Geometría II

Índice general

Tema 1. Diagonalización de endomorfismos

1. Introducción

Este primer tema está dedicado a encontrar expresiones matriciales 'sencillas' de endomorfismos, concretamente, matrices que sean diagonales. Dada una aplicación lineal, es natural buscar bases donde la expresión matricial sea sencilla, y así fácil de manipular. Podemos pensar, por ejemplo, que la matriz tenga ceros y unos. Ya sabemos que la respuesta es afirmativa, concretamente, si $f : V \to V'$ es una aplicación lineal, existen bases B y B' de V y V' respectivamente tales que

$$M(f,B,B') = \begin{pmatrix} 0 & 0 \\ 0 & I_r \end{pmatrix},$$

donde r es el rango de f. Concretamente, estas bases son la siguientes: sea B_1 una base de $Ker(f)$ y la extendemos a una base $B = B_1 \cup B_2$ de V. Sabemos que $B_2' = f(B_2)$ es una base de la imagen. Extendemos hasta tener una base $B' = B_3' \cup B_2'$ de V'.

Sin embargo, focalizamos el problema a *endomorfismos* y al caso que las dos bases que salen en las expresiones matriciales sean las mismas, es decir, $M(f,B,B) = M(f,B)$. Entre las matrices 'sencillas', nos centramos en el hecho de que sean diagonales, es decir, $a_{ij} = 0$ si $i \neq j$.

2. Vectores y valores propios

Definición 2.1 *Un endomorfismo $f : V \to V$ se dice que es diagonalizable si existe una base B tal que $M(f,B)$ sea una matriz diagonal.*

Ya anticipamos que hay endomorfimos que no son diagonalizables. Antes de poner algunos ejemplos de endomorfismos que nos muestren cuáles son los problemas que se plantean, digamos que todos los conceptos (y resultados) que vamos a dar se trasladan de manera inmediata a matrices cuadradas. Así, si $A \in \mathscr{M}_n(\mathbb{R})$ es una matriz de orden n, tomamos cualquier espacio vectorial V de dimensión n y B cualquier base, y definimos el endomorfismo f tal que $A = M(f,B)$. Entonces diremos que A es diagonalizable si f lo es. En tal caso, si B' es una base de V donde $M(f,B')$ es diagonal, la relación entre esta matriz y A es que son semejantes, es decir, existe una matriz regular tal que $A = P^{-1}M(f,B')P$. Este argumento nos permite dar la siguiente

Definición 2.2 *Una matriz cuadrada es diagonalizable si es semejante a una matriz diagonal.*

Sea f un endomorfismo de V y supongamos que es diagonalizable, es decir, existe una base $B = \{e_1, \ldots, e_n\}$ tal que

$$M(f,B) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Por tanto,

$$f(e_i) = \lambda_i e_i,\ 1 \leq i \leq n.$$

Por tanto el trabajo a realizar consiste en lo siguiente:

1. Hallar vectores $v \in V$, con $v \neq 0$, y escalares λ tales que $f(v) = \lambda v$.
2. De entre los vectores v, ser capaces de encontrar una base del espacio vectorial.

Veamos tres ejemplos que nos van a ilustrar los problemas que aparecen cuando queremos saber si un endomorfismo es diagonalizable.

Ejemplo 2.3 *Consideramos el endomorfismo de $\mathbb{R}^2$ dado por $f(x,y) = (y,x)$. Supongamos que existe $(x,y) \in \mathbb{R}^2$ y $\lambda \in \mathbb{R}$ tales que $f(x,y) = \lambda(x,y)$. Entonces $(y,x) = (\lambda x, \lambda y)$. Esta igualdad la escribimos como el siguiente sistema homogéneo:*

$$-\lambda x + y = 0$$

$$x - \lambda y = 0.$$

Ya que estamos buscando soluciones no triviales, entonces el determinante de la matriz de los coeficientes es cero, es decir,

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

es decir, $\lambda^2 - 1 = 0$, y así, $\lambda = 1$ y $\lambda = -1$. Hallamos, para cada valor anterior de λ, vectores (x,y) que satisfagan el sistema lineal de ecuaciones. Para $\lambda = 1$, tomamos $e_1 = (1,1)$ (o uno proporcional a él), y para $\lambda = -1$, $e_2 = (1,-1)$. Ya que $B = \{e_1, e_2\}$ es una base de $\mathbb{R}^2$, entonces

$$M(f,B) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

probando que f es diagonalizable.

Ejemplo 2.4 *Consideramos $f:\mathbb{R}^2\to\mathbb{R}^2$ dado por $f(x,y)=(x+y,y)$. Si existe $(x,y)\in\mathbb{R}^2$ y $\lambda\in\mathbb{R}$ tales que $f(x,y)=\lambda(x,y)$, entonces $(x+y,y)=(\lambda x,\lambda y)$. Esta igualdad la escribimos como*

$$(1-\lambda)x+y=0$$

$$(1-\lambda)y=0.$$

Habrá soluciones no triviales si

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix}=0,$$

es decir, $(\lambda-1)^2=0$, y así, $\lambda=1$. Hallamos vectores (x,y) que satisfagan el sistema lineal de ecuaciones anterior para $\lambda=1$. La ecuación a resolver es $y=0$, luego tomamos $e_1=(1,0)$ (o uno proporcional a él). Ya que no existen más vectores con la misma propiedad concluimos que, aunque hay vectores cuya imagen mediante f es proporcional a ellos mismos, no podemos encontrar una base del espacio formado por vectores de este tipo. Esto concluye que f no es diagonalizable.

Ejemplo 2.5 *Sea ahora $f(x,y)=(-y,x)$. Del mismo modo, existe $(x,y)\in\mathbb{R}^2$ y $\lambda\in\mathbb{R}$ tales que $f(x,y)=\lambda(x,y)$ si $(-y,x)=(\lambda x,\lambda y)$. Esta igualdad la escribimos como:*

$$-\lambda x-y=0$$

$$x-\lambda y=0.$$

Habrá soluciones no triviales si

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix}=0,$$

es decir, $\lambda^2+1=0$. Ya que no hay raíces, el endomorfismo no es diagonalizable.

Formalizamos todo lo anterior del siguiente modo.

Definición 2.6 *Dado un endomorfismo f de un espacio vectorial V, decimos que $\lambda\in\mathbb{R}$ es un valor propio (o autovalor) de f si existe $v\neq 0$ tal que $f(v)=\lambda v$. Al vector v lo llamamos un vector propio (o autovector) de λ.*

Si $A\in\mathscr{M}_n(\mathbb{R})$, decimos que $\lambda\in\mathbb{R}$ es un valor propio (o autovalor) de A si existe $x\in\mathbb{R}^n, x\neq 0$ tal que $Ax=\lambda x$. Al vector x lo llamamos un vector propio (o autovector) de λ.

Si $\lambda \in \mathbb{R}$, denotamos

$$V_\lambda = \{v \in V : f(v) = \lambda v\}.$$

Observemos que en dicho conjunto siempre está el vector 0, pero no quiere decir que λ sea un valor propio: habría que tener que V_λ tiene más vectores aparte del 0. Por otro lado, si λ es un valor propio, V_λ está formado por los vectores propios de λ *y* el vector 0.

A continuación detallamos algunas propiedades de los valores y vectores propios.

Proposición 2.7 *Con la notación anterior, tenemos:*

1. $f(V_\lambda) \subset V_\lambda$.
2. V_λ *es un subespacio propio.*
3. λ *es valor propio sii* $dim(V_\lambda) \geq 1$. *En tal caso, decimos que* V_λ *es el subespacio propio del valor* λ.
4. $V_\lambda = Ker(f - \lambda 1_V)$.
5. $dim(V_\lambda) = n - r(f - \lambda 1_V)$.
6. *Si* λ *es valor propio, entonces* $det(f - \lambda 1_V) = 0$.
7. *El número de valores propios es menor o igual que la dimensión de* V.
8. $Ker(f) = V_0$ *y* $\lambda = 0$ *es un valor propio sii la nulidad es al menos* 1.

Estas propiedades se extienden inmediatamente para matrices.

Para entender parte de las dificultades de saber cuándo un endomorfismo es diagonalizable, la encontramos en la propiedad 5 anterior. La ecuación $det(f - \lambda 1_V) = 0$ es una ecuación polinómica en λ de orden n (ya que en cada elemento de la diagonal principal de cualquier expresión matricial suya aparece el número $-\lambda$). Si f es diagonalizable tiene que haber n valores propios (que pueden estar repetidos). Esto en términos de la propiedad 5 quiere decir que dicha ecuación polinómica tiene que tener *exactamente n raíces reales*, lo cual no siempre es cierto. Así, un polinomio de orden 2 puede no tener ninguna raíz (aunque si tiene una, tiene dos), como sucede en el primer ejemplo anterior. Si es de orden impar, sabemos de Cálculo que al menos tiene una, pero no sabemos si hay más.

La segunda dificultad es que, aunque tengamos n raíces (que pueden estar repetidas), cuando hallemos vectores propios de cada valor propio, no sabemos si al juntarlos obtenemos una base del espacio vectorial. En el segundo ejemplo anterior, esto ha sido así porque es evidente a simple vista; por contra, en el tercer ejemplo, sólo encontramos un vector propio (linealmente independiente). El siguiente resultado responde en gran medida a esta cuestión.

Proposición 2.8 *Sean $\lambda_1,\ldots,\lambda_k$ valores propios de un endomorfismo y B_i base de V_{λ_i}, $1\leq i\leq k$. Entonces el conjunto de vectores $B_1\cup\ldots\cup B_k$ es linealmente independiente. En particular, si f tiene n valores propios diferentes, siendo n la dimensión del espacio vectorial, entonces f es diagonalizable.*

Demostración. La demostración la hacemos en diferentes pasos. Primero probamos que vectores propios de valores propios distintos son linealmente independientes. Lo hacemos por inducción, empezando por dos vectores. Efectivamente, si $e_i\in V_{\lambda_i}$, sea $a_1e_1+a_2e_2=0$. Aplicando f, tenemos $a_1\lambda_1e_1+a_2\lambda_2e_2=0$. Multiplicando la primera matriz por λ_1 y restando la segunda, tenemos $a_2(\lambda_2-\lambda_1)e_2=0$, luego $a_2=0$ y sustituyendo en la primera ecuación, $a_1=0$.

Tomamos ahora k valores propios diferentes y correspondientes vectores propios y probamos que son linealmente independientes. Si $a_1e_1+\ldots+a_ke_k=0$, aplicando f, tenemos $a_1\lambda_1e_1+a_2\lambda_2e_2+\ldots+a_k\lambda_ke_k=0$. Multiplicando la primera por λ_1 y restando la segunda, tenemos

$$a_2(\lambda_2-\lambda_1)e_2+\ldots+a_k(\lambda_k-\lambda_1)e_k=0.$$

Ya que hay $k-1$ vectores, aplicando inducción, concluimos $a_2=\ldots=a_k=0$. Sustituyendo en la primera ecuación, tenemos $a_1e_1=0$, es decir, $a_1=0$.

Probamos ahora la proposición. Sean $B_1=\{e_{11},\ldots,e_{1m_1}\}$, ..., $B_k=\{e_{k1},\ldots,e_{km_k}\}$ y tomamos una combinación lineal de todos ellos, que la escribimos así:

$$(a_{11}e_{11}+\ldots a_{1m_1}e_{1m_1})+\cdots+\left(a_{k1}e_{k1}+\ldots+a_{km_k}e_{km_k}\right)=0,$$

o $v_1+\ldots v_k=0$, donde cada v_i se corresponde a cada paréntesis. Ya que $v_i\in V_{\lambda_i}$, sabemos que son linealmente independientes. Sin embargo el coeficiente que acompaña a cada vector v_i es 1, luego la única posibilidad es que $v_i=0$ para todo i, es decir, cada paréntesis es 0. Pero como los vectores en cada una de estas combinaciones lineales son linealmente independientes, los coeficientes son cero, es decir, $a_{ij}=0$ para todo i,j.

La última afirmación de la proposición se prueba del siguiente modo: al haber al menos n valores propios diferentes, el número de vectores de V_{λ_i} es al menos 1, pero como hay n subespacios propios y no pueden haber más de n vectores linealmente independientes, entonces es que hay exactamente n vectores propios linealmente independientes, es decir, forman una base, y así f es diagonalizable.

3. El teorema de diagonalización

Definición 3.1 *Si f es un endomorfismo de un espacio vectorial, se llama el polinomio característico de f a*

$$P_f(\lambda) = |f - \lambda I|.$$

A la ecuación $P_f(\lambda) = 0$ se llama ecuación característica. de f

Si A es una matriz cuadrada, el polinomio característico es $P_A(\lambda) = |A - \lambda I|$.

Observemos que:

1. $P_f(\lambda)$ es un polinomio de orden n.
2. λ es un valor propio sii es una raíz de $P_f(\lambda)$.
3. Si f es diagonalizable, entonces $P_f(\lambda)$ tiene n raíces (que pueden ser iguales).

Sea λ un valor propio.

1. Se llama multiplicidad algebraica de λ, la multiplicidad de λ como raíz de $P_f(\lambda)$. La denotamos por a_λ
2. Se llama multiplicidad geométrica de λ a la dimensión de V_λ. La denotamos por g_λ.

Proposición 3.2 *Si λ es un valor propio de f, entonces $g_\lambda \leq a_\lambda$.*

Demostración. Supongamos que la dimensión de V_{λ_1} es m. Sea $\{e_1, \dots, e_m\}$ una base de V_{λ_1} que la extendemos a una B de V. Entonces

$$M(f,B) = \begin{pmatrix} \lambda_1 & & & * \\ & \ddots & & * \\ & & & * \\ 0 & \dots & 0 & * \end{pmatrix}.$$

Por tanto

$$P_f(\lambda) = (\lambda_1 - \lambda)^m Q(\lambda).$$

Corolario 3.3 *Si f tiene n valores propios diferentes, entonces es diagonalizable.*

Demostración. Si λ es un valor propio, entonces $g_\lambda \geq 1$. Por otro lado, $a_\lambda \geq g_\lambda \geq 1$, pero la suma de todas las multiplicidades aritméticas es n, luego $a_\lambda = 1$, luego $a_\lambda = g_\lambda = 1$, para cualquier valor propio λ.

Teorema 3.4 *Sea f un endomorfismo de un espacio vecotorial. Son equivalentes:*

1. *f es diagonalizable.*
2. a*) $P_f(\lambda)$ tiene n raíces (reales).*
   b*) $a_\lambda = g_\lambda$ para cualquier valor propio λ.*

Demostración. (1) $\Rightarrow$ (2) Si B es una base de vectores propios, entonces $M(f,B)$ es una matriz diagonal. Para esta matriz, $P_f(\lambda) = (\lambda_1 - \lambda)^{m_1} + \ldots + (\lambda_k - \lambda)^{m_k}$ con $m_1 + \ldots + m_k = n$. Aquí m_i es la multiplicidad aritmética a_i. Para hallar g_i, hallamos el rango de $f - \lambda_i I$. Lo hacemos para λ_1. Entonces la matriz $f - \lambda_1 I$ es una matriz diagonal donde los primeros m_1 elementos de la diagonal principal son ceros, y los demás no. Por tanto su rango es $n - m_1$, luego $g_1 = n - (n - m_1) = m_1$.

(2) $\Rightarrow$ (1) Al juntar bases de cada uno de los subespacios propios, el número de vectores que tenemos coincide con n, pues es la suma de las multiplicidades geométricas, donde cada una de éstas coinciden con la correspondiente aritmética. La suma de todas las aritméticas es n pues hay n raíces. Por la proposición 2.8, son base y está formada por vectores propios, luego f diagonaliza.

Si consideramos ahora matrices cuadradas, el teorema anterior se aplica del mismo modo, obteniendo que existe una matriz diagonal D y una matriz regular P tal que $A = P^{-1}DP$. Nos preguntamos qué matriz es P. Si B y B' son bases tales que $A = M(f,B)$ y $D = M(f,B')$, entonces $P = M(1_V, B, B')$. Ya que la base B (de $\mathbb{R}^n$) puede ser cualquiera, tomamos la base usual $B = B_u$. Entonces P^{-1} no es más que poner la base de los vectores propios en columnas (si B es otra base, la matriz P^{-1} es poner en coordenadas la base B' en términos de B, pero en el cálculo explícito de las bases de los subespacios propios, se trabaja en coordenadas respecto de B, luego al final dicha matriz es poner en columnas los vectores propios.

