

ÁLGEBRA LINEAL

Juan Diego Barrado Daganzo e Iker Muñoz Martínez¹
1º de Carrera

27 de mayo de 2023

¹Este documento se actualiza, para consultar las últimas versiones entrar en el enlace <https://github.com/JuanDiegoBarrado/AlgebraLineal>

QUIÉNES SOMOS

Somos un grupo de estudiantes de la Universidad Complutense de Madrid, concretamente del Doble Grado de Informática y Matemáticas que queremos compartir unos apuntes de calidad y, como mínimo, ordenados para que os sea más fácil llevar la asignatura al día (sobre todo a estudiantes de Doble Grado).

Estos apuntes son posibles gracias a la colaboración de más alumnos como tú que deciden aportar un granito de arena al proyecto. Puedes contribuir de la siguiente manera:

- Notificando erratas
- Modificando erratas
- Proponiendo mejoras
- Aportando ejemplos nuevos
- Aportando nuevas versiones

Para contribuir no tienes más que ponerte en contacto con juandbar@ucm.es o dejarnos un *Pull Request* en <https://github.com/JuanDiegoBarrado/AlgebraLineal>. Los detalles para que la contribución de todos sea lo más homogénea posible estarán en el fichero *Contribute.md* de dicho repositorio o, en caso de no aparecer correctamente, podéis poneros en contacto con el correo anteriormente mencionado.

Muchas gracias, esperamos que este documento te sea útil.

AGRADECIMIENTOS

Queremos dar las gracias al Profesor Juan Ramón Delgado, por ser el profesor que impartió la asignatura de *Álgebra Lineal* durante la elaboración de estos apuntes.

También queremos dar las gracias al Profesor Jose Manuel Gamboa, por elaborar otros manuales más formales sobre la asignatura que nos han permitido contrastar adecuadamente los nuestros y entender mejor los contenidos aquí expuestos.

Cálculo Diferencial © 2021 by Juan Diego Barrado & Iker Muñoz is licensed under Attribution-NonCommercial 4.0 International. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc/4.0/>

Índice general

1. Espacio vectorial y Aplicación lineal	5
1.1. Espacios Vectoriales	5
1.1.1. Dependencia lineal, sistemas generadores y base vectorial	8
1.1.1.1. Coordenadas y Cambio de Base	15
1.1.2. Espacio vectorial cociente	17
1.2. Aplicaciones Lineales	20
1.2.1. Isomorfía de Espacios Vectoriales	25
1.2.2. Matriz asociada a una aplicación lineal	27
1.2.2.1. Equivalencia de aplicaciones lineales y relación entre sus matrices	28
1.2.3. Rango de una aplicación lineal	30
2. Estructura de Endomorfismos	35
2.1. Polinomio mínimo y característico	36
2.1.1. Polinomio mínimo de un endomorfismo	36
2.1.1.1. Polinomio mínimo de un endomorfismo y un vector	38
2.1.1.2. Cálculo del polinomio mínimo	39
2.1.2. Polinomio característico de un endomorfismo	40
2.1.2.1. Teorema de Caley-Hamilton	41
2.2. Clases Notables de Endomorfismos	44
2.2.1. Diagonalizabilidad	45
2.2.1.1. Potencia enésima de matrices	48
2.2.2. Triangularizabilidad	49
2.3. Forma Canónica de Jordan	50
2.3.1. Teorema de Estructura	52

2.3.1.1.	Endomorfismos Nilpotentes	56
2.3.2.	Cálculo de Matrices y Bases de Jordan	58
2.3.2.1.	Conversión de trozos a endomorfismos nilpotentes	58
2.3.2.2.	Cálculo de la Matriz de Jordan	58
2.3.2.3.	Cálculo de la Base de Jordan	59
2.3.3.	Factores invariantes	59
3.	Formas Lineales y Bilineales	61
3.1.	Formas Lineales	61
3.1.1.	Espacio Dual	61
3.1.1.1.	Espacio Bidual	63
3.1.2.	Ortogonalidad	64
3.2.	Formas Bilineales	66
3.2.1.	Características de las formas bilineales	66
3.2.1.1.	Matriz asociada y Cambio de Base	68
3.2.2.	Simetría y antisimetría	69
3.2.2.1.	Forma cuadrática asociada	70
3.2.3.	Ortogonalidad	71
3.2.3.1.	Ortogonalidad en productos escalares	72
3.2.4.	Clasificación de formas bilineales	73
3.2.4.1.	Métodos de diagonalización de formas simétricas	75
3.2.4.2.	Clasificación de formas bilineales simétricas sobre \mathbb{C}	76
3.2.4.3.	Clasificación de formas bilineales simétricas sobre \mathbb{R} : Ley de inercia	77
4.	Espacio Vectorial Euclídeo	81
4.1.	Espacio Vectorial Euclídeo	81
4.2.	Aplicaciones Ortogonales	87
4.2.1.	Características de las aplicaciones ortogonales	88
4.2.2.	Clasificación de aplicaciones ortogonales	90
5.	Espacio Afín	93
5.1.	Espacio Afín	93
5.1.1.	Variedades Afines	95

5.1.1.1.	Ecuaciones cartesianas	99
5.1.1.2.	Ecuaciones paramétricas	99
5.1.1.3.	Paso de paramétricas a cartesianas	100
5.1.2.	Sistema de Referencia	100
5.1.2.1.	Cambio de sistema de referencia	101
5.2.	Aplicaciones afines	102
5.2.1.	Variedades ϕ -invariantes	104
5.2.2.	Aplicaciones notorias	106
5.2.2.1.	Aplicación constante	106
5.2.2.2.	Traslación de vector u	107
5.2.2.3.	Homotecia	108
5.2.2.4.	Proyección	109
5.2.2.5.	Simetrías	111
5.2.2.6.	Homología	112
5.2.3.	Clasificación de las Afinidades del Plano	113
5.2.3.1.	Clasificación en \mathbb{C}	113
5.2.3.2.	Clasificación en \mathbb{R}	113
5.3.	Isometrías	114
5.3.1.	Clasificación de Movimientos Rígidos	115
5.3.1.1.	Clasificación en el plano	116
5.3.1.2.	Clasificación en el espacio	117
6.	Cónicas	119
6.0.1.	Clasificación Afín	119
6.0.1.1.	Caracterización de las Cónicas	122
6.0.2.	Clasificación Afín-Euclídea	124

ESPACIO VECTORIAL Y APLICACIÓN LINEAL

Este capítulo define y estudia el concepto de los espacios vectoriales y los morfismos entre ellos: las aplicaciones lineales. Principalmente estudiaremos los elementos necesarios para trabajar cómodamente en dichos espacios y la representación matricial de las aplicaciones lineales que nos permitirá hacer lo propio, pero con ellas como protagonistas.

ESPACIOS VECTORIALES

El espacio vectorial es la representación algebraica en matemáticas de los objetos “cotidianos” (por las propiedades que le vamos a exigir) que son cerrados bajo la suma de elementos del conjunto y producto por un escalar.

Hay multitud de conjuntos que tienen estructura de espacio vectorial como son los polinomios de grado n concreto, las aplicaciones continuas de \mathbb{R} en \mathbb{R} , las sucesiones de números reales... Por lo que un conocimiento profundo de dicha estructura nos da ventaja a la hora de trabajar con todos los conjuntos anteriores. De ahí su importancia.

Definición (Espacio Vectorial)

Sea la terna $(V, +, \cdot)$ formada por las operaciones $+: V \times V \rightarrow V$, $\cdot: K \times V \rightarrow V$, donde K es un cuerpo conmutativo, decimos que es un **espacio vectorial sobre K** si y sólo si:

- $(V, +)$ es un grupo conmutativo.
- **Distributiva de la suma vectorial:** $\forall a \in K : \forall v, v' \in V : a(v + v') = av + av'$
- **Distributiva de la suma escalar:** $\forall a, a' \in K : \forall v \in V : (a + a')v = av + a'v$
- **Asociativa del producto por escalar:** $\forall a, a' \in K : \forall v \in V : a(a'v) = (aa')v$
- **Elemento Neutro del producto por escalar:** $\forall v \in V : 1_K \cdot v = v$

A los elementos de un espacio vectorial los llamaremos **vectores**.

En definitiva, vemos que se están pidiendo que el espacio vectorial tenga dos operaciones con propiedades que casi damos por hecho en multitud de conjuntos, lo que hace razonable la idea de que multitud de elementos en las matemáticas van a cumplir esta definición.

Proposición (Propiedades Algebraicas)

Sea $(V, +, \cdot)$ un espacio vectorial, denotando por 0_k al “cero” de K y 1_k al elemento neutro del producto de K y 0_v el elemento neutro del espacio vectorial con respecto a la suma, para $a, b \in K$ se tiene:

1. $ab = 0_k \Rightarrow a = 0_k \vee b = 0_k$
2. $a \cdot 0_k = 0_k$
3. $0_k \cdot v = 0_v$
4. $a \cdot 0_v = 0_v$
5. $a \cdot v = 0_v \Rightarrow a = 0_k \vee v = 0_v$

Demostración:

1.

$$ab = 0_k \text{ y } a \neq 0_k \xRightarrow{\exists a^{-1} \in K} a^{-1}ab = a^{-1}0 \Rightarrow b = 0_k$$

2.

$$a \cdot 0_k = a \cdot (0_k + 0_k) = a \cdot 0_k + a \cdot 0_k \xLeftrightarrow{-a \cdot 0_k} a \cdot 0_k - a \cdot 0_k = a \cdot 0_k + a \cdot 0_k - a \cdot 0_k \Leftrightarrow 0_k = a \cdot 0_k$$

3.

$$0_k \cdot v = (0_k + 0_k)v = 0_kv + 0_kv \xLeftrightarrow{-0_kv} 0_kv - 0_kv = 0_kv + 0_kv - 0_kv \Leftrightarrow 0_v = 0_kv$$

4.

$$av = 0_v \wedge a \neq 0_k \xRightarrow{\exists a^{-1}} a^{-1}av = a^{-1}0_v \Leftrightarrow (a^{-1}a)v = 0_v \Leftrightarrow v = 0_v$$

Ejemplo:

Como hemos comentado al principio, hay multitud de objetos matemáticos que tienen estructura de espacio vectorial, entre ellos:

- Sea $V = K^n = K \times \cdots \times K$ y sean las operaciones de suma y producto:

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$$

$$a \cdot (a_1, \dots, a_n) = (a \cdot a_1, \dots, a \cdot a_n)$$

es trivial comprobar que se verifican las propiedades de grupo conmutativo con respecto a la suma (que hay opuesto, que hay elemento neutro, etc.) y también que se verifican las relativas al producto. Por ejemplo, por demostrar alguna, vemos que:

$$(ab)(a_1, \dots, a_n) = (aba_1, \dots, aba_n) = a(ba_1, \dots, ba_n)$$

Este ejemplo concreto tiene nombre propio y es lo que conocemos por **enésimo espacio vectorial estándar o canónico**.

- Sea $V = K[x]$, es decir, el conjunto de los polinomios con coeficientes en K y la suma¹ y el producto por un escalar:

$$f(x) + g(x) = \sum_{i=0}^n a_i x_i + \sum_{i=0}^m b_i x_i = \sum_{i=0}^n (a_i + b_i) x_i$$

$$b \cdot f = b \sum_{i=0}^n a_i x^i = \sum_{i=0}^n (b \cdot a_i) x^i$$

de nuevo es trivial comprobar que tiene estructura de espacio vectorial.

¹En la suma de la definición se ha considerado que $n > m$ y que los b_i que se han añadido para completar el grado de $g(x)$ son todos cero.

- Sea $V = K_n[x] = \{f \in K[x] : \text{gr}(f) < n\}$ y las operaciones anteriores, este conjunto también verifica las propiedades de espacio vectorial. Cabe destacar que también es necesario comprobar que el conjunto es cerrado bajo la suma y el producto, esto es, que las operaciones no van a tener como resultado elementos de fuera del conjunto, es decir, polinomios de mayor grado.

- Sea $V = \text{Mat}_{m \times n}(K)$ y las operaciones suma y producto las definidas componente a componente como:

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

$$b(a_{ij}) = (b \cdot a_{ij})$$

también es trivial la comprobación de la estructura de espacio vectorial.

- Sea $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ donde } f \in C(\mathbb{R})\}$ y la suma y el producto:

$$\forall x \in \mathbb{R} : (f + g)(x) = f(x) + g(x)$$

$$(a \cdot f)(x) = a \cdot f(x)$$

también verifica las condiciones de espacio vectorial.

El ejemplo anterior sugiere una relación importante entre espacios vectoriales. Hemos visto que los polinomios de grado arbitrario son espacios vectoriales, pero que también los de grado acotado también lo son. El último conjunto está contenido en el primero, por lo que tenemos un espacio vectorial metido dentro de otro espacio vectorial y ambos definidos con las mismas operaciones.

Definición (Subespacio Vectorial)

Sea V un espacio vectorial sobre un cuerpo K y $W \subseteq V$ un subconjunto cuyo no vacío, decimos que W es un **subespacio vectorial** de V si es cerrado bajo las operaciones de V , es decir:

- $w, w' \in W : w + w' \in W$
- $a \in K \wedge w \in W : a \cdot w \in W$

Y lo denotamos por $W < V$.

Observación:

Uno podría pensar que habría que pedir que también tenga estructura de espacio vectorial con las operaciones definidas para V , pero ¡estas ya se verifican trivialmente! Como $W \subseteq V$ y las propiedades se cumplen $\forall v \in V$, en particular se cumplen para todos los $w \in W \subseteq V$ y, en consecuencia, con pedir que W sea cerrado bajo las operaciones de V cumple él mismo la definición de espacio vectorial.

Ejemplo:

Sea $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ donde } f \in C(\mathbb{R})\}$ que ya hemos visto que es un espacio vectorial, el conjunto $W = \{f \in V : f(-x) = f(x)\}$ es un subespacio vectorial pues es cerrado bajo la suma y el producto:

$$f, g \in W : (f + g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f + g)(x)$$

$$a \in \mathbb{K}, g \in W : (a \cdot g)(-x) = a \cdot g(-x) = a \cdot g(x) = (a \cdot g)(x)$$

Ver que un espacio vectorial puede contener otros espacios vectoriales más pequeños en su interior suscita en el matemático la curiosidad por estudiar cómo se comporta la relación de subespacio vectorial con respecto a las operaciones habituales conjuntistas: uniones, intersecciones, etc.

Proposición (Operaciones con subespacios)

Sea V un espacio vectorial y $W_1, W_2 \subset V$ subespacios vectoriales suyos, entonces:

- $W_1 \cap W_2 < V$

$$\blacksquare W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1 \wedge w_2 \in W_2\} < V$$

es decir, la estructura de espacio vectorial se conserva por la intersección y la suma de subespacios.

Observación:

De haber sido uno el que hubiera pensado el enunciado de la proposición anterior, lo más razonable al comenzar a escribirlo hubiera sido pensar que la unión $W_1 \cup W_2$ iba a ser un espacio vectorial. Sin embargo, en general, esta operación no conserva las propiedades de espacio vectorial. Para visualizarlo, pueden verse W_1 como el eje de abscisas y W_2 el de ordenadas. Es sencillo ver que la unión de ambos no es cerrada bajo la suma, puesto que $(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$.

Dependencia lineal, sistemas generadores y base vectorial

Una vez definida la estructura de espacio vectorial, vamos a ver una serie de resultados que nos van a permitir expresar cualquier vector del espacio en términos de solo unos pocos. Principalmente esto nos va a permitir razonar sobre el efecto que va a tener una transformación en cualquier vector simplemente observando que es lo que ocurre con dicha transformación en este selecto conjunto de vectores especiales.

Definición (Independencia lineal)

Sea el conjunto de vectores $\{v_1, v_2, \dots, v_n\}$, decimos que es **linealmente independiente** si y sólo si

$$a_1 v_1 + \dots + a_m v_m = 0_v \Rightarrow a_1 = a_2 = \dots = a_m = 0$$

es decir, que cualquier combinación lineal nula implica que los coeficientes son nulos.

Cuando esta condición no se verifica entonces decimos que el conjunto es **linealmente dependiente** y, en este caso, tenemos una combinación lineal con algún coeficiente no nulo:

$$\forall a_1, \dots, a_m \in K : \exists a_1 v_1 + \dots + a_m v_m = 0_v : \exists a_i \neq 0$$

Precisamente esta combinación es la que hace honor al nombre de *dependencia lineal* pues cualquier vector con coeficiente no nulo se puede expresar en términos de los demás:

$$v_i = -a_i^{-1} a_1 v_1 - \dots - a_i^{-1} a_{i-1} v_{i-1} - \dots - a_i^{-1} a_m v_m$$

Ejemplo:

Sea K^2 el espacio vectorial canónico y tomando los vectores $e_1 = (1, 0)$ y $e_2 = (0, 1)$, vemos que son linealmente independientes puesto que:

$$a_1 \cdot e_1 + a_2 \cdot e_2 = (0, 0) \Leftrightarrow (a_1, a_2) = (0, 0) \Leftrightarrow a_1 = a_2 = 0$$

pero si al conjunto $\{e_1, e_2\}$ añadimos el vector $v = (1, -2)$ se transforma en un conjunto linealmente dependiente pues:

$$-1 \cdot e_1 + 2 \cdot e_2 + 1 \cdot v = (0, 0)$$

Como hemos visto que la dependencia lineal permite expresar los vectores con coeficiente no nulo en términos de otros, nos interesa ahora estudiar qué vectores son dependientes de un conjunto de vectores dados. Es decir, lo que queremos saber es qué vectores son resultado de la combinación lineal del conjunto que tengamos entre manos.

Definición (Envoltura lineal)

Sea $S = \{v_1, \dots, v_m\}$ un conjunto de vectores, definimos su **envoltura lineal** como:

$$L(S) = \{a_1 v_1 + \dots + a_m v_m : a_1, \dots, a_m \in K\}$$

es decir, como el conjunto formado por todas las posibles combinaciones lineales.

Proposición

Sea $L(S)$ la envoltura lineal de un conjunto de vectores S , entonces:

- $L(S)$ tiene estructura de espacio vectorial
- $S \subset L(S)$
- $L(S)$ es el menor subespacio vectorial en V que contiene a S

Después de responder a la pregunta anterior, ahora surge otra: ¿habrá algún conjunto de vectores que me permitan expresar en términos de ellos todo un subespacio vectorial? ¿Y todo el espacio vectorial?

Definición (Sistema generador)

Sea $F \subset V$ un subespacio vectorial y S un conjunto de vectores, decimos que S es un **sistema generador**² de F si y sólo si $L(S) = F$.

Ejemplo:

Sea K^2 el espacio vectorial canónico, el conjunto $\{e_1, e_2\}$ es generador de K^2 porque un elemento cualquiera de $L(e_1, e_2)$ es de la forma $a_1e_1 + a_2e_2 = (a_1, a_2)$. Esto quiere decir que para cualquier $v \in V$ no tengo más que mirar sus coordenadas y dar esos valores a a_1 y a a_2 para poder ponerlo como combinación lineal de los vectores e_1 y e_2 .

Definición (Base vectorial)

Sea $B = \{v_1, \dots, v_n\}$ un conjunto de vectores, decimos que B es una **base vectorial** de V si y sólo si es linealmente independiente y sistema generador. Esto implica que los vectores de V se pueden expresar en términos de la base:

$$\forall v \in V : v = a_1v_1 + \dots + a_nv_n = a'_1v_1 + \dots + a'_nv_n$$

De hecho, los coeficientes son únicos, pues:

$$0 = (a_1 - a'_1)v_1 + \dots + (a_n - a'_n)v_n \xrightarrow{l.i.} a_1 - a'_1 = \dots = a_n - a'_n = 0 \Rightarrow a_i = a'_i$$

Ejemplo:

- Sea K^n el espacio vectorial canónico, llamamos a $B_c = \{e_1, \dots, e_n\}$ su base canónica.
- Sea $K[x]$ el espacio de los polinomios con coeficientes en K , éste no es finitamente generado porque su base no es finita, siendo una de ellas $B = \{1, x, x^2, \dots\}$.
- Sea $V = Mat_{3 \times 2}(K)$, una de las bases es:

$$B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

El concepto de base va a ser central y esencial en el desarrollo posterior de toda la asignatura. Por ello, vamos a perder un poco de tiempo estudiando y demostrando dos propiedades fundamentales: que la base es finita y que tiene tantos vectores como dimensión el espacio vectorial. Esto es esencial pues si hemos dicho que la virtud de este elemento es que permite estudiar en sus vectores las propiedades de una transformación y extender los resultados a los demás, entonces necesitamos que sea finita para que el estudio sea útil.

²En este documento, nos vamos a limitar a estudiar espacios vectoriales finitamente generados, esto es, aquellos en los que el sistema generador está formado por un conjunto finito de vectores.

Proposición

Sea $B = \{v_1, \dots, v_n\}$, este es base si y sólo si es un sistema de generadores **minimal**, es decir, que ningún subconjunto de B es sistema de generadores.

