\overline{a}_{11} + \cdots + \overline{x}_n\overline{a}_{1n})y_1 + \cdots + (\overline{x}_1\overline{a}_{n1} + \cdots + \overline{x}_n\overline{a}_{nn})y_n,$$
y el miembro derecho está dado por
$$\overline{x}_1(\overline{a}_{11}y_1 + \cdots + \overline{a}_{n1}y_n) + \cdots + \overline{x}_n(\overline{a}_{1n}y_1 + \cdots + \overline{a}_{nn}y_n).$$
Es fácil ver que ambos contienen los mismos términos. □

3.2. Grupos ortogonales

Consideremos el subconjunto de $M_n(\mathbb{F})$ de las matrices que preservan el producto interior:
$$\mathcal{O}(n, \mathbb{F}) = \{\, A \in M_n(\mathbb{F}) \mid \langle Ax, Ay\rangle = \langle x, y\rangle \quad \text{for all } x, y \in \mathbb{F}^n \,\}.$$

3.2 Proposición. *El subconjunto $\mathcal{O}(n,\mathbb{F})$ es un grupo.*

Demostración. En primer lugar, veamos que el producto de matrices restringido a $\mathcal{O}(n,\mathbb{F})$ es cerrado. Sean $A, B \in \mathcal{O}(n,\mathbb{F})$, entonces

$$\langle ABx, ABy\rangle = \langle Bx, By\rangle = \langle x, y\rangle,$$

por lo que

$$AB \in \mathcal{O}(n,\mathbb{F}).$$

Claramente la matriz identidad I pertenece a $\mathcal{O}(n,\mathbb{F})$.

Ahora veamos que si $A \in \mathcal{O}(n,\mathbb{F})$, entonces A tiene inversa y que esta pertenece a $\mathcal{O}(n,\mathbb{F})$.

Si $A \in \mathcal{O}(n,\mathbb{F})$, entonces

$$\langle Ae_i, Ae_j\rangle = \langle e_i, e_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Tenemos que Ae_i es la i-ésima columna de A que a su vez es el i-ésimo renglón de $\overline{A}^t$ y por lo tanto $\langle Ae_i, Ae_j\rangle$ es la entrada ij del producto

$$\overline{A}^t A.$$

Por lo tanto $\overline{A}^t A = I$. Además, también tenemos que $A\overline{A}^t = I$, ya que $\overline{A\overline{A}^t}^t = (A^t\overline{A})^t = \overline{A}^t A = I$. Por lo que tenemos que $\overline{A}^t = A^{-1}$. Ahora veamos que este inverso está en $\mathcal{O}(n,\mathbb{F})$, pero esto es cierto porque

$$\langle A^{-1}x, A^{-1}y\rangle = \langle AA^{-1}x, AA^{-1}y\rangle = \langle x, y\rangle,$$

demostrando con esto la proposición □

Por lo anterior, podemos definir tres nuevas familias de grupos

- Para $\mathbb{F} = \mathbb{R}$, escribimos $\mathcal{O}(n,\mathbb{F})$ como $O(n)$ y lo llamamos el **grupo ortogonal**.
- Para $\mathbb{F} = \mathbb{C}$, escribimos $\mathcal{O}(n,\mathbb{F})$ como $U(n)$ y lo llamamos el **grupo unitario**.
- Para $\mathbb{F} = \mathbb{H}$, escribimos $\mathcal{O}(n,\mathbb{F})$ como $Sp(n)$ y lo llamamos el **grupo simpléctico**.

3.3 Proposición. *Sea $A \in M_n(\mathbb{F})$. Entonces las siguientes condiciones son equivalentes:*

(a) $A \in \mathcal{O}(n,\mathbb{F})$,

(b) $\langle Ae_i, Ae_j\rangle = \delta_{ij}$,

(c) A manda bases ortonormales en bases ortonormales,

(d) Los renglones de A forman una base ortonormal,

(e) Las columnas de A forman una base ortonormal,

(f) $\overline{A}^t = A^{-1}$.

Demostración. Ejercicio. □

3.4 Proposición. *Sea $A \in M_n(\mathbb{R})$. Entonces $A \in O(n)$ si y sólo si A preserva la norma.*

Demostración. La matriz A preserva la longitud si y sólo si $\langle Ax, Ax\rangle = \langle x, x\rangle$ para toda $x \in \mathbb{R}^n$. Por lo tanto $\Rightarrow$ es trivial. Recíprocamente tenemos

$$\langle A(x+y), A(x+y)\rangle = \langle x+y, x+y\rangle$$
$$\langle Ax, Ax\rangle + \langle Ax, Ay\rangle + \langle Ay, Ax\rangle + \langle Ay, Ay\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle.$$