Acabamos con una observación. Entre otras aplicaciones que tiene saber si una matriz (o un endomorfismo) es diagonalizable es que permite hallar *de forma sencilla* potencias de una matriz. Concretamente, sea A una matriz y nos preguntamos por calcular la potencia A^n, donde n es un número anterior. *Si* A es diagonalizable, esta operación se puede computar del siguiente modo. Sabemos que $A = P^{-1}DP$ para cierta matriz regular P y matriz diagonal D. Entonces

$$A^n = (P^{-1}DP) \cdots (P^{-1}DP) = P^{-1}D^nP.$$

Pero la matriz D^n es sencilla de hallar: no es más que elevar a n los elementos de la diagonal principal de D, es decir,

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_k^n \end{pmatrix}.$$

Por tanto, la potencia n-ésima de A se reduce a computar tres productos de matrices.

Tema 2. Formas bilineales y formas cuadráticas

En este tema vamos a dar los primeros pasos para el estudio de las métricas euclídeas que se hará en profundidad en el tema siguiente. El objetivo de este tema es familiarizarnos con las aplicaciones que tienen propiedades comunes con la métrica euclídea. Introduciremos el concepto de forma bilineal y forma cuadrática y realizaremos una clasificación satisfactoria de ellas. Para ello usaremos ciertas técnicas de diagonalización para culminar en el teorema de Sylvester.

1. Forma bilineal

El primer ejemplo de forma bilineal y que nos va a motivar su definición es la métrica euclídea. Recordemos que en el plano $\mathbb{R}^2$, la métrica euclídea está definida por

$$g_E(x,y) = x_1y_1 + x_2y_2, \ \ x = (x_1,x_2), y = (y_1,y_2).$$

la cual permite manejar conceptos como distancias, módulos y ángulos. Así, el módulo de un vector x es

$$|x| = \sqrt{g_E(x,x)}$$

y el ángulo θ entre dos vectores no nulos $x,y \in \mathbb{R}^2$ es aquél cuyo coseno es

$$\cos\theta = \frac{g_E(x,y)}{|x||y|}.$$

La métrica euclídea es una aplicación cuyo dominio no es $\mathbb{R}^2$, sino el producto cartesiano $\mathbb{R}^2 \times \mathbb{R}^2$, concretamente, $g_E : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ y posee diferentes propiedades. Vamos a destacar éstas agrupándolas del siguiente modo. Las primeras tres son las siguientes: si $x,y,z \in \mathbb{R}^2$ y $\lambda \in \mathbb{R}$, entonces

1. $g_E(x+y,z) = g_E(x,z) + g_E(y,z)$.
2. $g_E(x,y+z) = g_E(x,y) + g_E(x,z)$.
3. $g_E(\lambda x,y) = \lambda g_E(x,y) + g_E(x,\lambda y)$.

Una cuarta propiedad, es la simetría de g_E, es decir,

$$g_E(x,y) = g_E(y,x) \qquad \text{(simetría)}$$

Y por último, una propiedad de 'definida positiva':

$$g(x,x) \geq 0, \ \forall x \in V \ \text{y} \ g(x,x) = 0 \Leftrightarrow x = 0.$$

Vamos a ver otros dos ejemplos de aplicaciones que satisfacen similares propiedades.

1. En $\mathbb{R}^2$, definimos la métrica de Lorentz-Minkowski como

$$g_L(x,y) = x_1y_1 - x_2y_2, \;\; x = (x_1,x_2), y = (y_1,y_2).$$

Es evidente que esta aplicación satisface las primeras cuatro propiedades que tiene g_E. Respecto de la última, hay que darse cuenta de que $g((0,1),(0,1)) = -1$, es decir, es negativo. Por otro lado, $g_L((1,1),(1,1)) = 0$, pero $x = (1,1)$ no es el elemento neutro.

2. De nuevo, en $\mathbb{R}^2$ definimos

$$g(x,y) = x_1y_1, \;\; x = (x_1,x_2), y = (y_1,y_2).$$

Esta aplicación g satisface también las cuatro primeras propiedades, pero no la última. Así, se tiene $g(x,x) = x_1^2 \geq 0$, pero el vector $x = (0,1)$ satisface $g(x,x) = 0$.

Un último ejemplo nos viene si usamos matrices. En general, "cualquier polinomio de grado dos en n-variables" define aplicaciones con propiedades similares a las que tiene g_E, al menos, las tres primeras. Por ejemplo, sea de nuevo $V = \mathbb{R}^2$ y

$$g(x,y) = x_1y_2 + x_2y_2.$$

Esta aplicación se puede escribir como

$$g(x,y) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x^tAy, \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Ahora la prueba de las tres primeras propiedades es una consecuencia del producto de matrices. Así, para la primera, tenemos

$$g(x+y,z) = (x+y)^tAz = (x^t + y^y)Az = x^tAz + y^tAz = g(x,z) + g(y,z).$$

Está claro que la cuarta propiedad no es cierta, pues $g((1,0),(0,1)) \neq g((0,1),(1,0))$, ni tampoco la quinta propiedad, pues $g((1,0),(1,0)) = 0$.

Definición 1.1 *Una forma bilineal en un espacio vectorial V es una aplicación $g : V \times V \to \mathbb{R}$, con las siguientes propiedades: para todo $u,v,w \in V$, y $\lambda \in \mathbb{R}$, tenemos,*

1. $g(u+v,w) = g(u,w) + g(v,w)$.
2. $g(u,v+w) = g(u,v) + g(u,w)$.
3. $g(\lambda u,v) = \lambda g(u,v) + g(u,\lambda u)$.

Si además se satisface la propiedad $g(u,v) = g(v,u)$ para todo $u,v \in V$, se dice que g es una métrica.

Cuando tengamos $g(u,v)$, diremos "u por v". También diremos que g es lineal en cada una de sus variables. Observemos entonces que si tenemos dos sumas de vectores en cada una de las variables, resultan cuatro suma de números. Podemos decir pues que una forma bilineal no tiene nada que ver con las formas lineales de un espacio vectorial, es decir, con el espacio dual.

Proposición 1.2 *Sea V un espacio vectorial y g una forma bilineal. Entonces*

1. $g(0,v) = g(v,0) = 0.$
2. $g(-u,v) = -g(u,v) = g(u,-v).$

Demostración. Ya que $0 = 0 \cdot 0$, entonces $g(0,v) = g(0 \cdot 0, v) = 0g(0,v) = 0$ Del mismo modo, como $-u = -1 \cdot u$, entonces $g(-u,v) = g(-1 \cdot u, v) = -1g(u,v) = -g(u,v)$.

□

Pongamos algunos ejemplos de formas bilineales.

1. La aplicación bilineal cero, $g = 0$, es decir, $g(u,v) = 0$ para todo $u,v \in V$.
2. Sea g una forma bilineal en V y sea $U \subset V$ un subespacio vectorial. Entonces
$$g_{|U} : U \times U \to \mathbb{R}_{|U}, \;\; g_{|U}(u,v) = g(u,v), \; (u,v \in U),$$
es una forma bilineal en U y se llama la forma bilineal inducida en U.
3. La métrica euclídea en $\mathbb{R}^n$:
$$g_E(x,y) = \sum_{i=1}^{n} x_i y_i.$$
4. La métrica de Lorentz-Minkowski en $\mathbb{R}^n$:
$$g_L(x,y) = \sum_{i=1}^{n-1} x_i y_i - x_n y_n.$$

5. Sea $A \in \mathscr{M}_n(\mathbb{R})$ y V un espacio vectorial. Fijamos una base $B = \{e_1, \ldots, e_n\}$ en V y se define

$$g(u, v) = x^t A y,$$

donde x e y son las coordenadas de u y v respecto de B, respectivamente, es decir, $u = \sum x_i e_i$, $y = \sum y_i e_i$.

El último ejemplo lo podemos 'extender' del siguiente modo.

Definición 1.3 *Sea g una forma bilineal en un espacio vectorial V. Si $B = \{e_1, \ldots, e_n\}$ es una base de V, se llama expresión matricial de g respecto de B a la matriz*

$$M_B(g) = (g(e_i, e_j)).$$

Como consecuencia, si $u, v \in V$, y x, y son las coordenadas respecto de B, entonces

$$g(u, v) = g(\sum x_i e_i, \sum y_i e_i) = \sum_{i,j} x_i y_j g(e_i, e_j) = x^t M_B(g) y.$$

Por tanto:

1. Dos formas bilineales g y g' son iguales si y sólo si $M_B(g) = M_B(g')$.
2. Si $A \in \mathscr{M}_n(\mathbb{R})$, la forma bilineal $g(x, y) = x^t A y$ definida en $\mathbb{R}^n$ es la que tiene como matriz $M_{B_u}(g) = A$.

Esta última propiedad se extiende a cualquier espacio vectorial del siguiente modo:

Proposición 1.4 *Sea $A \in \mathscr{M}_n(\mathbb{R})$, un espacio vectorial V de dimensión n y B una base de V. Entonces existe una única forma bilineal g tal que $A = M_B(g)$. Además, g es una métrica sii A es simétrica.*

Demostración. Basta definir g como $g(u, v) = x^t A y$ donde x e y representan las coordenadas de u y v respecto B, respectivamente. $\square$

Es natural preguntarse cómo cambia la expresión matricial de una forma bilineal al cambiar de bases. Sean pues B y B' dos bases en un espacio vectorial y sea P la matriz regular que es el cambio de bases de B' a B, es decir, es aquella matriz cuyas columnas son las coordenadas de cada uno de los elementos de B' respecto de B. Escribimos los vectores en coordenadas. Si X' e Y' son dos vectores u, v en coordenadas respecto de B', entonces sus coordenadas respecto de B son $X = PX'$ e $Y = PY'$. Entonces

$$g(u,v) = g(X,Y) = X^t M_B(g) Y = X'^t P^t M_B(g) P Y',$$

luego

$$M'_B(g) = P^t M_B(g) P.$$

En general, dadas dos matrices cuadradas del mismo orden A y C, se dice que A es *congruente* a C si existe una matriz regular P tal que $A = P^t C A$. Es inmediato que 'ser congruentes' es una relación de equivalencia, luego podemos decir que 'A y C son congruentes'. Como consecuencia, *dos matrices que son expresiones matriciales de una misma forma bilineal son matrices congruentes.*

Dotamos ahora de estructura de espacio vectorial a las formas bilineales. Si V es un espacio vectorial, denotamos por $L_2(V)$ al conjunto de todas las formas bilineales definidas en V. Para $g, h \in L_2(V)$ y $\lambda \in \mathbb{R}$, definimos

$$g + h : V \times V \to \mathbb{R}, \ \ (g+h)(u,v) = g(u,v) + h(u,v), \ (u, v \in V).$$

$$\lambda g : V \times V \to \mathbb{R}, \ \ (\lambda g)(u,v) = \lambda g(u,v), \ \ (u, v \in V).$$

Es inmediato probar que $g + h$ y λg son formas bilineales V. También es inmediato que si B es una base de V, entonces

$$M_B(g+h) = M_B(g) + M_B(h), \ \ M_B(\lambda g) = \lambda M_B(g).$$

Estas dos igualdades permite definir la aplicación

$$\phi : L_2(V) \to \mathscr{M}_n(\mathbb{R}), \ \ \phi(g) = M_B(g),$$

que es lineal. También evidente que es inyectiva. Sin embargo no podemos decir que sea un isomorfismo ya que no sabemos la dimensión de $L_2(V)$.

Teorema 1.5 *El espacio vectorial $L_2(V)$ tiene dimensión $dim(V)^2$. En particular, la aplicación phi es un isomorfismo.*

Demostración. Definimos una base de $L_2(V)$ del siguiente modo. Sea $B = \{e_1, \ldots, e_n\}$ una base de V, y para cada i, j, definimos

$$h_{ij} : V \times V \to \mathbb{R}, \ \ h_{ij}(u,v) = x_i y_j,$$

donde $u = (x_1, \ldots, x_n)$ y $v = (y_1, \ldots, y_n)$. Dicha aplicación la podemos escribir como

$$h(X,Y) = X^t E_{ij} Y,$$

donde E_{ij} es la matriz cuyos elementos son todos 0 excepto en el lugar (i,j), que es un 1. Por la proposición anterior, h_{ij} es una forma bilineal y $M_B(h) = E_{ij}$. Veamos que $\{h_{ij} : 1 \leq i,j \leq n\}$ es una base de $L_2(V)$. Esto finalizaría la demostración del teorema.

Independencia lineal. Sea $\sum_{ij} \lambda_{ij} h_{ij} = 0$. Fijamos k,l y aplicamos al par (e_j, e_k). Ya que $h_{ij}(e_j, e_k) = \delta_{ij}\delta_{jk}$, es decir, siempre es cero a no ser que el par sea justamente (e_i, e_j). Tenemos

$$0 = \sum_{ij} \lambda_{ij} h_{ij}(e_k, e_l) = \sum \lambda_{ij} \delta_{ij} \delta_{jk} = \lambda_{kl}.$$

Como k y l son arbitrarios, concluimos que $\lambda_{kl} = 0$ para todo k,l.

Sistema de generadores. Usando una cuenta parecida a la del párrafo anterior, es inmediato ver que si g es una forma bilineal, entonces

$$g = \sum_{ij} g(e_i, e_j) h_{ij}$$

pues al aplicar a ambas expresiones un par del tipo (e_k, e_l) da lo mismo. $\square$

2. Espacio métrico

Definición 2.1 *Un espacio métrico es un par (V,g) donde V es un espacio vectorial y g una métrica definida en V.*

Ya hemos visto varios ejemplos de métricas. Otros dos más son los siguientes:

1. En el espacio V de las matrices cuadradas de orden n, definimos

   $$g(A,C) = \text{traza}(AC^t).$$

   Es inmediato de las propiedades que g es una métrica en V. Otra forma de probarlo es detallar la expresión de $g(A,C)$. No es difícil darse cuenta que

   $$g(A,C) = \sum_{ij} a_{ij} c_{ij}, \ \ A = (a_{ij}), C = (c_{ij}).$$

   Por tanto, $g(A,C)$ no es más que considerar A y C como vectores de $\mathbb{R}^{n^2}$ y multiplicarlos con la métrica euclídea.

2. Sea $V = \mathscr{C}([0,1],\mathbb{R})$ el espacio vectorial de las funciones continuas. Entonces

$$g(f,h) = \int_0^1 f(x)h(x)dx$$

es una métrica definida positiva, es decir, $g(f,f) \geq 0$ y es 0 sii $f = 0$.

Dos vectores u y v son perpendiculares (u ortogonales o conjugados) si $g(u,v) = 0$. Si $g(u,u) = 0$, se dice que u es autoconjugado.

Definición 2.2 *Si (V,g) es un espacio métrico, se llama radical del espacio al conjunto*

$$rad(g) = \{v \in V : g(u,v) = 0\}.$$

Si $rad(g) = \{0\}$, se dice que g es no degenerada, y en caso contrario, que es degenerada.

Por tanto, el radical es el conjunto de vectores que son perpendiculares a todos los vectores del espacio. También que el vector 0 siempre pertenece al radical. En el espacio euclídeo, el radical es trivial. Lo mismo pasa en el espacio de Lorentz-Minkowski. Sin embargo si la métrica en $\mathbb{R}^2$ es $g(x,y) = x_1y_1$, entonces $rad(g) = \{(0,x) : x \in \mathbb{R}\}$.

Definición 2.3 *Si (V,g) es un espacio métrico, se llama cono de luz al conjunto*

$$C(g) = \{v \in V : g(v,v) = 0\}.$$

Por tanto el cono de luz es el conjunto formado por los vectores autoconjugados. En $(\mathbb{R}^2, g)$ con la métrica de Minkowski, $C(g) = \{(x,x) : x \in \mathbb{R}\} \cup \{(x,-x) : x \in \mathbb{R}\}$. El cono de luz no es un subespacio vectorial.