Demostración:

- “ \Rightarrow ”: supongamos que B no es minimal, entonces $\exists S \subsetneq B$ tal que S es sistema de generadores:

$$\exists v \in B : v \notin S : v = a_1 v_1 + \dots + a_m v_m : v_n, \dots, v_m \in S \Rightarrow 0_v = 1_k v - a_1 v_1 - \dots - a_m v_m \Rightarrow B \text{ no es l. i.}$$

- “ \Leftarrow ”: supongamos que B no es l.i., entonces $\exists v_1, \dots, v_m \in B : a_1 v_1 + \dots + a_m v_m = 0_v \Rightarrow$

$$\Rightarrow \exists i : a_i \neq 0_k \xrightarrow{p.ej.} a_m \neq 0 \Rightarrow v_m = -a_m^{-1} a_1 v_1 - \dots - a_m^{-1} a_{m-1} v_{m-1} \Rightarrow S = B \setminus \{v_m\} \text{ es sistm. gen.}$$

Lo que es una contradicción porque entonces B no es minimal.

Proposición

Sea $B = \{v_1, \dots, v_n\}$, este es base si y sólo si es un sistema linealmente independiente **maximal**, esto es, que cualquier conjunto que lo contiene no es linealmente independiente.

Teorema (Existencia de bases)

Sea V un espacio vectorial finitamente generado, entonces tiene alguna base.

Demostración:

Como V es finitamente generado, entonces: $\exists S = \{v_1, \dots, v_m\}$ sistema de generadores de V y entonces $\exists S' \subseteq S : S'$ es sistema de generadores minimal, por lo que S' es base. Esto es así porque si a un sistema de generadores se le puede quitar un elemento, entonces no es minimal y puede encontrarse en él uno que sí lo sea, entonces dicho sistema minimal ya es base.

Lema (de Intercambio)

Sea V un espacio finitamente generado, $L = \{v_1, \dots, v_s\}$ un conjunto linealmente independiente y $S = \{w_1, \dots, w_t\}$ un sistema de generadores de V , entonces:

1. $|L| \leq |S|$
2. $\exists i_1, \dots, i_s : \{v_1, \dots, v_s\} \cup S \setminus \{w_{i_1}, \dots, w_{i_s}\}$ es un sistema de generadores.

es decir, que se pueden intercambiar los elementos del conjunto linealmente independiente por cierto elementos del sistema generador.

Demostración:

Por inducción sobre el número de vectores linealmente independientes s , para $s = 1$ el conjunto formado por $L = \{v_1\}$ es l. i. (siempre que $v_1 \neq 0$). Como S es sistema de generadores, entonces v_1 se puede poner como combinación lineal de los elementos de S , es decir, $v_1 = a_1 w_1 + \dots + a_t w_t$ donde alguno de los $a_i \neq 0$.

Sin pérdida de generalidad, podemos suponer que es $a_1 \neq 0$ (pues de lo contrario bastaría con cambiar el orden de los vectores de S) y lo que implica que tiene inverso y entonces:

$$w_1 = a_1^{-1} v_1 - a_1^{-1} a_2 w_2 - \dots - a_1^{-1} a_t w_t \Rightarrow \{v_1, w_2, w_3, \dots, w_t\} \text{ es sist. gen.}$$

porque cualquier combinación lineal en la que participara w_1 puede ser modificada sustituyendo w_1 por su descomposición anterior que sólo depende de los vectores anteriores.

Demostrado para $s = 1$, suponemos que s es arbitrario, esto es, $L = \{v_1, \dots, v_s\}$ es l. i. y $S = \{w_1, \dots, w_t\}$, por la hipótesis de inducción sabemos que:

$$\{v_1, \dots, v_{s-1}\} \text{ es l.i.} \Rightarrow \{v_1, \dots, v_{s-1}, w_s, \dots, w_t\} \text{ es sistema de generadores}$$

Y de nuevo, podemos expresar v_s en términos de este nuevo sistema generador

$$v_s = a_1 v_1 + \dots + a_{s-1} v_{s-1} + a_s w_s + \dots + a_t w_t$$

si todos los términos a_s, \dots, a_t fueran cero, entonces $v_s = a_1 v_1 + \dots + a_{s-1} v_{s-1}$, pero esto entraría en contradicción con la independencia de L . Por tanto, como alguno de los a_s, \dots, a_t es no nulo, podemos suponer que $a_s \neq 0$ y entonces:

$$w_s = a_s^{-1} v_s - a_s^{-1} a_1 v_1 - \dots - a_s^{-1} a_{s-1} v_{s-1} - a_s^{-1} a_{s+1} v_{s+1} - \dots - a_s^{-1} a_t w_t$$

que por el mismo motivo de sustituir w_s por esta expresión en las combinaciones lineales donde interviniera demuestra que $\{v_1, \dots, v_{s-1}, v_s, w_{s+1}, \dots, w_t\}$ es sistema de generadores.

Proposición

Sea V un espacio vectorial finitamente generado y L un conjunto linealmente independiente, entonces L es finito.

Demostración:

Elegimos un sistema de generadores de s vectores $S = \{w_1, \dots, w_s\}$. Si L es linealmente independiente, por el lema de intercambio, el número de vectores que conforman L es menor que s , luego es finito.

Proposición

Si V es finitamente generado, entonces:

1. Todas las bases de V son finitas
2. Todas las bases de V tienen el mismo cardinal

Demostración:

1. Si B es base de V , entonces B es linealmente independiente \Rightarrow es finito.
2. Supongamos que B_1 y B_2 son ambas bases de V con un número de elementos s_1 y s_2 . Podemos decir que $s_1 \leq s_2$ por considerar B_1 linealmente independiente y B_2 sistema de generadores, pero del mismo modo, al revés, podemos decir que $s_2 \leq s_1$ haciendo las consideraciones al revés, en consecuencia: $s_1 = s_2$

Definición (Dimensión)

Sea V un espacio vectorial finitamente generado, definimos **la dimensión de V** como:

$$\dim V = |B|$$

donde B es una base cualquiera de V .

Ejemplo:

En algunos de los espacios vectoriales que hemos estudiado, las bases asociadas son:

- $\dim K^2 = 2 \Rightarrow B = \{e_1 = (1, 0), e_2 = (0, 1)\}$
- $\dim K^n = n \Rightarrow B = \{e_1, \dots, e_n\}$

$$\blacksquare \dim Mat_{m \times n}(K) = mn \Rightarrow B = \{E_{ij}\} \text{ donde } E_{ij} = (a_{rs}) \begin{cases} 1 & \text{si } (i,j) = (r,s) \\ 0 & \text{si } (i,j) \neq (r,s) \end{cases}$$

Observación:

Como todo espacio vectorial posee una base y toda base un cardinal (que es el mismo en todas), la dimensión de un espacio vectorial indica el número máximo de vectores linealmente independientes que puede haber y el mínimo número de vectores que lo generan.

Teorema (de extensión de la base)

Sea V un espacio vectorial finitamente generado y $W < V$ un subespacio del mismo, entonces:

1. W es finitamente generado
2. $\dim W \leq \dim V$
3. Si B_w es base de W , entonces $\exists B_v$ base de V tal que: $B_w \subseteq B_v$
4. $W = V \Leftrightarrow \dim W = \dim V$

Demostración:

1. Si L_w linealmente independiente en W , entonces también es linealmente independiente en V y, como V es finitamente generado, entonces $|L_w| \leq \dim V = |B_v|$. Por tanto, L_w es finito, luego $L_w \subset L \subset W$: L es conjunto linealmente independiente maximal.
2. Se demuestra junto con la 2 al demostrar que L_w es finito.
3. $\dim W > \dim V \Rightarrow |B_w| > |B_v| \Rightarrow B_v$ no es capaz de generar W pero sí V , pero como $W \subset V \Rightarrow \#$
4. $B_w = \{w_1, \dots, w_m\}$ base de W y $B = \{v_1, \dots, v_n\}$ base de V , sabemos que $m \leq n$. Consideramos B_w como linealmente independiente en V y a B como sistema generador en V , salvo el orden podemos, por el **lema de intercambio**, intercambiar vectores entre ambos conjuntos así: $\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$ el conjunto es linealmente independiente y de nuevo sistema de generadores, por lo que es base de V .

Ejemplo:

Consideremos K^2 como espacio vectorial y el vector $w_1 = (2, -3)$. El conjunto $\{w_1\}$ es linealmente independiente y es base del subespacio que genera (la envoltura lineal de $\{w_1\}$) que denotaremos por W_1 .

El teorema anterior nos dice que podemos extenderla a una base mayor, por ejemplo $\{e_1, (2, -3)\}$ o también $\{(2, -3), e_2\}$ y que queda demostrado que son bases de K^2 .

Observación:

El razonamiento del ejemplo anterior, nos permite enunciar la siguiente propiedad, sea $\dim V = n$ entonces:

- L es independiente y $|L| = n \Rightarrow L$ es base.
- S es sistema de generadores y $|S| = n \Rightarrow S$ es base.

en el primer caso porque además de linealmente independientes añadir cualquier vector lo hace dependiente, lo que implica que es sistema generador y en el segundo caso porque si alguno fuera dependiente de los demás el conjunto S menos ese vector también sería generador.

Teorema (de la Fórmula de Grassmann)

Sea V un espacio vectorial finitamente generado y $W_1, W_2 < V$ subespacios vectoriales, entonces se cumple:

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

Demostración:

Vamos a utilizar la notación $m_1 = \dim W_1$, $m_2 = \dim W_2$ y $p = \dim(W_1 \cap W_2)$ con la que observamos que $p \leq m_1$ y $p \leq m_2$.

Tomemos una base de la intersección $B_{W_1 \cap W_2} = \{v_1, \dots, v_p\}$, por el Teorema de extensión, sabemos que existen

$$B_{W_1} = \{v_1, \dots, v_p, w_{p+1}, \dots, w_{m_1}\} \quad B_{W_2} = \{v_1, \dots, v_p, w'_{p+1}, \dots, w'_{m_2}\}$$

Por lo tanto, para probar el resultado basta con probar que $B = \{v_1, \dots, v_p, w_{p+1}, \dots, w_{m_1}, w'_{p+1}, \dots, w'_{m_2}\}$ es base de $W_1 + W_2$ porque se puede apreciar que el número de elementos es $n + (n - p) + (m - p) = n + m - p$, justamente lo que dice la definición:

$$a_1 v_1 + \dots + a_p v_p + a_{p+1} w_{p+1} + \dots + a_{m_1} w_{m_1} + a'_{p+1} w'_{p+1} + \dots + a'_{m_2} w'_{m_2} = 0 \Leftrightarrow$$

$$\Leftrightarrow a_1 v_1 + \dots + a_p v_p + a_{p+1} w_{p+1} + \dots + a_{m_1} w_{m_1} = -(a'_{p+1} w'_{p+1} + \dots + a'_{m_2} w'_{m_2})$$

La primera parte pertenece a W_1 y la segunda a W_2 , pero como ambas son iguales sabemos que ambas pertenecen a $W_1 \cap W_2$. Esto último nos permite expresar dichos vectores en términos de la base de la intersección:

$$a'_{p+1} w'_{p+1} + \dots + a'_{m_2} w'_{m_2} = b_1 v_1 + \dots + b_p v_p \Leftrightarrow -b_1 v_1 - \dots - b_p v_p + a'_{p+1} w'_{p+1} + \dots + a'_{m_2} w'_{m_2} = 0$$

Precisamente estos vectores conforman la base B_{W_2} , así que son linealmente independientes, luego $b_1 = \dots = b_p = a'_{p+1} = \dots = a'_{m_2} = 0$ y esto implica que:

$$a_1 v_1 + \dots + a_p v_p + a_{p+1} w_{p+1} + \dots + a_{m_1} w_{m_1} = 0$$

De nuevo, estos vectores son la base B_{W_1} , así que su independencia lineal implica que $a_1 = \dots = a_p = \dots = a_{m_1} = 0$, lo que demuestra la independencia lineal de los vectores de la base B .

Demostrar que es sistema de generadores es trivial, puesto que cualquier vector $w_1 + w_2 = v \in W_1 + W_2$ con $w_1 \in W_1$ y $w_2 \in W_2$ o $v \in W_1$ o $v \in W_2$ se puede expresar en términos de la nueva base:

$$\begin{aligned} u_1 + u_2 &= \lambda_1 v_1 + \dots + \lambda_p v_p + \lambda_{p+1} w_{p+1} + \dots + \lambda_{m_1} w_{m_1} + \mu_1 v_1 + \dots + \mu_p v_p + \mu_{p+1} w'_{p+1} + \dots + \mu_{m_2} w'_{m_2} = \\ &= (\lambda_1 + \mu_1) v_1 + \dots + (\lambda_p + \mu_p) v_p + \lambda_{p+1} w_{p+1} + \dots + \lambda_{m_1} w_{m_1} + \mu_{p+1} w'_{p+1} + \dots + \mu_{m_2} w'_{m_2} \end{aligned}$$

Definición (Suma Directa)

Sea V un espacio vectorial y $W_1, W_2 < V$ dos subespacios vectoriales, decimos que V es **suma directa de W_1 y W_2** si y sólo si se verifica:

$$\forall v \in V : \exists! w_1 \in W_1, w_2 \in W_2 : v = w_1 + w_2$$

y lo expresamos como $V = W_1 \oplus W_2$.

Proposición (Caracterización de la Suma Directa)

Sea V un espacio vectorial y $W_1, W_2 < V$ dos subespacios vectoriales suyos, V es suma directa de W_1 y W_2 si y sólo si:

$$V = W_1 + W_2 \text{ y } W_1 \cap W_2 = \{0\}$$

Demostración:

■ \Leftarrow :

Supongamos que $v = w_1 + w_2 \in W_1 \cap W_2$, entonces:

$$w_1 + w_2 = w'_1 + w'_2 : w_1, w'_1 \in W_1 \text{ y } w_2, w'_2 \in W_2 \Rightarrow w_1 - w'_1 = w'_2 - w_2 \in W_1 \cap W_2 = 0 \Rightarrow \begin{cases} w_1 = w'_1 \\ w_2 = w'_2 \end{cases}$$

■ \Rightarrow :

Tomemos $w' \in W_1 \cap W_2$ no nulo, entonces:

$$v = w_1 + w_2 \Rightarrow v - w' = w_1 + w_2 - w' \Rightarrow v = \underbrace{w_1 + w'}_{\in W_1} + \underbrace{w_2 - w'}_{\in W_2}$$

Pero como la descomposición es única, entonces:

$$\begin{cases} w_1 = w_1 + w' \\ w_2 = w_2 - w' \end{cases} \Rightarrow w' = 0$$

El concepto de suma directa se entiende bien pensando en el espacio vectorial canónico K^2 . En este caso, si tomamos como $W_1 = K \times \{0\}$ (el eje de las x) y $W_2 = \{0\} \times K$ (el eje de las y) es sencillo ver que cualquier vector del plano es suma de dos vectores de ambos subespacios ¡porque son los ejes! y la intersección de ambos trivialmente es el $(0, 0)$ que es el 0 de K^2 .

Observación:

La caracterización final de la suma directa se complica cuando el número de sumandos aumenta. Si queremos pensar en la suma directa de tres tenemos que comprobar que se verifican las siguientes condiciones:

$$\begin{cases} V = W_1 + W_2 + W_3 \\ W_1 \cap W_2 = \{0\} \\ (W_1 + W_2) \cap W_3 = \{0\} \end{cases}$$

es decir, en el fondo es comprobar que todos los vectores son suma única de los vectores de cada subespacio.

Ejemplo:

Sea $V = Mat_2(\mathbb{R})$ el espacio vectorial de las matrices cuadradas de dos dimensiones, una posible base es:

$$B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Los conjuntos $S_2(\mathbb{R}) := \{A \in V : A = A^t\}$ y $A_2(\mathbb{R}) := \{A \in V : A = -A^t\}$ que denotan al conjunto de las matrices simétricas y antisimétricas, respectivamente son subespacios vectoriales de V porque son cerrados³ para la suma y el producto por un escalar:

$$\forall A, B \in A_2(\mathbb{R}) : A + B = -A^t - B^t = -(A^t + B^t) = -(A + B)^t \Rightarrow A + B \in A_2(\mathbb{R})$$

$$\forall a \in \mathbb{R}, \forall A \in A_2(\mathbb{R}) : a \cdot A = a \cdot (-A^t) = -a \cdot A^t = -(a \cdot A^t) \Rightarrow a \cdot A \in A_2(\mathbb{R})$$

Si determinamos las dimensiones de cada uno de estos subespacios:

$$A \in S_2(\mathbb{R}) \Rightarrow A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = a_{11} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \dim S_2(\mathbb{R}) = 3$$

$$A \in A_2(\mathbb{R}) \Rightarrow A = \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} = a_{12} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Rightarrow \dim A_2(\mathbb{R}) = 1$$

y vemos la intersección de ambos subespacios:

$$A \in A_2(\mathbb{R}) \cap S_2(\mathbb{R}) \Rightarrow \begin{cases} A = A^t \\ A = -A^t \end{cases} \Rightarrow A^t = -A^t \Leftrightarrow A^t = -A^t = 0 \Rightarrow A = 0$$

³Para las matrices de $S_2(\mathbb{R})$ la demostración es completamente análoga.

podemos darnos cuenta de que se cumplen todas las hipótesis de la caracterización de la suma directa. Por ello, la suma supone el total del espacio vectorial definido:

$$\dim(S_2 + A_2) = 3 + 1 - 0 = 4 = \dim \text{Mat}_2(\mathbb{R}) \Rightarrow V = S_2 \oplus A_2$$

Lo que demuestra que cualquier matriz cuadrada de dos dimensiones se puede poner como una única de dos matrices: una simétrica y otra antisimétrica.

$$A = \frac{A + A^t}{2} + \frac{A - A^t}{2}$$

puesto que $(A + A^t)^t = A^t + A^{tt} = A^t + A$ luego $A + A^t$ es simétrica y $(A - A^t)^t = A^t - A^{tt} = A^t - A$ luego $A - A^t$ es antisimétrica.

Coordenadas y Cambio de Base

En su momento vimos que un vector se puede expresar de forma única en términos de los vectores de una base vectorial.

$$v = a_1 v_1 + \cdots + a_n v_n$$

La relación que existe entre este vector y los de la base vectorial la marcan los coeficientes de la combinación lineal que tiene como resultado el vector inicial. Estos coeficientes son los que llamamos coordenadas y conocerlos permite trabajar con cualquier espacio vectorial como si se trata del espacio vectorial canónico en el que sólo escribimos (a_1, \dots, a_n) .

Definición (Coordenadas)

Sea $B = \{v_1, \dots, v_n\}$ una base de V , cada vector se escribe de modo único como:

$$\forall v \in V : v = x_1 v_1 + \cdots + x_n v_n : \exists! x_i \in K$$

definimos las **coordenadas de un vector** como los coeficientes x_i asociados a la combinación lineal de los vectores de la base que lo representa.

Ejemplo:

Sea $V = K^2$ y $B_c := \{e_1, e_2\}$ la base canónica, un vector cualquiera $v = (a_1, a_2)$ se puede expresar como:

$$v = (a_1, a_2) = a_1(1, 0) + a_2(0, 1) = a_1 e_1 + a_2 e_2$$

lo que implica que sus coordenadas con respecto a la base canónica son precisamente (a_1, a_2) . Sin embargo, si tomamos como base $B = \{e_1, e_1 + e_2\}$, el mismo vector se expresa ahora como:

$$v = (a_1, a_2) = a_1 e_1 + a_2 e_2 = a_1 e_1 - a_2 e_1 + a_2 e_1 + a_2 e_2 = (a_1 - a_2) e_1 + a_2 (e_1 + e_2)$$

lo que implica que respecto de la nueva base las coordenadas del mismo vector sean $(a_1 - a_2, a_2)$.

El ejemplo anterior pone de manifiesto que las coordenadas pueden cambiar en función de la base que escojamos para darlas. El cambio de base anterior era muy sencillo de calcular y no revestía mucha dificultad a la hora de especificar cuáles eran las nuevas coordenadas de v , pero ¿y si el cambio no es tan sencillo? ¿Hay alguna forma “canónica” de poder cambiar de base?

Proposición (Cambio de base)

Sea V un espacio vectorial y $B' = \{v'_1, \dots, v'_n\}$ y $B = \{v_1, \dots, v_n\}$ dos bases del mismo, entonces la relación entre las coordenadas de ambos viene dada por:

$$\underbrace{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}}_{\text{Coord. en } B'} = \underbrace{\begin{pmatrix} \overbrace{a_{11}}^{v_1 \text{ en } B'} & \cdots & \overbrace{a_{1n}}^{v_n \text{ en } B'} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{\text{Matriz de paso}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Coord. en } B}$$

donde las columnas son los vectores de la base de partida expresados en términos de la base de llegada.

