Apéndice del Curso Relatividad Especial.

Manuel Gutiérrez.
Departamento de Álgebra, Geometría y Topología.
Universidad de Málaga, 29071-Málaga.

Marzo 2009.

Índice general

Capítulo 1

El espacio vectorial $\mathbb{R}^n$.

Iniciamos este repaso al álgebra lineal con la estructura de espacio vectorial que es posiblemente la estructura más versátil y utilizada en la ciencia. Es usual tener un ejemplo de referencia a partir del cual se van dando definiciones más abstractas y propiedades. En este caso nuestro modelo será $\mathbb{R}^2$ y $\mathbb{R}^3$. Formalmente, se define

$$\mathbb{R}^n = \{(x_1, ..., x_n) \ / \ x_i \in \mathbb{R}, \ i = 1, ...n\}$$

como el conjunto formado por todas las n-uplas de números reales ordenados. Sus elementos se pueden sumar componente a componente, y también se pueden multiplicar por un número real, multiplicando cada componente por dicho número.

La primera operación se llama **ley de composición interna** y se define como

$$\begin{array}{rcl} + : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n & & \\ ((x_1, ..., x_n), (y_1, ..., y_n)) & \mapsto & (x_1, ..., x_n) + (y_1, ..., y_n) \end{array}$$

donde $(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$.

Propiedades.

1. Asociativa. Para cada $u, v, w \in \mathbb{R}^n$, $(u + v) + w = u + (v + w)$.
2. Elemento neutro. Existe un elemento $e \in \mathbb{R}^n$ tal que para cada $u \in \mathbb{R}^n$, $e + u = u + e = e$.

   En efecto, $e = (0, ..., 0)$, y se denotará por 0.
3. Inverso. Dado $u \in \mathbb{R}^n$, existe un $v \in \mathbb{R}^n$, tal que $u + v = v + u = 0$.

   En nuestro caso, $v = -u$ y se llama el opuesto de u.
4. Conmutativa. Para todo $u, v \in \mathbb{R}^n$, $u + v = v + u$.

Cualquier conjunto con las propiedades 1,2 y 3 se llama **grupo**, y si además tiene la propiedad 4 se llama **grupo abeliano o conmutativo**.

La segunda operación se llama **ley de composición externa** y se define como

$$\begin{array}{rcl} \cdot : \mathbb{R} \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ (\lambda, (x_1, ..., x_n)) & \mapsto & \lambda(x_1, ..., x_n) \end{array}$$

donde $\lambda(x_1, ..., x_n) = (\lambda x_1, ..., \lambda x_n)$.

Propiedades.

1. Distributivas. Para cada $\lambda, \mu \in \mathbb{R}$ y cada $u, v \in \mathbb{R}^n$,

   a) $(\lambda + \mu)u = \lambda u + \mu u$

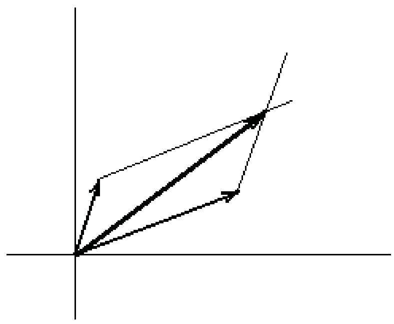
   b) $\lambda(u + v) = \lambda u + \lambda v$

2. $1u = u$

3. Seudoasociativa. Para cada $\lambda, \mu \in \mathbb{R}$ y cada $u \in \mathbb{R}^n$

   $(\lambda\mu)u = \lambda(\mu u)$

Los elementos de $\mathbb{R}^n$ se llaman **vectores** y los de $\mathbb{R}$ **escalares**. La tripleta $(\mathbb{R}^n, +, \cdot)$ formada por $\mathbb{R}^n$ con la suma de vectores y el producto por escalares se llama **espacio vectorial**. El nombre de los elementos de $\mathbb{R}^n$ proviene de su representación gráfica. Por ejemplo, $\mathbb{R}^2$ se representa como un plano con dos ejes perpendiculares que se cruzan en el 0 que se llamará **origen**, y si $u \in \mathbb{R}^2$, $u = (x_1, x_2)$, se representa como el vector que tiene origen en 0 y final en el punto de $\mathbb{R}^2$ de coordenadas (x_1, x_2). La suma de vectores está dada por la regla del paralelogramo.



Definición 1 *Una **combinación lineal** de vectores es cualquier suma finita del tipo*

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k.$$

Si un vector u se puede expresar como $u = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k$ se dice que u es combinación lineal de los vectores $v_1, v_2, ..., v_k$. Si todos los vectores de $\mathbb{R}^n$ pueden expresarse como combinación lineal de ellos, se dice que $(v_1, v_2, ..., v_k)$ es un **sistema generador**. De entre los sistemas generadores interesan los que estén formados por el mínimo número de elementos, para ello se introduce la siguiente noción.