Lo cual nos da

$$\langle Ax, Ay\rangle + \langle Ay, Ax\rangle = \langle x, y\rangle + \langle y, x\rangle,$$

y como $\langle\ ,\ \rangle$ es simétrico sobre $\mathbb{R}$, entonces tenemos que

$$\langle Ax, Ay\rangle = \langle x, y\rangle,$$

es decir, que $A \in O(n)$. □

3.5 Proposición. *La Proposición 3.4 se cumple para $\mathbb{C}$ y $\mathbb{H}$.*

Demostración. Como en la proposición anterior, calculemos

$$\langle A(e_i + e_j), A(e_i + e_j)\rangle = \langle e_i + e_j, e_i + e_j\rangle$$

para obtener

$$\langle Ae_i, Ae_j\rangle + \langle Ae_j, Ae_i\rangle = 0. \tag{3.1}$$

Por otro lado, consideremos $x = e_i x_i + e_j x_j$ y calculemos $\langle Ax, Ax\rangle = \langle x, x\rangle$ y obtenemos que

$$\overline{x_j}\langle Ae_j, Ae_i\rangle x_i + \overline{x_i}\langle Ae_i, Ae_j\rangle x_j = 0. \tag{3.2}$$

Ahora veamos los casos por separado. Si $A \in M_n(\mathbb{C})$ y $x_i, x_j \in \mathbb{C}$, entonces de (3.2) tenemos que

$$\overline{x_j} x_i \langle Ae_j, Ae_i\rangle + \overline{x_i} x_j \langle Ae_i, Ae_j\rangle = 0,$$

y usando (3.1)

$$\langle Ae_i, Ae_j\rangle(\overline{x_i}x_j - \overline{x_j}x_i) = 0,$$

lo que fuerza que $\langle Ae_i, Ae_j\rangle = 0$.

Para el caso $A \in M_n(\mathbb{H})$ y $x_i, x_j \in \mathbb{H}$, tenemos que (3.1) es equivalente a

$$\langle Ae_j, Ae_i\rangle + \overline{\langle Ae_j, Ae_i\rangle} = 0,$$

lo que implica que su parte real es cero, es decir $\langle Ae_j, Ae_i\rangle$ es de la forma

$$\langle Ae_j, Ae_i\rangle = a\mathsf{i} + b\mathsf{j} + c\mathsf{k}. \tag{3.3}$$

Substituyendo (3.3) en (3.2) y haciendo $x_i = \mathsf{i}$ y $x_j = \mathsf{j}$, tenemos

$$\begin{aligned} -\mathsf{j}(a\mathsf{i} + b\mathsf{j} + c\mathsf{k})\mathsf{i} - \mathsf{i}(-a\mathsf{i} - b\mathsf{j} - c\mathsf{k})\mathsf{j} &= 0 \\ -\mathsf{j}(-a - b\mathsf{k} + c\mathsf{j}) - \mathsf{i}(-a\mathsf{k} + b + c\mathsf{i}) &= 0 \\ 2c &= 0. \end{aligned}$$

Análogamente, haciendo $x_i = \mathsf{j}$ y $x_j = \mathsf{k}$ obtenemos que $a = 0$ y haciendo $x_i = \mathsf{i}$ y $x_j = \mathsf{k}$ obtenemos $b = 0$. Por lo tanto $\langle Ae_j, Ae_i\rangle = 0$.

□

Veamos como son los grupos $O(1), U(1)$ y $Sp(1)$. El grupo $O(1)$ es el conjunto de todos los números reales de norma uno, por lo tanto $O(1) = \{1, -1\}$. El grupo $U(1)$ es el conjunto de los números complejos de norma uno, es decir, el círculo unitario S^1. Finalmente, $Sp(1)$ es el grupo de los cuaternios de norma uno. Si definimos a la $(k-1)$-esfera unitaria como

$$S^{k-1} = \{\, x \in \mathbb{R}^k \mid \|x\| = 1 \,\},$$

tenemos que

$$O(1) = S^0, \quad U(1) = S^1 \quad Sp(1) = S^3.$$

Es un hecho interesante que estas son las únicas esferas que pueden ser grupos.

3.6 Proposición. *Si* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ *y* $A \in \mathcal{O}(n, \mathbb{F})$, *entonces*

$$(\det A)(\overline{\det A}) = 1.$$

Demostración. Como $A \in \mathcal{O}(n, \mathbb{F})$ tenemos que $A\overline{A}^t = I$, lo que implica que $(\det A)(\det \overline{A}^t) = 1$. Por otro lado tenemos que

$$\det \overline{A}^t = \det \overline{A} = \overline{\det A}.$$

□

Por lo tanto, si $A \in O(n)$, entonces $\det A \in \{1, -1\}$. Definimos

$$SO(n) = \{\, A \in O(n) \mid \det A = 1 \,\},$$

al cual llamaremos el **grupo ortogonal especial** o **grupo de rotaciones**.

De manera análoga, definimos

$$SU(n) = \{\, A \in U(n) \mid \det A = 1 \,\},$$

al cual llamaremos **grupo unitario especial**.

3.3. El problema del isomorfismo

Hasta ahora hemos definido varios grupos de matrices, a saber,

- El grupo general lineal real $GL_n(\mathbb{R})$, que tiene como subgrupo al grupo ortogonal $O(n)$ y este a su vez, contiene al grupo ortogonal especial $SO(n)$.
- El grupo general lineal complejo $GL_n(\mathbb{C})$, que tiene como subgrupo al grupo unitario $U(n)$ y este a su vez, contiene al grupo unitario especial $SU(n)$.
- El grupo general lineal cuaterniónico $GL_n(\mathbb{H})$, que tiene como subgrupo al grupo simpléctico $Sp(n)$.