Demostración:

En primer lugar, como ambas son bases del espacio V , podemos expresar en términos de ellas un vector v cualquiera:

$$v = x_1 v_1 + \cdots + x_n v_n \text{ en } B \qquad v = x'_1 v'_1 + \cdots + x'_n v'_n \text{ en } B'$$

Por tanto, como ambas expresiones hacen referencia al vector v podemos igualarlas:

$$\begin{aligned} x'_1 v'_1 + \cdots + x'_n v'_n &= x_1 v_1 + \cdots + x_n v_n \\ &= x_1 (a_{11} v'_1 + \cdots + a_{n1} v'_n) + \cdots + x_n (a_{1n} v'_1 + \cdots + a_{nn} v'_n) \\ &= (a_{11} x_1 + \cdots + a_{n1} x_n) v'_1 + \cdots + (a_{1n} x_1 + \cdots + a_{nn} x_n) v'_n \end{aligned}$$

donde la segunda igualdad surge de expresar $v_i = a_{1i} v'_1 + \cdots + a_{ni} v'_n$, es decir, los vectores de la base de partida, en términos de los vectores de la base de llegada.

Si pasamos todo a uno de los lados desde la expresión final anterior, entonces tenemos una combinación lineal de vectores independientes que vale 0. Precisamente por dicha independencia los coeficientes asociados tienen que ser nulos, es decir:

$$\begin{cases} x'_1 - a_{11} x_1 + \cdots + a_{n1} x_n = 0 \\ \vdots \\ x'_n - a_{1n} x_1 + \cdots + a_{nn} x_n = 0 \end{cases} \Rightarrow \begin{cases} x'_1 = a_{11} x_1 + \cdots + a_{n1} x_n \\ \vdots \\ x'_n = a_{1n} x_1 + \cdots + a_{nn} x_n \end{cases}$$

Basta para obtener la fórmula del enunciado con escribir este sistema lineal en forma matricial, de modo que queda como:

$$\underbrace{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}}_{\text{Coord. en } B'} = \underbrace{\begin{pmatrix} \underbrace{v_1 \text{ en } B'}_{a_{11}} & \cdots & \underbrace{v_n \text{ en } B'}_{a_{1n}} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{\text{Matriz de paso}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Coord. en } B}$$

Todo el razonamiento previo puede hacerse en sentido inverso, esto es, pasando de la base B' a la base B y entonces se obtiene el mismo resultado.

Proposición

Sea V un espacio vectorial y B y B' dos bases del mismo, las matrices de paso de ambas

$$\underbrace{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}}_{\text{Coord. en } B'} = M(B, B') \cdot \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Coord. en } B} \qquad \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\text{Coord. en } B} = M(B', B) \cdot \underbrace{\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}}_{\text{Coord. en } B'}$$

son inversas entre sí, es decir, $M(B, B') = M(B', B)^{-1}$.

Demostración:

Si denotamos por $P = M(B, B')$ y por $Q = M(B', B)$ y tenemos en cuenta la forma en que relacionan las coordenadas y que hemos visto anteriormente, entonces:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = P \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \cdot Q \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \Leftrightarrow 0 = (PQ - I) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \Leftrightarrow PQ - I = 0 \Leftrightarrow PQ = I$$

lo que quiere decir que P y Q son inversas la una de la otra.

Observación:

El resultado anterior expone que el conjunto de matrices cuadradas inversibles se puede identificar con el conjunto de posibles cambios de bases en un espacio vectorial.

Por este motivo, el conjunto al que $Mat_n(\mathbb{R})^x$ que hace referencia a las matrices cuadradas inversibles también se llama **grupo lineal** de n dimensión sobre (K) y se denota por $GL_n(K)$.

Definición (Característica de un Cuerpo)

Sea K un cuerpo, definimos su característica como el valor $p \in K$ que verifica que:

$$\chi(K) = \begin{cases} 0 & \text{si } 1_k \cdot n \neq 0 : \forall n \in \mathbb{N} \\ p > 0 & \text{si } p \cdot 1_k = 0 : 0 < n < p : n \cdot 1_k \neq 0 \end{cases}$$

Es decir, que el 0_k es el 0 cuando todos los múltiplos de 1 nos se anulan, pero cuando hay un valor para el que se anula y todos los anteriores no, entonces ese p es la característica del cuerpo.

Ejemplo:

Sea $V = K^2$ y $B' = \{e_1 + 2e_2, 2e_1 + e_2\}$ una base del mismo (puesto que tiene dos vectores y es trivial comprobar su independencia lineal, entonces las matrices paso entre B y B_c son:

$$M(B', B_c) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad M(B_c, B') = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}^{-1}$$

por lo visto en las proposiciones anteriores.

Sin embargo, invertir una matriz no es en general algo agradable y sencillo de hacer en dimensiones altas, luego podemos optar por expresar los vectores de la base canónica en términos de los vectores de la base B del enunciado:

$$\begin{cases} v'_1 = e_1 + 2e_2 \\ v'_2 = 2e_1 + e_2 \end{cases} \xrightarrow{E_1 = E_1 - 2E_2} \begin{cases} v'_1 - 2v'_2 = e_1 - 4e_1 = -3e_1 \\ e_2 = v'_2 - 2e_1 \end{cases}$$

De la primera ecuación del sistema sabemos que $e_1 = (\frac{-1}{3}, \frac{2}{3})_{B'}$ así que simplemente susituyendo en la segunda este valor de e_1 , tenemos que:

$$v'_2 - 2 \left(\frac{-1}{3}v'_1 + \frac{2}{3}v'_2 \right) \Rightarrow e_2 = \left(\frac{2}{3}, \frac{-1}{3} \right)_{B'}$$

Luego, sin necesidad de invertir la matriz "sencilla" que ya conocíamos, hemos sido capaces de encontrar la matriz de cambio de base que es

$$M(B_c, B') = \begin{pmatrix} \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1}{3} \end{pmatrix}$$

Por tanto, si ahora queremos expresar el vector $v = (7, -6)$ en términos de la nueva base B , basta con multiplicarlo por la matriz de cambio de base

$$v = (7, -6)_{B_c} \in \mathbb{R}^2 \Rightarrow v = \begin{pmatrix} \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1}{3} \end{pmatrix} \cdot \begin{pmatrix} 7 \\ -6 \end{pmatrix} = \begin{pmatrix} -19/3 \\ -20/3 \end{pmatrix} \Rightarrow v = \left(\frac{-19}{3}, \frac{-20}{3} \right)_{B'}$$

esto evita tener que estar continuamente calculando, como lo hemos hecho en el sistema anterior, las coordenadas del vector v en términos de B a mano.

Espacio vectorial cociente

Los conjuntos cocientes, en general, son un concepto abstracto complicado de entender al principio, pero con muy buenas propiedades. Estos permiten caracterizar y clasificar los elementos del conjunto inicial en un conjunto de subgrupos al que pertenecen todos los elementos relacionados entre sí a través de una relación de equivalencia.

Proposición

Sea $W < V$ un subespacio del espacio V , la siguiente relación:

$$v \equiv_W v' \Leftrightarrow v - v' \in W : v, v' \in V$$

es una relación de equivalencia que divide al espacio vectorial en particiones disjuntas dos a dos.

Demostración:

- Reflexiva: $v - v = 0_v = 0_w \in W \Rightarrow v \equiv_W v$
- Simétrica: $v \equiv_W v' \Rightarrow v - v' \in W \Rightarrow v' - v = -(v - v') \in W \Rightarrow v' \equiv_W v$
- Transitiva: $v \equiv_W v' \text{ y } v' \equiv_W v'' \Rightarrow v - v' \in W \text{ y } v' - v'' \in W \Rightarrow v - v'' = v - v' + v' - v'' \in W \Rightarrow v \equiv_W v''$

Definición (Clase Lateral)

Denotamos $[v]_w = \bar{v}$ y veamos que $\bar{v} = v + W = \{v + w : w \in W\}$:

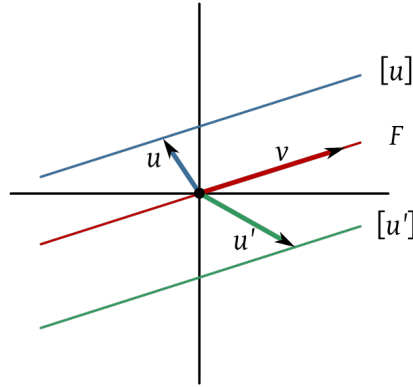
- Primero vemos que $\bar{v} \subset v + W$, consideremos:

$$v' \in \bar{v} \Rightarrow v' \equiv_W v \Rightarrow v' - v = w \in W \Rightarrow v' = v + w \in v + W$$

- Ahora vemos que $v + W \subset \bar{v}$, consideramos:

$$v' \in v + W \Rightarrow v' = v + w : w \in W \Rightarrow v' - v = w \in W \Rightarrow v' \equiv_W v \Rightarrow v' \in \bar{v}$$

A este espacio que hemos definido como \bar{v} se le conoce como **clase lateral**.



Definición (Conjunto Cociente)

Se define el **conjunto cociente** y se denota por $V/W = \{[v]_W : v \in V\}$ al conjunto de clases laterales resultantes de la relación de equivalencia generada por W y forma un **K-ESPACIO VECTORIAL** las siguientes propiedades:

- $\bar{v} + \bar{v'} = \overline{v + v'}$
- $a \cdot \bar{v} = \overline{a \cdot v}$

Observación:

Es notable destacar los siguientes aspectos:

- $0_{V/W} = [0]_W$ y que $-\bar{v} = \overline{-v}$.

- Hay que comprobar que la suma y el producto están bien definidos, es decir, que no dependen del representante escogido para operar:

- Sean $\bar{v}_1 = \bar{v}_2$ representantes de una clase y $\bar{v}'_1 = \bar{v}'_2$ otros de otra distinta, entonces:

$$\overline{v_1 + v'_1} = \overline{v_2 + v'_2} \Leftrightarrow (v_1 + v'_1) - (v_2 + v'_2) = 0 \Leftrightarrow \underbrace{(v_1 - v_2)}_{\in W} + \underbrace{(v'_1 - v'_2)}_{\in W} \in W$$

- El producto se demuestra de forma análoga.

Proposición

Sea V un espacio vectorial y $W < V$ un subespacio vectorial del mismo. Si denotamos $\dim V = n$ y $\dim W = m$, entonces $\dim V/W = n - m$.

De hecho, si $B = \{v_1, \dots, v_m\}$ es una base de W y $B' = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ es una ampliación suya para V , entonces $B'' = \{\bar{v}_{m+1}, \dots, \bar{v}_n\}$ es base de V/W .

Demostración:

Elegimos una base cualquiera de W : $B_W = \{w_1, \dots, w_m\}$ y ampliamos esta base para V : $B = \{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$, es decir, hemos añadido $n - m$ vectores a la base de B_W para ampliarla a B . Ahora probamos que $\{\bar{v}_{m+1}, \dots, \bar{v}_n\}$ es base de V/W .

- Linealmente independiente:

$$a_{m+1}\bar{v}_{m+1} + \dots + a_n\bar{v}_n = 0 \Leftrightarrow \overline{a_{m+1}v_{m+1} + \dots + a_nv_n} = 0 \Leftrightarrow a_{m+1}v_{m+1} + \dots + a_nv_n - 0 \in W$$

Como estos vectores pertenecen a W , entonces se podrán expresar en función de su base:

$$a_{m+1}v_{m+1} + \dots + a_nv_n = a_1w_1 + \dots + a_mw_m \Rightarrow a_{m+1}v_{m+1} + \dots + a_nv_n - a_1w_1 - \dots - a_mw_m = 0 \stackrel{B \text{ es base}}{\Rightarrow} \text{l. ind.}$$

- Sistema de generadores:

$$\forall v \in V : \bar{v} = \overline{a_1w_1 + \dots + a_mw_m + a_{m+1}v_{m+1} + \dots + a_nv_n} = \overline{a_1w_1 + \dots + a_mw_m} + a_{m+1}\bar{v}_{m+1} + \dots + a_n\bar{v}_n =$$

Como $a_1w_1 + \dots + a_mw_m \in W \Rightarrow \overline{a_1w_1 + \dots + a_mw_m} = 0$, luego tenemos que:

$$= a_{m+1}\bar{v}_{m+1} + \dots + a_n\bar{v}_n$$

Luego queda demostrado que para cualquier \bar{v} , este se puede expresar como combinación lineal de los vectores añadidos para ampliar la base, luego es sistema de generadores.

Ejemplo:

En \mathbb{R}^5 consideramos W como el subespacio⁴ de los vectores cuyas coordenadas se ajustan al siguiente

$$\text{sistema: } W = \begin{cases} x_1 + x_2 - x_5 = 0 \\ 2x_1 - x_4 = 0 \end{cases}, \text{ vamos a ver que } \dim W = 3 \text{ y vamos a calcular la base de } B_{V/W}.$$

Podemos ver que el sistema de ecuaciones forma un subespacio vectorial de V por ser un sistema homogéneo del que son vectores las soluciones de ese sistema. Su dimensión es el número de incógnitas menos el rango de la matriz de coeficientes, es decir, $\dim W = 5 - 2 = 3$. Esto es así porque según vimos en Bachillerato, es necesario que haya 3 incógnitas parametrizadas que en este caso van a ser x_3, x_4, x_5 y las otras toman valores conforme a ellas, así que esencialmente los vectores de W dependen de 3 valores, de ahí la dimensión 3.

Para hallar los vectores de la base de W , vemos que la solución del sistema es: $(x_1, x_2, x_3, x_4, x_5) = (\frac{x_4}{2}, x_5 - \frac{x_4}{2}, x_3, x_4, x_5)$. Por lo que cualquier vector solución se puede escribir como:

$$x_3(0, 0, 1, 0, 0) + x_4(1, -1, 0, 1, 0) + x_5(0, 1, 0, 0, 1)$$

⁴Más adelante se demostrará que esto es un subespacio

Otra forma es dar valores de vectores que sean linealmente independientes entre los tres parámetros, p. ej., $(x_3, x_4, x_5) = (1, 0, 0), (0, 1, 0), (0, 0, 1)$ y hallar los valores de x_1 y x_2 para completar con las dos variables que faltan cada uno de los vectores.

Con lo cual, para encontrar la base ahora de V/W , vemos que vectores es necesario añadir a la hallada para que sea base de V , un caso sencillo para ampliar la base es añadir vectores de la base canónica de \mathbb{R}^5 :

$$B = \{(0, 0, 1, 0, 0), (1, -1, 0, 1, 0), (0, 1, 0, 0, 1), v_1, v_2\} = \\ = \{(0, 0, 1, 0, 0), (1, -1, 0, 1, 0), (0, 1, 0, 0, 1), e_1 = (1, 0, 0, 0, 0), e_2 = (0, 1, 0, 0, 0)\}$$

Con lo cual, e_1 y e_2 forman la base de V/W .

APLICACIONES LINEALES

Al inicio del capítulo comentamos que íbamos a ver las características la estructura de espacio vectorial, pero que también íbamos a ver las características de los morfismos entre estructuras de este tipo.

Por ello, esta sección está dedicada al estudio de las propiedades que tienen estas transformaciones: las aplicaciones lineales. Es muy común y habitual en matemáticas tener que definir o utilizar aplicaciones (funciones) entre objetos matemáticos. Si además añadimos que hemos comentado que la estructura de espacio vectorial está muy presente en multitud de elementos en las matemáticas, con más razón será necesario sacar provecho y ventajas de todas las propiedades que podamos descubrir de las mismas.

Definición

Sea $\phi : V \longrightarrow V'$ una aplicación entre dos espacios vectoriales, decimos que ϕ es una **aplicación lineal** si y sólo si conserva la suma y el producto por un escalar, es decir, si verifica que:

- $\forall v_1, v_2 \in V : \phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$
- $\forall a \in K, \forall v \in V : \phi(a \cdot v) = a \cdot \phi(v)$

Al conjunto de aplicaciones lineales entre ambos espacios se le denota⁵ por $\text{Hom}(V, V')$, de homomorfismos de V en V' .

Ejemplo:

- Sea $\phi : K^3 \longrightarrow K^2$ y $\phi(x, y, z) = (x - y, x + y + z)$:

$$\phi[(x, y, z) + (x', y', z')] = \phi(x + x', y + y', z + z') = (x + x' - y - y', x + x' + y + y' + z + z') = \\ = (x - y, x + y + z) + (x' - y', x' + y' + z') = \phi(x, y, z) + \phi(x', y', z')$$
- Sea $\phi : V \longrightarrow V$ y sea ϕ K -lineal, entonces decimos que es un endomorfismo. Por ejemplo, $\phi : V \longrightarrow V$ siendo $\text{id}_V = \phi(v) = v$, es un endomorfismo.
- Sea $\phi : V \longrightarrow V$ y sea $\phi(v) = r \cdot v$ donde $r \in K$, también es un isomorfismo. Concretamente a esto se le conoce como **homotecia**⁶. ϕ cumple que $\phi = \phi^{-1}$ con lo que es fácil demostrar que es una biyección y por extensión un isomorfismo.
- Si tenemos $W < V$, podemos definir V/W y ahora $\phi : V \longrightarrow V/W$ siendo $\phi(v) = \bar{v}$, entonces ϕ es aplicación lineal.

⁵Si el espacio de llegada y de salida son el mismo, es decir, $V = V'$, entonces se denota por $\text{End}(V)$

⁶Y en este caso además es un isomorfismo dentro del propio V

- Si tenemos que $V = W \oplus U$ y $\phi : V \rightarrow V$, entonces $v = w + u$ de modo único y la aplicación $\phi(v) = w$ es lineal. A esta aplicación se la conoce como **proyección** de v con base W y dirección U .

- Vamos a ver que ocurre con los vectores de W :

$$v \in W \Leftrightarrow \phi(v) = v$$

Esto es así porque:

$$\begin{aligned} v \in W \Rightarrow v &= \underbrace{v}_{\in W} + \underbrace{0}_{\in U} \Rightarrow \phi(v) = v \\ \phi(v) = v \Rightarrow v &= \underbrace{w}_{=v} + u \Rightarrow u = 0 \Rightarrow v \in W \end{aligned}$$

- Vamos a ver que ocurre con los vectores de U :

$$v \in U \Leftrightarrow \phi(v) = 0$$

Esto es así porque:

$$\begin{aligned} v \in U \Rightarrow v &= \underbrace{0}_{\in W} + \underbrace{v}_{\in U} \Rightarrow \phi(v) = 0 \\ \phi(v) = 0 \Rightarrow v &= \underbrace{w}_{=0} + u \in U \end{aligned}$$

- Si tenemos que $V = W \oplus U$, entonces $v = w + u$ de modo único y la aplicación $\phi(v) = w - u$ es lineal. A esta aplicación se la conoce como **simetría** de v con base W y dirección U .

- Vamos a ver que ocurre con los vectores de W :

$$v \in W \Leftrightarrow \phi(v) = v$$

Esto es así porque:

$$\begin{aligned} v \in W \Rightarrow v &= \underbrace{v}_{\in W} + \underbrace{0}_{\in U} \Rightarrow \phi(v) = v - 0 = v \\ \phi(v) = v \Rightarrow w - u &= v = \underbrace{w}_{=v} + u \Rightarrow 2u = 0 \Rightarrow v = w \in W \end{aligned}$$

- Vamos a ver que ocurre con los vectores de U :

$$v \in U \Leftrightarrow \phi(v) = -v = -1_k \cdot v$$

Esto es así porque:

$$\begin{aligned} v \in U \Rightarrow v &= \underbrace{0}_{\in W} + \underbrace{v}_{\in U} \Rightarrow \phi(v) = 0 - v = -v \\ \phi(v) = -v \Rightarrow w - u &= -v = -w - u \Rightarrow 2w = 0 \Rightarrow v = u \in U \end{aligned}$$

Definición (Imagen y Núcleo)

Sea $\phi : V \rightarrow V'$ una aplicación lineal, llamamos **imagen de ϕ** al subespacio vectorial⁷ del espacio de llegada cuyos vectores son imagen de algún vector del espacio de salida:

$$Im(\phi) = \{\phi(v) : v \in V\} \leq V'$$

Del mismo modo, llamamos **núcleo de ϕ** al subespacio vectorial⁷ del espacio de partida cuyos vectores se transforman en el vector nulo del espacio de llegada:

$$ker(\phi) = \{v \in V : \phi(v) = 0_{v'}\} \leq V$$

⁷La demostración de que es un subespacio vectorial del espacio de llegada es trivial por la condición de que las aplicaciones lineales conservan sumas y productos por escalares.

Proposición

Sea $\phi : V \longrightarrow V'$ una aplicación lineal, entonces:

1. ϕ es suprayectiva $\Leftrightarrow \text{Im}(\phi) = V'$
2. ϕ es inyectiva $\Leftrightarrow \ker(\phi) = \{0_v\}$

Demostración:

1. trivial

2. De la primera se deduce que:

- “ \Rightarrow ”: supongamos que es inyectiva

Vemos primero que: $\phi(0_v) = \phi(0_v + 0_v) = \phi(0_v) + \phi(0_v) \Leftrightarrow 0_{v'} = \phi(0_v) \Rightarrow 0_v \in \ker(\phi)$, es decir, que el cero va al cero.

$$v \in \ker(\phi) \Rightarrow \phi(v) = 0_{v'} \Rightarrow \phi(v) = \phi(0) \Rightarrow \forall v \in \ker(\phi) : v = 0$$

por la inyectividad de ϕ .