Definición 2 *El sistema* $(v_1, v_2, ..., v_k)$ *se dice que es* ***linealmente independiente*** *cuando la única combinación lineal que es cero es la que tiene todos sus coeficientes cero, es decir*

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k = 0 \Longrightarrow \lambda_1 = \lambda_2 = ... = \lambda_k = 0.$$

Un sistema de vectores que es generador y linealmente independiente se llama **base**. Por ejemplo, el sistema $(e_1, e_2, ..., e_n)$, formado por los vectores

$$e_i = (0, ..., \overset{i)}{1}, ..., 0)$$

es una base. En efecto, si $u = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, entonces

$$\begin{aligned} u &= (x_1, x_2, ..., x_n) = (x_1, 0, ..., 0) + (0, x_2, ..., 0) + ... + (0, 0, ..., x_n) \\ &= x_1(1, 0, ..., 0) + x_2(0, 1, ..., 0) + ... + x_n(0, ..., 1) = x_1 e_1 + x_2 e_2 + ... x_n e_n \end{aligned}$$

lo que prueba que genera todo $\mathbb{R}^n$. Para ver que son linealmente independientes, si $\lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n = 0$, entonces

$$\begin{aligned} (0, 0, ..., 0) &= \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n \\ &= \lambda_1(1, 0, ..., 0) + \lambda_2(0, 1, ..., 0) + ... + \lambda_n(0, ..., 1) \\ &= (\lambda_1, \lambda_2, ..., \lambda_n) \end{aligned}$$

luego $\lambda_1 = \lambda_2 = ... = \lambda_n = 0$. La base $(e_1, e_2, ..., e_n)$ se llama **base canónica**. Se prueba que todas las bases tienen el mismo número de vectores, y ese número común se llama **dimensión** del espacio $\mathbb{R}^n$. Por tanto la dimensión de $\mathbb{R}^n$ es n.

Una base siempre se considera como un sistema ordenado de vectores, de modo que cambiando dos vectores entre sí da lugar a otra base distinta.

Si $B = (v_1, ..., v_n)$ es una base de $\mathbb{R}^n$, y u es un vector cualquiera, entonces $u = x_1 v_1 + ... + x_n v_n$. A la n-upla $(x_1, ..., x_n)$ se le llama **componentes de** u **en la base** B.

Ejercicio 16 *Probar que* $(1, 0, -1)$, $(0, 2, 1)$ *y* $(1, 1, 2)$ *forman una base de* $\mathbb{R}^3$.

Ejercicio 17 *Probar que la expresión de un vector en una base es única.*

Definición 3 *Una* ***aplicación lineal*** *de* $\mathbb{R}^m$ *a* $\mathbb{R}^n$ *es una aplicación* $f : \mathbb{R}^m \to \mathbb{R}^n$ *que respeta las combinaciones lineales de vectores, es decir,*

$$f(\lambda_1 u_1 + ... + \lambda_k u_k) = \lambda_1 f(u_1) + .. + \lambda_k f(u_k).$$

Un ***isomorfismo*** *es una aplicación lineal que además es una biyección.*

Ejercicio 18 *Probar que un isomorfismo lleva una base en otra base. Por tanto los isomorfismos sólo se pueden definir en espacios vectoriales de igual dimensión.*

Ejercicio 19 (*) *Probar que toda base se puede interpretar como un isomorfismo que lleva dicha base a la base canónica.*

Ejercicio 20 *Probar que una aplicación lineal queda determinada conociendo las imágenes de los vectores de una base.*

Ejercicio 21 *Probar que si $f : \mathbb{R}^n \to \mathbb{R}^n$ es lineal y además es inyectiva o sobreyectiva, entonces es biyectiva.*

1.1. Subespacios vectoriales.

Definición 4 *Un subconjunto no vacío S del espacio vectorial $\mathbb{R}^n$ se dice que es un **subespacio** cuando contiene todas las combinaciones lineales que se pueden formar con miembros de S.*

Consecuencias inmediatas.

1. El conjunto $\{0\}$ es un subespacio de $\mathbb{R}^n$.
2. Todo subespacio de $\mathbb{R}^n$ contiene al elemento 0.
3. Si $B = (v_1, ..., v_k)$ es un conjunto no vacío de vectores de $\mathbb{R}^n$, entonces la familia $L(v_1, ..., v_k)$ formada por todas las combinaciones lineales de elementos de B es un subespacio vectorial generado por B.
4. Todo subespacio admite una base, esto es, un sistema generador formado por vectores linealmente independientes.
5. El número de vectores de una base de un subespacio es fijo, y se llama **dimensión del subespacio**. Por convenio se establece que la dimensión del subespacio $\{0\}$ es cero.
6. Si S_1, S_2 son subespacios de $\mathbb{R}^n$ con $S_1 \subset S_2$ y $\dim S_1 = \dim S_2$, entonces $S_1 = S_2$.