Una pregunta natural es si todos estos grupos son diferentes o si algunos de ellos son isomorfos entre si. Para probar que dos grupos son isomorfos, tenemos que dar un isomorfismo entre ellos, lo cual puede ser difícil. Por otro lado, para probar que no lo son, tenemos que demostrar que no existe *ningún* isomorfismo entre ellos, lo cual puede ser aún más difícil. Otra forma más fácil de atacar el problema, consiste en buscar propiedades que sean invariantes bajo isomorfismo, de ésta manera, si un grupo posee dicha propiedad y otro no, entonces no pueden ser isomorfos.

A continuación daremos un isomorfismo entre dos de los grupos de matrices que hemos definido: $Sp(1)$, el grupo de todos los cuaternios de norma uno; y $SU(2)$, el grupo de todas las matrices A, 2×2 complejas, tales que $A\overline{A}^t = I$ y $\det A = 1$. La operación en $Sp(1)$ es la multiplicación de cuaternios y en $SU(2)$ es la multiplicación de matrices.

3.7 Proposición. *El homomorfismo* $\Psi\colon M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$ *definido por* (2.4) *en la Sección 2.4 induce un isomorfismo*

$$\Psi\colon Sp(1) \to SU(2).$$

Demostración. Hemos visto que Ψ induce un homomorfismo inyectivo de $GL_n(\mathbb{H})$ en $GL_{2n}(\mathbb{C})$, por lo tanto su restricción a $Sp(1)$ es también inyectiva. Entonces sólo tenemos que demostrar que

(a) si $A \in Sp(1)$ entonces $\Psi(A) \in SU(2)$, y

(b) toda $B \in SU(2)$ es la imagen $\Psi(A)$ para alguna $A \in Sp(1)$.

Si $A = a + \mathsf{i}b + \mathsf{j}c + \mathsf{k}d$ entonces $\Psi(A) = \left(\begin{smallmatrix} a+ib & -c-id \\ c-id & a-ib \end{smallmatrix}\right)$, por lo que

$$\Psi(A)\overline{\Psi(A)}^t = \begin{pmatrix} a+ib & -c-id \\ c-id & a-ib \end{pmatrix}\begin{pmatrix} a-ib & c+id \\ -c+id & a+ib \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

ya que $a^2+b^2+c^2+d^2 = 1$. Además $\det \Psi(A) = 1$ por lo que $\Psi(A) \in SU(2)$.

Sea $B = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right) \in SU(2)$. Como $\det B = 1$ y del hecho de que los renglones de B son vectores unitarios ortogonales tenemos que

$$\delta = \overline{\alpha} \quad \text{y} \quad \gamma = -\overline{\beta}.$$

Por lo tanto, si $\alpha = a + ib$ y $\beta = -c - id$, podemos tomar $A = a + \mathsf{i}b + \mathsf{j}c + \mathsf{k}d$ y tendremos $\Psi(A) = B$ y $a^2 + b^2 + c^2 + d^2 = 1$. □

Capítulo 4

Homomorfismos

4.1. Curvas en un espacio vectorial

En esta sección definiremos un invariante de los grupos de matrices, definiremos su **dimensión**. Dos grupos de matrices que no tengan la misma dimensión no podrán ser isomorfos. Para poder definir la dimensión necesitamos definir primero el espacio de vectores tangentes de un grupo de matrices.

Sea V un espacio vectorial de dimensión finita. Una **curva** γ en V es una función continua $\gamma\colon (a,b) \to V$, donde (a,b) es un intervalo abierto en $\mathbb{R}$.

Para $c \in (a,b)$ decimos que γ es **diferenciable** en c si el límite

$$\lim_{h\to 0} \frac{\gamma(c+h)-\gamma(c)}{h}$$

existe. En caso de existir, dicho límite es un vector en V al cual denotaremos por $\gamma'(c)$ y lo llamaremos el **vector tangente** a γ en $\gamma(c)$.

Si escogemos una base para V y representamos a la curva γ como $(\gamma_1,\ldots,\gamma_n)$, donde las γ_i son funciones con valores reales, entonces $\gamma'(c)$ existe si y sólo si cada $\gamma_i'(c)$ existe y

$$\gamma'(c) = (\gamma_1'(c),\ldots,\gamma_n'(c)).$$

Tenemos que $M_n(\mathbb{R})$, $M_n(\mathbb{C})$ y $M_n(\mathbb{H})$ pueden ser considerados espacios vectoriales reales de dimensiones n^2, $2n^2$ y $4n^2$ respectivamente, por lo que podemos considerar curvas en dichos espacios. Si G es un grupo de matrices en $M_n(\mathbb{F})$ entonces una curva en G es una curva en $M_n(\mathbb{F})$ con todos sus valores $\gamma(u)$ para $u \in (a,b)$ en G.