- “ \Leftarrow ”: supongamos cierto que el $\ker(\phi) = \{0_v\}$, entonces:

$$\phi(v_1) = \phi(v_2) \Rightarrow \phi(v_1 - v_2) = 0'_v \Rightarrow v_1 - v_2 \in \ker(\phi) \Rightarrow v_1 - v_2 = 0_v \Rightarrow v_1 = v_2$$

Observación:

La demostración del enunciado anterior para la inyectividad pone de manifiesto que en las aplicaciones lineales SIEMPRE $\phi(0) = 0$.

Proyecciones

Para las proyecciones resulta que la base es la imagen y la dirección es el núcleo de la misma, es decir, $\text{Im}(\phi) = W$ y $\ker(\phi) = U$, siendo $V = W \oplus U$. Además, en las proyecciones $\phi^2 = \phi$.

Demostración:

- $W = \text{Im}(\phi)$

- “ \subset ”:

$$w \in W \Rightarrow w = \phi(v) \Rightarrow \phi(w) = \phi(\phi(v)) = \phi(v) = w \Rightarrow w \in \text{Im}(\phi)$$

- “ \supset ”:

$$w \in \text{Im}(\phi) \Rightarrow w = \phi(v) \Rightarrow \phi(v) \in W$$

- $U = \ker(\phi)$:

$$v \in U \Leftrightarrow \phi(v) = 0 \Rightarrow v \in \ker(\phi)$$

- ϕ proyección $\Leftrightarrow \phi^2 = \phi$

- “ \Rightarrow ”:

$$\phi^2(v) = \phi^2(w + u) = \phi(\phi(w + u)) = \phi(w) = w = \phi(v)$$

Por ello se las llama idempotencias.

- “ \Leftarrow ”: tenemos que ser capaces de componer, sabiendo que $\phi^2 = \phi$, que $V = W \oplus U$ siendo $W = \{v \in V : \phi(v) = v\}$ base de la simetría y $U = \{v \in V : \phi(v) = 0\}$ dirección de la simetría. Es decir, partimos de la hipótesis y de la definición de ambos subespacios y hay que demostrar la suma directa.

$$v \in W \cap U \Leftrightarrow 0 = \phi(v) = v \Leftrightarrow v = 0 \Rightarrow W \cap U = \{0\}$$

$$\phi(v) \in W \wedge v - \phi(v) \in U \Rightarrow v = \phi(v) + v - \phi(v)$$

Esto es así porque $\phi(\phi(v)) = \phi^2(v) = \phi(v) \Rightarrow \phi(v) \in W$ y $\phi(v - \phi(v)) = \phi(v) - \phi^2(v) = 0 \Rightarrow 0 \in U$.

Solo falta ver que existe una proyección de V con base W y dirección U que como es única, esta debe coincidir necesariamente con ϕ :

$$\begin{aligned} \varphi : V &\longrightarrow V \text{ proyección con base } W \text{ y dirección } U \Rightarrow W = \{v \in V : \varphi(v) = v\} \wedge U = \{v \in V : \varphi(v) = 0\} \Rightarrow \\ &\Rightarrow v \in W \vee v \in U : \phi(v) = \varphi(v) \Rightarrow v \in V : v = w + u \Rightarrow \phi(v) = \phi(w) + \phi(u) = \varphi(w) + \varphi(u) = \varphi(v) : \forall v \in V \end{aligned}$$

Además, ϕ es suprayectiva:

$$\Leftrightarrow \text{Im}(\phi) = V \Leftrightarrow W = V \wedge U = \{0\} \Leftrightarrow id_v = \phi$$

Además, ϕ es inyectiva:

$$\Leftrightarrow \ker(\phi) = \{0\} \Leftrightarrow V = \{0\} \Leftrightarrow W = V \Leftrightarrow \phi = id_v$$

Simetrías

Para las simetrías resulta que $\text{Im}(\phi) = V$ y $\ker(\phi) = \{0\}$, por lo tanto todas las simetrías son inyectivas y suprayectivas y en consecuencia, isomorfismos. Además, en las simetrías $\phi = \phi^{-1}$.

Demostración:

$$\blacksquare V = \text{Im}(\phi)$$

$$v \in V \Leftrightarrow v = w + u \Rightarrow w - (-u) = \phi(w - u) \in \text{Im}(\phi)$$

$$\blacksquare \{0\} = \ker(\phi):$$

$$v \in \ker(\phi) \wedge v = w + u \Rightarrow w + u = v = 0 = \phi(v) = w - u \Rightarrow u = w \Rightarrow u = w \in U \cap W = \{0\} \Rightarrow u = w = 0 = v$$

$$\blacksquare \phi \text{ simetría} \Leftrightarrow \phi^{-1} = \phi$$

$$\bullet \text{ “}\Rightarrow\text{”}:$$

$$(\phi \circ \phi)(v) = \phi(\phi(v)) = \phi(w - u) = w - (-u) = w + u = v \Rightarrow \phi = id_v$$

Por ello se las llama involuciones

- “ \Leftarrow ”: tenemos que ser capaces de componer, sabiendo que $\phi^2 = id_V$, que $V = W \oplus U$ siendo $W = \{v \in V : \phi(v) = v\}$ base de la simetría y $U = \{v \in V : \phi(v) = -v\}$ dirección de la simetría. Es decir, partimos de la hipótesis y de la definición de ambos subespacios y hay que demostrar la suma directa.

$$v \in W \cap U \Leftrightarrow -v = \phi(v) = v \Leftrightarrow v = 0 \Rightarrow W \cap U = \{0\}$$

$$v + \phi(v) \in W \wedge v - \phi(v) \in U \Rightarrow v = \frac{v + \phi(v)}{2} + \frac{v - \phi(v)}{2}$$

Esto es así porque $\phi(v + \phi(v)) = \phi(v) + v \Rightarrow \phi(v) + v \in W$ y $\phi(v - \phi(v)) = \phi(v) - v \Rightarrow \phi(v) - v \in U$.

Solo falta ver que existe una simetría de V con base W y dirección U que como es única, esta debe coincidir necesariamente con ϕ :

$$\begin{aligned} \varphi : V &\longrightarrow V \text{ simetría con base } W \text{ y dirección } U \Rightarrow W = \{v \in V : \varphi(v) = v\} \wedge U = \{v \in V : \varphi(v) = -v\} \Rightarrow \\ &\Rightarrow v \in W \vee v \in U : \phi(v) = \varphi(v) \Rightarrow v \in V : v = w + u \Rightarrow \phi(v) = \phi(w) + \phi(u) = \varphi(w) + \varphi(u) = \varphi(v) : \forall v \in V \end{aligned}$$

Por último, la relación que existe, dada la misma base y la misma dirección, entre las simetrías y las proyecciones es: Sean ϕ simetría y φ la proyección asociada:

$$\begin{cases} \phi(v) = w - u \\ \varphi(v) = w \\ id(v) = w + u \end{cases} \Rightarrow (\phi + id)(v) = 2w = 2\varphi(v) \Rightarrow \frac{1}{2} \cdot (\phi + id) = \varphi$$

Proposición (Operaciones con aplicaciones lineales)

Sean $\phi_1, \phi_2 : V \longrightarrow V'$ aplicaciones K – lineales y las operaciones:

- $(\phi_1 + \phi_2)(v) \stackrel{def}{=} \phi_1(v) + \phi_2(v)$
- $(a \cdot \phi_1)(v) \stackrel{def}{=} a \cdot \phi_1(v)$

entonces la terna $(Hom(V, V'), +, \cdot)$ tiene estructura de espacio vectorial.

Observación:

Del mismo modo, se puede ver que $(\phi_2 \circ \phi_1)(v)$ es también una aplicación lineal si ambas aplicaciones lo son. Con lo que $(End(V), +, \circ)$ conforma un anillo conmutativo y unitario.

Teorema

Sean V y V' espacios vectoriales, $B := \{v_1, \dots, v_n\}$ una base de V y $v'_1, \dots, v'_n \in V'$ vectores arbitrarios de V' , entonces:

$$\exists! \phi : V \longrightarrow V' \text{ tal que } \phi(v_i) = v'_i : \forall i = 1, \dots, n$$

es decir, una aplicación lineal queda únicamente determinada por la imagen de los vectores de una base del espacio de partida.

Demostración:

- Existencia:

Cualquier $v = (a_1, \dots, a_n)_B \in V$ se expresa en términos de la base B como $v = \sum_{i=1}^n a_i v_i$, luego parece razonable que definamos la aplicación ϕ como $\phi(v) := \sum_{i=1}^n a_i v'_i$. Para probar que efectivamente esta definición cumple el enunciado, no hay más que demostrar su linealidad:

$$v + w = (a_1 + b_1, \dots, a_n + b_n)_B \Rightarrow \phi(v + w) = \sum_{i=1}^n (a_i + b_i) v'_i = \sum_{i=1}^n a_i v'_i + \sum_{i=1}^n b_i v'_i = \phi(v) + \phi(w)$$

$$a \cdot v = (a \cdot a_1, \dots, a \cdot a_n)_B \Rightarrow \phi(a \cdot v) = \sum_{i=1}^n (a \cdot a_i) v'_i = a \cdot \sum_{i=1}^n a_i v'_i = a \cdot \phi(v)$$

y que conserva la propiedad $\phi(v_i) = v'_i$ del enunciado:

$$v_i = (0, \dots, 1, \dots, 0)_B \Rightarrow \phi(v_i) = 1 \cdot v'_i = v'_i$$

- Unicidad:

Supongamos que hay dos y veamos que son la misma. Como hay dos, sabemos que $\phi(v_i) = \varphi(v_i) = v'_i$, luego:

$$\begin{aligned} \phi(v) &= \phi \left(\sum_{i=1}^n a_i v_i \right) = \sum_{i=1}^n a_i \phi(v_i) = \sum_{i=1}^n a_i v'_i \\ \varphi(v) &= \varphi \left(\sum_{i=1}^n a_i v_i \right) = \sum_{i=1}^n a_i \varphi(v_i) = \sum_{i=1}^n a_i v'_i \end{aligned}$$

Ejemplo:

Sean $V = K^2$ y $V' = K^3$, solo hay una aplicación lineal que cumpla que

$$\phi(e_1) = (2, 1, -3) \qquad \phi(e_2) = (0, 1, 0)$$

porque como cada vector está definido en términos de la base, la aplicación es lineal y los vectores de la base tienen determinada su imagen, entonces para un vector arbitrario ocurre que:

$$\phi(a_1, a_2) = \phi(a_1 e_1 + a_2 e_2) = a_1(2, 1, -3) + a_2(0, 1, 0) = (2a_1, a_1 + a_2, -3a_1)$$

Isomorfía de Espacios Vectoriales

Ya hemos visto que en general si conozco las coordenadas de un vector conozco prácticamente todo lo relevante para los resultados que hemos expuesto sobre cambios de base y aplicaciones lineales. Como trabajar con el vector simplemente conociendo sus coordenadas es lo mismo que trabajar en K^n , esto nos da una pista sobre que trabajar en un espacio vectorial cualquiera va a ser muy parecido, si no igual, a trabajar en K^n .

Cuando dos espacios vectoriales tienen la misma estructura, se dice que son isomorfos. Esto tiene grandes ventajas pues por muy complicado que sea el espacio en el que me muevo, si conozco un espacio vectorial que sea isomorfo a él y donde me maneje bien, entonces podré trabajar indistintamente en este último y aplicar los resultados del mismo en el primero (a través de una biyección entre los espacios adecuada).

Teorema

Dos espacios vectoriales son isomorfos si y solo si la dimensión de ambos es la misma.

$$V \simeq V' \Leftrightarrow \dim_k V = \dim_k V'$$

Demostración:

- “ \Rightarrow ”: vamos a ver que $B' = \{\phi(v_1), \dots, \phi(v_n)\}$ es base de V'

Sea $B = \{v_1, \dots, v_n\}$ base de V y $\phi: V \rightarrow V'$ un isomorfismo, entonces:

$$v' \in V' \stackrel{\text{isom}f}{\Rightarrow} v' = \phi(v) = \phi(a_1 v_1 + \dots + a_n v_n) = a_1 \phi(v_1) + \dots + a_n \phi(v_n) \Rightarrow \{\phi(v_1), \dots, \phi(v_n)\} \text{ es sist. gener.}$$

Ahora vemos que son linealmente independientes:

$$0 = a_1 \phi(v_1) + \dots + a_n \phi(v_n) = \phi(a_1 v_1 + \dots + a_n v_n) \Rightarrow a_1 v_1 + \dots + a_n v_n \in \ker(\phi) = \{0\} \Rightarrow 0 = a_1 v_1 + \dots + a_n v_n = 0$$

- “ \Leftarrow ”:

Sea $B = \{v_1, \dots, v_n\}$ base de V y $B' = \{v'_1, \dots, v'_n\}$ base de V' , entonces:

$$\exists! \phi: V \rightarrow V' : \phi(v_i) = v'_i : \forall i = 1, \dots, n$$

$$\exists! \varphi: V' \rightarrow V : \varphi(v'_i) = v_i : \forall i = 1, \dots, n$$

$$\varphi \circ \phi: V \rightarrow V \Rightarrow (\varphi \circ \phi)(v_i) = \varphi(v'_i) = v_i \Rightarrow \varphi \circ \phi = id_V$$

Pero del mismo modo se sigue que:

$$\phi \circ \varphi: V' \rightarrow V' \Rightarrow (\phi \circ \varphi)(v'_i) = \phi(v_i) = v'_i \Rightarrow \phi \circ \varphi = id_{V'}$$

Lo que implica que ambas son biyectivas y como son lineales, ambas son isomorfismos.

Corolario

Si la dimensión de un espacio cualquiera es n , entonces es isomorfo al producto cartesiano de n veces K .

$$\dim_k V = n \Rightarrow V \simeq K^n$$

Demostración:

$$\dim_k V = n = \dim_k K^n \Rightarrow V \simeq K^n$$

Sin embargo, no podemos identificar un isomorfismo canónico que verifique este teorema, solo podemos afirmar que existe porque para poder identificarlo es necesario definir una base de V' y cambia en función de la base elegida.

Teorema (de la dimensión)

Sea $\phi : V \longrightarrow V'$ una aplicación lineal y V un espacio vectorial con dimensión finita $\dim_k V = n$, entonces se cumple que:

1. $\ker(\phi)$ y $\text{Im}(\phi)$ son de dimensión finita.
2. $\dim_k \ker(\phi) + \dim_k \text{Im}(\phi) = n$

Demostración:

Como $\ker(\phi) < V$ y $\dim_k V = n$ entonces sabemos que $\dim_k \ker(\phi) \leq n$, por tanto, escogemos una base del $\ker(\phi)$ como $\{v_{r+1}, \dots, v_n\}$ y la extendemos a V de modo que la base de V es $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ con estos últimos $n - r$ vectores del $\ker(\phi)$.

Vamos a ver que $\{\phi(v_1), \dots, \phi(v_r)\}$ es una base de la $\text{Im}(\phi)$:

$$\begin{aligned} v' \in \text{Im}(\phi) &\Rightarrow v' = \phi(v) = \phi(a_1 v_1 + \dots + a_r v_r + a_{r+1} v_{r+1} + \dots + a_n v_n) = \\ &= a_1 \phi(v_1) + \dots + a_r \phi(v_r) + a_{r+1} v_{r+1} + \dots + a_n \phi(v_n) = a_1 \phi(v_1) + \dots + a_r \phi(v_r) \end{aligned}$$

Y ahora vemos que estos son linealmente independientes:

$$\begin{aligned} 0 &= a_1 \phi(v_1) + \dots + a_r \phi(v_r) = \phi(a_1 v_1 + \dots + a_r v_r) \Rightarrow a_1 v_1 + \dots + a_r v_r \in \ker(\phi) \Rightarrow a_1 v_1 + \dots + a_r v_r = \\ &= a_{r+1} v_{r+1} + \dots + a_n v_n \Leftrightarrow a_1 v_1 + \dots + a_r v_r - a_{r+1} v_{r+1} - \dots - a_n v_n = 0 \Rightarrow a_1 = \dots = a_r = a_{r+1} = \dots = a_n = 0 \end{aligned}$$

Teorema (Primer Teorema de Isomorfía)

Supongamos que $\phi : V \longrightarrow V'$ es una aplicación K -lineal se cumple que:

$$V / \ker(\phi) \simeq \text{Im}(\phi)$$

Este teorema vale siempre, sea cual sea la dimensión de V

Demostración:

Supongamos que V es de dimensión finita, entonces $\ker(\phi)$ y $\text{Im}(\phi)$ también son finitos por el teorema de la dimensión, además sabemos que:

$$\begin{aligned} \dim V &= \dim \ker(\phi) + \dim \text{Im}(\phi) \Leftrightarrow \dim V - \dim \ker(\phi) = \dim \text{Im}(\phi) \Leftrightarrow \dim V / \ker(\phi) = \dim \text{Im}(\phi) \\ &\Rightarrow V / \ker(\phi) \simeq \text{Im}(\phi) \end{aligned}$$

Vamos ahora a ver la aplicación lineal que relaciona ambos conjuntos y establece el isomorfismo:

$$\begin{aligned} \bar{\phi} : V / \ker(\phi) &\longrightarrow \text{Im}(\phi) \\ \bar{\phi}(\bar{v}) &= \phi(v) \end{aligned}$$

De este modo ya se cumple el isomorfismo, pero hay que comprobar que está bien definida probando que:

$$\bar{v} = \bar{w} \Rightarrow \phi(v) = \phi(w)$$

Y por último también que:

ϕ es lineal, inyectiva y suprayectiva

Teorema (Segundo Teorema de Isomorfía)

Consideramos $W, U < V$ y definimos los subespacios:

$$(W + U)/W \simeq U/(W \cap U)$$

Demostración:

Considerando W y U de dimensión finita, entonces $W + U$ es de dimensión finita y $W \cap U$ también lo es. Ahora vemos que:

$$\dim(W + U/W) = \dim(W + U) - \dim W \stackrel{\text{Grassman}}{=} \dim U - \dim(W \cap U) = \dim(U/W \cap U)$$

Matriz asociada a una aplicación lineal

El hecho de que una aplicación lineal quede determinada de modo único por su imagen a través de los vectores de una base sugiere que la información necesaria para conocer a una aplicación lineal es finita. Por tanto, existirá una representación de la misma que contenga esta información finita y que simplifique la forma de trabajar con ella.