Dados dos subespacios, S_1, S_2 con bases B_1 y B_2 respectivamente, la suma de S_1 y S_2 es el subespacio generado por $B_1 \cup B_2$

$$S_1 + S_2 = L(B_1 \cup B_2).$$

Ejercicio 22 *Probar que $S_1 + S_2 = \{u + v \ / \ u \in S_1, \ v \in S_2\}$.*

Ejercicio 23 *Probar que $S_1 + S_2$ es el menor subespacio que contiene a S_1 y a S_2.*

Ejercicio 24 *Probar que $S_1 \cap S_2$ es el mayor subespacio que está contenido a la vez en S_1 y en S_2.*

Las dimensiones de estos subespacios están relacionadas por la **fórmula de Grassmann**

$$\dim(S_1 + S_2) = \dim S_1 + \dim S_2 - \dim(S_1 \cap S_2).$$

Si se tiene $S_1 + S_2 = \mathbb{R}^n$, y $S_1 \cap S_2 = \{0\}$, entonces se dice que $\mathbb{R}^n$ es suma directa de S_1 y S_2, y se denota $\mathbb{R}^n = S_1 \oplus S_2$.

Capítulo 2

Matrices.

Una **matriz** $n \times m$ es una tabla formada por n filas y m columnas de números reales ordenados, del tipo

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix}$$

La matriz anterior se representa a veces abreviadamente por (a_{ij}). Si $B = (b_{ij})$ es otra matriz $m \times r$, es decir, el número de columnas de A coincide con el número de filas de B, entonces podemos multiplicarlas en el orden AB de la siguiente manera

$$AB = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & ... & b_{1r} \\ b_{21} & b_{22} & ... & b_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{mn2} & ... & b_{mr} \end{pmatrix} = (c_{ij})$$

siendo $c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$, es decir, el elemento ij-ésimo del producto se obtiene multiplicando la i-ésima fila de A con la j-ésima columna de B componente a componente y sumando.

Nótese que si multiplicamos A por la derecha con el vector $e_i = (0, ..., \overset{i)}{1}, ..., 0)$ de la base canónica de $\mathbb{R}^m$, visto como una columna, se tiene

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

Manuel Gutiérrez. Universidad de Málaga.

es decir, lo que hace es seleccionar la i-ésima columna de A. Si la multiplicación se hace por la izquierda (considerando ahora e_i vector de $\mathbb{R}^n$, lo que hace es seleccionar la i-ésima fila de A

$$\begin{pmatrix} 0 & ... & 1 & ... & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ a_{21} & a_{22} & ... & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nm} \end{pmatrix} = \begin{pmatrix} a_{i1} & a_{i2} & ... & a_{im} \end{pmatrix}$$

Finalmente, si multiplicamos por la izquierda con el i-ésimo vector $e_i \in \mathbb{R}^n$, y por la derecha con el j-ésimo vector $e_j \in \mathbb{R}^m$, lo que se obtiene es el elemento a_{ij} de la matriz A.

Las matrices constituyen una herramienta de primer orden en álgebra lineal. Por ejemplo, fijada una base de $\mathbb{R}^n$ representan biunívocamente a las aplicaciones lineales. En efecto, sea $B = (v_1, v_2, ..., v_n)$ una base. Cada vector $u \in \mathbb{R}^n$ lo representamos como una n-upla formada por sus componentes en la base B, $u = (x_1, x_2, ..., x_n)$. Esto significa que

$$u = x_1 v_1 + x_2 v_2 + ... + x_n v_n.$$

Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal. Supongamos que las imágenes de los vectores de la base B, expresadas en la propia base B, están dadas por

$$f(v_j) = \sum_{i=1}^{n} a_{ij} v_i.$$

Las componentes forman una matriz $n \times n$ que llamaremos **matriz de la aplicación f en la base B** y denotaremos $A = (a_{ij})$. Entonces podemos conocer $f(u)$ haciendo

$$\begin{aligned} f(u) &= \sum_{j=1}^{n} x_j f(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} a_{ij} v_i \\ &= \sum_{i,j=1}^{n} x_j a_{ij} v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right) v_i \end{aligned}$$

lo que significa que las componentes $(y_1, y_2, ..., y_n)$ de $f(u)$ en la base B estás dadas por por el producto matricial

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Como se puede ver, la matriz de f tiene por j-ésima columnas las componentes de $f(v_j)$ en la base B.

Ejercicio 25 (*) *Fijada una base B, probar que la matriz de la composición de dos aplicaciones lineales en B está dada por el producto de las matrices que representan a dichas aplicaciones.*

Si ahora consideramos que la aplicación f es un isomorfismo, entonces $B' = (f(v_1), f(v_2), ..., f(v_n))$ es también una base, y podemos considerarlo como un cambio de base. Entonces a la matriz de f en la base B se le llama **matriz del cambio de base de B' a B** porque nos permite calcular las componentes de un vector en la base B a partir de sus componentes en B'. En efecto, supongamos que conocemos la expresión de cada vector de B' en la base B

$$f(v_j) = \sum_{i=1}^{n} a_{ij} v_i.$$

Si $u \in \mathbb{R}^n$ tiene componentes $(x_1, x_2, ..., x_n)$ en B y $(x_1', x_2', ..., x_n')$ en B', entonces

$$\begin{aligned} \sum_{i=1}^{n} x_i v_i = u = \sum_{j=1}^{n} x_j' f(v_j) &= \sum_{j=1}^{n} x_j' \sum_{i=1}^{n} a_{ij} v_i \\ &= \sum_{i,j=1}^{n} x_j' a_{ij} v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j' \right) v_i \end{aligned}$$

Por tanto las relación entre las componentes está dado por el producto matricial

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$$

Esto a veces se expresará de forma más compacta como $x = Ax'$.