Supongamos que tenemos curvas $\gamma,\sigma\colon (a,b) \to G$. Definimos una nueva curva, la **curva producto** por

$$(\gamma\sigma)(u) = \gamma(u)\sigma(u).$$

4.1 Proposición. *Sean $\gamma,\sigma\colon (a,b) \to G$ curvas diferenciables en $c \in (a,b)$. Entonces la curva producto $\gamma\sigma$ es diferenciable en c y*

$$(\gamma\sigma)'(c) = \gamma(c)\sigma'(c) + \gamma'(c)\sigma(c).$$

Demostración. Sea $\gamma(u) = (\gamma_{ij}(u))$, $\sigma(u) = (\sigma_{ij}(u))$. Entonces

$$(\gamma\sigma)(u) = (\sum_k \gamma_{ik}(u)\sigma_{kj}(u)),$$

por lo tanto

$$\begin{aligned}(\gamma\sigma)'(u) &= (\sum_k \{\gamma'_{ik}(u)\sigma_{kj}(u) + \gamma_{ik}(u)\sigma'_{kj}(u)\}) \\ &= \gamma'(u)\sigma(u) + \gamma(u)\sigma'(u).\end{aligned}$$

□

4.2 Proposición. *Sea G un grupo de matrices en $M_n(\mathbb{F})$. Sea T el conjunto de todos los vectores tangentes $\gamma'(0)$ a curvas $\gamma\colon (a,b) \to G$, con $\gamma(0) = I$, $(0 \in (a,b))$. Entonces T es un subespacio de $M_n(\mathbb{F})$.*

Demostración. Si $\gamma'(0)$ y $\sigma'(0)$ están en T, entonces $(\gamma\sigma)(0) = \gamma(0)\sigma(0) = II = I$ y

$$(\gamma\sigma)'(0) = \gamma'(0)\sigma(0) + \sigma(0)\sigma'(0) = \gamma'(0) + \sigma'(0),$$

por lo tanto T es cerrada bajo adición.

Si $\gamma'(0) \in T$ y $r \in \mathbb{R}$ y hacemos

$$\sigma(u) = \gamma(ru),$$

entonces σ es diferenciable y $\sigma(0) = \gamma(0) = I$ y tenemos que $\sigma'(0) = r\gamma'(0)$. Por lo tanto T es cerrado bajo multiplicación escalar.

Como $M_n(\mathbb{F})$ es un espacio vectorial de dimensión finita, T también lo es. □

Definición. Si G es un grupo de matrices su **dimensión** es la dimensión del espacio vectorial T de vectores tangentes de G en I.

4.3 Ejemplo. El grupo $U(1)$ tiene dimensión 1, pues consiste en los complejos de módulo 1, los cuales forman el círculo unidad.

4.4 Ejemplo. La dimensión de $Sp(1)$ es 3. Sea $\gamma\colon (a,b) \to Sp(1)$ una curva suave con $\gamma(0) = 1$. Entonces $\gamma'(0)$ será un elemento de $\mathbb{H} = \mathbb{R}^4$. En primer lugar,veremos que $\gamma'(0)$ está en el espacio generado por $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$; es decir, es un cuaternio con parte real 0.

Sea $\gamma(t) = x(t) + \mathsf{i}y(t) + \mathsf{j}z(t) + \mathsf{k}w(t)$ con $x(0) = 1$ y $y(0) = 0$, $z(0) = 0$ $w(0) = 0$. Tenemos que $x(0)$ es un máximo para la función x, por lo tanto $\gamma'(0) = 0 + \mathsf{i}y'(0) + \mathsf{j}z'(0) + \mathsf{k}w'(0)$.

Recíprocamente, sea $q = \mathsf{i}\mu + \mathsf{j}\nu + \mathsf{k}\lambda$ un cuaternio con parte real cero. Afirmamos que existe una curva suave γ en $Sp(1)$ tal que $\gamma'(0) = q$. Ciertamente, ya que la curva

$$\gamma(t) = \sqrt{1 - \sin^2 \mu t - \sin^2 \nu t - \sin^2 \lambda t} + \mathsf{i}\sin \mu t + \mathsf{j}\sin \nu t + \mathsf{k}\sin \lambda t,$$

cumple la condición anterior, para $t \in [0, \epsilon)$, con $\epsilon > 0$ pequeña.

4.5 Ejemplo. La dimensión de $GL_n(\mathbb{R})$ es n^2. La función determinante $\det\colon M_n(\mathbb{R}) \to \mathbb{R}$ es continua y $\det I = 1$. Entonces existe alguna bola de radio ϵ alrededor de I en $M_n(\mathbb{R})$ tal que para cada A en esta bola, $\det A \neq 0$; es decir, $A \in GL_n(\mathbb{R})$. Si v es un vector en $M_n(\mathbb{R})$ define una curva σ en $M_n(\mathbb{R})$ de la siguiente manera

$$\sigma(t) = tv + I.$$

Entonces $\sigma(0) = I$, $\sigma'(0) = v$ y pata t pequeña, $\sigma(t) \in GL_n(\mathbb{R})$. Entonces el espacio tangente T es todo $M_n(\mathbb{R})$ cuya dimensión es n^2.

Un argumento similar muestra que $\dim GL_n(\mathbb{C}) = 2n^2$.