Sea $\phi : V \longrightarrow V'$ una aplicación K -lineal, $B = \{v_1, \dots, v_n\}$ una base de V y $B' = \{v'_1, \dots, v'_m\}$ una de V' , sabemos que ϕ está determinada de modo único por su imagen a través de los vectores de la base B :

$$\begin{cases} \phi(v_1) = a_{11}v'_1 + \dots + a_{m1}v'_m \\ \vdots \\ \phi(v_n) = a_{1n}v'_1 + \dots + a_{mn}v'_m \end{cases}$$

Por otro lado, si calculamos la imagen de un vector v arbitrario tenemos que:

$$\begin{aligned} \phi(v) &= \phi(x_1v_1 + \dots + x_nv_n) \\ &= x_1\phi(v_1) + \dots + x_n\phi(v_n) \\ &= x_1(a_{11}v'_1 + \dots + a_{m1}v'_m) + \dots + x_n(a_{1n}v'_1 + \dots + a_{mn}v'_m) \\ &= \underbrace{(a_{11}x_1 + \dots + a_{1n}x_n)}_{x'_1}v'_1 + \dots + \underbrace{(a_{m1}x_1 + \dots + a_{mn}x_n)}_{x'_m}v'_m \end{aligned}$$

Por tanto, podemos observar que la relación con las coordenadas (x'_1, \dots, x'_m) se puede ver como un sistema lineal en el que interviene la información sobre la imagen de los vectores de la base inicial. Dicho sistema lineal, si se representa en su expresión matricial, tiene la forma: Con lo cual para transformar un vector en otro tenemos que:

$$\begin{pmatrix} \phi(v) \\ \underbrace{x'_1} \\ \vdots \\ x'_m \end{pmatrix} = \underbrace{\begin{pmatrix} \underbrace{\phi(v_1)}_{a_{11}} & \dots & \underbrace{\phi(v_n)}_{a_{1n}} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_{M_{B,B'}(\phi)} \cdot \begin{pmatrix} \underbrace{v}_{x_1} \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \phi(v) = M_{B,B'}(\phi) \cdot v^t$$

Teorema (de la Matriz Asociada)

Sea $\phi : V \longrightarrow V'$ una aplicación K -lineal, $B = \{v_1, \dots, v_n\}$ una base de V y $B' = \{v'_1, \dots, v'_m\}$ una de V' , podemos expresar la aplicación lineal como:

$$\forall v \in V : \phi(v) = M_{B,B'}(\phi) \cdot v^t$$

Proposición (Operaciones con matrices asociadas)

- Sean $\phi_1, \phi_2 : V \longrightarrow V'$ dos aplicaciones lineales y B y B' bases respectivas de los espacios V y V' , entonces:

$$M_{B,B'}(\phi_1 + \phi_2) = M_{B,B'}(\phi_1) + M_{B,B'}(\phi_2)$$

- Sea $\phi : V \longrightarrow V'$, B y B' bases respectivas y $a \in \mathbb{K}$, entonces:

$$M_{B,B'}(a \cdot \phi) = a \cdot M_{B,B'}(\phi)$$

- Sean las aplicaciones $V \xrightarrow{\phi} V' \xrightarrow{\varphi} V''$ y B y B' bases respectivas, entonces:

$$M_{B,B''}(\varphi \circ \phi) = M_{B',B''}(\varphi) \cdot M_{B,B'}(\phi)$$

Proposición

Sea $(\text{Hom}_k(V, V'), +, \cdot)$ el espacio vectorial de los homomorfismos de V' en V y suponiendo $\dim V = n$ y $\dim V' = m$, entonces

$$\begin{aligned} \Phi : (V, V') &\longrightarrow \text{Mat}_{m \times n}(K) \\ \phi &\longmapsto M_{B,B'}(\phi) \end{aligned}$$

es un isomorfismo de espacios vectoriales, es decir:

$$\text{Hom}_k(V, V') \simeq \text{Mat}_{m \times n}(K)$$

Demostración:

- Está bien definida porque ya comprobamos que toda aplicación lineal está definida de modo único por la transformación de los vectores de una base y además cada aplicación podía escribirse en forma de matriz, que está asociada a cada ϕ .
- La suma y el producto se respetan de forma trivial.
- Que es inyectiva:

$$\phi \in \ker(\Phi) \Rightarrow \Phi(\phi) = 0 \Rightarrow \Phi(\phi) = M_{B,B'}(\phi) \Rightarrow \phi(v_1) = \cdots = \phi(v_n) = 0 \Rightarrow \phi = 0 \Rightarrow \text{Inyectiva}$$

- Ahora que es suprayectiva:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \text{Mat}_{m \times n}(K) \Rightarrow \exists \phi : V \rightarrow V' \text{ tal que } \phi(v_i) = a_{1i}v'_1 + \cdots + a_{mi}v'_m \Rightarrow \exists! \phi \Rightarrow \text{Sobre}$$

Como para que la imagen de ϕ sea A tiene que definirse a partir de los vectores de V' , entonces es única porque ya demostramos que estos la determinan.

Observación:

Como el resultado anterior es cierto para cualesquiera V y V' , entonces $\text{End}_k(V) \simeq \text{Mat}_n(K)$ a través del mismo isomorfismo.

Equivalencia de aplicaciones lineales y relación entre sus matrices

Después de identificar cada aplicación lineal con su matriz asociada, puede surgir la siguiente pregunta: ¿cómo cambia la matriz si cambio las bases respecto de las cuales está definida ϕ ?

Además, si la matriz asociada cambia al elegir bases distintas, tiene sentido definir un conjunto cociente donde las matrices que representan a la misma aplicación lineal desde bases distintas sean todas equivalentes.

Definición (Equivalencia de matrices)

Sea $A_1, A_2 \in \text{Mat}_{m \times n}(K)$ y V, V' espacios vectoriales de dimensiones n y m respectivamente, decimos que son **matrices equivalentes** si y sólo si:

$$A_1 \approx A_2 \Leftrightarrow \exists \phi : V \rightarrow V' \text{ lineal y bases } B_1, B_2 \subset V \text{ y } B'_1, B'_2 \subset V' : \begin{cases} A_1 = M_{B_1, B'_1}(\phi) \\ A_2 = M_{B_2, B'_2}(\phi) \end{cases}$$

es decir, si representan a la misma aplicación lineal respecto de bases distintas.

Por tanto, el estudio de este conjunto cociente ahora se debería centrar en ver si cada clase de equivalencia tiene más de un elemento y ver cómo se relacionan los elementos de la misma clase de equivalencia. Para ello, supongamos una aplicación $\phi : V \rightarrow V'$ y B_1, B_2 bases de V y B'_1, B'_2 bases de V' , con matrices⁸ asociadas:

$$A_1 = \text{Mat}_{B_1, B'_1}(\phi) \qquad A_2 = \text{Mat}_{B_2, B'_2}(\phi)$$

A través del siguiente diagrama, se puede ver que considerar la aplicación expresada en términos de B_1 y B'_1 es lo mismo que considerarla en términos de B_2 y B'_2 pero cambiando de base antes y después para que todo se exprese en las bases iniciales:

$$\underbrace{V}_{B_1} \xrightarrow{id_V} \underbrace{V}_{B_2} \xrightarrow{\phi} \underbrace{V'}_{B'_2} \xrightarrow{id_{V'}} \underbrace{V'}_{B'_1}$$

es decir, que podemos ver $M_{B_1, B'_1}(\phi) = A_1$ como la composición⁹ de las funciones que en el diagrama anterior aparecen, luego:

$$A_1 = M_{B_1, B'_1}(\phi) = M_{B'_2, B'_1}(id_{V'}) \cdot M_{B_2, B'_2}(\phi) \cdot M_{B_1, B_2}(id_V) = \underbrace{M(B'_2, B'_1)}_Q \cdot A_2 \cdot \underbrace{M(B_1, B_2)}_P = Q \cdot A_2 \cdot P$$

Como P y Q son matrices de cambio de base, pertenecen al grupo lineal, es decir, tienen inversa. Esto permite expresar, simplemente multiplicando por la inversa en ambos lados, A_2 en término de A_1 por matrices cuadradas del mismo modo que lo estábamos haciendo para A_1 . Con este último resultado estamos en condiciones de poder reformular la definición de equivalencia de matrices de la siguiente manera.

Definición (Equivalencia de matrices)

Sean $A_1, A_2 \in \text{Mat}_{m \times n}(K)$, se define la **equivalencia** como:

$$A_1 \approx A_2 \Leftrightarrow \exists P \in GL_n(K), Q \in GL_m(K) : A_1 = Q \cdot A_2 \cdot P$$

donde el conjunto $GL(K)$ se refiere a las matrices inversibles de los tamaños adecuados.

Observación:

El desarrollo anterior permite relacionar todas las matrices que expresan a una aplicación lineal a través de una clase de equivalencia en el conjunto cociente que induce la relación de equivalencia entre matrices.

Proposición

Sea V un espacio vectorial, $B = \{v_1, \dots, v_n\}$ una base del mismo y $P \in GL_n(K)$ de la forma $P = (p_{ij})$, entonces el conjunto de vectores:

$$B' := \left\{ v'_1 = \sum_{i=1}^n p_{i1} v_i, \dots, v'_n = \sum_{i=1}^n p_{in} v_i \right\}$$

es base de V y P es la matriz de cambio de base a B' .

⁸Nótese que pueden considerarse ambas matrices como matrices de una aplicación lineal porque anteriormente hemos demostrado que existía un isomorfismo entre el conjunto de matrices y el de aplicaciones.

⁹Nótese que la composición de funciones se hace en orden inverso al producto de las mismas.

Demostración:

1. Veamos la independencia lineal, pues como hay tantos vectores como dimensión del espacio, se cumpliría trivialmente que es base

$$\begin{aligned} 0 &= a_1 v'_1 + \cdots + a_n v'_n \\ &= a_1 (p_{11} v_1 + \cdots p_{n1} v_n) + \cdots + a_n (p_{1n} v_1 + \cdots p_{nn} v_n) \\ &= (p_{11} a_1 + \cdots p_{n1} a_n) v_1 + \cdots + (p_{1n} a_1 + \cdots + p_{nn} a_n) v_n \end{aligned}$$

Como los vectores v_i formaban una base, los coeficientes asociados a cada uno en la expresión anterior son nulos. Esto, expresado de forma matricial, se puede escribir como:

$$\begin{cases} p_{11} a_1 + \cdots p_{n1} a_n = 0 \\ \vdots \\ p_{1n} a_1 + \cdots + p_{nn} a_n \end{cases} \Rightarrow P \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = P^{-1} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \Rightarrow a_i = 0$$

2. Probado lo anterior, esto es trivial pues si lo primero es cierto, coger cada vector de B' y expresarlo en base B , hace que los coeficientes de P sean sus coordenadas.

Observación:

Del mismo modo que pudimos identificar cada matriz con una aplicación lineal de la que es asociada, el teorema anterior nos permite considerar que toda matriz inversible es una matriz de algún cambio de base dentro de un espacio vectorial.

Rango de una aplicación lineal

Una característica fundamental tanto de las aplicaciones lineales como del álgebra matricial es el rango de dicho objeto. Esta característica nos permite desde discutir sistemas lineales para elucubrar sus posibles soluciones como conocer la dimensión de la imagen del subespacio de llegada de una aplicación lineal.

Definición (Rango de una aplicación)

Sea $\phi : V \rightarrow V'$ una aplicación lineal, definimos el **rango de la aplicación** como:

$$rg(\phi) := \dim_k Im(\phi)$$

es decir, es la dimensión del subespacio imagen de la aplicación lineal.

Si consideramos una matriz $A \in Mat_{m \times n}(K)$ definida como $A = (a_{ij})$, vimos que esta se podía interpretar como la matriz asociada a una aplicación lineal y, en ese caso, el rango de la aplicación lineal se correspondería con la dimensión de la envoltura lineal de los vectores columna (imagen de los de la base) que definen la imagen de la aplicación lineal. Por esto, tiene sentido dar la siguiente definición.

Definición (Rango de una matriz)

Sea $A \in Mat_{m \times n}(K)$ una matriz definida como $A = (a_{ij})$, definimos el **rango de la matriz** como:

$$rg(A) := \dim_k L(a_{11}e_1 + \cdots + a_{m1}e_m, \cdots, a_{1n}e_1 + \cdots + a_{mn}e_m)$$

es decir, $rg(A)$ es la dimensión del subespacio generado por los vectores columnas de la matriz A .

Observación:

Como la dimensión de la envoltura lineal es el número de vectores linealmente independientes, esto nos proporciona un método para calcular la dimensión de un subespacio o la independencia lineal de un conjunto de vectores.

Claramente la definición anterior que se ha dado para el rango de una matriz solo es interesante si dicho rango se conserva entre todas las matrices que pueden representar una misma aplicación lineal. Por tanto, nuestros pasos se encaminan a poder decir que dicho rango depende exclusivamente de la aplicación lineal y no de la representación escogida para la misma.

Lema

Sea $A = (a_{ij}) \in \text{Mat}_{m \times n}(K)$ y W a un espacio vectorial cualquiera de dimensión $\dim W = m$ escogemos una base suya $B = \{w_1, \dots, w_m\}$ y la definición de rango de la matriz A no cambia.

$$rg(A) = \dim L(a_{11}w_1 + \dots + a_{m1}w_m, \dots, a_{1n}w_1 + \dots + a_{mn}w_m)$$

Demostración:

En el fondo, la dimensión de $L()$ no es más que el número de vectores linealmente independientes de ese sistema de generadores, por lo que yo si encuentro un conjunto de vectores linealmente independientes maximal de los que forman ese sistema, entonces serán una base. Por ello podemos considerar que hay una cantidad $p \leq n$ de vectores linealmente independientes siendo n el número de vectores que hay dentro de $L()$. Y lo mismo ocurre con el $L()$ de la condición 2) del rango de una aplicación lineal, en el que designaremos su p por q . Ahora se trata de ver que $p = q$

$$\{a_{11}e_1 + \dots + a_{m1}e_m, \dots, a_{1q}e_1 + \dots + a_{mq}e_m\} \text{ supongamoslos l. ind.}$$

Tenemos que llegar a que $\{a_{11}w_1 + \dots + a_{m1}w_m, \dots, a_{1q}w_1 + \dots + a_{mq}w_m\}$ son linealmente independientes.

$$\begin{aligned} 0 &= \lambda_1(a_{11}w_1 + \dots + a_{m1}w_m) + \dots + \lambda_q(a_{1q}w_1 + \dots + a_{mq}w_m) = \\ &= (\lambda_1 a_{11} + \dots + \lambda_q a_{1q})w_1 + \dots + (\lambda_1 a_{m1} + \dots + \lambda_q a_{mq})w_m \Rightarrow \end{aligned}$$

Como $B = \{w_1, \dots, w_m\}$ es base de W , entonces:

$$\begin{aligned} 0 &= \underbrace{(\lambda_1 a_{11} + \dots + \lambda_q a_{1q})}_{=0} w_1 + \dots + \underbrace{(\lambda_1 a_{m1} + \dots + \lambda_q a_{mq})}_{=0} w_m \Rightarrow \\ &\Rightarrow 0 = (\lambda_1 a_{11} + \dots + \lambda_q a_{1q})e_1 + \dots + (\lambda_1 a_{m1} + \dots + \lambda_q a_{mq})e_m = \\ &= \lambda_1(a_{11}e_1 + \dots + a_{m1}e_m) + \dots + \lambda_q(a_{1q}e_1 + \dots + a_{mq}e_m) \Rightarrow \lambda_1 = \dots = \lambda_q = 0 \end{aligned}$$

Porque $B = \{e_1, \dots, e_m\}$ es base.

Observación:

Este lema lo que nos resume es que para la definición del rango de una matriz, se pueden utilizar los vectores de la base de un espacio de dimensión m cualquiera, es decir, que no es necesario que sean los del espacio K^m ni tampoco los de la base canónica de un espacio.

Proposición

Sea $\phi : V \rightarrow V'$ una aplicación lineal y $A = M_{B,B'}(\phi)$ su representación matricial en bases arbitrarias, entonces:

$$\forall A = M_{B,B'}(\phi) \in \text{Mat}_{m \times n}(K) : rg(\phi) = rg(A)$$

es decir, el rango de la aplicación y de la matriz coincide para todas las matrices equivalentes.

Demostración:

Simplemente basta con ver que

$$rg(\phi) = \dim \text{Im}(\phi) = \dim L(\phi(v_1), \dots, \phi(v_n)) = \dim L(a_{11}v'_1 + \dots + a_{m1}v'_m, \dots, a_{1n}v'_1 + \dots + a_{mn}v'_m)$$

y precisamente el lema anterior nos dice que este último término es igual al rango de A .

Proposición

Sea $A \in \text{Mat}_{m \times n}(K)$ con rango r , entonces A es equivalente a una matriz¹⁰ con esta forma:

$$A \approx A' = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) \in \text{Mat}_{m \times n}(K)$$

Demostración:

$$A = M_{B, B'}(\phi), \phi: \underbrace{V}_{B_1} \rightarrow \underbrace{V'}_{B'_1}$$

Ahora definimos $B_2 = \{w_1, \dots, w_r, \underbrace{w_{r+1}, \dots, w_n}_{B \text{ del } \ker(\phi)}\}$ base de V y esto puede hacerse por el teorema

de la dimensión. Y $B'_2 = \{\phi(w'_1), \dots, \phi(w'_r), w'_{r+1}, \dots, w'_m\}$ base de V' , es decir, los primeros son imágenes de los primeros de B'_1 y después se amplía la base hasta V' , pero para ello primero hay que comprobar que las imágenes son linealmente independientes.

$$0 = a_1\phi(w_1) + \dots + a_r\phi(w_r) = \phi(a_1w_1 + \dots + a_rw_r) = \phi(a_1w_1 + \dots + a_rw_r) \Rightarrow v = a_1w_1 + \dots + a_rw_r \in \ker(\phi) \Rightarrow$$

$$\Rightarrow v = a_{r+1}w_{r+1} + \dots + a_nw_n \Rightarrow a_1w_1 + \dots + a_rw_r = a_{r+1}w_{r+1} + \dots + a_nw_n \Leftrightarrow$$

$$\Leftrightarrow a_1w_1 + \dots + a_rw_r - a_{r+1}w_{r+1} - \dots - a_nw_n = 0$$

Pero como son base: $a_1 = \dots = a_r = 0 \Rightarrow$ demostrado que son linealmente independientes. Con lo cual, quien es esta matriz?:

$$Mat_{B_2, B'_2}(\phi) = \begin{pmatrix} \underbrace{\phi(w_1)_{B'_2}}_{\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}} & \underbrace{\phi(w_2)_{B'_2}}_{\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}} & \underbrace{\dots}_{\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}} & \dots & \underbrace{\phi(w_r)_{B'_2}}_{\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}} & \underbrace{\phi(w_{r+1}), \dots}_{=0} & \underbrace{\phi(w_n)=0}_{\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}} \\ \vdots & \vdots & \vdots & \ddots & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

Proposición

Sean $A_1, A_2 \in \text{Mat}_{m \times n}(K)$, entonces:

$$A_1 \approx A_2 \Leftrightarrow rg(A_1) = rg(A_2)$$

es decir, la equivalencia se caracteriza por rangos.

Demostración:

■ “ \Rightarrow ”:

$$A_1 = M_{B_1, B'_1}(\phi) \text{ y } A_2 = M_{B_2, B'_2}(\phi) \Rightarrow rg(A_1) = rg(\phi) = rg(A_2)$$

■ “ \Leftarrow ”:

$$rg(A_1) = rg(A_2) = r \Rightarrow A_1 \approx \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) \text{ y } A_2 \approx \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) \Rightarrow A_1 \approx A_2$$

¹⁰Esta notación quiere decir que la matriz se puede dividir en 4 cuadrantes desiguales, la parte superior izquierda corresponde a una matriz identidad cuadrada de tamaño r y el resto de cuadrantes a matrices nulas que no tienen por qué ser cuadradas.

Observación:

Los enunciados anteriores permiten, no solo caracterizar la relación de equivalencia entre matrices a través del rango, sino conocer el número de clases de equivalencia y el mejor representante de cada una en el conjunto cociente $Mat_{m \times n}(K)/\approx$. Por ejemplo, si tenemos que $m = 4$ y $n = 3$, entonces $rg(A) \leq 3 \Rightarrow r = 0, 1, 2, 3$, es decir, hay cuatro clases de equivalencia y cada matriz puede ser equivalente a cada una de las matrices con caja I_r comentadas anteriormente.

Teorema

Sea $A \in Mat_{m \times n}(K)$ y denotando por A^t a su traspuesta, entonces:

$$rg(A) = rg(A^t)$$

es decir, ambas matrices son equivalentes.

Demostración:

$$\begin{aligned} A &\approx \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) \Rightarrow \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right) = Q \cdot A \cdot P : P \in Gl_n(K), Q \in Gl_m(K) \Rightarrow \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right)^t = (QAP)^t \\ &= P^t A^t Q^t : P^t \in Gl_n(K), Q^t \in Gl_m(K) \Rightarrow \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right), A \in Mat_{n \times m} : \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right)^t \approx A^t \Rightarrow \\ &\Rightarrow rg \left(\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right)^t \right) = rg(A^t) \Leftrightarrow rg(A^t) = r = rg(A) \end{aligned}$$

Proposición

Sea $P \in Mat_n(K)$, entonces dicha matriz es inversible si y sólo si su rango es máximo:

$$P \in Gl_n(K) \Leftrightarrow rg(P) = n$$

Demostración:

■ “ \Rightarrow ”:

$$\begin{aligned} P \in Gl_n(K) &\Rightarrow P = M(B', B) : B = \{v_1, \dots, v_n\} \text{ y } B' = \{v'_1, \dots, v'_n\} \text{ bases} \Rightarrow \\ &\Rightarrow \{v'_1, \dots, v'_n\} \text{ l. ind. pero cada } v_i = p_{1i}v_1 + \dots + p_{ni}v_n \text{ por ser } P \text{ matriz de paso} \Rightarrow rg(P) = n \end{aligned}$$

Por aquello de que el número de columnas linealmente independientes era el mismo respecto de bases arbitrarias.

■ “ \Leftarrow ”:

$$\begin{aligned} rg(P) = n &\Rightarrow B = B_c \text{ y } B' = \{p_1e_1 + \dots + p_ne_n\} \text{ siendo los vectores de } B' \text{ los vectores columnas de } P \Rightarrow \\ &\Rightarrow \text{ como } rg(P) = n \text{ } B' \text{ es l. ind.} \Rightarrow B' \text{ base} \Rightarrow P = M(B', B) \in Gl_n(K) \end{aligned}$$

ESTRUCTURA DE ENDOMORFISMOS

Este capítulo está dedicado a unos homomorfismos concretos: los que van de un cuerpo en sí mismo o endomorfismos. En el capítulo anterior, habíamos visto que para cualquiera aplicación lineal podíamos encontrar una matriz concreta más sencilla (por rango) que también representaba a la aplicación lineal. En este caso, esto no va a ser tan sencillo puesto que generalmente en los endomorfismos se suele dar la misma base de llegada que de partida.

En esta línea, el capítulo va a tratar de desarrollar la teoría necesaria para poder utilizar la matriz más sencilla posible y va a introducir ciertos subespacios muy especiales, relacionados con el estudio de la matriz sencilla anterior.