En la práctica se suelen dar las dos bases B y B' sin mencionar el isomorfismo que las relaciona, tan sólo se dice que P es la matriz del cambio de base de B a B'. En las cuentas de antes es claro que $P = A^{-1}$ y que $x' = Px$.

Definición 5 *Sea $A = (a_{ij})$ una matriz $k \times k$. Se define el **adjunto de elemento** a_{ij} como la matriz A_{ij} de orden $(k-1) \times (k-1)$ formada por la matriz A suprimiendo la fila y la columna que contienen al elemento a_{ij}. El **determinante de** A es el número real $|A|$ definido inductivamente como :*

1. *Si $k = 1$, $|A| = a_{11}$ el único elemento que tiene.*
2. *Si $k > 1$, entonces*

$$|A| = a_{11} |A_{11}| - a_{12} |A_{12}| + ... + (-1)^{1+k} |A_{1k}|.$$

Propiedades inmediatas.

1. $|A| = |A^t|$ donde A^t es la matriz transpuesta de A, es decir, $A^t = (a_{ji})$.
2. Si se cambian dos filas consecutivas de A, el determinante cambia de signo.
3. Si se cambian dos columnas consecutivas de A, el determinante cambia de signo.
4. $|A| \neq 0$ si y sólo si los vectores fila son linealmente independientes.
5. $|A| \neq 0$ si y sólo si los vectores columna son linealmente independientes.
6. Si a un vector fila de A se le suma una combinación lineal del resto de los vectores fila y se dejan las demás filas igual, se obtiene otra matriz con el mismo determinante que A.
7. Si a un vector columna de A se le suma una combinación lineal del resto de los vectores columna y se dejan las demás columnas igual, se obtiene otra matriz con el mismo determinante que A.
8. Si B es otra matriz de igual orden que A, entonces $|AB| = |A|\,|B|$.

Una matriz con determinante no nulo se llama **matriz regular**.

Capítulo 3

El Grupo General Lineal.

Una aplicación $f : U \to V$ entre dos espacios vectoriales es un **isomorfismo** cuando es lineal y biyectiva. Si además $U = V$, entonces se llama **automorfismo**. Es fácil ver que el conjunto de todos los automorfismos de un espacio vectorial V tiene una estructura de grupo para la composición. Se llama **grupo de los automorfismos de** V, $Aut(V)$.

Ejercicio 26 *Probar que $Aut(\mathbb{R}^n)$ tiene estructura de grupo.*

Siguiendo el programa Erlangen de F. Klein, este grupo determina una geometría que es el objeto de estudio del **Álgebra lineal**.

Proposición 6 *Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ un automorfismo, entonces se tienen las siguientes propiedades:*

1. $f(0) = 0$.
2. *Si $(v_1, ..., v_n)$ es una base, entonces $(f(v_1), ..., f(v_n))$ es otra base.*
3. *Si V es un subespacio, entonces $f(V) = \{f(v) \ / \ v \in V\}$ es un subespacio de la misma dimensión.*
4. *Si U, V son subespacios, entonces $f(U+V) = f(U)+f(V)$ y $f(U \cap V) = f(U) \cap f(V)$.*

Puesto que, fijada una base, las aplicaciones lineales están representadas por matrices, podemos estudiar el grupo $Aut(\mathbb{R}^n)$ a partir de ellas.

Ejercicio 27 *Usando las propiedades de los determinantes probar que, fijada una base, existe una correspondencia biyectiva entre los automorfismos de $\mathbb{R}^n$ y las matrices regulares.*

Así pues el conjunto de las matrices regulares de orden n forman un grupo llamado **grupo general lineal** y se denota $GL(n, \mathbb{R})$. Conviene conocer un procedimiento operativo para calcular la inversa de una matriz. Si $A = (a_{ij})$, y $A^{-1} = (b_{ij})$, entonces

$$b_{ij} = (-1)^{i+j} \frac{|A_{ji}|}{|A|}$$

Otro procedimiento, que evita tener que calcular determinantes, consiste en resolver el sistema $AA^{-1} = I$, tomando los elementos de A^{-1} como las incógnitas, y resolverlo por el método de eliminación de Gauss-Jordan. Por ejemplo, para invertir la matriz regular