A continuación obtendremos cotas superiores para las dimensiones de $O(n)$, $U(n)$ y $Sp(n)$, pero para ello, necesitamos algunos preliminares.

Definición. La matriz $A \in M_n(\mathbb{R})$ es **antisimétrica** si $A + A^t = 0$; es decir, si $(a_{ij}) = (-a_{ji})$ para toda ij. En particular, los términos de la diagonal deben ser cero.

4.6 Proposición. *Sea $so(n)$ el conjunto de todas las matrices antisimétricas en $M_n(\mathbb{R})$. Entonces $so(n)$ es un subespacio vectorial de $M_n(\mathbb{R})$, y su dimensión es $\frac{n(n-1)}{2}$.*

Demostración. Tenemos que $0 \in so(n)$ y si $A, B \in so(n)$ entonces

$$(A + B) + (A + B)^t = A + A^t + B + B^t = 0,$$

por lo que $so(n)$ es cerrada bajo la adición. También es cerrada bajo la multiplicación escalar, ya que si $A \in so(n)$ y $r \in \mathbb{R}$, entonces $(rA)^t = rA^t$ por lo que $rA + (rA)^t = r(A + A^t = 0$.

Para verificar la dimensión $so(n)$ necesitamos una base. Sea E_{ij} la matriz cuyas entradas son todas cero excepto en la entrada ij, la cual es 1, y la entrada ji es -1. Si se definimos E_{ij} solamente para $i < j$, es fácil ver que forman una base para $so(n)$ que tiene

$$(n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$$

elementos. □

Definición. Una matriz $B \in M_n(\mathbb{C})$ es **antihermitiana** si

$$B + \overline{B}^t = 0.$$

Entonces si $b_{jk} = c + id$, tenemos que $\overline{b}_{kj} = -b_{jk} = c - id$ y $b_{kj} = -c + id$. n particular, si $j = k$ se tiene $c + id = -c + id$, por lo que los términos de la diagonal de una matriz antihermitiana son imaginarios puros.

Sea $su(n)$ el conjunto de las matrices antihermitianas en $M_n(\mathbb{C})$. De acuerdo a la observación anterior, se ve que $su(n)$ no es un espacio vectorial sobre $\mathbb{C}$.

4.7 Proposición. *$su(n) \subset M_n(\mathbb{C})$ es un espacio vectorial real de dimensión $n + 2\frac{n(n-1)}{2} = n^2$.*

Definición. Una matriz C en $M_n(\mathbb{H})$ es **antisimpléctica** si

$$C + \overline{C}^t = 0.$$

El conjunto $sp(n)$ de las matrices antisimplécticas en $M_n(\mathbb{H})$ es un espacio vectorial real de dimensión

$$3n + 4\frac{n(n-1)}{2} = n(2n+1).$$

4.8 Proposición. *Si β es una curva que pasa por la identidad, $\beta(0) = I$,*

en $O(n)$, entonces $\beta'(0)$ es antisimétrica,
en $U(n)$, entonces $\beta'(0)$ es antihermitiana,
en $Sp(n)$, entonces $\beta'(0)$ es antisimpléctica.

Demostración. En cada caso, se tiene que la curva producto es constante

$$\beta(u)\overline{\beta}^t(u) = I.$$

Por lo tanto su derivada es cero. □

4.9 Corolario.

- $\dim O(n) \leq \frac{n(n-1)}{2}$,
- $\dim U(n) \leq n^2$,
- $\dim Sp(n) \leq n(2n+1)$.

4.2. Homomorfismos suaves

Sea $\phi\colon G \to H$ un homomorfismo de grupos de matrices. Ya que G y H están contenidos en espacios vectoriales, es claro lo que significa que ϕ sea continua. Siempre que hagamos referencia a un homomorfismo los consideraremos continuo.

Una curva

$$\rho\colon (a,b) \to G$$

define una curva $\phi \circ \rho\colon (a,b) \to H$ mediante $(\phi \circ \rho)(u) = \phi(\rho(u))$ en H.

Definición. Un homomorfismo $\phi\colon G \to H$ de grupos de matrices es **suave** si para toda curva diferenciable ρ en G, $\phi \circ \rho$ es diferenciable.

Definición. Sea $\phi\colon G \to H$ un homomorfismo suave de grupos de matrices. Si $\gamma'(0)$ es un vector tangente a G, en I, definimos un vector tangente $d\phi(\gamma'(0))$ a H en I por

$$d\phi(\gamma'(0)) = (\phi \circ \gamma)'(0).$$

La aplicación resultante $d\phi\colon T_G \to H_H$ se llama la **diferencial** de ϕ.

4.10 Proposición. *La aplicación $d\phi\colon T_G \to T_H$ es una aplicación lineal.*

Demostración. Sea $\rho'(0)$ y $\sigma'(0)$ en T_G. Para $a, b \in \mathbb{R}$ definimos las curvas ρ_a, σ_b por $\rho_a(u) = \rho(au)$ y $\sigma_b(u) = \sigma(bu)$. Entonces $\rho_a'(0) = a\rho'(0)$ y $\sigma_b'(0) = b\sigma'(0)$.