Proposición 2.4 *En un espacio métrico (V,g), el radical es un subespacio vectorial y*

$$dim(rad(g)) = dim(V) - r(M_B(g)),$$

donde B es una base de V. Se llama nulidad de g a la dimensión de $rad(g))$ y rango de g a la dimensión de $M_B(g)$.

Observemos que el rango de la matriz $M_B(g)$ no depende de la base B elegida ya que al tomar otra, la matriz cambia por congruencia, en particular, por equivalencias, donde el rango de la matriz no cambia.

Demostración. El hecho de ser un subespacio vectorial es fácil. Para su dimensión, trabajamos en coordenadas respecto de B, el vector $X \in \mathbb{R}^n$ está en el radical sii $g(Y,X) = 0$ para todo $Y \in \mathbb{R}^n$, es decir, $Y^t M_B(g)X = 0$. Ya que es para todo vector Y de $\mathbb{R}^n$, entonces es equivalente a decir que $M_B(g)X = 0$, luego X está en el núcleo de la matriz $M_B(g)$. $\square$

De la demostración anterior, tenemos:

Corolario 2.5 *Sea $B = \{e_1, \ldots, e_n\}$ una base de una espacio métrico (V,g). Si escribimos $X \in \mathbb{R}^n$ las coordenadas de un vector de V, se tiene:*

$$rad(g) = \{X \in \mathbb{R}^n : M_B(g)X = 0\}.$$

Como consecuencia, tenemos un *método para hallar una base y las ecuaciones cartesianas del radical* de una métrica: basta hallar la expresión matricial de la métrica respecto de una base, $M_B(g)$, y considerar esta matriz como la matriz de un endomorfismo. Entonces el núcleo de dicha matriz coincide con el radical de la métrica.

Una clasificación de las métricas es la siguiente.

Definición 2.6 *Sea (V,g) un espacio métrico. Se dice que g es:*

1. *semidefinida positiva si $g(v) \geq 0$ para todo $v \in V$.*
2. *semidefinida negativa si $g(v,v) \leq 0$ para todo $v \in V$.*
3. *definida positiva si es semidefinida positiva y $g(v,v) = 0$ sii $v = 0$.*
4. *definida negativa si es semidefinida negativa y $g(v,v) = 0$ sii $v = 0$.*
5. *indefinida si no es ni semidefinida positiva ni semidefinida negativa.*

El siguiente resultado es inmediato:

Proposición 2.7 *Sea (V,g) un espacio métrico y $U \subset V$ un subespacio vectorial.*

1. *Si g es semidefinida positiva (resp. semidefinida negativa), entonces $g_{|U}$ es semidefinida positiva (resp. semidefinida negativa).*
2. *Si g es definida positiva (resp. definida negativa), entonces $g_{|U}$ es definida positiva (resp. definida negativa).*

Para mostrar algunos ejemplos, volvemos a considerar expresiones matriciales de métricas. Sea V un espacio vectorial y B una base del mismo. Sabemos que para cada matriz simétrica $A \in S_n(\mathbb{R})$, existe una única métrica g tal que $A = M_B(g)$. Tomamos una matriz A sencilla, por ejemplo, que sea diagonal:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Entonces

$$g(x,x) = \sum_{i=1}^{n} a_{ii} x_i^2.$$

Por tanto es evidente que g es:

1. semidefinida positiva sii $a_{ii} \geq 0$ para todo i.
2. semidefinida negativa sii $a_{ii} \leq 0$ para todo i.
3. definida positiva sii $a_{ii} > 0$ para todo i.
4. definida negativa sii $a_{ii} < 0$ para todo i.
5. indefinida si existen i, j tales que $a_{ii} < 0 < a_{jj}$.

3. Ortogonalidad. Subespacio ortogonal

Definición 3.1 *Sea $S \subset V$ un subconjunto de un espacio métrico (V, g). El subespacio ortogonal $S^\perp$ es el conjunto*

$$S^\perp = \{u \in V : g(u,v) = 0, \forall v \in S\}.$$

Observemos que S no tiene porqué ser un subespacio vectorial. Probaremos que $S^\perp$ es un subespacio vectorial, luego tiene sentido preguntarse por su dimensión y es natural relacionarla con la de S. Ya que S no es, en general, un subespacio vectorial, esta pregunta tiene sentido cambiando S por el subespacio vectorial generado por S.

Proposición 3.2 *Sea (V, g) un espacio métrico y $S \subset V$.*

1. $S^\perp$ es un subespacio vectorial.

2. Si $U \subset W$, entonces $W\perp \subset U^\perp$.

3. Si $U =< u_1,\ldots,u_m >$, entonces $U^\perp =< \{u_1,\ldots,u_m\} >^\perp$.

4. $V^\perp = rad(g)$ y $\{0\}^\perp = V$.

5. $U \subset (U^\perp)^\perp$.

Proposición 3.3 *Sea (V,g) un espacio métrico y $u \in V$. Si $g(u,u) \neq 0$, entonces $V =< u > \oplus < u >^\perp$. Como consecuencia, $dim(< u >^\perp) = dim(V) - 1$.*

Demostración. Veamos las dos propiedades de la suma directa.

1. Probamos que $< u > \cap < u >^\perp = \{0\}$. Si v es un vector en la intersección, existe $\lambda \in \mathbb{R}$ tal que $v = \lambda u$ y como $v \in < u >^\perp$, entonces $g(v,u) = 0$. Usando la bilinealidad de g, $\lambda g(u,u) = 0$, y como $g(u,u) \neq 0$, entonces $\lambda = 0$, probando que $v = \lambda u = 0$.

2. Probamos que $V =< u > + < u >^\perp$. Sea $v \in V$. Hay que hallar $\lambda \in \mathbb{R}$ y $w \in < u >^\perp$ tal que $v = \lambda u + w$. Si multiplicamos por u en esta igualdad, tenemos que λ tiene que ser
$$\lambda = \frac{g(v,u)}{g(u,u)}.$$
Por tanto, definimos λ como dicho número y definimos $w = v - \lambda u$. Queda por probar que, efectivamente, $w \in < u >^\perp$. Para ello es suficiente con que $g(w,u) = 0$. Pero es evidente que
$$g(w,u) = g(v,u) - \lambda g(u,u) = g(v,u) - \frac{g(v,u)}{g(u,u)} g(u,u) = 0.$$

□

Extendemos el resultado anterior del siguiente modo

Teorema 3.4 *Sea (V,g) un espacio no degenerado y $U \subset V$ un subespacio vectorial. Entonces*

1. $dim(U^\perp) = dim(V) - dim(U)$.

2. $(U^\perp)^\perp = U$.

Demostración.

1. Supongamos que $dim(U)=m$ y sea $U=<e_1,\ldots,e_m>$. Extendemos a una base de V: $B=\{e_1,\ldots,e_m,e_{m+1},\ldots,e_n>$ y trabajamos en coordenadas respecto de B. Sea $A=M_B(g)$. Entonces

$$
\begin{aligned}
U^{\perp} &= \{x\in\mathbb{R}^n : x^tAe_i=0, 1\leq i\leq m\} = \left\{x\in\mathbb{R}^n : x^tA\begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix}=0,\ldots,x^tA\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots 0\end{pmatrix}=0\right\}\\
&= \{x\in\mathbb{R}^n : a_{11}x_1+\ldots+a_{n1}x_n=0,\ldots,a_{1m}x_1+\ldots+a_{nm}x_n=0\}\\
&= \left\{x\in\mathbb{R}^n : x^t\begin{pmatrix}a_{11} & \ldots & a_{1m}\\ a_{21} & \ldots & a_{2m}\\ \vdots & \ldots & \vdots\\ a_{n1} & \ldots & a_{nm}\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}\right\}
\end{aligned}
$$

Ya que g es no degenerada, el rango de A es n. Por tanto

$$
r\begin{pmatrix}a_{11} & \ldots & a_{1m}\\ a_{21} & \ldots & a_{2m}\\ \vdots & \ldots & \vdots\\ a_{n1} & \ldots & a_{nm}\end{pmatrix}=m.
$$

Esto implica que la dimensión de $U^{\perp}$ es $n-m$.

2. Es consecuencia de la anterior y que $U\subset(U^{\perp})^{\perp}$.

$\square$

Después del teorema 3.4, uno espera obtener $V=U\oplus U^{\perp}$. Esto sería así si probamos que $U\cap U^{\perp}=\{0\}$. Pero ya hemos visto que esto no es cierto, incluso si la métrica es no degenerada. Por ejemplo, para el espacio de Lorentz-Minkowski $(\mathbb{R}^2,g_L)$, si $U=<(1,1)>$, entonces $U^{\perp}=U$, luego no puede darse $\mathbb{R}^2=U\oplus U^{\perp}$. Esto se debe a que aunque $(\mathbb{R}^2,g_L)$ es no degenerado, la métrica inducida en U es degenerada. El resultado satisfactorio nos lo da el siguiente teorema:

Teorema 3.5 *Sea un espacio métrico no degenerado (V,g) y $U\subset V$ un subespacio vectorial. Son equivalente los siguientes enunciados:*

1. $V = U \oplus U^{\perp}$.

2. El subespacio U es no degenerado.

Como consecuencia, U es no degenerado si y sólo si $U^{\perp}$ es no degenerado.

Demostración. Supongamos que $V = U \oplus U^{\perp}$. Para probar que U es no degenerado, hay que ver que la métrica inducida en U es no degenerada, o dicho de otro modo, que $R(g_{|U}) = \{0\}$. Sea pues $v \in R(g_{|U})$. Por un lado $g(v,u) = 0$ para todo $u \in U$, por estar en el radical. Por otro, $g(v,w) = 0$ para todo vector $w \in U^{\perp}$. Ya que $V = U + U^{\perp}$, si $p \in V$ y $p = u + w$ es su descomposición como suma de un vector de U y otro de $U^{\perp}$, entonces

$$g(v,p) = g(v,u) + g(v,w) = 0 + 0 = 0.$$

Esto prueba que $v \in R(g)$, pero como g es no degenerada, $R(g) = \{0\}$ y $v = 0$, como se quería probar.

Supongamos ahora que U es no degenerado. Para probar que $V = U \oplus U^{\perp}$, veamos que $U \cap U^{\perp} = \{0\}$ y aplicar posteriormente el teorema 3.4. Si $v \in U \cap U^{\perp}$, veamos que $v \in R(g)$. Dado $w \in V$, consideramos su descomposición $w = u + w'$, con $u \in U$ y $w' \in U^{\perp}$. Entonces es inmediato que $g(v,w) = g(v,u) + g(v,w') = 0$. Como $R(g) = \{0\}$ ya que g es no degenerada, entonces $v = 0$, como se quería probar.

La última parte del teorema es consecuencia de que $(U^{\perp})^{\perp} = U$. □

4. Formas cuadráticas

Hay que una manera 'equivalente' de trabajar con métricas, y es con formas cuadráticas.

Definición 4.1 *Sea (V,g) un espacio métrico. Se llama forma cuadrática asociada a g a la aplicación*

$$\phi_g : V \to \mathbb{R}, \ \ \phi_g(u) = g(u,u).$$

Cuando se sobreentienda la métrica g, escribiremos simplemente ϕ. Las siguientes propiedades son inmediatas:

Proposición 4.2 *Si ϕ es la forma cuadrática de una métrica g, entonces*

1. $\phi(\lambda u) = \lambda^2\phi(u)$*,* $\lambda \in \mathbb{R}$*.*

2. $\phi(0) = 0$*.*

3. $\phi(u+v) = \phi(u)+\phi(v)+2g(u,v)$*.*

4. $g(u,v) = \frac{1}{2}(\phi(u+v)-\phi(u)-\phi(v))$ *(forma polar de g).*

La última propiedad anterior nos dice que la métrica g se conoce a partir de la forma cuadrática. Veamos esto con el siguiente ejemplo.

Sea V un espacio de dimensión n. Sabemos que una matriz simétrica $A \in S_n(\mathbb{R})$ define una métrica en V del siguiente modo[1]: si B es una base, entonces g es la métrica cuya expresión matricial $M_B(g)$ es A, es decir, $g(x,y) = x^tAy$, cuando escribimos los vectores de V en coordenadas respecto de la base B. Hallamos la forma cuadrática de g. Si $u \in V$ tiene coordenadas $x \in \mathbb{R}^n$ respecto de B, es decir, $u = \sum x_ie_i$, entonces

$$\phi(u) = \phi(x) = x^tAx = \sum_{i,j=1}^{n} a_{ij}x_ia_{ij}x_j = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j.$$

Por tanto, toda expresión polinómica de grado 2 y homogénea en sus variables, es decir, no aparecen términos de grado 1 ni de grado 0, determina una forma cuadrática y también una métrica. Así, si consideramos en $\mathbb{R}^3$,

$$\phi(x) = x_1x_3 - 4x_2x_3 + 5x_3^2$$

entonces ϕ podemos escribirla como

$$\phi(x) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & -2 \\ \frac{1}{2} & -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Como consecuencia, la expresión de la métrica g asociada a ϕ es

$$M_B(g) = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & -2 \\ \frac{1}{2} & -2 & 5 \end{pmatrix}$$

y g se expresa como

$$g(x,y) = \frac{1}{2}x_1y_3 + \frac{1}{2}x_3y_1 - 2x_2y_3 - 2x_3y_2 + 5x_3y_3.$$

Todas las definiciones que hagamos para una métrica g se traslada a su forma cuadráticas. Así tenemos:

[1] En verdad, hay tantas métricas como bases, ya que cada base define una métrica diferente.

Definición 4.3 *Sea (V,g) un espacio métrico.*

1. *Si B es una base de V, se llama expresión matricial de ϕ respecto de B a $M_B(\phi) = M_B(g)$.*

2. *Se dice que la forma cuadrática ϕ_g es no degenerada (resp. degenerada) si g es una métrica no degenerada (resp. degenerada).*

También clasificamos las formas cuadráticas diciendo que son definidas positivas, definidas negativas, etc.

5. Clasificación de métricas. El teorema de Sylvester

Es evidente que el hecho de que la expresión matricial sea diagonal ha sido clave para distinguir la métrica. Por tanto, caben hacerse las siguientes preguntas:

1. ¿existe un método 'fácil' para poder clasificar una métrica?

2. ¿existe una base B tal que $M_B(g)$ sea una matriz diagonal?

3. dada una matriz simétrica $A \in S_n(\mathbb{R})$, ¿es posible diagonalizarla por congruencias?

El objetivo ahora va a ser dar una respuesta a las anteriores preguntas. Anticipamos cuál va a ser. Efectivamente, vamos a probar que existen tales bases, es decir, que toda matriz simétrica se diagonaliza por congruencias. Pero es más, y a la vista del ejemplo anterior, vamos a probar, que cuando se diagonaliza dicha matriz (o dicha métrica, para cierta base), el número de números positivos, negativos y 0, en la diagonal es independiente de la diagonalización que se haga, es decir, independiente de la base que se consiga. Éste es el llamado teorema de Sylvester.

Para la demostración del teorema de Sylvester, hay que probar dos cosas. Primero, la existencia de tales bases, o dicho de otro modo, que una matriz simétrica se puede diagonalizar por congruencias. Por otro, la unidad de la cantidad de los números positivos, negativos y cero que va a haber en la diagonal principal de la matriz diagonal que se consiga.

Para la primera parte, es decir, para la existencia, vamos a hacer tres demostraciones diferentes. Ambas son constructivas, es decir, nos van a dar un método para hallar dicha base. El primer método va a consistir en ir simplificando el problema, al ir reduciendo el orden de la matriz de la métrica. El segundo usará transformaciones elementales que ya se emplearon en la definición del rango de una matriz. El tercero consiste en reducir la expresión de la forma cuadrática asociada a una forma más sencilla mediante 'cuadrados perfectos'.

Antes de todo este proceso, podemos probar que el número de ceros es independiente de la base, y como uno sospecha, corresponde con la dimensión del radical de la métrica.