Definición (Semejanza de Aplicaciones)

Sea $\phi : V \rightarrow V$ una aplicación lineal de un espacio vectorial V en sí mismo, con dimensión $\dim V = n$, definimos la **semejanza entre matrices** como:

$$A \sim A' \Leftrightarrow \exists B, B' \subset V : \begin{cases} A = M_B(\phi) \\ A' = M_{B'}(\phi) \end{cases}$$

De nuevo, y haciendo el desarrollo del capítulo anterior, la definición previa impone que sean matrices que representan a la misma aplicación lineal respecto de bases distintas. Este hecho da pie a la posibilidad de pasar de una a otra a través de una matriz de cambio de base, de modo que ambas quedan relacionadas a través de la expresión:

$$A' = P^{-1}AP : P = M(B', B) \in Gl_n(K)$$

lo que permite dar una definición equivalente a la primera prescindiendo (al menos en apariencia) de las aplicaciones lineales.

Definición (Semejanza de Matrices)

Sean $A, A' \in Mat_n(K)$, decimos que ambas son **semejantes** si y sólo si existe una matriz en el grupo lineal que las relacione como:

$$A \sim A' \Leftrightarrow \exists P \in Gl_n(K) : A' = P^{-1}AP$$

Observación:

Al igual que ocurría con la equivalencia de matrices, esta relación de equivalencia (más fuerte que la anterior) permite definir un conjunto cociente y agrupar las matrices semejantes en la misma clase.

POLINOMIO MÍNIMO Y CARACTERÍSTICO

Posteriormente cobrará más sentido esta noción, pero por lo pronto podemos decir que es interesante desde el punto de vista algebraico preguntarse por si una matriz es raíz de un polinomio.

Polinomio mínimo de un endomorfismo

Consideramos $A \in \text{Mat}_n(K)$ y el conjunto de vectores $\{I_n, A, A^2, \dots, A^{n^2}\} \subset \text{Mat}_n(K)$. Como $\dim \text{Mat}_n(K) = n^2$ y en ese conjunto hay $n^2 + 1$ vectores, necesariamente ese conjunto debe ser linealmente dependiente. Con lo cual, $\exists a_0, \dots, a_{n^2} \in K : a_0 I + \dots + a_{n^2} A^{n^2} = 0$ donde $\exists a_i \neq 0$, es decir, en cierta manera se puede interpretar que A es la raíz del polinomio:

$$(a_0 + a_1 x + \dots + a_{n^2} x^{n^2})(A) = f(A) : f \neq 0$$

Por tanto, es lógico preguntarse cuál es el polinomio mínimo para el cuál esto ocurre.

Definición (Polinomio Mínimo de un Endomorfismo)

Sea $\phi \in \text{End}(V)$ y dada la aplicación

$$\begin{aligned} \Phi_\phi : K[x] &\longrightarrow \text{End}(V) \\ p(x) &\longmapsto p(\phi) \end{aligned}$$

se define el **polinomio mínimo** $q_\phi(x)$ de un endomorfismo como el mínimo polinomio en el $\ker \Phi_\phi$.

La definición anterior formaliza la idea inicial de encontrar el polinomio de menor grado del cual el endomorfismo es raíz. Sin embargo, la idea inicial de la subsección proponía el razonamiento para matrices y no para endomorfismos ¿El polinomio mínimo del endomorfismo será el mismo que el polinomio mínimo de las representaciones matriciales del mismo? ¿Será el mismo entre representaciones matriciales? ¿Existe y hay alguna forma de calcularlo?

Definición (Polinomio Mínimo de una Matriz)

Sea $A \in \text{Mat}_n(K)$ y dada la aplicación

$$\begin{aligned} \Phi_A : K[x] &\longrightarrow \text{Mat}_n(K) \\ p(x) &\longmapsto p(A) \end{aligned}$$

se define el **polinomio mínimo** $q_A(x)$ de una matriz como el mínimo polinomio del $\ker \Phi_A$.

Proposición (Existencia)

Sea $A \in \text{Mat}_n(K)$, siempre existe un polinomio no constante del que es raíz.

$$\forall A \in \text{Mat}_n(K) : \exists f \in K[x] : f(A) = 0$$

Proposición (Unicidad)

Sea $A \in \text{Mat}_n(K)$, existe un único polinomio $q_A \in K[x]$ que cumple:

- $q_A \neq 0$
- q_A mónico
- $q_A(A) = 0$

$$\blacksquare f(A) = 0 \Rightarrow q_A \mid f$$

Demostración:

Las tres primeras condiciones ya se han visto donde se explica la parte de como descubrir dicho polinomio, por lo que vamos a demostrar la última:

$$f(A) = 0 \Rightarrow f(A) = 0 = g(A)q_A(A) + r(A) \Rightarrow r(A) = 0$$

Como r es de grado menor que q_A , no puede ser no nulo porque en ese caso habríamos encontrado un polinomio de grado menor que q_A y que se anula en A , es decir, que q_A no sería mínimo.

Veamos que es único, para ello supongamos que existen dos. Por el razonamiento anterior, $q_A \mid q'_A$, pero también ocurre que $q'_A \mid q_A$. Por tanto, podemos deducir que $q_A = q'_A$.

Proposición

Sean $A, A' \in \text{Mat}_n(K)$ dos matrices semejantes, ambas tienen el mismo polinomio mínimo.

Demostración:

La demostración es puramente algebraica, basta con ver que:

$$\begin{aligned} q_A(A') &= q_A(P^{-1}AP) \\ &= (a_0 + a_1x + \cdots + x^m)(P^{-1}AP) \\ &= a_0I + a_1P^{-1}AP + \cdots + (P^{-1}AP)^m \\ &= a_0I + P^{-1}a_1AP + \cdots + P^{-1}A^mP \\ &= P^{-1}(a_0I + a_1A + \cdots + A^m)P \\ &= P^{-1}\underbrace{q_A(A)}_{=0}P = 0 \end{aligned}$$

Con este último resultado ya podemos afirmar que la definición de polinomio mínimo de una matriz coincide entre todas las matrices semejantes, esto es, entre todas las posibles representaciones matriciales de una misma aplicación lineal. Para concluir con las preguntas que planteamos al inicio de todo este desarrollo, basta con ver que el endomorfismo y las representaciones matriciales tienen en común el mismo polinomio mínimo.

Para un polinomio cualquiera $f(x) = a_0 + \cdots + a_mx^m$ podemos considerar $f(\phi)$ como el endomorfismo resultante de sustituir este endomorfismo en el polinomio y operar según las operaciones para aplicaciones lineales vistas anteriormente. Cuando analizamos la matriz de dicho endomorfismo $f(\phi)$, entonces:

$$f(\phi) = a_0id_v + \cdots + a_m\phi^m \Rightarrow M_B(f(\phi)) = a_0I + a_1A + \cdots + a_mA^m = f(A)$$

nos damos cuenta de que dicha matriz es la misma que obtendríamos sustituyendo $A = M_B(\phi)$ en el mismo polinomio f , es decir, $f(A) = M_B(f(\phi))$.

Por tanto, del razonamiento anterior podemos ver que

$$\begin{aligned} q_\phi(\phi) = 0 &\Leftrightarrow M_B(q_\phi(\phi)) = 0 \Leftrightarrow q_\phi(A) = 0 \Rightarrow q_\phi \mid q_A \\ q_A(A) = 0 &\Leftrightarrow M_B(q_A(\phi)) = 0 \Leftrightarrow q_A(\phi) = 0 \Rightarrow q_A \mid q_\phi \end{aligned}$$

donde cabe destacar que en la segunda ecuación la igualdad $M_B(q_A(\phi)) = 0$ es para cualquier base B (porque las matrices semejantes tienen el mismo polinomio mínimo q_A) lo que justifica poder afirmar que el endomorfismo $q_A(\phi)$ es el nulo.

Por tanto, visto que el razonamiento de polinomio mínimo se aplica para cualquier matriz y que dos matrices semejantes conservan su polinomio mínimo, podemos hacer referencia a el polinomio mínimo del endomorfismo **sea cual sea la matriz que lo representa**, puesto que este se conserva independientemente de la base en la que esté expresada su matriz.

Ejemplo:

- Sea $f \in K[x]$ mónico, donde f es no constante $\Rightarrow C_f \Rightarrow q_{C_f} = f$ donde C_f es la matriz compañera del polinomio f .
- Si ϕ es proyección, entonces:

$$\phi^2 = \phi \Leftrightarrow (x^2 - x)(\phi) = 0 \Leftrightarrow q_\phi = \begin{cases} x(\phi) = 0 & \Rightarrow \phi = 0 \\ (x-1)(\phi) = 0 & \Rightarrow \phi = id \\ (x^2 - x)(\phi) = 0 & \Rightarrow \phi \neq 0, id \end{cases}$$

- Si ϕ es simetría, entonces:

$$\phi^2 = id \Leftrightarrow (x^2 - 1)(\phi) = 0 \Leftrightarrow \begin{cases} (x-1)(\phi) = 0 & \Rightarrow \phi = id \\ (x+1)(\phi) = 0 & \Rightarrow \phi = -id \\ (x^2 - x)(\phi) = 0 & \Rightarrow \phi \neq 0, -id \end{cases}$$

- Supongamos que $q_\phi = (x-1)(x-r) : r \neq 0, 1, -1$ donde decimos que este endomorfismo es una homología general de razón r .

Polinomio mínimo de un endomorfismo y un vector

Lógicamente, $q_\phi(\phi) = 0$ implica que $\forall v \in V : q_\phi(\phi)(v) = 0$ puesto que hemos comentado que $q_\phi(\phi)$ es la aplicación nula, pero esto genera una nueva pregunta: ¿cuál es el polinomio mínimo $q_v(\phi)$ para que $q_v(\phi)(v) = 0$?, es decir, ¿habrá otro polinomio distinto que se anule en un cierto vector, pero que no haga ϕ la aplicación nula?

Definición (Polinomio mínimo de un Endomorfismo y un Vector)

Sea $\phi \in \text{End}(V)$ y $v \in V$ no nulo, dada la aplicación

$$\begin{aligned} \Phi_{v,\phi} : K[x] &\longrightarrow V \\ p(x) &\longmapsto p(\phi)(v) \end{aligned}$$

se define el **polinomio mínimo** $q_{v,\phi}$ de un endomorfismo y un vector como el mínimo polinomio del $\ker \Phi_{v,\phi}$.

Proposición

Sea $\phi \in \text{End}(V)$ y $v \in V$, el polinomio mínimo del vector $q_{v,\phi}$ divide al polinomio mínimo del endomorfismo q_ϕ .

Demostración:

Como $q_\phi(\phi)(v) = 0$ tiene que ser un polinomio de grado mayor que el que hemos definido como mínimo $q_{v,\phi}$ y, por tanto, este último debe dividirlo.

Ejemplo:

Supongamos que el endomorfismo ϕ es una proyección cuya matriz respecto de la base canónica es:

$$M_{B_c}(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

¿cuál es el polinomio mínimo de esta proyección? Por todo lo desarrollado antes, es fácil ver que $q_\phi(\phi) = x^2 - x$, pero ¿quién es $q_{\phi, e_1}(\phi)$ y $q_{\phi, e_2}(\phi)$?

En primer lugar, para e_1 nos percatamos podemos ver que $\phi(e_1) = e_1$. Esto quiere decir que $(x-1)(\phi)(e_1) = 0$ y como es de grado uno, no hay otro menor que él donde se anule. Utilizando el mismo razonamiento para e_2 , vemos que $\phi(e_2) = 0$ y entonces $q_{\phi, e_2} = x$.

Proposición

Sea $\phi \in \text{End}(V)$ y $B = \{v_1, \dots, v_n\}$ una base de V , entonces:

$$q_\phi = \text{mcm}\{q_{\phi, v_1}, \dots, q_{\phi, v_n}\}$$

por lo que el mínimo global se puede recuperar a partir de los mínimos para cada vector de la base.

Demostración:

Para demostrarlo basta con ver que q_ϕ divide al polinomio $\bar{q} = \text{mcm}\{q_{\phi, v_1}, \dots, q_{\phi, v_n}\}$ y viceversa:

- $\bar{q} \mid q_\phi$: simplemente vemos que $\forall i \in I : q_i \mid q_\phi \Rightarrow \bar{q} \mid q_\phi$.
- $q_\phi \mid \bar{q}$:

El polinomio \bar{q} puede escribirse como $\forall i \in I : \bar{q} = g_i q_i$. Con lo cual, vemos que le pasa a $\bar{q}(\phi)$ cuando se le aplica a v_i :

$$[\bar{q}(\phi)](v_i) = [(g_i q_i)(\phi)](v_i) = [g_i(\phi) \circ q_i(\phi)](v_i) = g_i(\phi) \left(\underbrace{q_i(\phi)(v_i)}_{=0} \right) = g_i(\phi)(0) = 0$$

Por tanto, $\bar{q}(\phi)$ es la aplicación nula porque se anula sobre cada uno de los vectores de la base, luego $q_\phi \mid \bar{q}$.

Cálculo del polinomio mínimo

Ya visto el concepto de polinomio mínimo, en la práctica basta con calcular el polinomio mínimo de cada uno de los vectores de la base del espacio vectorial y saber que el buscado es el m.c.m. de ellos.

Además, existe un método canónico para la búsqueda del polinomio mínimo de cada vector. Como lo que se busca es una expresión $a_0 v + a_1 \phi(v) + \dots + a_n \phi^n(v)$ de tamaño mínimo y con algún $a_i \neq 0$, entonces basta calcular $v, \phi(v), \phi^2(v), \dots \in V$ hasta que el conjunto sea linealmente dependiente¹ y los coeficientes a_i vendrían dados por la combinación lineal no trivial nula de la dependencia lineal.

Ejemplo:

Sea $\phi \in \text{End}(V)$ representado en la base canónica por

$$M_B(\phi) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

para calcular el polinomio mínimo para e_1 , seguimos los pasos anteriores:

$$\begin{aligned} \phi(e_1) &= e_2 + e_3 \\ \phi^2(e_1) &= 2e_1 + e_2 + e_3 \end{aligned}$$

En este punto paramos, pues hemos encontrado un conjunto dependiente a través de la relación

$$\phi^2(v) - 2e_1 - 2\phi(e_1) = 0 \Rightarrow [(x^2 - x - 2)(\phi)](e_1) = 0$$

lo que implica que el polinomio mínimo es de la forma

$$q_{\phi, e_1} = x^2 - x - 2$$

¹Realmente este polinomio es el mínimo porque si existiera otro de grado menor para el cual se anulase el vector evaluado en el endomorfismo resultante, entonces habría un conjunto más pequeño linealmente independiente y eso no puede ser.

Polinomio característico de un endomorfismo

Aunque el polinomio mínimo goza de buenas propiedades, tenemos una herramienta más potente: el polinomio característico. Este especial polinomio nos permitirá determinar desde el número y dimensión de los subespacios invariantes por ϕ hasta poder expresar la matriz de ϕ de la forma más sencilla posible.

Definición (Autovalores y Vectores Propios)

Sea $\phi \in \text{End}(V)$, $\lambda \in K$ y $v \neq 0$, definimos un **vector propio** del endomorfismo como aquel que verifica:

$$\phi(v) = \lambda \cdot v$$

y llamamos **autovalor** al λ asociado.

Definición (Polinomio Característico)

Sea $\phi \in \text{End}(V)$ y $A = M_B(\phi)$, definimos el **polinomio característico** de ϕ como:

$$p_\phi(x) = |A - x \cdot I_n|$$

Proposición

Sea $\phi \in \text{End}(V)$, $A = M_B(\phi)$ y $A' = M_{B'}(\phi)$, el polinomio característico de ambas coincide, es decir, matrices semejantes tienen el mismo polinomio característico.

Demostración:

Sean $A = M_B(\phi)$ y $A' = M_{B'}(\phi)$, ambas matrices por ser semejantes cumplen que $A' = P^{-1}AP$ y el resto es algebraico:

$$\begin{aligned} p_{A'} &= |A' - xI_n| \\ &= |P^{-1}AP - P^{-1}xP| \\ &= |P^{-1}(A - xI_n)P| \\ &= |P^{-1}|(A - xI_n)|P| \\ &= |(A - xI_n)| = p_A \end{aligned}$$

Observación:

De nuevo, igual que en el polinomio mínimo, el resultado anterior prueba que la definición es consistente pues no depende de la representación matricial utilizada para ello. Por tanto, tiene sentido hablar de $p_\phi(x)$ en vez de $p_A(x)$.

Observación:

Un desarrollo exhaustivo del polinomio característico de grado n reflejaría que los coeficientes siguen la expresión:

$$(-1)^n x^n + (-1)^{n-1}(a_{11} + \dots + a_{nn})x^{n-1} + \dots + A_r x^r + \dots + \det A$$

esto revela que propiedades matriciales² como la traza $\text{tr}(A) = a_{11} + \dots + a_{nn}$ o el determinante $\det A$ se mantienen invariantes entre matrices equivalentes.

Definición (Espectro)

Sea $\phi \in \text{End}(V)$, definimos el **espectro** de un endomorfismo

$$\sigma(\phi) = \{\lambda_1, \dots, \lambda_t\} : q_\phi(\lambda_i) = 0$$

como el conjunto de todos sus autovalores.

² A_r denota la suma de los determinantes de los menores de orden $n - r$ que tienen la diagonal principal sobre la diagonal principal de A .

Teorema de Caley-Hamilton

Aunque hemos adelantado que el polinomio característico tendrá muy buenas propiedades y que el mínimo y el característico coinciden, es fundamental probar que el polinomio característico se anula en el endomorfismo porque la gran mayoría de resultados provechosos de esta definición utilizarán dicho resultado.

Proposición

Sea $\phi \in \text{End}(V)$, entonces λ es autovalor si y solo si:

- $q_\phi(\lambda) = 0$
- $|A - \lambda I| = 0$

Demostración:

- $2 \Rightarrow 1$

$$\phi(v) - \lambda v = 0 \Rightarrow q_{\phi,v} = x - \lambda \Rightarrow x - \lambda \mid q_\phi \Rightarrow q_\phi(\lambda) = 0$$

- $1 \Rightarrow 2$

$$q_\phi(\lambda) = 0 \Rightarrow x - \lambda \mid q_\phi \Rightarrow q_\phi = (x - \lambda)g(x)$$

Sea $B = \{v_1, \dots, v_n\}$ una base consideremos $q_i = q_{\phi,v_i}$, entonces si $x - \lambda \nmid q_1, \dots, q_n \Rightarrow x - \lambda \nmid \text{mcm}\{q_1, \dots, q_n\}$ pero ocurre que $q_\phi = \text{mcm}\{\dots\}$ luego necesariamente $x - \lambda \mid q_i$ para algún $i \in \mathbb{N}$. Por tanto, se tiene (podemos escoger que es q_1 porque el orden no importa) que $q_1 = q_{\phi,v_1} = (x - \lambda)h(x)$. Esto necesariamente conlleva a que $(\phi - \lambda \cdot \text{id})(h(\phi)(v_1)) = 0$, $h(\phi)(v_1)$ no puede ser 0 porque sería el polinomio mínimo de v_1 y no lo es luego precisamente $h(\phi)(v_1) = v$, es decir, es el vector donde se cumple la desigualdad.

- $2 \Rightarrow 3$

Si expresamos en términos de una base $B = \{v_1, \dots, v_n\}$ de forma que queda $v = a_1v_1 + \dots + a_nv_n$ vemos que como $\phi(v) = \lambda \cdot v$ se ve fácilmente que $(A - \lambda \cdot I_n)(v) = 0$, luego como ese sistema tiene solución no nula el determinante de la matriz del sistema es 0, por lo tanto tenemos que $|A - \lambda \cdot I_n| = 0$

Lema

Sea $\phi \in \text{End}(V)$, el polinomio mínimo y característico tienen las mismas raíces.

Demostración:

Tenemos la siguiente cadena de equivalencias

$$q_\phi(\lambda) = 0 \Leftrightarrow |A - \lambda I_n| = 0 \Leftrightarrow p_\phi(\lambda) = 0$$

por tanto, λ es raíz de q_ϕ si y sólo si lo es de p_ϕ .

Lema

Sea $f \in K[x]$ un polinomio irreducible, si f^a es la potencia exacta que divide a q_ϕ , entonces

$$f^a \mid q_\phi, f^a \mid q_\phi, f^{a+1} \nmid q_\phi \Rightarrow f^a \mid p_\phi$$

Demostración:

Sea $B = \{v_1, \dots, v_n\}$ base de V . Llamamos $q_i = q_{\phi,v_i}$

Sabemos que $q_\phi = \text{mcm}\{q_1, \dots, q_n\}$. Queremos ver que hay vectores en V cuyo polinomio mínimo es exactamente f^a . Razonamos por reducción al absurdo.

Supongamos $q_i = f^{a_i} g_i$ tal que

$$f^{a_i} || q_i : a_1 \leq a_2 \leq \dots \leq a_n \leq a$$

Supongamos que $a_n < a$. Entonces $mcm\{q_1, \dots, q_n\} = f^{a_n} \cdot mcm\{g_1, \dots, g_n\}$. Sin embargo, esto es una contradicción ya que $mcm\{q_1, \dots, q_n\} = q_\phi$, y por hipótesis $f^a | q_\phi$, pero $a_n < a$.