Por definición

$$d\phi[(\rho_a\sigma_b)'(0)] = [\phi \circ (\rho_a\sigma_b)]'(0).$$

Como ϕ es un homomorfismo

$$\phi \circ (\rho_a\sigma_b) = (\phi \circ \rho_a)(\phi \circ \sigma_b).$$

Entonces

$$\begin{aligned}(\phi \circ (\rho_a\sigma_b))'(0) &= (\phi \circ \rho_a)'(0)(\phi \circ \sigma_b)(0) + (\phi \circ \rho_a)(0)(\phi \circ \sigma_b)'(0) \\ &= ad\phi(\rho'(0)) + bd\phi(\sigma'(0)).\end{aligned}$$

□

4.11 Proposición. *Si $\phi\colon G \to H$ y $\psi\colon H \to K$ son homomorfismos suaves, entonces $\psi \circ \phi$ es suave y*

$$d(\psi \circ \phi) = d\psi \circ d\phi.$$

Demostración. La primera parte es obvia. Para la segunda, sea $\gamma'(0)$ un vector tangente de G. Entonces,

$$d(\psi \circ \phi)(\gamma'(0)) = (\psi \circ \phi \circ \gamma)'(0) = d\psi(\phi \circ \gamma)'(0) = d\psi \circ d\phi(\gamma'(0)).$$

□

4.12 Corolario. *Si $\phi\colon G \to H$ es un isomorfismo suave, entonces $d\phi\colon T_G \to T_H$ es un isomorfismo lineal y $\dim G = \dim H$.*

Demostración. Tenemos que $\phi^{-1} \circ \phi = I$, entonces $d\phi^{-1} \circ d\phi\colon T_G \to T_G$, es la identidad. Por lo que $d\phi$ es inyectiva y $d\phi^{-1}$ es suprayectiva. De donde $\phi \circ \phi^{-1}$ es la identidad, por lo que $d\phi \circ d\phi^{-1}\colon T_H \to T_H$ es la identidad. Entonces $d\phi^{-1}$ es inyectiva y $d\phi$ es suprayectiva. □

Capítulo 5

Exponencial y Álgebras de Lie

5.1. Exponencial de una matriz

Dado un grupo de matrices G hemos definido un espacio vectorial T_G, el espacio tangente a G en I. En este capítulo definiremos aplicaciones que mandan a T en G y viceversa. Trabajaremos únicamente con matrices reales, los casos de $\mathbb{C}$ y $\mathbb{H}$ son análogos. Necesitamos dichas aplicaciones para calcular las dimensiones de algunos grupos de matrices.

Definición. Sea A una matriz real $n \times n$. Definimos

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

donde A es la matriz producto AA, etc. Decimos que esta serie converge si cada una de las n^2 series de números reales

$$(I)_{ij} + (A)_{ij} + (\frac{A^2}{2!})_{ij} + (\frac{A^3}{3!})_{ij} + \dots$$

convergen.

5.1 Proposición. *Para toda matriz A real $n \times n$, la serie*

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

converge.

Demostración. Sea m el mayor $|a_{ij}|$ en A. Entonces
El mayor elemento en el primer término es 1.
El mayor elemento en el segundo término es m
El mayor elemento en el tercer término es $\leq \frac{nm^2}{2!}$.

El mayor elemento en el cuarto término es $\leq \frac{n^2m^3}{3!}$, etc.

Por lo tanto cualquiera de las series ij está dominada por 1, m, $\frac{n}{2!}^2$, $\frac{n^2m^3}{3!}$,..., $\frac{n^{k-2}m^{k-1}}{(k-1)!}$,...

Aplicando la prueba de la razón a esta serie tenemos que

$$\frac{n^{k-1}m^k}{k!}\frac{(k-1)!}{n^{k-2}m^{k-1}} = \frac{nm}{k}.$$

Como m y n son fijos, la razón tiende a 0 cuando k tiende a infinito, por lo tanto la serie converge absolutamente. □

Esta función exponencial se comporta de manera similar a la función exponencial de números reales e^x, con $x \in \mathbb{R}$. Si 0 denota a la matriz cero, entonces tenemos que

$$e^0 = I.$$

5.2 Proposición. *Si las matrices A y B conmutan, entonces tenemos*

$$e^{A+B} = e^A e^B.$$

5.3 Corolario. *Para toda matriz A de tamaño $n \times n$, e^A es no singular.*

Demostración. Las matrices A y $-A$ conmutan, por lo tanto $I = e^0 = e^{A-A} = e^A e^{-A}$ y entonces $1 = (\det e^A)(\det e^{-A})$, por lo que $\det e^A \neq 0$. □

Del corolario anterior, tenemos que la aplicación $\exp\colon M_n(\mathbb{R}) \to M_n(\mathbb{R})$, en realidad manda a $M_n(\mathbb{R})$ en $GL_n(\mathbb{R})$.