Teorema 5.1 *Sea un espacio métrico (V,g). Si B es una base tal que la métrica diagonaliza con B, es decir,*

$$M_B(g) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

entonces el número de 0 *en la diagonal principal coincide con la nulidad.*

Demostración. Hallamos el radical de g a partir de $M_B(g)$. Sabemos que el radical, en coordenadas respecto de B, es el núcleo de la matriz $M_B(g)$. O dicho de otro modo, su dimensión, es $n - r(M_B(g))$. Pero es evidente que dicho rango es el número de números no nulos de la diagonal principal, probando el resultado. □

Lema 5.2 *Sea un espacio métrico (V,g) y $U \subset V$ un subespacio vectorial. Si $V = U \oplus R(g)$, entonces $g_{|U}$ es no degenerada.*

Demostración. Sea $u \in U$ tal que $u \in R(g_{|U})$. Entonces u es ortogonal a todos los vectores de U. Pero como también es ortogonal a todos los de $R(g)$, entonces es ortogonal a todos los de V, ya que $V = U + R(g)$. Esto prueba que $u \in R(g)$. Pero como $U \cap R(g) = \{0\}$, entonces $u = 0$, como se quería probar. □

Una consecuencia del anterior resultado es la siguiente. Sea $\{e_1, \ldots, e_m\}$ una base de $R(g)$, y ampliamos hasta una base de V: $B = \{e_1, \ldots, e_n\}$. Llamamos $U =< e_{m+1}, \ldots, e_n >$ y $B'\{e_{m+1}, \ldots, e_n\}$. Entonces al hallar la expresión matricial $M_B(g)$, tenemos

$$M_B(g) = \left(\begin{array}{c|c} 0 & 0 \\ 0 & C \end{array} \right),$$

donde C es una matriz simétrica de orden $n - m$, concretamente, $C = M_{B'}(g_{|U})$. Por el lema anterior, C es una matriz regular.

El recíproco es cierto en el siguiente sentido. Supongamos que $B = \{e_1, \ldots, e_n\}$ es una base de V de manera que

$$M_B(g) = \left(\begin{array}{c|c} 0 & 0 \\ 0 & C \end{array} \right),$$

donde C es una matriz regular de orden $n-m$. Entonces sabemos que $n(g) = n - r(M_B(g))$, que aquí es $n-m$. Por tanto, la nulidad es m. También sabemos que el radical es el núcleo de la matriz $M_B(g)$, pero es evidente por la forma que tiene dicha matriz que $R(g) = < e_1, \ldots, e_m >$. Finalmente es evidente que $U = < e_{m+1}, \ldots, e_n >$ está en suma directa con $R(g)$ y que la métrica inducida en U es no degenerada, ya que expresión matricial respecto de la base anterior es C.

Teorema 5.3 (existencia: prueba 1. Reduciendo la dimensión) *Sea (V, g) un espacio métrico. Entonces existe una base donde la métrica diagonaliza.*

Demostración. Tomamos $R(g)$ y un subespacio U que esté en suma directa con $R(g)$. El lema anterior nos dice que la métrica $g_{|U}$ es no degenerada. Veamos que existe una base de U donde diagonaliza la métrica $g_{|U}$, y por tanto, al juntar esta base con una base de $R(g)$ y del hecho de que $V = R(g) \oplus U$, se obtiene una base de V donde diagonaliza la métrica g. O dicho de otro modo, basta probar el teorema para espacios métricos no degenerados.

Tomamos un vector $u_1 \in U$ tal que $g(u_1, u_1) \neq 0$. Este vector existe, pues en caso contrario, es decir, si para todo $u \in U$, $g(u, u) = 0$, entonces para todo $u, v \in U$, tenemos

$$0 = g(u+v, u+v) = g(u,u) + g(v,v) + 2g(u,v) = 0 + 0 + 2g(u,v),$$

es decir, $g(u, v) = 0$, y la métrica $g_{|U}$ sería nula, y así, degenerada: contradicción.

Una vez que tenemos u_1, sabemos que $U = < u_1 > \oplus < u_1 >^{\perp}$, y que $< u_1 >^{\perp}$ es no degenerado. Ahora hacemos el mismo argumento para $< u_1 >^{\perp}$, es decir, si llamamos $U_2 = < u_1 >^{\perp}$, entonces U_2 es no degenerado y existe $u_2 \in U_2$ tal que $g(u_2, u_2) \neq 0$. Es evidente que $g(u_1, u_2) = 0$. De esta manera, hemos reducido el problema a un subespacio de dimensión una menos que la que tenía U. Y así hacemos sucesivamente, hasta finalizar con un subespacio de dimensión 1. Después de este proceso, hemos encontrado n vectores $B' = \{u_1, \ldots, u_n\}$ con la siguientes propiedades:

1. B' es una base: esto es consecuencia que cada nuevo vector que se toma está en un subespacio que está en suma directa con el generador por los anteriores.

2. $g(u_i, u_i) \neq 0$, para todo i.

3. $g(u_i, u_j) = 0$ si $i \neq 0$.

Esto no es más que decir que la expresión matricial de $g_{|U}$ respecto de la base B' es diagonal, finalizando la demostración[2] $\square$

Observación. Según el método, primero hay que hallar el radical de la métrica, luego un subespacio U tal que $V = \mathrm{rad}(g) \oplus U$, hallar una base conjugada de U por el método anterior y añadirle una base del radical. Sin embargo, uno puede empezar sin necesidad de hallar el radical. Concretamente, empezamos con un vector $u \in V$ tal que $g(u,u) \neq 0$. Entonces hallamos $< u >^{\perp}$, que está en suma directa. Por tanto, tomamos u y le añadimos una base conjugada de $< u >^{\perp}$. Pudiera ocurrir que $< u >^{\perp}$ es ya el radical. Nos daríamos cuenta de ellos cuando al calcular la métrica en dicho subespacio es 0. Si no fuera el radical, existiría un segundo vector $u_2 \in < u >^{\perp}$ tal que $g(u_2, u_2) \neq 0$, y entonces seguiríamos con el método.

Observación. El problema que nos podemos encontrar con este método es que no sea 'fácil' hallar un vector tal que $g(u,u) \neq 0$. Si tenemos la expresión matricial de la métrica, basta mirar la diagonal principal. Si para algún i, $a_{ii} \neq 0$, entonces $g(e_i, e_i) = a_{ii} \neq 0$. ¿Qué sucede si todos los elementos de la diagonal principal son 0? Entonces nos fijamos en una entrada (i, j) de dicha matriz que no sea nula. Entonces el vector $v = e_i + e_j$ satisface

$$g(v, v) = 2g(e_i, e_j) = 2a_{ij} \neq 0,$$

y ya hemos encontrado el vector.

Realizamos ahora la segunda demostración. Como ya hemos comentado, si $A = M_B(g)$, tenemos que encontrar una matriz regular P tal que $P^t AP = D$ es una matriz diagonal. Dicha matriz P se escribirá como producto de matrices elementales: $P = E_1 \ldots E_k$. Entonces $D = (E_k)^t \ldots (E_1)^t AE_1 \ldots E_k$. Lo que estamos haciendo a la matriz A cuando la multiplicamos por la izquierda por $(E_1)^t$ y por la derecha por E_1 es realizar una transformación elemental por filas, y luego *la misma* por columnas. De esta manera,

> *Si queremos diagonalizar por congruencias, toda operación elemental que se haga a la matriz A por filas, hay que hacer la misma por columnas.*

[2] Uno podría hacer una demostración más formal, haciendo un argumento de inducción. Concretamente, ya hemos visto que basta probar el resultado para espacios métricos no degenerados. Entonces la demostración se hace por inducción *sobre la dimensión del espacio vectorial*. Para $n = 1$, es evidente. Si suponemos cierto para $n-1$, lo probamos para n. Dado el espacio vectorial U, y con la notación de la demostración, tomamos el vector u_1 y luego el subespacio U_2. Ya que U_2 es no degenerado y tiene dimensión $n-1$, aplicamos inducción sobre U_2, obteniendo el resultado.

Por tanto, la cuestión es: dada una matriz simétrica A ¿existen transformaciones por filas, y por columnas (las mismas) de manera que resulte una matriz diagonal? En tal caso, decimos que A es diagonalizable por congruencias.

Nota 5.4 *Obsérvese que diagonalizar por congruencias no es diagonalizar por semejanzas, a no ser que $P^t = P^{-1}$, lo cual no es cierto en general. Además, diagonalizar por semejanzas era lo mismo que diagonalizar endomorfismos, pero diagonalizar por congruencias es 'diagonalizar' formas bilineales.*

Teorema 5.5 (existencia: prueba 2. Transformaciones elementales) *Sea (V,g) un espacio métrico. Entonces existe una base donde la métrica diagonaliza.*

Demostración. Sea B una base cualquiera de V y $A = M_B(g)$ la expresión matricial de g respecto de una base B. Veamos que A diagonaliza por congruencias, y por tanto, la base buscada B' es la que satisface $P = M(I_V, B', B)$, donde P^tAP es una matriz diagonal.

Podemos suponer que la matriz A es no nula, pues en caso contrario, no hay nada que hacer ($g = 0$ y la matriz $D = 0$ para cualquier base).

1. Es posible tomar $a_{11} \neq 0$. Si a_{11} es cero, miramos en el resto de los elementos de la diagonal principal.

   a) Caso 1. Supongamos que existe un $a_{ii} \neq 0$. Para simplificar, suponemos que es a_{22}. Vamos a colocar a_{22} en el lugar $(1,1)$ mediante transformaciones elementales por congruencias. Sea $b = a_{22}$. Hacemos para ello la operación F_{12}, y luego C_{12}:
   $$\begin{pmatrix} 0 & a \\ a & b \end{pmatrix} \xrightarrow{F_{12}} \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \xrightarrow{C_{12}} \begin{pmatrix} b & a \\ a & 0 \end{pmatrix}.$$

   b) Caso que todos los $a_{ii} = 0$. Sea $a_{ij} \neq 0$. Entonces tenemos una submatriz 2×2 del tipo $\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$. Sumamos a la primera fila, la segunda, y lo mismo para las columnas, es decir, hacemos $F_{12}(1)$ y luego $C_{12}(1)$:
   $$\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} \xrightarrow{F_{12}(1)} \begin{pmatrix} a & a \\ a & 0 \end{pmatrix} \xrightarrow{C_{12}(1)} \begin{pmatrix} 2a & a \\ a & 0 \end{pmatrix}.$$

2. Una vez que tenemos $a_{11} \neq 0$, usamos este valor como pivote, y hacemos cero en la primera fila y columna, de la manera habitual, pero haciendo las mismas operaciones con las columnas.

3. Ahora pasamos a la matriz de orden uno menos que la anterior, que consiste en quitar la primera fila y la primera columna, y repetir el proceso.

$\square$

Tanto para el cálculo del rango de una matriz, como para diagonalizar por congruencias, es bueno tener como pivote el número 1. Para tener ese 1, hay hacer operaciones del tipo $F_i(\lambda)$. Como hay que hacer la operación tanto por fila como por columna, no basta con dividir por a_{ii}. Lo vemos en los dos siguientes ejemplos significativos. Supongamos que tenemos $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

1. Caso $a > 0$. Hacemos $F_1(1\sqrt{a})$ y luego $C_1(1/\sqrt{a})$:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \xrightarrow{F_1(1/\sqrt{a})} \begin{pmatrix} \sqrt{a} & \frac{b}{\sqrt{a}} \\ b & c \end{pmatrix} \xrightarrow{C_1(1/\sqrt{a})} \begin{pmatrix} 1 & \frac{b}{\sqrt{a}} \\ \frac{b}{\sqrt{a}} & c \end{pmatrix}.$$

2. Caso $a < 0$. Hacemos $F_1(1\sqrt{-a})$ y luego $C_1(1/\sqrt{-a})$:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \xrightarrow{F_1(1/\sqrt{-a})} \begin{pmatrix} -\sqrt{-a} & \frac{b}{\sqrt{-a}} \\ b & c \end{pmatrix} \xrightarrow{C_1(1/\sqrt{-a})} \begin{pmatrix} -1 & \frac{b}{\sqrt{-a}} \\ \frac{b}{\sqrt{-a}} & c \end{pmatrix}.$$

Como consecuencia de la demostración, la matriz P que se busca es la obtenida a partir de la matriz identidad I_n mediante las operaciones elementales. Concretamente, sabemos que

$$P^t = (E_k)^t \dots (E_1)^t I_n, \ \ P = I_n E_1 \dots E_k.$$

Por tanto,

La matriz P es:

1. *la que resulta de hacer a la matriz identidad I_n las mismas operaciones elementales por columnas, o,*
2. *la matriz traspuesta que resulta de hacer a la matriz identidad I_n las mismas operaciones elementales por filas.*

Teorema 5.6 (existencia: prueba 3. Cuadrados perfectos) *Sea (V, g) un espacio métrico. Entonces existe una base donde la métrica diagonaliza.*

Demostración. Respecto de una base $B=\{e_1,\ldots,e_n\}$ hallamos la expresión de la forma cuadrática asociada $\phi(x)=\sum_{ij}a_{ij}x_ix_j$. Distinguimos dos casos.

1. Existe algún elemento $a_{ii}\neq 0$. Sin perder generalidad, podemos suponer que es a_{11}, es decir, ϕ tiene en su expresión un término de la forma ax_1^2, con $a\neq 0$. Tomamos todos los términos de ϕ donde aparezca x_1, que llamamos $T(x_1)$, escribimos $\phi(x)=T_1(x_1)+T_2(x_2,\ldots,x_n)$ y trabajamos con $T_1(x_1)$. Estos términos se escribirán de la forma $T(x_1)=ax_1^2+x_1b$, donde b depende de las variables $x_2,\ldots,x_n$. Entonces,
$$T(x_1)=a(x_1^2+\frac{b}{a}x_1)=a((x_1+\frac{b}{2a})^2-\frac{b^2}{4a^2})=a(x_1+\frac{b}{2a})^2-\frac{b^2}{4a}.$$
Ahora $T_2-b^2/(4a)$ depende sólo de $x_2,\ldots,x_n$ y volvemos a aplicar el método.

2. Todos los elementos a_{ii} son cero. Entonces tomamos en ϕ todos los términos $S(x_i,x_j)$ donde aparezca x_i y x_j para algún para i,j fijos. Podemos escribir $S=ax_ix_j+bx_i+cx_j$, donde $a\neq 0$, b y c depende sólo del resto de variables. Escribimos:
$$\begin{aligned} S &= a(x_ix_j+\frac{b}{a}x_i+\frac{c}{a}x_j)=a(x_i+\frac{c}{a})(x_j+\frac{b}{a})-\frac{bc}{a} \\ &= \frac{a}{4}\left((x_i+x_j+\frac{b}{a}+\frac{c}{a})^2-(x_i-x_j+\frac{c}{a}-\frac{b}{a})^2\right)-\frac{bc}{a}. \end{aligned}$$
Aquí hemos usado la identidad $xy=(x+y)^2-(x-y)^2)/4$. Ahora el término bc/a junto $\phi-S$ sólo depende de todas las variables menos $x-i$ y x_j y se repite el método.

Hacemos un cambio de variable. En el caso (1), llamamos $x_1'=x_1+b/2a$ y $x_j'=x_j$ para $j>1$. Si $X'=(x_1,\ldots,x_n')$, las anteriores igualdades la escribimos como $X'=QX$, donde Q es una matriz triangular superior, concretamente, la primera fila es $(1,b/2a$ y el resto coincide las filas de la identidad. Esto muestra que Q es regular y es la matriz de cambio de base de una base B' definida por $Q=M(1_V,B,B')$. La base B' se calcula hallando la inversa de Q. La expresión de ϕ en esa base es $\phi(X')=ax_1'^2+\sum_{ij>1}a_{ij}'x_i'x_j'$.

En el caso (2), $x_i'=x_i+x_j+(b+c)/a$ y $x_j'=x_i-x_j+(c-b)/a$. Para simplificar la notación, suponemos $i=1$ y $j=2$. Hacemos el cambio de coordendas
$$x_1'=x_1+x_2+\frac{b+c}{a}$$
$$x_2'=x_1-x_2+\frac{c-b}{a}.$$

Ahora la matriz Q dada por $X' = QX$, donde $x'_j = x_j$ para $j > 2$, es

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

donde las últimas $(n-2)$ filas coincide con las de la identidad. Esto muestra que el rango de Q es n, luego determina una base B'. $\square$

Una vez probado que existe una base donde diagonaliza la métrica, la parte final para encontrar bases conjugadas es ahora sencilla.