Por tanto, $a_n = a$

$$\Rightarrow q_{\phi, v_n} = f^a \cdot g_n \Rightarrow 0 = f^a(\phi)(\underbrace{g_n(\phi)(v_n)}_{v \neq 0}) \Rightarrow q_{\phi, v} = f^a$$

Ya que si $v = 0$, q_{ϕ, v_n} sería como mucho g_n , pero g_n es menor que el mínimo.

Por tanto, $\exists v \in V, v \neq 0 : q_{\phi, v} = f^a$

Llamamos $m = gr(f^a)$. Por las propiedades del polinomio mínimo, sabemos que $\nexists f \in K[x], gr(f) < m : f_{\phi, v} = 0$.

Veamos que los vectores $\{v, \phi(v), \dots, \phi^{m-1}(v)\}$ son linealmente independientes. Supongamos

$$0 = a_0 v + a_1 \phi(v) + \dots + a_{m-1} \phi^{m-1}(v)$$

$$0 = [(a_0 + a_1 x + \dots + a_{m-1} x^{m-1})(\phi)](v)$$

Sin embargo, estamos diciendo que el polinomio $a_0 + a_1 x + \dots + a_{m-1} x^{m-1}$ que sobre $\phi(v)$ vale 0, pero su grado es menor que m, que es el grado del mínimo, lo cual es imposible salvo que el polinomio sea nulo.

$$a_1 = \dots = a_{m-1} = 0$$

Extendemos ahora estos m vectores independientes a una base de todo el espacio.

$$B = \{v, \phi(v), \dots, \phi^{m-1}(v), v_{m+1}, \dots, v_n\} \text{ base de } V$$

Veamos que $f^a | p_\phi$.

$$v \mapsto \phi(v), \phi(v) \mapsto \phi^2(v), \dots, \phi^{m-1}(v) \mapsto \phi^m(v)$$

Los primeros vectores de la base van al siguiente de la base, excepto el $\phi^{m-1}(v)$, que va a $\phi^m(v)$. Estudiamos dicha imagen.

$$m = gr(f^a) \Rightarrow f^a = x^m + b_{m-1} x^{m-1} + \dots + b_0 \Rightarrow \phi^m(v) + b_{m-1} \phi^{m-1}(v) + b_1 \phi(v) + b_0 v = 0$$

$$\underbrace{\phi(\phi^{m-1}(v))}_{v_m} = \phi^m(v) = -b_0 v - b_1 \phi(v) - \dots - b_{m-1} \phi^{m-1}(v)$$

$$A = M_B(\phi) = \left(\begin{array}{c|c} C_{f^a} & M \\ \hline 0 & N \end{array} \right)$$

$$p_\phi = |A - xI_n| = \left| \begin{array}{c|c} C_{f^a} - xI_m & M \\ \hline 0 & N - xI_{n-m} \end{array} \right| \stackrel{\text{Triangular inferior}}{=} |C_{f^a} - xI_m| \cdot |N - xI_{n-m}| = p_{C_{f^a}} \cdot p_N = \pm f^a \cdot p_N$$

Por tanto, $f^a | p_\phi$ □

A partir de aquí, tenemos $\begin{cases} q_\phi = f_1^{a_1} \cdot \dots \cdot f_t^{a_t} \\ \pm p_\phi = f_1^{e_1} \cdot \dots \cdot f_t^{e_t} \cdot f_{t+1}^{e_{t+1}} \cdot \dots \cdot f_s^{e_s} \end{cases} \quad 1 \leq a_i \leq e_i$

Veamos que los factores $f_{t+1}^{e_{t+1}} \cdot \dots \cdot f_s^{e_s}$ no pueden existir, es decir, que si $f \in K[x]$ irreducible y $f|p_\phi \Rightarrow f|q_\phi$. La idea de la demostración se basa en la reducción al caso de que f fuera de grado 1, utilizando la técnica de extensión del cuerpo a un cuerpo mayor donde sabemos que f descompone completamente.

Sea $K \subset \tilde{K} : p_\phi$ descompone completamente en $\tilde{K}[x]$. Sea $A = M_B(\phi) \in \text{Mat}_n(K) \subseteq \text{Mat}_n(\tilde{K})$

Sea $\tilde{V} = \tilde{K}^n$ un \tilde{K} -espacio vectorial. En este espacio \tilde{K}^n fijamos la base \tilde{B}_c .

Definimos $\tilde{\phi} \in \text{End}(\tilde{K}^n)$ tal que $M_{\tilde{B}_c}(\tilde{\phi}) = A$. Se cumple que:

- $p_{\tilde{\phi}} = |A - xI_n| = p_\phi$
- $\tilde{K}[x] \ni q_{\tilde{\phi}} | q_\phi \in K[x]$

Razonamos por reducción al absurdo. Supongamos que $f \nmid q_\phi \xRightarrow{f \text{ irreducible}} 1 = \text{mcd}(f, q_\phi)$. Por Bézout:

$$1 = a \cdot f + b \cdot q_\phi$$

Por otra parte, ya hemos probado que $p_\phi = p_{\tilde{\phi}}$ y $q_{\tilde{\phi}}$ tienen las mismas raíces. Sea

$$\tilde{\lambda} \in \tilde{K} : \underset{p_{\tilde{\phi}}(\tilde{\lambda})=0}{f(\tilde{\lambda})} = 0 \Rightarrow x - \tilde{\lambda} | f$$

Sea $\tilde{v} \neq 0 : \tilde{\phi}(\tilde{v}) = \tilde{\lambda} \cdot \tilde{v}$, es decir, que \tilde{v} es vector propio del autovalor $\tilde{\lambda}$. Descomponiendo $f = g(x)(x - \tilde{\lambda})$ tenemos que:

$$1 = a(x)g(x)(x - \tilde{\lambda}) + b(x)q_\phi(x)$$

Usando el endomorfismo $\tilde{\phi}$:

$$id_{\tilde{V}} = a(\tilde{\phi}) \circ g(\tilde{\phi}) \circ (\tilde{\phi} - \tilde{\lambda} \cdot id_{\tilde{V}}) + b(\tilde{\phi}) \circ q_\phi(\tilde{\phi})$$

OBS: $q_\phi(\tilde{\phi}) = 0$

$$id_{\tilde{V}} = a(\tilde{\phi}) \circ g(\tilde{\phi}) \circ (\tilde{\phi} - \tilde{\lambda} \cdot id_{\tilde{V}})$$

Aplicamos esta identidad entre aplicaciones lineales al vector \tilde{v} elegido previamente:

$$\tilde{v} = id_{\tilde{V}}(\tilde{v}) = a(\tilde{\phi}) \left(g(\tilde{\phi})(\tilde{\phi}(\tilde{v}) - \tilde{\lambda}\tilde{v}) \right)$$

Por construcción de $\tilde{v} : \tilde{\phi}(\tilde{v}) - \tilde{\lambda}\tilde{v} = 0$

$$\tilde{v} = a(\tilde{\phi}) \left(g(\tilde{\phi})(\tilde{\phi}(0)) \right) = 0(!)$$

Por tanto, todos los irreducibles de q_ϕ y de p_ϕ son los mismos, cambiando únicamente los exponentes. Así,

$$q_\phi | p_\phi$$

□

Teorema (de Caley-Hamilton)

Todo endomorfismo es raíz de su polinomio característico, es decir,

$$p_\phi(\phi) = 0$$

Demostración:

La demostración de este teorema consiste en probar que el polinomio mínimo siempre divide al polinomio característico:

$$q_\phi | p_\phi$$

Porque implica directamente que $p_\phi(\phi) = 0$

Proposición

Sea q_ϕ el polinomio mínimo y p_ϕ el polinomio característico, entonces poseen los mismos irreducibles en su descomposición salvo la multiplicidad de estos, que es mayor en la del característico.

$$1 \leq a_i \leq e_i : q_\phi = f_1^{a_1} \cdots f_t^{a_t} \wedge \pm p_\phi = f_1^{e_1} \cdots f_t^{e_t} \cdot f_{t+1}^{e_{t+1}} \cdots f_s^{e_s}$$

Demostración:

Sea $q_\phi = f_1^{a_1} \cdots f_t^{a_t}$ y $\pm p_\phi = f_1^{e_1} \cdots f_t^{e_t} \cdot f_{t+1}^{e_{t+1}} \cdots f_s^{e_s}$. Vamos a ver que estos factores adicionales no pueden existir, y en consecuencia solo cambian los exponentes de cada factor irreducible, pero no los factores irreducibles. Para ello basta con demostrar que $f | p_\phi \Rightarrow f | q_\phi$

Sea $f \in K[x] : f$ es irreducible y $f | p_\phi \Rightarrow f | q_\phi$. Vamos a extender el cuerpo K a un cuerpo \tilde{K} de manera que $K \subset \tilde{K}$ y p_ϕ descompone en $\tilde{K}[x]$. Entonces tenemos que la matriz de A también tiene sentido en $\tilde{K} : A = M_B(\phi) \in Mat_n(K) \subset Mat_n(\tilde{K})$. Si fijamos una base en un espacio del nuevo cuerpo, entonces ahora A puede representar la matriz de otro endomorfismo dentro de ese cuerpo, por lo que ahora si llamamos $\tilde{V} = \tilde{K}^n$ donde es \tilde{K} -espacio vectorial y definimos \tilde{B}_c podemos considerar $\tilde{\phi} \in End(\tilde{K}^n) : M_{\tilde{B}_c}(\tilde{\phi}) = A$.

Ahora tenemos que $p_{\tilde{\phi}} = |A - x \cdot I_n|$, pero este mismo polinomio es el de ϕ , luego $p_{\tilde{\phi}} = |A - x \cdot I_n| = p_\phi$. Del mismo modo, tenemos que $q_{\tilde{\phi}} \in \tilde{K}[x]$ y $q_\phi \in K[x]$, pero por el cuerpo al que pertenece cada uno se tiene que $q_{\tilde{\phi}} \in \tilde{K}[x] | q_\phi \in K[x]$

Por reducción a lo absurdo, supongamos que $f \nmid q_\phi$, este caso como f es irreducible, entonces $1 = mcd(f, q_\phi)$ y por el Teorema De Bezout se tiene que $1 = a \cdot f + b \cdot q_\phi$.

Por otro lado, sabemos que $p_{\tilde{\phi}}$ y $q_{\tilde{\phi}}$ tienen las mismas raíces, por lo que sea $\tilde{\lambda} \in \tilde{K} : f(\tilde{\lambda}) = 0$. Como $p_{\tilde{\phi}} = p_\phi$ entonces $f | p_{\tilde{\phi}}$ y como $gr(f) \geq 1$ entonces para alguno de estos landa $f(\tilde{\lambda}) = 0$, lo que implica que $x - \tilde{\lambda} | f$ y por tanto $p_{\tilde{\phi}}(\tilde{\lambda}) = 0$, es decir, $\tilde{\lambda}$ es un autovalor. Con lo cual tenemos vectores propios, así que sea $\tilde{v} = 0 : \tilde{\phi}(\tilde{v}) = \tilde{\lambda} \cdot \tilde{v}$, tenemos recuperando la identidad de Bezout anterior que:

$$1 = a(x)g(x)(x - \tilde{\lambda}) + b(x)q_\phi(x) \Rightarrow id_{\tilde{v}} = a(\tilde{\phi}) \circ g(\tilde{\phi}) \circ (\tilde{\phi} - \tilde{\lambda}id_{\tilde{v}}) + b(\tilde{\phi}) \circ \underbrace{q_\phi(\tilde{\phi})}_{=0} = a(\tilde{\phi}) \circ g(\tilde{\phi}) \circ (\tilde{\phi} - \tilde{\lambda}id_{\tilde{v}})$$

Si ahora lo aplicamos al $\tilde{v} \neq 0$ elegido previamente tenemos que:

$$\tilde{v} = a(\tilde{\phi}) \left(g(\tilde{\phi}) \left(\underbrace{\tilde{\phi} - \tilde{\lambda}\tilde{v}}_{=0} \right) \right) \Rightarrow \tilde{v} = 0 \Rightarrow \#$$

CLASES NOTABLES DE ENDOMORFISMOS

Tal y como dijimos al comienzo del capítulo, el objetivo de este desarrollo es poder hacer una clasificación exhaustiva de los endomorfismos a través de la clase de equivalencia a la que pertenezcan. Además, nos gustaría que para escoger el representante de su clase de equivalencia escojamos la matriz más sencilla posible, esto es, aquella que permita un cálculo computacional más sencillo a la hora de trabajar con la matriz.

Definición (Subespacios Propios)

Sea V un espacio vectorial y $\phi \in \text{End}(V)$, se define un **subespacio propio** para el autovalor λ como

$$V_{\phi,\lambda} = \{v \in V : \phi(v) = \lambda v\}$$

aquellos generados por vectores propios.

Observación:

La comprobación de que es un subespacio es trivial, puesto que se trata de $V_{\phi,\lambda} = \ker(\phi - \lambda \text{id})$ y ya demostramos en su momento que esto se trata de subespacios cuando ϕ eran aplicaciones lineales.

Definición (Multiplicidades Algebraicas y Dimensión geométrica)

Sea $\phi \in \text{End}(V)$ cuyo polinomio característico es

$$\pm p_\phi = (x - \lambda_1)^{e_1} \cdots (x - \lambda_t)^{e_t}$$

llamamos

- **multiplicidades algebraicas** a los e_i de los autovalores
- **dimensión geométrica** a $d_i = \dim_K V_{\phi,\lambda_i}$ de los autovalores.

Observación:

En el fondo, la dimensión geométrica puede verse como el número de vectores propios linealmente independientes dentro del subespacio $V_{\phi,\lambda}$ de cada autovalor λ_i .

Proposición

Sea $\phi \in \text{End}(V)$ y λ un autovalor suyo, entonces:

$$\forall i \in \mathbb{N} : d_i \leq e_i$$

Demostración:

Como d es la dimensión de un subespacio consideramos una base del mismo: $B = \{v_1, \dots, v_d\}$ base de $V_{\phi,\lambda}$ y ahora completamos dicha base a una base de V siendo esta $B' = \{v_1, \dots, v_d, v_{d+1}, \dots, v_n\}$ así que vamos a expresar la matriz de ϕ en base B' :

$$M_{B'}(\phi) = \left(\begin{array}{c|c} \lambda I_d & M \\ \hline 0 & N \end{array} \right)$$

Esto ocurre así porque los primeros d vectores de la base van a ellos mismos multiplicados por λ (son autovectores) y las otras cajas no son tan relevantes.

Si calculamos ahora su polinomio característico vemos que:

$$p_\phi = \left| \begin{array}{c|c} (\lambda - x)I_d & M \\ \hline 0 & N - xI_{n-d} \end{array} \right| = (\lambda - x)^d p_N$$

Pero como sabemos que $p_\phi = (x - \lambda)^e g(x)$ vemos claramente que $e \geq d$.

Diagonalizabilidad

La primera clase de endomorfismos que nos encontramos son los diagonalizables. Estos tienen una propiedad excelente a la hora de trabajar con ellos: sólo tienen valores distintos de 0 en la diagonal principal. Dicha característica confiere una gran ventaja a trabajar con esta matriz adecuada, por ejemplo, para calcular las potencias de matrices o las imágenes sucesivas de la aplicación.

Definición (Diagonalizabilidad)

Sea $\phi \in \text{End}(V)$, decimos que es **diagonalizable** si existe alguna base B respecto de la cual su matriz $M_B(\phi)$ sea diagonal.

Ejemplo:

Sea $\phi : V \rightarrow V$ una simetría, vimos en su momento que $V = W \oplus U$ con $\dim V = n$, $\dim W = m$ y $\dim U = m - n$. De este modo, una base $B = \underbrace{\{v_1, \dots, v_m\}}_{B_W} \cup \underbrace{\{v_{m+1}, \dots, v_n\}}_{B_U}$ determina de modo único

la matriz del endomorfismo que quedaría, con los vectores escogidos en B , de la forma

$$M_B(\phi) = \left(\begin{array}{c|c} I_n & 0 \\ \hline 0 & -I_{n-m} \end{array} \right)$$

Proposición

Sea $\phi \in \text{End}(V)$, entonces es diagonalizable si y sólo si V posee alguna base formada por vectores propios.

Demostración:

Por un lado, tenemos que \Rightarrow se demuestra como:

$$\exists B = \{v_1, \dots, v_n\} : M_B(\phi) = D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \Rightarrow \phi(v_i) = \lambda_i \cdot v_i \Rightarrow v_i \text{ autovector}$$

Por otro, la implicación contraria \Leftarrow es viendo que:

$$B \text{ base : } \phi(v_i) = \lambda_i v_i : i = 1, \dots, n \Rightarrow M_B(\phi) = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Observación:

Este resultado es uno de los primeros que nos permite discutir rápidamente si un endomorfismo es diagonalizable. Basta con calcular sus vectores propios y comprobar si se puede formar una base con ellos.

Ejemplo:

Supongamos que tenemos el endomorfismo definido por:

$$M_{B_c}(\phi) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \Rightarrow \begin{cases} p_\phi = (x+1)^2(x-2) \\ q_\phi = (x+1)(x-2) \end{cases}$$

los autovalores son las raíces del polinomio mínimo y característico, que son $\lambda_1 = -1$ y $\lambda_2 = 2$. El primer subespacio propio $V_{\phi, \lambda_1} = \ker(\phi + id) \equiv x + y + z = 0$ es un plano, lo que permite encontrar dos vectores linealmente independientes y el segundo $V_{\phi, \lambda_2} = \ker(\phi - 2id) \equiv \begin{cases} x - 2y + z = 0 \\ x + y - 2z = 0 \end{cases}$ es una recta, que aporta un único vector linealmente independiente.

Por tanto, si unimos los tres vectores comentados se tiene que $B' = \{(e_1 - e_2), (e_1 - e_3), (e_1 + e_2 + e_3)\}$ y basta comprobar su independencia para poder afirmar que es base. En dicha base, la expresión matricial de ϕ sería:

$$A' = M_{B'}(\phi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Proposición

Sea $\phi \in \text{End}(V)$ y λ_1, λ_2 autovalores distintos, los vectores propios de V_{ϕ, λ_1} son independientes de los de V_{ϕ, λ_2} .

Demostración:

$$a_1 v_1 + a_2 v_2 = 0 \Rightarrow \phi(a_1 v_1 + a_2 v_2) = \phi(0) = 0 \Rightarrow a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 = 0$$

Por otro lado tenemos que:

$$a_1 v_1 + a_2 v_2 = 0 \Rightarrow a_1 \lambda_1 v_1 + a_2 \lambda_1 v_2 = 0$$

Restando ambas:

$$a_2(\lambda_2 - \lambda_1)v_2 = 0 \Rightarrow a_2 = 0$$

Y tenemos que $a_1 v_1 = 0 \Rightarrow a_1 = 0$

Proposición

Sea $\phi \in \text{End}(V)$, entonces ϕ es diagonalizable si y sólo si p_ϕ descompone completamente en $K[x]$ y $d_i = e_i$.

Demostración:

■ \Rightarrow

Supongamos que ϕ es diagonalizable, entonces existe una base donde su matriz es diagonal, es decir,

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \Rightarrow p_\phi = |A - x \cdot I_n| = (\lambda_1 - x) \cdots (\lambda_n - x) = (-1)^n \cdot (x - \lambda_1) \cdots (x - \lambda_n)$$

Como es diagonalizable por hipótesis, entonces existe una base formada por vectores propios: $B = \{\underbrace{v_{11}, \dots, v_{1m_1}}_{V_{\phi, \lambda_1}}, \dots, \underbrace{v_{t1}, \dots, v_{tm_t}}_{V_{\phi, \lambda_t}}\}$. Entonces tenemos que $m_1 \leq d_1, \dots, m_t \leq d_t$. Como

es base tiene que haber n vectores y como sabemos que $gr(p_\phi) = \dim V = n = e_1 + \dots + e_t = m_1 + \dots + m_t \leq d_1 + \dots + d_t \Rightarrow (d_1 - e_1) + \dots + (d_t - e_t) = 0 \Rightarrow d_i = e_i$.

■ \Leftarrow

Si el polinomio descompone completamente en $K[x]$, entonces $gr(p_\phi) = n = e_1 + \dots + e_t$ y si suponemos que $d_i = e_i$, entonces se tiene que $n = e_1 + \dots + e_t = d_1 + \dots + d_t$. Pero como d_1 indica que hay d_1 vectores linealmente independientes para el primer autovalor, podemos ir construyendo con cada conjunto de vectores de cada d_i una base de n y como es una base formada por autovectores, entonces es diagonalizable.

Proposición

Sea $\phi \in \text{End}(V)$, entonces ϕ es diagonalizable si y sólo si q_ϕ descompone en $K[x]$ y las raíces son distintas.

Demostración:

■ \Rightarrow

Ya es conocido porque si p_ϕ lo es y $q_\phi \mid p_\phi$ ya está.