$$\begin{pmatrix} 2 & -2 & 0 \\ 1 & 1 & -2 \\ -1 & 2 & 1 \end{pmatrix}$$

se plantea el sistema

$$\begin{pmatrix} 2 & -2 & 0 \\ 1 & 1 & -2 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

que desplegado sería

$$\begin{aligned} 2b_{11} - 2b_{21} \qquad &= 1 \qquad & 2b_{12} - 2b_{22} \qquad &= 0 \qquad & 2b_{13} - 2b_{23} \qquad &= 0 \\ b_{11} + b_{21} - 2b_{31} &= 0 & b_{12} + b_{22} - 2b_{32} &= 1 & b_{13} + b_{23} - 2b_{33} &= 0 \\ -b_{11} + 2b_{21} + b_{31} &= 0 & -b_{12} + 2b_{22} + b_{32} &= 0 & -b_{13} + 2b_{23} + b_{33} &= 1 \end{aligned}$$

ahora hay que despejar los coeficientes b_{ij} y para ello podemos trabajar con cada bloque de ecuaciones puesto que comparten los coeficientes. Las operaciones que podemos hacer son

1. Multiplicar una ecuación por un número no nulo.
2. Sumar a una ecuación, miembro a miembro, una combinación lineal del resto de ecuaciones.

Necesitamos repetir estas operaciones hasta que la matriz de los coeficientes sea la identidad. Puesto que son operaciones que se basan en los coeficientes, podemos hacerlo prescindiendo de las incógnitas b_{ij}, y además podemos hacer los tres bloques simultáneamente.

Debemos usar repetidamente los pasos 1 y 2 sobre las filas de la matriz

$$\begin{pmatrix} 2 & -2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{pmatrix}$$

para convertirla en otra que tenga en el primer bloque la identidad. Entonces la matriz que queda en el segundo bloque es precisamente la inversa que buscamos. Vamos a hacerlo paso a paso.

$$\begin{pmatrix} 2 & -2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ -1 & 2 & 1 & 0 & 0 & 1 \end{pmatrix} \equiv$$
$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & -3 & -1 & 1 & -1 \\ 0 & 2 & 2 & 1 & 0 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & -3 & -1 & 1 & -1 \\ 0 & 0 & 8 & 3 & -2 & 4 \end{pmatrix} \equiv$$
$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & -3 & -1 & 1 & -1 \\ 0 & 0 & 1 & \frac{3}{8} & -\frac{2}{8} & \frac{4}{8} \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 0 & \frac{5}{8} & \frac{2}{8} & \frac{4}{8} \\ 0 & 1 & 0 & \frac{1}{8} & \frac{2}{8} & \frac{4}{8} \\ 0 & 0 & 1 & \frac{3}{8} & -\frac{2}{8} & \frac{4}{8} \end{pmatrix}$$

Supongamos que A es la matriz de la aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^n$ en la base B. Podemos saber cuál es la matriz A' de f en otra base B' sabiendo la matriz del cambio de base P de B a B'. En efecto, todo vector $u \in \mathbb{R}^n$ de componentes x en B y x' en B', tiene sus componentes relacionadas por $x' = Px$. En particular la imagen de u por f, tiene componentes Ax en la base B y $A'x' = A'Px$ en B', por tanto $A'Px = PAx$ para todo x, de donde se obtiene

$$A' = PAP^{-1}.$$

Obsérvese en particular que $|A'| = \left|PAP^{-1}\right| = |P|\,|A|\,\left|P^{-1}\right| = |A|$, es decir, el determinante de la matriz que representa a f es invariante frente a un cambio de bases.

Capítulo 4

Espacio euclídeo.

Definición 7 *Un* **producto escalar** *en* $\mathbb{R}^n$ *es una aplicación* $\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *que verifica:*

1. *Es bilineal, es decir, es lineal en cada variable.*

2. *Es simétrica:* $\langle u, v\rangle = \langle v, u\rangle$ *para todo* $u, v \in \mathbb{R}^n$*.*

3. *Es definida positiva:* $\langle u, u\rangle \geq 0$ *para todo* $u \in \mathbb{R}^n$*. La igualdad se alcanza si y sólo si* $u = 0$*.*

Fijada una base $B = (u_1, ..., u_n)$, el producto escalar determina una matriz $A = (a_{ij})$ definiendo $a_{ij} = \langle u_i, u_j\rangle$. Se le llama **matriz del producto escalar en la base** B. Esta matriz determina completamente al producto escalar. En efecto, si $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n)$ son las componentes de u y v respectivamente en la base B, entonces

$$\begin{aligned}\langle u, v\rangle &= \left\langle \sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j \right\rangle = \sum_{i,j=1}^{n} x_i y_j \langle u_i, u_j\rangle \\ &= \sum_{i,j=1}^{n} x_i y_j a_{ij} = x^t A y.\end{aligned}$$