Sea $B = \{e_1, \ldots, e_n\}$ una base donde diagonaliza la métrica. Después de una reordenación de los elementos de la base, colocamos en primer lugar aquéllos donde $g(e_i, e_i) > 0$, luego los que satisface $g(e_i, e_i) < 0$ y finalmente aquéllos que pertenecen al radical. Antes de probar la unicidad, cambiamos la base por otra, de manera que los elementos de la diagonal sean 1, -1 y 0. Para ello, basta con:

1. Si $g(e_i, e_i) > 0$, se cambia e_i por $e_i / \sqrt{g(e_i, e_i)}$.
2. Si $g(e_i, e_i) < 0$, se cambia e_i por $e_i / \sqrt{-g(e_i, e_i)}$.
3. Si $g(e_i, e_i) = 0$, se deja el mismo vector.

Observemos que este cambio de base por otra no cambia el número de elementos positivos y elementos negativos de la diagonal principal, sino que aquél que era positivo se cambia por 1 y el que era negativo, por -1.

Definición 5.7 *Una base conjugada o base ortonormal B en un espacio métrico es una base donde $M_B(g)$ es una matriz diagonal y los elementos de la diagonal principal son* 1, -1 *o* 0.

O dicho de otra manera, una base tal que $g(e_i, e_i)$ es 1, -1 o 0, y $g(e_i, e_j) = 0$ si $i \neq j$.

Como consecuencia del teorema de existencia anterior, tenemos:

Corolario 5.8 *Todo espacio vectorial métrico tiene bases conjugadas.*

Finalmente, probamos que el número de 1, -1 y 0 no depende de la base conjugada que se tome en (V, g).

Teorema 5.9 (unicidad) *Sea (V,g) un espacio vectorial métrico. Entonces el número de* 1*,* -1 *y* 0 *de la diagonal principal de cualquier expresión matricial de g respecto de una base conjugada, es independiente de la base conjugada que se tome en (V,g).*

Demostración. Ya hemos probado que el número de 0 corresponde con la nulidad de la métrica, que es un invariante. La demostración es por reducción al absurdo. Sean

$$B = \{e_1, \dots, e_m, e_{m+1}, \dots, e_{m+k}, e_{m+k+1}, \dots, e_{m+k+s}\}$$

$$B' = \{e_1, \dots, e_m, v_{m+1}, \dots, v_{m+k'}, v_{m+k'+1}, \dots, v_{m+k'+s'}\}$$

dos bases conjugadass, ordenadas para que los elementos de la expresión matricial sean 0, 1 y -1 en dicho orden, correspondientes a los números m, k y k', y s y s' respectivamente. Observemos que podemos tomar en ambas bases los mismos vectores que generan el radical ya que los 0 se corresponde con dicho subespacio

La demostración es por reducción al absurdo. Supongamos que $k > k'$ luego $s < s'$, pues $k+s = k'+s'$. Consideramos el conjunto de vectores

$$\{e_1, \dots, e_m, e_{m+1}, \dots, e_{m+k}, e'_{m+k'+1}, \dots, e'_{m+k'+s'}\}.$$

Los primeros m se corresponden con el radical, los siguientes k, con 1, y los últimos con -1. Veamos que estos vectores son linealmente independientes: una vez probado esto, la contradicción está en que el número de vectores es mayor que n, pues dicho número es $m+k+s' > m+k'+s' = n$.

Tomamos una combinación lineal nula:

$$0 = \sum_{i=1}^{m} c_i e_i + \sum_{i=1}^{k} a_i e_{m+i} + \sum_{i=1}^{s'} b_i e'_i.$$

Si

$$v = \sum_{i=1}^{m} c_i e_i + \sum_{i=1}^{k} a_i e_{m+i} = -\sum_{i=1}^{s'} b_i e'_i$$

no es difícil darse cuenta que

$$g(v,v) = \sum_{i=1}^{k} a_i^2 = -\sum_{i=1}^{s'} b_i^2.$$

Esto implica que $a_i = 0$ y $b_i = 0$ para todo i, luego $\sum_{i=1}^{m} c_i e_i = 0$, y como son independientes, entonces $c_i = 0$ para todo i. $\square$

Una vez probado los teoremas 5.3, 5.5 y 5.9, resumimos todo los anterior en el el teorema de Sylvester.

Teorema 5.10 (Sylvester) *En todo espacio vectorial métrico existen bases conjugadas. Además, el número de* 1, -1 *y* 0 *de la diagonal principal de cualquier expresión matricial de* g *respecto de una base conjugada, es independiente de la base conjugada que se tome en* (V,g).

Recordemos que un enunciado análogo lo podemos hacer para formas cuadráticas y matrices simétricas. Así tenemos:

Corolario 5.11 *Sea* $A \in S_n(\mathbb{R})$ *una matriz simétrica. Entonces existe una matriz regular* P *tal que*

$$P^tAP = \begin{pmatrix} 0 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 0 & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{pmatrix} = \begin{pmatrix} O_p & & \\ & I_k & \\ & & -I_m \end{pmatrix}.$$

Además, el número de cero p*, el número de* 1*,* k*, y el números de* -1*,* m*, es independiente de la matriz regular* P *que diagonalice por congruencias la matriz* A.

Del teorema de Sylvester, podemos dar *ahora* la siguiente definición.

Definición 5.12 *Sea* (V,g) *un espacio métrico. Se llama:*

1. *signatura de* g *a* $\sigma(g) = (k,m)$*, donde* k *y* m *es el número de* 1 *y* -1*, respectivamente, obtenidos de una base conjugada.*
2. *índice de* g *al número de* -1.

Por tanto:

1. $r(g) = k+m$.
2. $n(g)+r(g) = n(g)+k+m = dim(V)$.

3. si se conoce la signatura, se conoce la nulidad y el rango.
4. si se conoce la nulidad, se conoce el rango.

Corolario 5.13 *Sean $A, C \in S_n(\mathbb{R})$. Entonces A y C son congruentes si y sólo si tienen la misma signatura.*

Corolario 5.14 *Sea un espacio métrico (V, g) de dimensión n. Sea $\sigma(g) = (k, m)$.*

1. *La métrica es definida positiva sii $\sigma(g) = (n, 0)$.*
2. *La métrica es definida negativa sii $\sigma(g) = (0, n)$.*
3. *La métrica es semidefinida positiva sii $\sigma(g) = (k, 0)$.*
4. *La métrica es definida negativa sii $\sigma(g) = (0, m)$.*
5. *La métrica es no degenerada sii $k + m = n$.*
6. *La métrica es no degenerada sii $k + m < n$.*

Sabemos que la ecuación $\sum_{i=1}^{n} a_i x_i^2 = 0$ tiene solución trivial si y sólo si todos los a_i son positivos, o todos los a_i son negativos. Por tanto:

Corolario 5.15 *Sea una forma cuadrática ϕ en $\mathbb{R}^n$ y consideramos la ecuación $\phi(x) = 0$. Son equivalentes los siguientes enunciados:*

1. *La única solución es la trivial.*
2. *ϕ es definida positiva o definida negativa.*
3. *$\sigma(\phi) = (n, 0)$ o $\sigma(\phi) = (0, n)$.*

Una consecuencia del teorema de Sylvester, es el siguiente resultado que se usa en cálculo para saber si una matriz cuadrada y simétrica es definida positiva o definida negativa.

Teorema 5.16 *Sea (V, g) un espacio métrico y $A = M_B(g)$ para una base de B de V. Llamamos A_{ii} la submatriz que resulta de tomar en A las primera i filas y las primeras i columnas. Si denotamos por $|A_{ii}|$ su determinante, tenemos:*

1. *La métrica g es definida positiva sii $|A_{ii}| > 0$ para todo i.*

2. *La métrica g es definida negativa sii $(-1)^i|A_{ii}| > 0$ para todo i .*

Demostración. Hacemos la demostración para el primer apartado (el otro es análogo). Supongamos que g es definida positiva. Ya que todo subespacio vectorial de un espacio métrico definido positivo es definido positivo, cada subespacio $U_i =< B_i >=< e_1,\ldots,e_i >$ es definido positivo, donde $B = \{e_1,\ldots,e_n\}$. Por otro lado, el determinante de $M_{B_i}(g_{|U_i})$, que coincide con $|A_{ii}|$ es positivo porque esta matriz es congruente a la matriz identidad, luego es positivo.

El recíproco se hace por inducción sobre la dimensión del espacio vectorial. El caso $n = 1$ es evidente ya que $M_B(g)$ es una matriz 1×1, es diagonal, y por tanto, al tener determinante positivo, su signatura es $\sigma(g) = (1,0)$. Supongamos cierto para $n-1$ y sea $dim(V) = n$. Con la notación del apartado anterior, la expresión matricial $M_{B_{n-1}}(g_{|U_{n-1}})$ es la matriz $A_{n-1,n-1}$. Por hipótesis de inducción, $g_{|U_{n-1}}$ es definida positiva y por el teorema de Sylvester, existe una base conjugada $\{v_1,\ldots,v_{n-1}\}$. Como estos vectores son de U_{n-1}, entonces $B' = \{v_1,\ldots,v_{n-1},e_n\}$ es una base de V y la expresión matricial de g en dicha base es

$$M_{B'}(g) = \begin{pmatrix} 1 & & & a_{1,n} \\ & \ddots & & \vdots \\ & & 1 & a_{n-1,n} \\ a_{n,1} & \ldots & a_{n,n-1} & a_{nn} \end{pmatrix}.$$

Tenemos ahora dos maneras diferentes para acabar la demostración (de forma análoga al teorema de Sylvester).

1. Sea sabe que g es no degenerada pues el determinante de A no es cero. Como $g_{|U}$ tampoco es degenerada, $V = U\oplus U^\perp$. Si $U^\perp =< v_n >$, entonces el conjunto $B'' = \{v_1,\ldots,v_{n-1},v_n\}$ es base y

$$M_{B''}(g) = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ 0 & \ldots & 0 & \lambda \end{pmatrix}, \tag{1}$$

   con $\lambda = g(v_{n-1},v_{n-1})$. Luego B'' es una base donde diagonaliza la métrica. Además, el signo del determinante de esta matriz es el mismo que $|A|$, por ser congruentes, en particular, es positivo. Pero $|M_{B''}(g)| = \lambda$, luego $\lambda > 0$ y así, $\sigma(g) = (n,0)$, es decir, g es definida positiva.

2. Diagonalizamos por congruencias la matriz $M_{B'}(g)$. Para ello hacemos ceros en la última fila, haciendo transformaciones elementales del tipo $F_{n1}(-a_{n,1}),\ldots,F_{n,n-1}(-a_{n,n-1})$, y luego las mismas para las columnas. Al final nos queda una matriz del tipo (1), y el argumento finaliza del mismo modo que en el apartado anterior.

$\square$

Corolario 5.17 *Sea (V,g) es un espacio vectorial métrico y una base B. Si g es definida positiva (resp. definida negativa), entonces $|M_B(g)|>0$ (resp. $(-1)^n|M_B(g)|>0$).*

Acabamos esta parte generalizando a nuestro contexto el resultado de espacios vectoriales el resultado de espacios vectoriales que decía que un conjunto de vectores linealmente independientes se podía extender (añadir vectores) hasta obtener una base del espacio vectorial.

Teorema 5.18 *Sea un espacio métrico (V,g) y $\{e_1,\ldots,e_m\}$ vectores linealmente independientes tales que $g(e_i,e_j)=\delta_{ij}\varepsilon_i$, donde $\varepsilon_i=1$ o -1. Entonces existe una base conjugada de (V,g) que contiene a $\{e_1,\ldots,e_m\}$.*

Demostración. Por hipótesis, el subespacio $U=<e_1,\ldots,e_m\}$ es un subespacio no degenerado y $B_U=\{e_1,\ldots,e_m\}$ es una base conjugada. Extendemos B_u hasta conseguir una base B de V. Entonces la expresión matricial de la métrica g respecto de B es

$$M_B(g)=\left(\begin{array}{ccc|c} \varepsilon_1 & & & * \\ & \ddots & & * \\ & & \varepsilon_m & * \\ * & * & * & M_{B'}(g) \end{array}\right),$$

donde $M_{B'}(g_W)$ y $W=<B'>$, con $B'=\{e_{m+1},\ldots,e_n\}$. Se diagonaliza por congruencias. Como la matriz de los ε_i está formada por 1 y -1 (y no por 0), al diagonalizar, los primeros m vectores no cambian. Por tanto, al hacer ceros, y obtener la base conjugada, se obtiene una base de la forma $\{e_1,\ldots,e_m,e'_{m+1},\ldots,e'_n\}$ $\square$

6. Ejemplo de cálculo de bases conjugadas

Mostramos un ejemplo de cómo se calcula dicha base con los dos métodos.

Ejemplo 6.1 (método 1) Consideramos en $\mathbb{R}^3$ una métrica cuya expresión matricial respecto de la base usual es

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -2 \\ 3 & -2 & -7 \end{pmatrix}.$$

El primer paso es hallar el radical de la métrica. Ya que el rango de A es 2, la nulidad de g es 1. Entonces

$$R(g) = \{x : Ax = 0\} = \{(x,y,z) : x+2y+3z = 0, x-z = 0\} = < (1,-2,1) > .$$

Hallamos un subespacio en suma directa con $R(g)$. Para ello ampliamos a una base de $\mathbb{R}^3$ y tomamos el subespacio generado por los dos últimos vectores. Por utilidad posterior, ampliamos con los elementos de la base usual, por ejemplo, $B = \{(1,-2,1),(1,0,0),(0,1,0)\}$ y sea $U =< e_1, e_2 >$. Sabemos que U es no degenerado y la expresión de la métrica restringida a U respecto de la base $B' = \{e_1, e_2\}$ es

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

Buscamos un vector $u_1 \in U$ tal que $g(u_1,u_1) \neq 0$: ésta es la principal dificultad de este método. Para no complicarse, tomamos un vector de la base (si la hubiera) cuyo elemento (i,i) en la matriz no es cero. En este caso nos basta tomar $u_1 = e_1 = (1,0,0)$. Sabemos que $g(u_1,u_1) = 1$.

El segundo paso es hallar $< u_1 >^\perp$. Recordemos que esto se hace en el espacio métrico $(U, g_{|U})$. Como $u_2 \in U$, entonces se escribe $u_2 = xe_1 + ye_2 = (x,y,0)$. Ya que $g(u_2,u_1) = 0$, entonces

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0,$$

es decir, $x+2y = 0$. Por tanto, $x = -2y$ y $u_2 = y(-2,1)$. Luego tomamos $u_2 = (-2,1)_{B'} = (-2,1,0)$. Calculamos $g(u_2,u_2)$:

$$\begin{pmatrix} -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -4.$$

Por tanto la base $B'' = \{(1,-2,1),(1,0,0),(-2,1,0)\}$ diagonaliza la métrica con

$$M_{B''}(g) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

Esto nos dice que la nulidad es 1, la signatura es $(1,1)$, el rango es 2 y el índice es 1. Hallamos una base conjugada dividiendo por $\sqrt{\pm g(u_i,u_i)}$, según el caso. Sólo hay que hacerlo para el tercer vector. Como $g(u_2,u_2) = -4$, dividimos por 2. Por tanto una base conjugada es $B''' = \{(1,-2,1),(1,0,0),(-1,1/2,0)\}$ y la expresión matricial de la métrica es

$$M_{B'''}(g) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Ejemplo 6.2 (método 2) Consideramos la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -2 \\ 3 & -2 & -7 \end{pmatrix}.$$

Como tenemos un pivote en el lugar $(1,1)$, hacemos ceros en la primera columna y fila.