5.4 Proposición. *Si A es una matriz real antisimétrica, entonces e^A es ortogonal.*

Demostración. Tenemos que $I = e^0 = e^{A+A^t} = e^A e^{A^t} = (e^A)(e^A)^t$, probando con esto que la matriz e^A es ortogonal. □

Por lo tanto, si $so(n) \subset M_n(\mathbb{R})$ es el subespacio de matrices antisimétricas, entonces tenemos que

$$\exp\colon so(n) \to O(n).$$

Es importante notar dos cosas que la proposición anterior NO dice:

(I) No dice que toda matriz ortogonal es de la forma e^A para alguna A antisimétrica, es decir, no dice que $\exp\colon so(n) \to O(n)$ es suprayectiva.

(II) No dice que si e^A es ortogonal, entonces A es antisimétrica.

Veamos algunos casos para $n = 2$.

Una matriz 2×2 real antisimétrica general es de la forma

$$\alpha = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \qquad x \in \mathbb{R}.$$

Para calcular e^α calculemos las potencias de α.

$$\alpha^2 = \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix},\ \alpha^3 = \begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix},\ \alpha^4 = \begin{pmatrix} x^4 & 0 \\ 0 & x^4 \end{pmatrix}, \dots$$

Entonces

$$e^\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix} + \cdots$$

De la posición $1,1$ tenemos que

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \cos x, \text{ etc.}$$

Por lo que obtenemos

$$e^\alpha = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}$$

que es una rotación en el plano de x radianes. Por lo que para toda matriz real 2×2 antisimétrica α tenemos que

$$\det e^\alpha = 1, \text{ es decir } \ e^\alpha \in SO(2).$$

Entonces, por ejemplo, la reflexión $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in O(2)$ no puede ser obtenida de esta manera.

Nótese también que $e^\alpha = I$ no implica que α es la matriz cero, ya que si $\alpha = \left(\begin{smallmatrix} 0 & 2\pi \\ -2\pi & 0 \end{smallmatrix}\right)$ entonces $e^\alpha = I$.

5.2. Subgrupos a un parámetro

Definición. Un **subgrupo a un parámetro** γ en un grupo de matrices G es un homomorfismo suave del grupo aditivo $\mathbb{R}$ a G,

$$\gamma\colon \mathbb{R} \to G.$$

Nótese que es suficiente con conocer a γ para una vecindad U de 0 en $\mathbb{R}$. Para $x \in \mathbb{R}$, tenemos que para alguna $n \in \mathbb{N}$, $\frac{1}{n}x \in U$ y $\gamma(x) = n(\gamma(\frac{1}{n}x))$.

5.5 Ejemplo. Sean $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ y $A \in M_n(\mathbb{F})$. Entonces

$$\gamma(u) = e^{uA} = I + uA + u^2\frac{A^2}{2!} + \cdots$$

es un subgrupo a un parámetro de $GL_n(\mathbb{F})$ y $\gamma'(0) = A$.

5.6 Proposición. *Sea γ un subgrupo a un parámetro de $GL_n(\mathbb{F})$. Entonces existe $A \in M_n(\mathbb{F})$ tal que*

$$\gamma(u) = e^{uA}.$$

Por lo tanto, todo vector tangente a $GL_n(\mathbb{F})$ es la derivada en 0 de algún subgrupo a un parámetro. Veremos ahora que esto también es cierto para los grupos $\mathcal{O}(n,\mathbb{F})$.

5.7 Proposición. *Sea A un vector tangente a $\mathcal{O}(n,\mathbb{F})$. Entonces existe un único subgrupo a un parámetro γ en $\mathcal{O}(n,\mathbb{F})$ tal que*

$$A = \gamma'(0).$$

Demostración. Por definición, tenemos que $A = \rho'(0)$ donde ρ es una curva suave en $\mathcal{O}(n,\mathbb{F})$. Entonces

$$\rho(u)\overline{\rho(u)}^t = I,$$

por lo tanto

$$\rho'(0) + \overline{\rho'(0)}^t = 0,$$

es decir, $A + \overline{A}^t = 0$.

Por otro lado, $\gamma(u) = e^{uA}$ es un subgrupo a un parámetro de $GL_n(\mathbb{F})$, pero pertenece a $\mathcal{O}(n,\mathbb{F})$ ya que

$$\gamma(u)\overline{\gamma(u)}^t = e^{uA}e^{u\overline{A}^t} = e^{u(A+\overline{A}^t)} = I.$$

□

Por lo tanto tenemos para $GL_n(\mathbb{F})$ y $\mathcal{O}(n,\mathbb{F})$ una correspondencia uno a uno entre vectores tangentes y subgrupos a un parámetro.

Tomando $\mathbb{F} = \mathbb{R}$ tenemos que el espacio tangente a $O(n)$ es $so(n)$, el espacio vectorial de todas las matrices antisimétricas $n \times n$. Por lo tanto $\dim O(n) = \dim so(n) = \frac{n(n-1)}{2}$.

Tomando $\mathbb{F} = \mathbb{C}$ tenemos que el espacio tangente a $U(n)$ es $su(n)$, el espacio vectorial de todas las matrices complejas antihermitianas $n \times n$. Por lo tanto $\dim U(n) = \dim su(n) = n^2$

Tomando $\mathbb{F} = \mathbb{H}$ tenemos que $\dim Sp(n) = n(2n+1)$.