Si fijamos otra base $B' = (u_1', ..., u_n')$, entonces podemos encontrar una relación entre las matrices del producto escalar en ambas bases. Sea P la matriz del cambio de base de B a B'. Entonces si x y x' son las componentes de u en B y B', e y, y' las de v, sabemos que

$$\begin{aligned}x' &= Px \\ y' &= Py\end{aligned}$$

Manuel Gutiérrez. Universidad de Málaga.

por tanto se tiene

$$\begin{aligned}\langle u,v\rangle &= x^t A y\\ \langle u,v\rangle &= x'^t A' y' = (Px)^t A' P y\\ &= x^t P^t A' P y\end{aligned}$$

y como u, v son arbitrarios, se tiene

$$A = P^t A' P.$$

Si fijamos la base canónica, la aplicación

$$\langle x,y\rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

con $x = (x_1, ..., x_n)$ $y = (y_1, ..., y_n)$ es un producto escalar y su matriz en la base canónica es I. Se llama **producto escalar euclídeo** o **métrica euclídea**. El par $(\mathbb{R}^n, \langle\ ,\ \rangle)$ se llama **espacio euclídeo**.

Dos vectores u, v se dicen **ortogonales** cuando $\langle u,v\rangle = 0$. El ortogonal de un vector u es el conjunto $u^\perp = \{v \in \mathbb{R}^n \ /\ \langle u,v\rangle = 0\}$.

Ejercicio 28 (*) *Probar que si $u \neq 0$, entonces $u^\perp$ es un subespacio tal que $\mathbb{R}^n = L(u) \oplus u^\perp$.*

Un producto escalar siempre tiene asociado una **norma** que es la función

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

definida por

$$|u| = (\langle u,u\rangle)^{\frac{1}{2}}$$

Por ejemplo, la norma asociada a la métrica euclídea es la longitud del vector

$$|u| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Un vector u se dice **unitario** cuando $|u| = 1$. Una **base ortonormal** es una base formada por vectores ortogonales dos a dos y unitarios.

La base canónica es una base ortonormal para el producto euclídeo.

Propiedades.

Si $u, v \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$, entonces

1. $|u| = 0$ si y sólo si $u = 0$.
2. $|\lambda u| = |\lambda|\,|u|$.
3. Desigualdad de Schwarz. $|\langle u,v\rangle| \leq |u|\,|v|$ y la igualdad se da si y sólo si u y v son colineales.
4. Desigualdad triangular. $|u+v| \leq |u| + |v|$ y la igualdad se da si y sólo si $u = kv$ con $k \in \mathbb{R}^+$.

Ejercicio 29 (*) *Usando el ejercicio 28, probar la desigualdad de Schwarz.*

Ejercicio 30 (*) *Usando la desigualdad de Schwarz, probar la desigualdad triangular.*

Debido a la desigualdad de Schwarz, cuando u y v no son cero se tiene

$$-1 \leq \frac{\langle u, v\rangle}{|u|\,|v|} \leq 1$$

de modo que existe un único número $\alpha \in [0, \pi]$ tal que

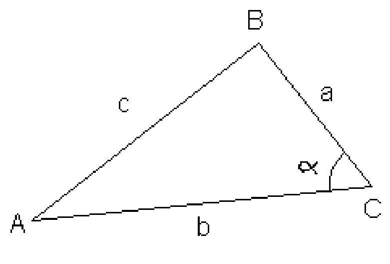
$$\cos\alpha = \frac{\langle u, v\rangle}{|u|\,|v|}. \tag{4.1}$$

Dicho número se llama **ángulo entre** u **y** v. La capacidad de un producto escalar de medir longitud de vectores y ángulos entre ellos es lo que le confiere la importancia que tiene.

Obsérvese que se elige el coseno y no el seno por que es la única elección compatible con la trigonometría. Por ejemplo, en el caso euclídeo dos vectores serán ortogonales cuando el ángulo entre ellos sea $\frac{\pi}{2}$, que corresponde a la perpendicularidad clásica.

Con esta herramienta en la mano muchos resultados de la geometría Griega clásica se hacen evidentes. Por ejemplo, si en un triángulo de lados a, b, c, el ángulo que forma a y b es α, entonces el teorema del coseno afirma que

$$c^2 = a^2 + b^2 - 2ab\cos\alpha$$



Si los vértices opuestos a a, b, c son A, B, C respectivamente, y tomamos C como el origen, se tiene

$$|A - B|^2 = \langle (A - B), (A - B)\rangle = |A|^2 + |B|^2 - 2\langle A, B\rangle$$

que es justo el teorema del coseno, usando (4.1).

En adelante supondremos fijada la métrica euclídea.

Definición 8 *Una aplicación* $f : \mathbb{R}^n \to \mathbb{R}^n$ *se llama* ***isometría*** *cuando es un isomorfismo que conserva la métrica. Esto es, para todo* $u, v \in \mathbb{R}^n$

$$\langle f(u), f(v) \rangle = \langle u, v \rangle .$$

Ejercicio 31 (*) *Probar que si* $f : \mathbb{R}^n \to \mathbb{R}^n$ *conserva la métrica entonces es un isomorfismo.*

Ejercicio 32 (*) *Probar que* $f : \mathbb{R}^n \to \mathbb{R}^n$ *es una isometría si y sólo si lleva bases ortonormales en bases ortonormales.*

El conjunto de todas las isometrías forma un grupo para la composición que se llama **grupo de las isometrías lineales de** $\mathbb{R}^n$.