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -2 \\ 3 & -2 & -7 \end{pmatrix} \xrightarrow{F_{21}(-2),F_{31}(-3)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & -8 & -16 \end{pmatrix} \\ &\xrightarrow{C_{21}(-2),C_{31}(-3)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -8 \\ 0 & -8 & -16 \end{pmatrix} \xrightarrow{F_{32}(-2)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -8 \\ 0 & 0 & 0 \end{pmatrix} \\ &\xrightarrow{C_{32}(-2)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Hallamos la base donde diagonaliza. Tomamos la matriz identidad y hacemos las mismas operaciones por columnas:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_{21}(-2),C_{31}(-3)} \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_{32}(-2)} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

Por tanto la base donde diagonaliza es $B' = \{(1,0,0),(-2,1,0),(1,-2,1)\}$ con

$$M_{B'}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Cambiamos ahora la base por $B'' = \{(1,0,0),(-1,1/2,0),(1,-2,1)\}$, que es conjugada, y la matriz ahora es

$$M_{B''}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Entonces $R(g) =< (1,-2,1) >$, y se obtiene el mismo resultado que en el método anterior.

Ejemplo 6.3 (método 3) Consideramos la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -2 \\ 3 & -2 & -7 \end{pmatrix}.$$

La forma cuadrática es

$$\phi(x,y,z) = x^2 - 7z^2 + 4xy + 6xz - 4yz.$$

Entonces

$$x^2 + 4xy + 6xz = (x+2y+3z)^2 - (2y+3z)^2.$$

Por tanto,

$$\phi = (x+2y+3z)^2 - 4y^2 - 16z^2 - 16yz.$$

Trabajamos ahora con $-4y^2 - 16z^2 - 16yz$, fijando y. Entonces

$$-4y^2 - 16yz = -4(y^2 + 4yz) = -4((y+2z)^2 - 4z^2) = -4(y+2z)^2 + 16z^2.$$

Entonces

$$\phi = (x+2y+3z)^2 - 4(y+2z)^2.$$

Hacemos el cambio de variable

$$x' = x+2y+3z, y' = y+2z, z' = z$$

y escribimos

$$X' = QX = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} X.$$

Entonces B' viene dada por

$$M(1_{\mathbb{R}^3}, B'B_u) = Q^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

luego $B' = \{(1,0,0), (-2,1,0), (5,-4,1)\}$ y la expresión de la métrica es

$$M_{B'}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Una base conjugada es $B'' = \{(1,0,0), (-2,1,0)/2, (1,-2,1)\}$ y

$$M_{B''}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

De nuevo, $rad(g) =< (1,-2,1) >$, y se obtiene el mismo resultado que en el método anterior.

Tema 3. Espacios vectoriales euclídeos

1. Métricas euclídeas

Definición 1.1 *Una métrica euclídea en un espacio vectorial es una métrica definida positiva. Se dice que el espacio es un espacio (métrico) euclídeo.*

Por tanto:

1. Es equivalente a decir que la signatura es $\sigma(g) = (n, 0)$.
2. Es equivalente a que los determinantes encajados en una expresión matricial suya son todos positivos.
3. Es una métrica no degenerada.
4. Si $U \subset V$ es un subespacio vectorial, $V = U \oplus U^\perp$.
5. $dim(U^\perp) = dim(V) - dim(U)$.
6. Una base ortonormal es aquella base donde $g(e_i, e_j) = \delta_{ij}$.
7. Si B y B' son bases ortonormales, entonces la matriz de cambio de base P satisface $I = P^t I P$, es decir, $P^t P = I$.

Una matriz ortogonal A es aquella matriz cuadrada que satisface $A^t A = I$, y por tanto, también $AA^t = I$. En particular, es regular. El conjunto de todas las matrices ortogonales, que denotamos por $O(n)$, es un grupo para la multiplicación de matrices.

Proposición 1.2 *El espacio de matrices ortogonales $(O(n), \cdot)$ es un grupo. El conjunto $SO(n) = \{A \in O(n) : det(A) = 1\}$ se llama subgrupo especial ortogonal y también es un grupo.*

Sabemos cómo construir bases ortonormales. Para ello, hallamos primero una *base ortogonal*, es decir, una base donde la métrica diagonaliza, y luego, dividimos cada vector v de dicha base por $\sqrt{g(v,v)}$. Si $B = \{e_1, \ldots, e_n\}$ es una base y $A = M_B(g)$ es la expresión matricial de g en dicha base, sabemos que los elementos de la diagonal principal son positivos, y lo mismo sucede después de hacer transformaciones elementales. Llamamos *base de Gram-Schmidt* obtenida a partir de B a la base que se deduce de las siguientes operaciones: se hace ceros en la primera columna y fila mediante operaciones de la forma $F_{i1}(-a_{i1}/a_{11})$; se hace ceros en la segunda columna y fila mediante operaciones de la forma $F_{i2}(-a'_{i2}/a'_{11})$; y así sucesivamente. A la base resultante, después de hacerla conjugada la denotamos por $B' = \{e'_1, \ldots, e'_n\}$. Sea

$$U_k =< e'_1, \ldots, e'_k >, \;\; B'_k = \{e'_1, \ldots, e'_k\}.$$

La base de Gram-Schmidt tiene las dos siguientes propiedades:

1. $U_k =< e'_1, \ldots, e'_k >=< e_1, \ldots, e_k >$, para $1 \leq k \leq n$.
2. Si $A_k = M(1_{U_k}, B'_k, B_k)$, entonces $det(A_k) = 1$.

También podemos hallar la base ortogonal $\{\bar{e}_1, \ldots, \bar{2}_n\}$ correspondiente a la base de Gram-Schmidt mediante el siguiente proceso recurrente.

1. Sea $\bar{e}_1 = e_1$.
2. Se halla $\bar{e}_2 = e_2 + \lambda \bar{e}_1$ de manera que $g(\bar{e}_1, \bar{e}_2) = 0$. Con esta condición obtenemos

$$0 = g(e_2, \bar{e}_1) + \lambda g(\bar{e}_1, \bar{e}_1) \Rightarrow \lambda = -\frac{g(e_2, \bar{e}_1)}{g(\bar{e}_1, \bar{e}_1)}.$$

3. Se halla $\bar{e}_3 = e_3 + \lambda \bar{e}_2 + \mu \bar{e}_1$ de manera que

$$g(\bar{e}_3, \bar{e}_2) = g(\bar{e}_3, \bar{e}_1) = 0.$$

   De aquí,

$$\lambda = -\frac{g(e_3, \bar{e}_1)}{g(\bar{e}_1, \bar{e}_1)}, \mu = -\frac{g(e_3, \bar{e}_2)}{g(\bar{e}_2, \bar{e}_2)}.$$

y así se va siguiendo el proceso, obteniendo en cada paso

$$\bar{e}_k = e_k - \sum_{i=1}^{k-1} \frac{g(e_k, \bar{e}_i)}{g(\bar{e}_i, \bar{e}_i)} \bar{e}_i.$$

Recordemos también que ya que una métrica euclídea es no degenerada tenemos el siguiente resultado: si B' es una base ortonormal de un subespacio U, entonces existe una base ortonormal $B' \cup B''$ de V, es decir, podemos *ampliar* una base ortonormal de un subespacio hasta obtener una base ortonormal del espacio entero.

Proposición 1.3 *Sea $B = \{e_1, \ldots, e_n\}$ una base ortonormal.*

1. *Si $v \in V$, entonces $v = \sum_{i=1}^{n} g(v, e_i) e_i$.*
2. *Si $u, v \in V$, entonces $g(u, v) = \sum_{i=1}^{n} g(u, e_i) g(v, e_i)$.*
3. *Si $v \in V$, entonces $|v|^2 = \sum_{i=1}^{n} g(v, e_i)^2$. (ver definición posterior).*
4. *Sea $U \subset V$ un espacio de dimensión $n-k$ con ecuaciones cartesianas respecto de B:*
$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0, \;\; 1 \leq i \leq k.$$
Entonces una base de $U^\perp$ es $\{(a_{11}, \ldots, a_{1n}), \ldots, (a_{k1}x_1 + \ldots + a_{kn}x_n = 0)\}$.

Definición 1.4 *Sea (V, g) un espacio euclídeo. Se define el módulo de un vector $v \in V$ como*
$$|v| = \sqrt{g(v, v)}.$$

Tenemos las siguientes propiedades:

1. $|v| = 0$ si y sólo si $v = 0$.
2. $|\lambda v| = |\lambda||v|$.
3. $|u+v|^2 = |u|^2 + |v|^2 + 2g(u, v)$.
4. $|u+v|^2 = |u|^2 + |v|^2$ si y sólo si $u \perp v$.

Teorema 1.5 (desigualdad de Cauchy-Schwarz) *Si $u, v \in V$, entonces*
$$|g(u, v)| \leq |u||v|$$
y la igualdad se da si y sólo si $\{u, v\}$ son linealmente dependientes.

Demostración. Si uno de los vectores es 0, es evidente el resultado. Supongamos que no son cero. Para $\lambda \in \mathbb{R}$, se considera
$$|u + \lambda v|^2 \geq 0.$$
Entonces $|u|^2 + \lambda^2 |v|^2 + 2\lambda g(u, v) \geq 0$. Por tanto, el discriminante de este polinomio de segundo grado debe de ser no positivo, obteniendo la desigualdad.

Si son linealmente dependientes u y v, se da la igualdad. Si se da la igualdad, se da la igualdad en las desigualdades anteriores, en particular, $u + \lambda v = 0$ para cierto λ, probando que son linealmente dependientes. $\square$

Si u y v no son ceros, la desigualdad de Cauchy-Schwarz se puede escribir como

$$-1 \leq \frac{g(u,v)}{|u||v|} \leq 1.$$

Como la aplicación coseno es biyectiva de $[0,\pi]$ en $[-1,1]$, el número del centro es coseno de un *único* número $\theta \in [0,\pi]$.

Definición 1.6 *Si $u,v \in V$ y no son ceros, se define el ángulo entre u y v como $\angle(u,v) \in [0,\pi]$ tal que*

$$\cos\angle(u,v) = \frac{g(u,v)}{|u||v|}.$$

Tenemos las siguientes propiedades:

1. $\angle(u,v) = \angle(v,u)$.
2. $\angle(u,v) \in \{0,\pi\}$ si y sólo si $\{u,v\}$ son linealmente dependientes.
3. $g(u,v) = |u||v|\cos\angle(u,v)$.
4. $u \perp v$ si y sólo si $\angle(u,v) = \pi/2$.

2. Simetrías ortogonales y proyecciones ortogonales

Sea U un subespacio vectorial de un espacio euclídeo (V,g). Sabemos entonces que $V = U \oplus U^\perp$ y por tanto, todo vector $v \in V$ se escribe de forma única como suma de uno de U y otro de $U^\perp$: $v = u + w$.

Definición 2.1 *Se llama simetría ortogonal respecto de U a la aplicación*

$$S_U = S : V \to V, \;\; S(v) = u - w, \;\; (v = u + w).$$

Proposición 2.2 *Con la notación anterior,*

1. *S es un isomorfismo con $S \circ S = 1_V$.*
2. *$S_{|U} = 1_U$.*

3. $S_{|U^\perp} = -1_{U^\perp}$.

4. *S es diagonalizable, con valores propios* 1 *y* -1 *y los subespacios propios son* $V_1 = U$, $V_{-1} = U^\perp$.

Definición 2.3 *Se llama proyección ortogonal sobre* U *a la aplicación*

$$\pi : V \to V, \ \ \pi(v) = u, \ \ (v = u + w)$$

Proposición 2.4 *Con la notación anterior,*

1. π *es una aplicación lineal.*

2. $\pi \circ \pi = \pi$.

3. $Ker(\pi) = U^\perp$ *e* $Im(\pi) = U$. .

4. $\pi_{|U} = 1_U$.

5. π *es diagonalizable, con valores propios* 1 *y* 0 *y* $V_1 = U$, $V_0 = U^\perp$.

Si ahora denotamos con subíndices el subespacio correspondiente, tenemos:

Proposición 2.5 *1.* $S_U \circ S_{U^\perp} = -1_V$.

2. $\pi_U \circ \pi_{U^\perp} = 0$.

3. Endomorfismo autoadjunto

Definición 3.1 *Un endomorfismo* $f : (V, g) \to (V, g)$ *es autoadjunto si*

$$g(f(u), v) = g(u, f(v)), \ \ \forall u, v \in V.$$

Para una base B de V y si denotamos $A = M(f, B)$ y $G = M_B(g)$, la igualdad anterior se escribe

$$(AX)^t GY = X^t GAY \Leftrightarrow X^t A^t GY = X^t GAY, \ \ u = AX, v = AY.$$

Por tanto, $A^t G = GA$. En particular, si B es ortonormal, $A = A^t$.

Proposición 3.2 *Sea un endomorfismo $f:(V,g)\to(V,g)$ y $A=M(f,B)$. Son equivalentes los siguientes enunciados:*

1. *f es autoadjunto.*
2. *$A^tG=GA$ para cualquier base.*
3. *La matriz GA es simétrica para cualquier base.*
4. *$A^t=A$ para cualquier base ortonormal.*

Esto nos permite construir muchos endomorfismos autoadjuntos del siguiente modo: sea A una matriz simétrica de orden n. Cogemos un espacio euclídeo (V,g) de dimensión n y B una base ortonormal. Entonces el endomorfismo dado por

$$f(X)=AX$$

donde X representa coordenadas respecto de B, es un endomorfismo autoadjunto. En verdad, es un si y sólo si en el siguiente sentido.

Proposición 3.3 *Sea (V,g) un espacio euclídeo de dimensión n y B una base ortonormal. Si $S_n(\mathbb{R})$ es el espacio vectorial de las matrices simétricas, la aplicación*

$$\{\textit{Endomorfismos autoadjuntos de } (V,g)\}\longrightarrow S_n(\mathbb{R})$$

$$f\mapsto M(f,B)$$

es un isomorfismo de espacios vectoriales.

Proposición 3.4 *Sea $f:(V,g)\to(V,g)$ es un endomorfismo autoadjunto.*

1. *Si $U\subset V$ es un espacio vectorial tal que $f(U)\subset U$, entonces $f(U^\perp)\subset U^\perp$.*
2. *Si λ,μ son dos valores propios distintos de f, entonces $V_\lambda\perp V_\mu$.*

Demostración.

1. Trivial.

2. Si $u \in V_\lambda$ y $v \in V_\mu$, entonces $f(u) = \lambda u$ y $f(v) = \mu v$. Por tanto

$$g(f(u),v) = g(u,f(v)) \Rightarrow g(\lambda u,v) = g(u,\mu v) \Rightarrow (\lambda - \mu)g(u,v) = 0,$$

obteniendo el resultado.

□

La importancia de los endomorfismos autoadjuntos radica en el siguiente

Teorema 3.5 *Todo endomorfismo autoadjunto es diagonalizable. Además, es posible encontrar una base ortonormal de vectores propios.*

En términos de matrices, el enunciado anterior es equivalente a:

Teorema 3.6 *Toda matriz simétrica es diagonalizable. Además, es posible diagonalizar a la vez por semejanzas y congruencias, es decir, si $A \in S_n(\mathbb{R})$, existe $P \in O(n)$ tal que $P^t AP$ es una matriz diagonal.*

La clave en la demostración radica en probar que toda matriz simétrica real tiene tantos valores propios reales como indica su orden.