5.3. Álgebras de Lie

Es fácil ver que $so(n)$, $su(n)$ y $sp(n)$ no son cerradas bajo multiplicación de matrices. Por ejemplo, si

$$\alpha = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \quad \text{entonces} \quad \alpha^2 = \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix},$$

la cual no es antisimétrica.

5.8 Proposición. *Para $\mathbb{F} \in \{\mathbb{R},\mathbb{C},\mathbb{H}\}$ y $A,B \in M_n(\mathbb{F})$ definimos*

$$[A,B] = AB - BA.$$

Entonces $so(n)$, $su(n)$ y $sp(n)$ son cerrados bajo $[\ ,\ ]$.

Demostración. Tenemos que probar que

$$(AB - BA) + \overline{(AB - BA)}^t = 0.$$

El miembro izquierdo es

$$\begin{aligned} AB - BA + \overline{B}^t\overline{A}^t - \overline{A}^t\overline{B}^t &= AB + (A\overline{B}^t - A\overline{B}^t) - BA \\ &\quad + (-B\overline{A}^t + B\overline{A}^t) + \overline{B}^t\overline{A}^t - \overline{A}^t\overline{B}^t \\ &= A(B + \overline{B}^t) - (A + \overline{A}^t)\overline{B}^t - B(A + \overline{A}^t) + (B + \overline{B}^t)\overline{A}^t \\ &= 0. \end{aligned}$$

□

Por lo tanto, $so(n)$, $su(n)$ y $sp(n)$ son álgebra sobre $\mathbb{R}$ con el producto dado por este corchete. Este producto satisface las siguientes propiedades:

(I) $[A, B] = -[B, A]$,

(II) $[A, B + C] = [A, B] + [A, C]$,

(III) $[A + B, C] = [A, C] + [B, C]$,

(IV) Para $r \in \mathbb{R}$, $r[A, B] = [rA, B] = [A, rB]$,

(V) $[A, [B, C]] + [B, [A, C]] + [C, [A, B]] = 0$.

La última propiedad es conocida como la **identidad de Jacobi**.

Definición. Un espacio vectorial real con un producto que satisfaga las propiedades anteriores es llamada un **álgebra de Lie**.

Consideremos las álgebras de Lie de dimensiones bajas.
Para dimensión 1, el espacio vectorial es $\mathbb{R}$ y si $x, y \in \mathbb{R}$ tenemos que

$$[x, y] = x[1, y] = xy[1, 1] = 0,$$

por la primera propiedad. Por lo tanto tenemos el producto trivial.
Consideremos a $\mathbb{R}^2$ con base $\{e_1, e_2\}$. De las propiedades debemos tener:

$$[e_1, e_1] = 0, \quad [e_2, e_2] = 0 \quad \text{y} \quad [e_1, e_2] = -[e_2, e_1].$$

Sea $[e_1, e_2] = ae_1 + be_2$. Entonces, por ejemplo,

$$\begin{aligned} [e_1, [e_1, e_2]] &= [e_1, (ae_1 + be_2)] \\ &= a[e_1, e_1] + b[e_1, e_2] \\ &= b(ae_1 + be_2). \end{aligned}$$

Por la identidad de Jacobi tenemos

$$[e_1, [e_1, e_2]] + [e_1, [e_2, e_1]] + [e_2, [e_1, e_1]] = 0,$$

por lo que

$$b(ae_1 + be_2) + [e_1, (-ae_1 - be_2)] = 0$$

lo cual es cierto para toda a y b. Si tomamos $a = b = 0$ obtenemos el álgebra de Lie trivial. Para cualesquiera otros valores de a y b se obtienen álgebras de Lie no triviales.

Ejemplos de álgebras de Lie de dimensión 3 son los siguientes:

$$so(3) = \left\{ \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} | a, b, c \in \mathbb{R} \right\}.$$

Por otro lado, si tomamos a $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$ como base de $\mathbb{R}^3$ y definimos

$$[\mathsf{i}, \mathsf{j}] = \mathsf{k} \qquad [\mathsf{j}, \mathsf{k}] = \mathsf{i} \qquad [\mathsf{k}, \mathsf{i}] = \mathsf{j},$$

obtenemos un álgebra de Lie de dimensión 3.

Bibliografía

[1] Theodor Bröcker and Tammo tom Dieck. *Representations of Compact Lie Groups.* Graduate Texts in Mathematics 98. Springer-Verlag, 1985.

[2] Morton L. Curtis. *Matrix Groups.* Universitext. Springer-Verlag, 1984.

[3] J. Dieudonné. *Panorama de las matemáticas puras: la elección bourbakista.* Editorial Reverté, Encarnación 86, 08024 Barcelona, 1987.

[4] Alcebíades Rigas. *Grupos de Lie via Exemplos.* IMPA, 1993.

[5] Michael Spivak. *A comprehensive introduction to differential geometry 1*, volume I. Publish or Perish, Houston, Texas (U.S.A.), second edition, 1979.

[6] Frank W. Warner. *Foundations of Differentiable Manifolds and Lie Groups.* Graduate Texts in Mathematics 94. Springer-Verlag, 1983.