Este grupo está representado por un subgrupo del grupo general lineal, llamado **grupo ortogonal** y se denota $O(n)$. Sus elementos se llaman **matrices ortogonales**. Para caracterizarlo fijamos una base $B = (u_1, u_2, ..., u_n)$, tal que la matriz de la métrica euclídea en ella sea C y la de la isometría A. Podemos escribir la condición $\langle f(u), f(v) \rangle = \langle u, v \rangle$ en la base B, como

$$x^t A^t C A y = x^t C y$$

pero si además la base B es ortonormal, entonces $C = I$, y nos queda

$$x^t A^t A y = x^t y$$

para todo x, y, de donde se deduce que

$$O(n) = \{A \in GL(n, \mathbb{R}) \ / \ A^t A = I\}$$

Si tomamos determinantes a una matriz ortogonal se observa que $|A| = \pm 1$. Las matrices ortogonales de determinante 1 forman un subgrupo de $O(n)$ llamado **grupo especial ortogonal** $SO(n)$.

El grupo ortogonal tiene dos componentes conexas formadas por las matrices de determinante 1 y por las de determinante -1.

Capítulo 5

Espacio Afín.

En esta sección estudiaremos $\mathbb{R}^n$ desde el punto de vista de su estructura afín. Aunque las ideas de espacio afín ya aparecen en la obra de Descartes (1596-1659) la noción fue introducida formalmente por H. Weyl en su libro "Space, time and matter"[6].

Definición 9 *Sea A un conjunto cuyos elementos llamaremos puntos, y V un espacio vectorial sobre $\mathbb{R}$. Se dice que A es un **espacio afín asociado a** V cuando existe una ley de composición externa*

$$\begin{aligned} + \quad &: A \times V \to A \\ &(p, v) \mapsto p + v \end{aligned}$$

que verifica:

1. *Seudoasociatividad. Para todo $p \in A$ y todo $u, v \in V$, $(p + u) + v = p + (u + v)$.*

2. *Para cada $p, q \in A$ existe un único $v \in V$ con $p = q + v$.*

A ceces se denota $v = \overrightarrow{pq}$ y se llama **vector de origen** p **y extremo** q. Se define la **dimensión de** A como la de V.

En nuestro caso es claro que tomaremos $A = \mathbb{R}^n$ y $V = \mathbb{R}^n$, y que la operación $+$ es la suma de vectores de $\mathbb{R}^n$. Esto facilita el cálculo, pues identificando un punto de $\mathbb{R}^n$ con el vector de origen 0 y extremo el punto, se tiene $\overrightarrow{pq} = q - p$.

La idea es por tanto poder trasladar los conceptos relacionados con el espacio vectorial a cualquier punto del espacio.

Definición 10 *Sea F un subespacio vectorial de $\mathbb{R}^n$. Se llama **variedad lineal** o **subespacio afín que pasa por** $p \in \mathbb{R}^n$ **con dirección** F, al subconjunto de $\mathbb{R}^n$*

$$p + F = \{p + v \ / \ v \in \mathbb{R}^n\}.$$

La **dimensión de** $p + F$ es la de F. Las variedades lineales de dimensión 0 son los puntos. Las de dimensión 1 y $n-1$ se llaman **rectas** e **hiperplanos** respectivamente.

Los problemas de incidencia entre variedades lineales no son triviales. Por ejemplo la noción de suma de variedades lineales es

$$(p+U)+(q+W) = p + (U + W + L(\overrightarrow{pq})).$$

Definición 11 *Un* ***sistema de referencia*** *es un conjunto* $R = \{p; u_1, ...u_n\}$ *donde* $p \in \mathbb{R}^n$ *es el* ***origen*** *del sistema y* $(u_1, ...u_n)$ *es una base de* $\mathbb{R}^n$*.*

Fijado un sistema de referencia podemos dar coordenadas a los puntos del espacio de la siguiente forma. Sea $q \in \mathbb{R}^n$, entonces las **coordenadas de** q **en** R son las componentes de $\overrightarrow{pq}$ en la base $(u_1, ...u_n)$ y se denota q_R.

Definición 12 *Cada vector* $a \in V = \mathbb{R}^n$ *define una* ***traslación*** *definida por*

$$\begin{aligned} \tau_a : \mathbb{A} &\to \mathbb{A} \\ p &\mapsto p + a \end{aligned}$$

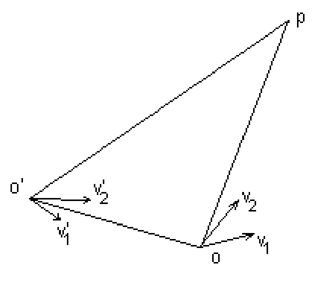
Definición 13 *Una* ***aplicación afín*** $f : \mathbb{A} \to \mathbb{A}$ *es una composición de una aplicación lineal y una traslación. Se llama* ***transformación afín*** *cuando la aplicación lineal es un isomorfismo.*

Esta definición no es la más general y tiene sentido gracias a que en nuestro caso $A = V = \mathbb{R}^n$, pero es simple y sir a nuestros propósitos.