Lema 3.7 *Sea $A \in S_n(\mathbb{R})$. Entonces A tiene n valores propios.*

Demostración. Sea $P_A(\lambda)$ el polinomio característico de A. El teorema fundamental del Álgebra dice que dicho polinomio tiene n raíces complejas. Además, como los coeficientes del polinomio son reales, si λ es una raíz, su conjugada también lo es. Sea $\lambda = a + ib$ una raíz de $P_A(\lambda)$. Sea $C = (A - \lambda I)(A - \bar{\lambda} I)$, donde $\bar{\lambda} = a - ib$ es el conjugado de λ. Sabemos que el determinante de C es 0. Además, desarrollando, y usando que A es simétrica obtenemos:

$$C = (A - aI)^t(A - aI) + b^2 I.$$

Como el determinante de C es cero, existe $X \in \mathbb{R}^n$, $X \neq 0$ tal que $CX = 0$. Multiplicando a izquierda y derecha por X^t y X, tenemos

$$X^t(A - aI)^t(A - aI)X + b^2 X^t X = 0.$$

Observemos que si $X = (x_1, \ldots, x_n)$, entonces $X^t X = \sum_{i=1}^n x_i^2$. Por tanto, y llamando $Y = (A - aI)$, la igualdad anterior se escribe como

$$\sum_{i=1}^n y_i^2 + b^2 \sum_{i=1}^n x_i^2 = 0.$$

Como $X \neq 0$, entonces $b = 0$, probando que $\lambda \in \mathbb{R}$. □

Demostración. [del teorema 3.6.] Tomamos $U = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m}$, donde $\lambda_1, \ldots, \lambda_m$ son los valores propios de f. Falta por probar que $V = U$. En caso contrario, $V = U \oplus U^{\perp}$, con $dim(U^{\perp}) \geq 1$. Es evidente que $f(U) \subset U$, luego $f(U^{\perp}) \subset U^{\perp}$ por ser f autoadjunto. Entonces $f : U^{\perp} \to U^{\perp}$ es un endomorfismo autoajunto de U^t, en particular, tiene valores propios por el lema anterior y porque la dimensión de $U^{\perp}$ es mayor o igual que 1: contradicción, ya que todos los vectores propios de f ya estaban en U.

Además, la demostración nos da un método para hallar la base ortonormal.

> *Método:* Vamos tomando una base de cada subespacio propio, B_i. Entonces $B = B_1 \cup \ldots \cup B_m$ es una base de vectores propios de V. Además los vectores de B_i son perpendiculares a los de B_j. Ahora vamos cambiando cada base B_i por una base B'_i de vectores ortonormales (por ejemplo, Gram-Schmidt). Luego $B' = B'_1 \cup \ldots \cup B'_m$ es la base buscada.

□

Pongamos dos ejemplos:

1. Diagonalizamos por semejanzas y congruencias la matriz
$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$
Como A es simétrica, sabemos que lo que hay que hacer primero es hallar una base de vectores propios. El polinomio característico es $-\lambda^3 + 2\lambda^2$ y los valores propios son $\lambda = 2$ y $\lambda = 0$, que es doble. Aunque podríamos pensar si se satisface que la multiplicidad geométrica de $\lambda = 0$ coincide con su algebraica, el hecho de que A sea simétrica nos asegura que coinciden. Hallamos los subespacios propios: $V_2 =< (0,0,1) >$ y $V_0 =< (1,1,0),(1,0,0) >$. Hallamos bases ortonormales de cada uno de los subespacios propios. En este caso, la base de V_2 se deja como está, y la de V_0, es cambiarla a $\{(1,0,0),(0,1,0)\}$. Por tanto $P^t AP = D$, donde $P \in O(3)$ y
$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

2. Consideramos $(\mathbb{R}^2, g)$ el espacio euclídeo donde g es
$$M_{B_u}(g) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Definimos el endomorfismo $f:\mathbb{R}^2\to\mathbb{R}^2$ dado por

$$M(f,B_u)=\begin{pmatrix}0&0\\1&2\end{pmatrix}.$$

Probamos que f es autoadjunto y hallamos una base donde diagonaliza.

Observemos, en primer lugar, que $M(f,B_u)$ no es simétrica: pero también que B_u no es una base ortonormal. Para probar que f es autoadjunto hay que probar que la matriz

$$M_{B_u}(g)M(f,B_u)=\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}0&0\\1&2\end{pmatrix},$$

es simétrica. Pero

$$M_{B_u}(g)M(f,B_u)=\begin{pmatrix}1&2\\2&4\end{pmatrix}.$$

Para hallar la base ortonormal, diagonalizamos $M(f,B_u)$ como siempre se ha hecho: los valores propios de $P(\lambda)$ son $\lambda=0$ y $\lambda=2$ (probando que es diagonalizable, aunque ya se sabía esto de antemano). Hallamos base de los subespacios propios:

$$V_0=\{(x,y)\in\mathbb{R}^2:x+2y=0\}=<(-2,1)>.$$

Para hallar V_2 se haría de la manera estándard, pero aquí usamos que $V_2\perp V_0$, y por tanto, $V_2=(V_0)^\perp$. Entonces $(x,y)\in V_2$ si

$$(x\ \ y)\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix}=0,$$

obteniendo $x=0$. Por tanto, $V_2=<(0,1)>$. Entonces una base donde diagonaliza f es $\{(-2,1),(0,1)\}$ y ahora hallamos una ortonormal. Como hay dos vectores en la base que son perpendiculares al pertenecer a diferentes subespacios propios, entonces sólo hay que dividir cada uno por su módulo:

$$(-2\ \ 1)\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}-2\\1\end{pmatrix})=6,\ (0\ \ 1)\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix})=2.$$

Por tanto la base que se pide es $\{(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}}),(0,\frac{1}{\sqrt{2}})\}$.

Nos preguntamos si un endomorfismo definido en un espacio vectorial es autoadjunto para alguna métrica del mismo. Concretamente, sea $f:V\to V$ un endomorfismo en un espacio vectorial V, el cual no tiene en principio una estructura de espacio métrico. La pregunta anterior se formula del siguiente modo: ¿existe una métrica g en V tal que $f:(V,g)\to(V,g)$ sea autoadjunto? Si fuera así, es decir, es condición necesaria, el endomorfismo sería diagonalizable. Veamos que el recíproco es inmediato.

Teorema 3.8 *Sea $f : V \to V$ un endomorfismo. Son equivalentes los siguientes enunciados:*

1. *f es diagonalizable.*
2. *Existe una métrica g tal que $f : (V,g) \to (V,g)$ es autoadjunto.*

Demostración. Sólo hay que probar (1)$\Rightarrow$ (2). Sea $B = \{e_1, \ldots, e_n\}$ una base de vectores propios. Se define g de manera que B es una base ortonormal, es decir,

$$M_B(g) = I.$$

Para probar que es autoadjunto, y ya que B es una base ortonormal, es suficiente con ver que $A = M(f,B)$ es una matriz simétrica. Pero dicha matriz A es diagonal, luego simétrica. □

Una consecuencia importante de la teoría de endomorfismos autoadjuntos es que podemos hallar ahora fácilmente la signatura de una métrica.

Corolario 3.9 *Sea (V,g) un espacio métrico y $A = M_B(g)$ una expresión matricial suya. Entonces la signatura de g es $\sigma(g) = (k,s)$, donde*

$$k = \textit{número de valores propios de } A \textit{ que son positivos}$$

$$s = \textit{número de valores propios de } A \textit{ que son negativos}$$

Demostración. Sabemos que existe $P \in O(n)$ tal que $D = P^tAP$ es diagonal. Pero $P^t = P^{-1}$, luego dicha matriz D de valores propios es donde diagonaliza la métrica. □

4. Isometrías lineales

Definición 4.1 *Una isometría $f : (V,g) \to (V',g')$ entre dos espacios euclídeos es una aplicación biyectiva que es lineal y*

$$g'(f(u), f(v)) = g(u,v), \ \ \forall u,v \in V.$$

Por tanto, toda isometría es un isomorfismo. Las isometrías lineales preservan todas las propiedades métricas, como muestra el siguiente resultado.

Proposición 4.2 *Sea $f : (V,g) \to (V',g')$ una isometría.*

1. *f preserva módulos: $|f(v)| = |v|$, $\forall v \in V$.*
2. *f preserva ángulos: $\angle(u,v) = \angle(f(u), f(v))$, $\forall u, v \in V$.*
3. *f preserva la ortogonalidad: si $u \perp v$, entonces $f(u) \perp f(v)$.*
4. *f preserva los subespacios ortogonales: si $U \subset V$, entonces $f(U)^{\perp} = f(U^{\perp})$.*
5. *f preserva bases ortogonales: si B es una base ortogonal de (V,g), entonces $f(B)$ es una base ortogonal de $(V'g')$.*
6. *f preserva bases ortonormales: si B es una base ortonormal de (V,g), entonces $f(B)$ es una base ortonormal de $(V'g')$.*

También tenemos las siguientes caracterizaciones:

Proposición 4.3 *Sea $f : (V,g) \to (V',g')$ un isomorfismo.*

1. *f es una isometría si y sólo si f preserva módulos.*
2. *f es una isometría si y sólo si f preserva bases ortonormales.*

Un ejemplo destacado de isometrías apareció con las simetría ortogonales. Efectivamente, si U es un subespacio de un espacio euclídeo, y escribimos $V = U \oplus U^{\perp}$, entonces la simetría ortogonal respecto de U era la aplicación

$$S_U : V \to V, \;\; S(u+w) = u - w.$$

Recordemos también que S_U es diagonalizable, siendo $V_1 = U$ y $V_{-1} = U^{\perp}$. En el caso que U sea un hiperplano, a tal simetría la llamaremos *reflexión respecto de U*.

Sea una isometría f, B y B' bases ortonormales de V y V' respectivalmente. Como $f(B)$ es una base ortonormal de V', la matriz de cambio de base es una base ortogonal, es decir, $A = M(1_{V'}, f(B), B') \in O(n)$. Pero observemos que $A = M(f, B, B')$. Del mismo modo se tiene el recíproco, es decir,

Teorema 4.4 *Sea $f : (V,g) \to (V',g')$ un isomorfismo. Son equivalentes:*

1. *f es una isometría.*

2. *$M(f,B,B') \in O(n)$ para bases ortonormales B y B' de V y V' respectivamente.*

Demostración. Para bases cualesquiera, llamando $G = M_B(g)$ y $G' = M_{B'}(g')$, y escribiendo $A = M(f,B,B')$, entonces escribiendo en coordenadas los vectores, $g(X,Y) = X^t G Y$ y $g'(f(X), f(Y)) = X^t A^t G' A y$. Luego f será isometría sii $A^t G' A = G$. El resultado se sigue cuando al tomar bases ortonormales, las matrices G y G' son las matrices identidades. □

Observación. Viendo una matriz cuadrada de orden n como n vectores columnas, entonces esta matriz es ortogonal si y sólo si dichos vectores son una base ortonormal para la métrica euclídea de $\mathbb{R}^n$, es decir, dos a dos son perpendiculares y cada columna tiene módulo 1.

Esto permite establecer un isomorfismo: fijando una base ortonormal de (V,g),

$$\{\text{Isometrías de } (V,g) \text{ en } (V,g)\} \longrightarrow O(n)$$

$$f \longmapsto M(f,B).$$

Por tanto el conjunto de todas las isometrías de un espacio euclídeo (V,g) en sí mismo constituye un grupo para la composición. De manera sencilla, se tiene también que la composición de isometrías es una isometría y la inversa de una isometría también lo es (no estamos considerando aquí endomorfismos).

Proposición 4.5 *Sea $f : (V,g) \to (V,g)$ una isometría. Entonces:*

1. *$det(f) = \pm 1$.*
2. *Si λ es un valor propio, entonces $\lambda = \pm 1$.*
3. *$V_1 \perp V_{-1}$.*
4. *Si U es un subespacio vectorial tal que $f(U) \subset U$, entonces $f(U^\perp) \subset U^\perp$.*

Demostración.

1. Es consecuencia de que $M(f,B) \in O(n)$.
2. Si $f(u) = \lambda u$, entonces $g(f(u), f(u)) = g(u,u)$, es decir, $\lambda^2 g(u,u) = g(u,u)$, luego $\lambda^2 = 1$.

3. Si $u \in V_1$ y $v \in V_{-1}$, entonces para todo $u \in U$ tenemos:

$$g(u,v) = g(f(u), f(v)) = g(1 \cdot u, -1 \cdot v) = -g(u,v) \Rightarrow g(u,v) = 0.$$

Si $v \in U^{\perp}$, entonces

$$g(f(v),u) = g(f(v), f(f^{-1}(u))) = g(v, f^{-1}(u)) = 0.$$

□

Hemos visto que las simetrías ortogonales son diagonalizables. Damos la vuelta a este resultado en la familia de las isometrías lineales.

Teorema 4.6 *Toda isometría diagonalizable es una simetría ortogonal.*

Demostración. Sabemos que si f es una isometría, entonces los valores propios son 1 y -1. Como f es diagonalizable, entonces $V = V_1 \oplus V_{-1}$. Entonces es inmediato que $f = S_U$, donde $U = V_1$. □

Finalmente, extendemos el resultado que se tenía de espacios vectoriales que decía que dos espacios son isomorfos si y sólo si tienen la misma dimensión. Es claro que la identidad es una isometría, que la inversa de una isometría es otra isometría y que la composición de isometrías es otra isometría (esto se puede probar también llevando el problema a matrices ortogonales, y probando que $O(n)$ es un grupo). Por tanto, podemos establecer una relación de equivalencia en el conjunto de todos los espacios euclídeos: dos espacios euclídeos están relacionados si existe una isometría entre ellos. En tal caso se dice que son *isométricos*.

Tenemos ahora el siguiente resultado.

Teorema 4.7 *Dos espacios son isométricos si y sólamente tienen la misma dimensión.*

Demostración. Sólo hay que probar que si tienen la misma dimensión, es posible establecer una isometría entre ellos. Tomamos bases ortonormales en cada uno de los espacios. Como tienen la misma dimensión, las dos bases tienen el mismo número de elementos. Por tanto, podemos definir un isomorfismo que me lleve una base en la otra. Como además son ortonormales, entonces es una isometría. □

5. Formas canónicas de isometrías en $\mathbb{R}^2$ y $\mathbb{R}^3$.

Vamos a clasificar todas las isometrías de un espacio euclídeo (V,g) en el sentido que vamos a hallar expresiones matriciales *sencillas* de la isometría respecto de *adecuadas bases ortonormales*. Para los casos concretos en que la dimensión de V es 2 y 3, daremos nombres a cada una de las isometrías de la clasificación.

Una primera clasificación, aunque no determina de forma unívoca la isometría, atiende al determinante: si es 1 se llama *directa*, y si es -1, *inversa*. Además:

Proposición 5.1 *Respecto de bases ortonormales, una isometría es:*

1. *directa si y sólo si la matriz asociada es de $SO(n)$.*
2. *inversa si y sólo si la matriz asociada es de $O(n)-SO(n)$.*

Tomamos una base ortonormal B y sea $A=M(f,B)$. Sabemos que $A\in O(n)$. Luego lo que nos estamos preguntando es cómo son las matrices ortogonales de orden n. Denotamos por V_1 y V_{-1} los subespacios propios de f y escribamos $V=V_1\oplus V_{-1}\oplus U$. Sabemos que $f:U\to U$ y que no tiene valores propios. En particular, la dimensión de U es *par*, ya que todo polinomio de grado impar tiene al menos una raíz. Por otro lado, y como consecuencia del teorema 4.6, tenemos:

Corolario 5.2 *Una isometría es una simetría ortogonal si y sólo si su expresión matricial respecto de una base ortonormal es una matriz simétrica.*

Concretamos ahora el caso $n=2$. Sea $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in O(2)$. Multiplicando $A^tA=I$, tenemos las siguientes ecuaciones

$$a^2+c^2=1$$

$$b^2+d^2=1$$

$$ab+cd=0.$$

De las dos primeras, existe $\theta,\varphi\in[0,2\pi)$ tal que

$$a=\cos\theta, c=\sin\theta \quad b=\cos\varphi,\ d=\sin\varphi.$$

Y de la tercera, $\cos(\theta-\varphi)=0$. Por tanto, $\theta-\varphi=\pi/2$ o $\theta-\varphi=3\pi/2$. Concluimos entonces que A es uno de los dos siguientes tipos:

1.
$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

2.
$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Puede uno darse cuenta, hallando los valores propios que:

1. Las matrices del tipo (1) son diagonalizables con $dim(V_1) = dim(V_{-1}) = 1$. Además
$$V_1 =< (\sin\theta, 1-\cos\theta) > .$$
Respecto de una base $B = \{v_1, v_2\}$ con $v_1 \in V_1$ y $v_2 \in V_{-1}$, la matriz de la isometría es
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
Por tanto es una *simetría respecto de una recta*, o *reflexión respecto de una recta*.

2. Si $\theta = 0$, es la matriz identidad. Si $\theta = \pi$, es $-I$ y se llama *simetría central*. En cualquier otro caso, la matriz no es diagonalizables y se llama *giro de ángulo* θ o *rotación de ángulo* θ.

Teorema 5.3 *Las clasificación de las isometrías en un espacio euclídeo de dimensión* 2 *y sus formas canónicas son:*

1. *La identidad,* $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
2. *Simetría central,* $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.
3. *Reflexión respecto de una recta,* $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
4. *Giro o rotación de ángulo* $\theta \in [0, 2\pi)$, $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $\theta \neq 0, \pi$.

En la clasificación anterior, y para el giro, tenemos que probar que el ángulo θ es único: observemos que para las matrices ortogonales de orden 2 que son directas, y salvo la identidad y la simetría central, la expresión matricial de la isometría respecto de una base ortonormal es de la forma $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, pero puede ocurrir que al cambiar a otra base ortonormal, la expresión matricial sea del mismo tipo *pero para otro ángulo* ϕ. Para ello sólo que hay que hacer un cambio de base. Concretamos del siguiente modo.

Sea $f : (V,g) \to (V,g)$ una isometría en un espacio de dimensión 2 cuya expresión matricial respecto de una base ortonormal B es $M(f,B) = A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Sea B' otra base ortonormal, y sabemos que

$$M(f,B') = M(1_V,B,B')AM(1_V,B',B),$$

con $P = M(1_V,B'.B) \in O(2)$. Por la clasificación anterior, sabemos que la matriz P puede ser directa o inversa. Además, $P^{-1} = P^t$. Es inmediato que:

1. Si $P = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$, entonces $P^{-1}AP = A$.
2. Si $P = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$, entonces $P^{-1}AP = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

La diferencia entre un caso y otro viene en el hecho de si P es directa o inversa. Antes de precisar ahora la definición de giro, necesitamos la siguiente definición.

Definición 5.4 *Sea (V,g) un espacio euclídeo y B una base ortonormal que se fija en él. Se dice que una base ortonormal B' es una base positivamente orientada (resp. negativamente orientada) si $det M(1_V,B',B) = 1$ (resp. $det M(1_V,B',B) = -1$).*

Realizamos las siguientes observaciones:

1. El concepto de estar una base positivamente (o negativamente) orientada depende de la base ortonormal B elegida inicialmente. En el caso de $(\mathbb{R}^n, g_u)$ tomaremos siempre la base usual.
2. Estar positivamente (o negativamente) orientada establece una relación de equivalencia en el conjunto de las bases ortornormales.
3. El concepto se puede generalizar a bases cualesquiera, sin necesidad de que sean ortonormales, cambiando el hecho de que el determinante sea 1 o -1 por positivo o negativo, respectivamente.

A partir de ahora, y especialmente cuando tengamos que trabajar con giros, supondremos que en un espacio euclídeo hay una base ortonormal previamente fijada.

Definición 5.5 *Un giro de ángulo $\theta \in [0,2\pi)$ en un espacio euclídeo de dimensión 2 es una isometría no diagonalizable cuya expresión matricial respecto de una base ortonormal es*
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Observemos que de las formas canónicas, las trazas de las cuatro matrices son diferentes, a saber, 2, -2, 0 y $2\cos\theta \in (-1,1)$. Ya que la traza es un invariante de un endomorfismo, tenemos el siguiente método explícito para saber de qué tipo es una isometría en un espacio de dimensión 2.

Corolario 5.6 (método de clasificación) *Sea A la expresión matricial de una isometría f en un espacio de dimensión 2. Entonces:*

1. *Si traza$(A) = 2$ entonces f es la identidad.*
2. *Si traza$(A) = -2$ entonces f es la simetría central.*
3. *Si traza$(A) = 0$ entonces f es una reflexión respecto de la recta V_1.*
4. *Si traza$(A) \in (-1,1)$ f es un giro de ángulo θ con $\cos\theta =$ traza$(A)/2$. En este caso, el seno de θ es, salvo un signo, $\sqrt{4-\text{traza}(A)^2}/2$. El ángulo vendrá dado una vez que hayamos fijado una base ortonormal positivamente orientada.*

Pasamos ahora al caso $n = 3$ y sea $A \in O(3)$. Como el orden de la matriz es impar, entonces A tiene al menos un valor propio. Además, si tiene otro más, tiene entonces exactamente tres valores propios. Distingamos las dos posibilidades.

1. Si hay tres valores propios, estos pueden ser: tres valores 1, y es la identidad; tres valores -1, y es la menos identidad; dos valores 1 o un valor 1, teniendo una simetría respecto de un plano o una simetría respecto de una recta.
2. Si sólo hay un valores propio, sea U el correspondiente subespacio propio. Entonces $f(U) = U$, luego $f : U^\perp \to U^\perp$ es una isometría sin valores propios (salvo que $f_{|U^\perp}$ sea la identidad o menos la identidad. Atendiendo a la clasificación en el caso $n = 2$, debe ser un giro en $U^\perp$.

Teorema 5.7 *Las clasificación de las isometrías en un espacio euclídeo de dimensión* 3 *y sus formas canónicas son:*

1. a) *La identidad,* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

   b) *Simetría central,* $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

   c) *Reflexión respecto de un plano,* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

   d) *Simetría respecto de una recta, o simetría axial,* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

2. a) *Giro o rotación de ángulo* θ *respecto de una recta,* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in (0,2\pi) - \{\pi\}$.

   b) *Reflexión respecto de un plano* U *seguido de un giro respecto de* $U^{\perp}$ *de ángulo* θ, $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in (0,2\pi) - \{\pi\}$.

En cualquier de los dos casos de dimensiones, la isometría queda determinada por el conocimiento de los subespacios V_1 y V_{-1}.

Isometrías de V^2		
Nombre	dim (V_1)	dim (V_{-1})
Identidad	2	0
Reflexión respecto de una recta	1	1
Simetría central	0	2
Giro	0	0

Cuadro 1: Isometrías de un espacio euclídeo de dimensión 2.

De nuevo observamos que las trazas de las formas canónicas son todas diferentes en los siguientes casos: en el giro respecto de una recta o en el giro seguido de una reflexión,

Isometrías de V^3			
Nombre	dim (V_1)	dim (V_{-1})	dim $(V_1 \oplus V_{-1})^\perp$
Identidad	3	0	0
Reflexión respecto de un plano	2	1	0
Simetría axial	1	2	0
Simetría central	0	3	0
Giro respecto de una recta	1	0	1
Reflexión + giro	0	1	1

Cuadro 2: Isometrías de un espacio euclídeo de dimensión 3.

o cuando es uno de estos tipos y $1+2\cos\theta = \pm 1$ o $-1+2\cos\theta = \pm 1$. En el primer caso, el determinante, otro invariante, permite distinguirlos. En el segundo, sólo cuando $\cos\theta = 0$, es decir, $\theta = \pi/2$ o $\theta = 3\pi/2$, y el determinante de nuevo distingue los casos.

Corolario 5.8 (método de clasificación) *Sea A la expresión matricial de una isometría f en un espacio de dimensión 3. Sea $T = traza(A)$ y $D = det(A)$. Entonces:*

1. *Si $T = 3$, entonces f es la identidad.*
2. *Si $T = -3$, entonces f es la simetría central.*
3. *Si $T = 1$, $D = -1$, entonces f es una reflexión respecto del plano V_1.*
4. *Si $T = 1$, $D = 1$, entonces f es un giro de ángulo $\theta \in \{\pi/2, 3\pi/2\}$ respecto de V_1.*
5. *Si $T = -1$, $D = 1$, entonces f es una simetría axial respecto del plano V_1.*
6. *Si $T = -1$, $D = -1$, entonces f es un giro de ángulo $\theta \in \{\pi/2, 3\pi/2\}$ respecto de V_{-1} seguido de la reflexión respecto del plano $V_{-1}^\perp$.*
7. *Si $T \notin \{\pm 1, \pm 3\}$ y $D = 1$, entonces f es un giro respecto de la recta V_1 de ángulo θ con $\cos\theta = (T-1)/2$.*
8. *Si $T \notin \{\pm 1, \pm 3\}$ y $D = -1$, entonces f es un giro respecto de la recta V_{-1} de ángulo θ con $\cos\theta = (T+1)/2$, seguido de la reflexión respecto del plano $V_{-1}^\perp$.*

En los tres últimos casos, el ángulo vendrá dado una vez que hayamos fijado una base ortonormal positivamente orientada.

Del resultado anterior, y también para el caso de giros en un espacio de dimensión 2, determinamos ahora el cálculo de $\theta \in [0, 2\pi)$.

Objetivo. La finalidad de esta sección es doble:

1. Dada una isometría (por ejemplo, una expresión matricial), clasificar la isometría determinando los elementos geométricos que la definen.
2. Hallar la expresión de una isometría conociendo de qué tipo es.

Hacemos dos ejemplos ilustrativos:

1. Se considera el endomorfismo de $\mathbb{R}^3$ dado por

$$M(f,B_u) = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}.$$

Probar que es una isometría, clasificarla determinando los elementos geométricos que la definen y hallar una base ortonormal que determine su forma canónica.

Es fácil darse cuenta que $M(f,B_u)^t M(f,B_u) = I$, luego al ser B_u una base ortonormal, nos dice que f es una isometría. Como la matriz no es simétrica, entonces ya sabemos que no es una simetría ortogonal. Más aún, ya que la traza es $-1/3$ y el determinante es 1, sabemos que es un giro respecto de la recta V_1 con $\cos\theta = (-1/3-1)/2 = -2/3$. Hallamos V_1, es decir, $\text{Ker}(M(f,B_u) - 1\cdot I) =< (1,2,0) >$. Por otro lado, $(V_1)^\perp$ tiene ecuación $x+2y=0$: $(V_1)^\perp =< (0,0,1),(2,-1,0) >$.

Para hallar el ángulo, calculamos una base ortonormal positivamente orientada de $\mathbb{R}^3$ y para ello vamos a elegir una en la que intervengan los vectores de V_1 y de $V_1^\perp$. Una base ortonormal de $V_1^\perp$ es, por ejemplo, $B'' = \{v_1 = e_1, v_2 = e_2/\sqrt{5}\}$. Comprobamos si $B = \{(1,2,0)/\sqrt{5}, v_1, v_2\}$ está positivamente orientada, es decir, si su determinante es 1, o lo que es equivalente, a que $det((1,2,0),e_1,e_2)$ sea positivo. Como sí lo es, entonces es positivamente orientada. *Importante:* En caso de que no, basta con cambiar el orden de dos vectores, por ejemplo, los dos últimos.

Para el cálculo del $\sin\theta$ (que sabemos que es $\sqrt{5}/3$ o $-\sqrt{5}/3$), sabemos que

$$f(v_1) = \cos\theta v_1 + \sin\theta v_2 \Rightarrow \sin\theta = g(f(v_1),v_2).$$

Ya que $f(v_1) = (2,2,1)/3$, se tiene que $\sin\theta = 2/\sqrt{5}v_2$. De aquí se deduce el ángulo θ es aquel único número $\theta \in [0,2\pi)$ tal que $\cos\theta = -\frac{2}{3}$ y $\sin\theta = \frac{\sqrt{5}}{3}$.

Resumiendo, f es un giro de ángulo θ respecto de la recta $< (1,2,0) >$, con $\cos\theta = -\frac{2}{3}$ y $\sin\theta = \frac{\sqrt{5}}{3}$. La forma canónica es

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{2}{3} & -\frac{\sqrt{5}}{3} \\ 0 & \frac{\sqrt{5}}{3} & -\frac{2}{3} \end{pmatrix},$$

respecto de la base $\{(1,2,0)/\sqrt{5},(0,0,1),(2,-1,0)/\sqrt{5}\}$.

2. Hallar la matriz respecto de la base usual de un giro en $\mathbb{R}^3$ de ángulo $\pi/2$ respecto de la recta $L =< (1,0,1 >$.

   Hallamos una base de L, $L =< (1/\sqrt{2},0,1/\sqrt{2}) >$ y otra de
   $$L^{\perp} =< (0,1,0),(1,0,-1) >=< (0,1,0),(1/\sqrt{2},0,-1/\sqrt{2}) > .$$
   Consideramos la base ortonormal
   $$B = \{(1/\sqrt{2},0,1/\sqrt{2}),(0,1,0),(1/\sqrt{2},0,-1/\sqrt{2})\}.$$
   Nos preguntamos si está positivamente orientada. Como su determinante es -1, entonces no lo es. Por tanto, la base que tomamos es:
   $$B' = \{(1/\sqrt{2},0,1/\sqrt{2}),(1/\sqrt{2},0,-1/\sqrt{2}),(0,1,0)\}.$$
   Sabemos entonces que
   $$M(f,B') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\pi/2 & -\sin\pi/2 \\ 0 & \sin\pi/2 & \cos\pi/2 \end{pmatrix}.$$
   Entonces $M(f,B_u) = PM(f,B')P^t$, donde $P^t = M(1_{\mathbb{R}^3},B_u,B')$. Así
   $$\begin{aligned} M(f,B_u) &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\pi/2 & -\sin\pi/2 \\ 0 & \sin\pi/2 & \cos\pi/2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix}. \end{aligned}$$

Últimas observaciones. Sea f una isometría y A la expresión matricial respecto de una base B.

1. Para comprobar que f es una isometría hay que comprobar que $A^tGA = G$, donde $G = M_B(g)$. Si B es una base ortonormal, se reduce a comprobar que A es ortogonal.

2. El conocimiento de V_1 y V_{-1} permite saber *exactamente* qué isometría es: ver cuadro 2.

3. El cálculo del polinomio característico permite saber *exactamente* qué isometría es, pues podemos comprobar si 1 y -1 son valores propios y con qué multiplicidad.

4. El determinante de A permite una primera clasificación de la isometría en directa o inversa.

6. Clasificación de isometrías en dimensión arbitraria

Siguiendo la idea de la sección anterior, y para una isometría dada, escribimos $U = V_1 \oplus V_{-1}$. Entonces $f(U) \subset U$ y así $f(U^\perp) \subset U^\perp$, con $\dim(U^\perp) = 2k$. Restringimos f a $U^\perp$, donde sabemos que no hay vectores propios. Probamos las siguientes propiedades:

1. La aplicación $h := f + f^{-1} : U^\perp \to U^\perp$ es autoadjunta.

   Demostración. Evidente. □

2. Existe $e \in U^\perp$ tal que $h(e) = \lambda e$, es decir, $f^2(e) = \lambda f(e) - e$.

3. Sea $P_1 =< e, f(e) >$. Entonces P_1 es un plano vectorial y es invariante por f.

   Demostración. Si $ae + bf(e) = 0$, entonces $b = 0$, pues en caso contrario, e es un valor propio de f. Y si $b = 0$, entonces $a = 0$.

   Veamos ahora que $f(P_1) = P_1$. Por un lado, $f(e) \in P_1$. Por otro, $f(f(e)) = \lambda f(e) - e \in P_1$. □

4. Como $f_{|P_1}$ es una isometría y no tiene valores propios, entonces es un giro, es decir, de la forma
$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}.$$

5. Ahora en $U^\perp$, tomamos $U^\perp = P_1 \oplus P_1^\perp$, y hacemos el mismo razonamiento, volviendo a encontrar un plano vectorial donde la restricción de f es un giro.

Teorema 6.1 *Si $f : (V, g) \to (V, g)$ una isometría, entonces para cierta base ortonormal B de V, la expresión matricial de f es*

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & G_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & G_k \end{pmatrix}, \quad G_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

donde $\theta_i \in (0, 2\pi) - \{\pi\}$. A esta expresión matricial se le llama forma canónica.

Observemos que para una simetría ortogonal, es decir, una isometría donde $V = V_1 \oplus V_{-1}$, se tiene que si se toma una base B_1 de U y otra de $U^{\oplus}$, *no necesariamente ortonormal*, entonces la expresión matricial de la simetría respecto de B es:

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}.$$

De la demostración del teorema 6.1, concluimos:

Corolario 6.2 *Sea $f : (V,g) \to (V,g)$ una isometría. Si B es una base ortonormal, la matriz*

$$M(f + f^{-1}, B)$$

es simétrica.