El conjunto de todas las transformaciones afines tiene estructura de grupo llamado **grupo afín**. Supongamos que $f = \tau_a \circ g$ es una aplicación afín. Si fijamos una referencia $R = \{o; v_1, ..., v_n\}$, supongamos que g tiene matriz A en la base $B = (v_1, ..., v_n)$, y que $a_B = c$. Sea $p \in \mathbb{R}^n$ con coordenadas $x = p_R$, $x' = f(p)_R$. Entonces, la aplicación afín está dada por la expresión matricial

$$x' = f(p)_R = (\tau_a \circ g(p))_R = (g(p) + a)_R = Ax + c$$

Sea $R = \{o; v_1, ..., v_n\}$ y $R' = \{o'; v'_1, ..., v'_n\}$ dos sistemas de referencia con bases asociadas $B = (v_1, ..., v_n)$ y $B' = (v'_1, ...v'_n)$ respectivamente, $p \in \mathbb{R}^n$ un punto de coordenadas $x = (x_1, ..., x_n)$ en R y $x' = (x'_1, ..., x'_n)$ en R', y supongamos que $\overrightarrow{o'o}_{B'} = c \in \mathbb{R}^n$



Sea $C = (c_{ij})$ la matriz del cambio de base de B a B'. Las coordenadas de p en R y en R' las denotamos

$$p_R = \overrightarrow{op}_B = x$$
$$p_{R'} = \overrightarrow{o'p}_{B'} = x'$$

Puesto que $\overrightarrow{op}$ es un vector de $\mathbb{R}^n$ podemos ver sus componentes en la base B' en función de las componentes en B,

$$\overrightarrow{op}_{B'} = Cx$$

y como $\overrightarrow{o'p} = \overrightarrow{o'o} + \overrightarrow{op}$, las componentes en la base B' serán $\overrightarrow{o'p}_{B'} = \overrightarrow{o'o}_{B'} + \overrightarrow{op}_{B'}$, es decir,

$$x' = c + Cx.$$

Fijada la referencia R, se dice que la referencia R' se obtiene aplicando a R una transformación afín.

Así que un cambio de referencia afín está dado por una transformación afín.

Recíprocamente, dada una referencia afín $R = \{o; v_1, ..., v_n\}$ y una transformación afín $f : A \to A$ dada por $f(u) = \tau_a \circ g$, entonces $R' = \{f(0); g(v_1), ..., g(v_n)\}$ es una referencia afín.

En definitiva, el conjunto de todas las referencias afines se puede representar por el grupo afín.

Cuando consideramos $\mathbb{R}^n$ con su estructura afín y su estructura euclídea, se llama **espacio afín euclídeo**. Cuando la base del sistema de referencia es ortonormal, se dice que es un **sistema de referencia ortonormal**.

Definición 14 *Las transformaciones afines $f : \mathbb{R}^4 \to \mathbb{R}^4$ que en una referencia ortonormal se escriben como $x' = Ax + c$ con*

$$A = \begin{pmatrix} 1 & 0 \\ v & F \end{pmatrix}$$

*siendo $v \in \mathbb{R}^3$, $c \in \mathbb{R}^4$ y $F \in O(3, \mathbb{R})$ se llaman **transformaciones de Galileo**. Tienen una estructura de grupo llamado **grupo de Galileo**.*

Ejercicio 33 *Probar que las transformaciones de Galileo tienen efectivamente estructura de grupo.*

La importancia del grupo de Galileo radica en que fijada una referencia ortonormal en $\mathbb{R}^4$, todas las referencias que se obtienen a partir de ella por miembros del grupo describen una familia distinguida de referencias llamadas **referencias inerciales**, que son aquellas en las que son válidas las leyes de la mecánica de Newton.

Índice alfabético

Bibliografía

[1] A. Einstein, *El significado de la Relatividad.* Espasa-Calpe, S. A. 1971.

[2] Hoffmann, Banesh. *La relatividad y sus orígenes.* Editorial Labor S. A. 1985.

[3] B. O'Neill, *Semi-riemannian geometry with application to Relativity.* Academic Press, inc. 1983.

[4] W. Rindler, *Relativity. Special, General and Cosmological.* Oxford University Press, 2001.

[5] J. M. Sánchez Ron, *El origen y desarrollo de la relatividad.* Alianza Editorial, 1983.

[6] H. Weyl, *Space, time and matter*, Dover, 1952.

[7] N.M.J. Woodhouse, *Special Relativity*, Springer, 2003.