■ \Leftarrow

$$q_\phi = (x - \lambda_1) \cdots (x - \lambda_t) : \lambda_1 \neq \cdots \neq \lambda_t = (x - \lambda_i)g_i(x)$$

Vemos que $1 = \text{mcd}(g_1, \dots, g_t)$, luego $1 = g_1(x)a_1(x) + \dots + g_t(x)a_t(x)$. Tenemos pues que

$$id_v = g_1(\phi) \circ a_1(\phi) + \dots + g_t(\phi) \circ a_t(\phi)$$

y entonces ocurre:

$$\forall v \in V : v = g_1(\phi)(a_1(\phi)(v)) + \cdots + g_t(\phi)(a_t(\phi)(v))$$

Es decir, podemos expresar cada v como $v = v_1 + \cdots + v_t$ donde cada $v_i = g_i(\phi)(a_i(\phi)(v))$. Veamos que cada uno de esos sumandos es un vector propio:

$$\begin{aligned} v_i \in V_{\phi, \lambda_i} : \phi(v_i) - \lambda_i v_i &= (x - \lambda_i)(\phi)(v_i) = (x - \lambda_i)(\phi)(g_i(\phi)(a_i(\phi)(v))) = \\ &= [(x - \lambda_i)g_i(x)a_i(x)](\phi)(v) = q_\phi(\phi)(a_i(\phi)(v)) = 0(a_i(\phi)(v)) = 0 \end{aligned}$$

Entonces lo que tenemos es que cada vector v se puede expresar como la suma de v_1, \dots, v_t vectores, si consideramos B_i base de V_{ϕ, λ_i} (que genera v_i), entonces tenemos que $B_1 \cup \cdots \cup B_t$ es sistema de generadores de V y además como vimos que vectores propios de autovalores distintos son independientes, la unión de todas las bases es independiente y por tanto es base de V .

Ejemplo:

Vamos a aplicarlo al caso de las homologías: supongamos que $q_\phi = (x - 1)(x - r) : r \neq 0, 1, -1$ y vemos que como se cumplen las condiciones anteriores, entonces es diagonalizable. Entonces, por lo probado antes tenemos que $v = v_1 + v_r$ cada uno en uno de los subespacios propios. Vamos a ver que la intersección es 0; $v_1 = v_r \Rightarrow \phi(v_1) = v_1 \wedge r \cdot v_1 = 0 \Leftrightarrow (1 - r)v_1 = 0 \Rightarrow v_1 = 0$. Tenemos entonces que $V = V_{\phi, 1} \oplus V_{\phi, r}$, luego los primeros vectores van a ellos mismos y los segundos se multiplican por r .

Por ejemplo, en dimensión K^2 tenemos que las únicas aplicaciones lineales diagonalizables posibles responden a la siguiente afirmación:

$$\begin{cases} x - \lambda \\ (x - \lambda_1)(x - \lambda_2) : \lambda_1 \neq \lambda_2 \end{cases}$$

En el primer caso son homotecias de razón λ y en el segundo la matriz resultante es:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \Rightarrow \phi = \phi_1 \circ \phi_2$$

Es decir, cualquier aplicación en el plano que sea diagonalizable y no sea una homotecia es producto de dos homologías generales. Y esto es aplicable a cualquier dimensión con el correspondiente aumento de producto de homologías.

Potencia enésima de matrices

Una aplicación bastante buena de esta parte de teoría es hallar la fórmula general de la sucesión geométrica cuyo término base es la matriz A . Para ello hay que distinguir dos casos:

■ Diagonalizable

Este caso es el más sencillo puesto que sabemos que si una matriz es diagonalizable, es posible convertirla en su matriz diagonal asociada a través de una matriz de cambio de base.

Da la casualidad de que la potencia enésima de una matriz diagonal es muy fácil de calcular por lo que es relativamente sencillo calcular el término general de la matriz diagonal y después transformar a la matriz original a través de esa matriz de paso.

Ej.:

Sea $A = \begin{pmatrix} -7 & -6 \\ 12 & 10 \end{pmatrix}$ su polinomio característico es $p_\phi = (x - 2)(x - 1)$, luego su matriz diagonal es: $D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. El término general para la matriz diagonal es $D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 1^n \end{pmatrix}$

luego a través de la matriz de cambio de base $P = M(B, B') = \begin{pmatrix} -4 & -3 \\ 3 & 2 \end{pmatrix}$ se tiene que:

$$A = P^{-1}DP \Rightarrow A^n = P^{-1}D^nP = \begin{pmatrix} -8 \cdot 2^n + 9 & -6 \cdot 2^n + 6 \\ 12 \cdot 2^n - 12 & 9 \cdot 2^n - 8 \end{pmatrix}$$

■ NO diagonalizable

Este caso es más complejo y se explicará el procedimiento general a través de un ejemplo (aunque explicando los pasos a seguir y las partes).

Ejemplo:

Tomemos por ejemplo la matriz:

matriz

Triangularizabilidad

Las condiciones para que una matriz sea diagonalizable son muy restrictivas pues obligan al endomorfismo a poder formar una base de vectores propios. En general, estas condiciones tan específicas no se pueden dar y debemos buscar otra forma “sencilla” para las matrices de aquellos endomorfismos no diagonalizables. Una de las posibles soluciones es poder triangularizar la matriz, lo que no logra la simplicidad de la diagonalización pero sí simplifica los cálculos al obtener muchos 0 dentro de la matriz.

Definición (Triangularizabilidad)

Sea $\phi \in \text{End}(V)$, decimos que ϕ es **triangularizable** si existe alguna base B de V respecto de la cual la matriz de ϕ sea triangular.

Observación:

En el documento se va a desarrollar la teoría para conseguir matrices triangulares inferiores, esto es, con los 0 bajo la diagonal principal. Si se desea una matriz triangular superior, basta con invertir el orden de los vectores de la base en la que es triangular inferior.

Teorema

Sea $\phi \in \text{End}(V)$, entonces ϕ es triangularizable si y sólo si p_ϕ descompone en $K[x]$.

Demostración:

■ \Rightarrow

Supongamos que:

$$M_B(\phi) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nm} \end{pmatrix} \Rightarrow p_\phi = |A - xI| = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1m} \\ 0 & a_{22} - x & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nm} - x \end{vmatrix} =$$

$$= (a_{11} - x) \cdots (a_{nm} - x)$$

Y cabe destacar que cada elemento de la diagonal, es un autovalor correspondiente del endomorfismo.

■ \Leftarrow

Partimos ahora de que el polinomio característico descompone completamente, luego:

$$p_\phi = (x - \lambda_1) \cdots (x - \lambda_n)$$

Por inducción sobre n :

$$n = 1 \Rightarrow M_B(\phi) = (a_{11}) \Rightarrow Ok$$

Ahora vemos $n - 1 \stackrel{?}{\Rightarrow} n$:

$$p_\phi = (x - \lambda_1) \cdots (x - \lambda_n)$$

Como λ_1 es autovalor, $\exists v_1 \neq 0 : \phi(v_1) = \lambda_1 v_1$ y entonces puedo ampliar este vector a una base de V que será $B = \{v_1, w_2, \dots, w_n\}$ con w_i arbitrarios. Llamamos $W < V$ al subespacio generado por $\{w_2, \dots, w_n\}$ y se ve que la dimensión de W es $n - 1$. Por tanto la matriz resultante queda:

$$A = M_B(\phi) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Consideramos ahora esta matriz una matriz por cajas de forma que:

$$M_B(\phi) = \begin{pmatrix} \lambda_1 & M \\ 0 & N \end{pmatrix}$$

Puedo considerar N como $M_{\{w_2, \dots, w_n\}}(\psi)$ de forma que $\psi : W \rightarrow W$. El polinomio característico de la matriz anterior es $p_\phi = |A - xI| = (\lambda_1 - x)p_N = (\lambda_1 - x)p_\psi$ luego tenemos que $p_\psi = (x - \lambda_2) \cdots (x - \lambda_n)$ y aplicando la hipótesis de inducción tenemos que ψ es triangularizable con lo cual como en la columna de λ_1 de la matriz A ya tiene todo ceros debajo ya está.

FORMA CANÓNICA DE JORDAN

De nuevo, aunque las clases anteriores nos dan unas muy buenas representaciones para cierto tipo de endomorfismos, no son válidas para todos en general. Nuestro objetivo en esta sección es desarrollar una estructura general, canónica, para todos los endomorfismos que encontremos. Esta forma será diagonal o triangular para los endomorfismos que sí cumplan las condiciones de la sección anterior, pero tendrá una nueva representación para aquellos que no encajen en los conceptos vistos.

Definición (Subespacios Invariantes)

Sea $\phi : V \rightarrow V$ y $W < V$ se dice que W es un **subespacio invariante** si es cerrado bajo ϕ , es decir, $\phi(W) \subset W$.

Observación:

El resultado de esta definición es que si W es un subespacio ϕ -invariante (que denotaremos por $W <_\phi V$), entonces la restricción de ϕ al subespacio es de nuevo un endomorfismo, pero esta vez sobre W :

$$\phi|_W : W \rightarrow W \Rightarrow \phi|_W \in \text{End}(W)$$

Además, podemos ver que $\phi(L(v_1)) \subset L(v_1) \Leftrightarrow \phi(v_1) = \lambda v_1$, luego las rectas invariantes son las generadas por vectores propios.

Proposición

Sea $\phi \in \text{End}(V)$, entonces los subespacios propios $V_{\phi, \lambda}$ y $\ker \phi$ son subespacios invariantes del mismo.

$$\phi(V_{\phi, \lambda}) \subset V_{\phi, \lambda} \qquad \ker \phi <_\phi V$$

Proposición

Sea $\phi \in \text{End}(V)$ un endomorfismo y $W <_\phi V$ un subespacio invariante del mismo, entonces:

- El polinomio característico de la restricción a W divide al polinomio característico del endomorfismo inicial, es decir:

$$p_{\phi|_W} \mid p_\phi$$

- El polinomio mínimo de la restricción a W divide al polinomio mínimo del endomorfismo inicial, es decir:

$$q_{\phi|_W} \mid q_{\phi}$$

Observación:

El resultado anterior es fundamental, puesto que nos permite caracterizar los subespacios invariantes de la dimensión que queramos. Si tenemos un polinomio característico concreto, los subespacios invariantes de dimensión 2 solo podrán ser aquellos para los cuales el polinomio característico de la restricción a ellos mismos tengan 2 factores del polinomio característico original, etc.

Lema

Sea $\phi \in \text{End}(V)$ y $g \in K[x] : g(\phi) = 0$ que podemos factorizar como:

$$g = g_1 \cdot \dots \cdot g_t : g_i \in K[x] : \text{mcd}(g_i, g_j) = 1$$

entonces se cumple que:

$$V =_{\phi} \ker g_1(\phi) \oplus \dots \oplus \ker g_t(\phi)$$

Demostración:

Llamamos a $v_i = \ker g_i(\phi)$ vamos a ver que todo vector de V se expresa como suma de dichos v_i y que la intersección de todos es nula.

Nos fijamos solo en uno de los factores:

$$\begin{aligned} g = g_i h_i &\Rightarrow \text{mcd}(h_1, \dots, h_t) = 1 \Rightarrow 1 = h_1(x)a_1(x) + \dots + h_t(x)a_t(x) \Rightarrow id_v = h_1(\phi) \circ a_1(\phi) + \dots + h_t(\phi) \circ a_t(\phi) \Rightarrow \\ &\Rightarrow v = [h_1(\phi) \circ a_1(\phi)](v) + \dots + [h_t(\phi) \circ a_t(\phi)](v) \Rightarrow v = v_1 + \dots + v_t : v_i = [h_i(\phi) \circ a_i(\phi)](v) \end{aligned}$$

Ahora tenemos que ver que cada uno de esos vectores está en su núcleo correspondiente:

$$\begin{aligned} g_i(\phi)(v_i) &= [g_i(\phi) \circ h_i(\phi) \circ a_i(\phi)](v) = [g_i(x)h_i(x)a_i(x)(\phi)](v) = [a_i(x)g_i(x)](\phi)(v) = a_i(\phi)(g(\phi)(v)) = a_i(\phi)(0) = 0 \\ &\Rightarrow v_i \in \ker g_i(\phi) \end{aligned}$$

Ahora vamos a ver que si la suma de todos es el vector nulo, entonces necesariamente todos son el vector nulo (eso es que la intersección es nula):

$$v_1 + \dots + v_t = 0 \Rightarrow 0 = h_i(\phi)(v_1 + \dots + v_t) = h_i(\phi)(v_i)$$

Por otra parte tenemos que $g_i(\phi)(v_i) = 0$ pero habíamos definido que:

$$\text{mcd}(g_i, h_i) = 1 \Rightarrow 1 = a_i(x)g_i(x) + b_i(x)h_i(x) \Rightarrow v_i = [a_i(\phi) \circ g_i(\phi)](v_i) + [b_i(\phi) \circ h_i(\phi)](v_i) = 0 \Rightarrow v_i = 0$$

Definición (Matriz diagonal por cajas)

Sea V un espacio vectorial de modo que $V =_{\phi} V_1 \oplus \dots \oplus V_t$ y $\phi \in \text{End}(V)$. Si escogemos B de forma que esté formada por unión de las bases de cada subespacio, es decir, $B = B_1 \cup \dots \cup B_t$, entonces la matriz de ϕ queda³ como:

$$M_B(\phi) = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_t \end{pmatrix}$$

donde cada $A_i = M_{B_i}(\phi|_{V_i})$.

Además, para denotar una matriz diagonal por cajas utilizaremos la siguiente notación:

$$A = A_1 \oplus \dots \oplus A_t$$

³Los 0 de la matriz denotan matrices nulas de las dimensiones adecuadas

Ejemplo:

Supongamos que $q_\phi = \underbrace{(x - \lambda_1) \cdots (x - \lambda_t)}_{g_1} : \lambda_i \neq \lambda_j$ con lo cual sabemos que es diagonalizable.

Vemos que $\text{mcd}(g_i, g_j) = 1$ osea que se cumple la condición del lema porque $q_\phi(\phi) = 0$ por lo que si llamo $V_i = \ker\{(x - \lambda_i)(\phi)\} = \ker(\phi - \lambda_i \text{id}_V) = V_{\phi, \lambda_i}$ y esto quiere decir que $V = V_1 \oplus \cdots \oplus V_t$ y eso es diagonalizable.

Teorema de Estructura

Aplicando el lema anterior al polinomio mínimo y característico, vemos que el endomorfismo se puede descomponer como la suma directa de los subespacios generados por el ker de cada factor del polinomio mínimo y característico⁴, es decir, podemos simplificar el estudio de un endomorfismo al estudio de cada uno de los trozos en los que se divide su polinomio característico y después juntar todas las bases para obtener la matriz escalonada por cajas de antes.

$$q_\phi = f_1^{a_1} \cdots f_t^{a_t} \quad p_\phi = f_1^{e_1} \cdots f_t^{e_t}$$

Por el lema anterior, podemos dividir V en suma directa de los subespacios del ker de cada factor:

$$V = {}_\phi V_1 \oplus \cdots \oplus V_t : V_i = \ker f_i^{a_i}(\phi) \quad V = {}_\phi V'_1 \oplus \cdots \oplus V'_t : V'_i = \ker f_i^{e_i}(\phi)$$

Ahora mismo, estamos tratando de ver que la descomposición según uno u otro es indiferente, puesto que generan los mismos subespacios:

$$\ker f_i(\phi) \subset \ker f_i^2(\phi) \subset \cdots \subset \ker f_i^{a_i}(\phi) \subset \cdots \subset \ker f_i^{e_i}(\phi) \Rightarrow V_i \subset V'_i$$

Como la suma directa de ambas descomposiciones son el mismo espacio vectorial V y cada una está contenida en su homóloga, entonces $\forall i \in \mathbb{N} : \dim V_i = \dim V'_i$ porque si no la suma de las dimensiones daría resultados distintos, por tanto, por estar uno metido en el otro y tener la misma dimensión $V_i = V'_i$. Y con este razonamiento ya tendríamos dividido de forma única el espacio V :

$$V = {}_\phi V_1 \oplus \cdots \oplus V_t : \forall i \in \mathbb{N}, V_i = \ker f_i^{a_i}(\phi) = \ker f_i^{e_i}(\phi)$$

Como se tratan del núcleo de una aplicación lineal, sabemos que la suma directa es de subespacios ϕ -invariantes y, por tanto, se induce un endomorfismo en cada subespacio generado por la restricción $\phi_i = \phi|_{V_i} : V_i \rightarrow V_i$. Una vez hecho esto, definimos B_i base de V_i y llamamos $A_i = M_{B_i}(\phi_i)$ y generamos la base total uniendo las de cada trozo $B = B_1 \cup \cdots \cup B_t$ y llamando a $A = M_B(\phi)$ quedaría la matriz del total como: $A = A_1 \oplus \cdots \oplus A_t$.

Tratamos ahora de estudiar estos endomorfismos generados por la restricción del inicial a los subespacios ϕ -invariantes. Para ello, vamos a intentar ver que el polinomio mínimo y característico de la restricción coincide precisamente con el factor del polinomio sobre el que estamos estudiando. En primer lugar, tenemos⁵ que

$$q_{\phi_i} \mid f_i^{a_i} \Rightarrow q_{\phi_i} = f_i^{b_i} : b_i \leq a_i$$

y hay que ver que exponente es precisamente a_i , para lo que suponemos $q_{\phi_1} = f_1^{b_1} : b_1 < a_1$. De esta forma, consideramos que:

$$f_1^{a_1-1} \cdots f_2^{a_2} \cdots f_t^{a_t}$$

y entonces $f_1^{a_1-1} \cdots f_2^{a_2} \cdots f_t^{a_t}(\phi) = 0$, pero eso diría que $q_\phi \mid f_1^{a_1-1} \cdots f_2^{a_2} \cdots f_t^{a_t}$ que es absurdo, luego:

$$\boxed{q_{\phi_i} = f_i^{a_i}}$$

En segundo lugar, con el característico que ocurre lo mismo, es el mismo polinomio que el mínimo pero elevado a un exponente que puede ser igual o mayor:

$$q_{\phi_i} \mid p_{\phi_i} \Rightarrow \pm p_{\phi_i} = f_i^{c_i} : a_i \leq c_i$$

⁴Porque ya vimos que coinciden

⁵Ya que sabemos que la restricción se anula en el trozo de polinomio que le toca.

Como sabemos que A es diagonal por cajas, entonces $A = A_1 \oplus \cdots \oplus A_t$ y sucede:

$$\begin{aligned} p_\phi &= |A - xI_n| = |(A_1 - xI_{m_1}) \oplus \cdots \oplus (A_t - xI_{m_t})| = |A_1 - xI_{m_1}| \cdots |A_t - xI_{m_t}| = p_{\phi_1} \cdots p_{\phi_t} \Rightarrow \\ &\Rightarrow p_\phi = f_1^{c_1} \cdots f_t^{c_t} = p_{\phi_1} \cdots p_{\phi_t} \Rightarrow gr(p_{\phi_i}) = c_i \Rightarrow e_i = c_i \\ &\boxed{p_{\phi_i} = f^{e_i}} \end{aligned}$$

Observación:

Del hecho anterior, deducimos que $\dim V_i = gr(f_i^{e_i})$, es decir, que el grado del polinomio característico siempre coincide con la dimensión del espacio V sobre el que actúa ϕ .

Teorema (Subespacios Cíclicos)

Sea $v \in V$ y $\phi \in \text{End}(V)$, definimos el conjunto

$$L_\phi(v) = \{g(\phi)(v) : g \in K[x]\} = \{g(\phi)(v) : g \in K[x], gr(g) < gr(q_{\phi,v})\}$$

es decir, todas las posibles combinaciones lineales⁶ de potencias de ϕ evaluadas en v y que verifica las siguientes propiedades:

- $L_\phi(v) <_\phi V$
- $B_v = \{v, \phi(v), \dots, \phi^{m-1}(v)\}$ es una base siendo $m = gr(q_{\phi,v})$.
- $M_{B_v}(\phi|_{L_\phi(v)}) = C_{q_{\phi,v}}$
- $L_\phi(v)$ es el subespacio ϕ -invariante **más pequeño** que contiene al vector v

Demostración:

- Veamos que es cerrado bajo ϕ :

$$g(\phi)(v) \in L_\phi(v) \Rightarrow \phi(g(\phi)(v)) = x(\phi)(g(\phi)(v)) = [(x \cdot g(x))(\phi)](v) = h(\phi)(v) : h \in K[x] \Rightarrow \in L_\phi(v)$$

- Veamos que son independientes y sistema de generadores:

$$a_0v + a_1\phi(v) + \cdots + a_{m-1}\phi^{m-1}(v) = 0 \Rightarrow (a_0 + a_1x + \cdots + a_{m-1}x^{m-1})(\phi)(v) \stackrel{gr(\cdots) \leq gr(q)}{\Rightarrow} 0$$

$$\Rightarrow a_0 + a_1x + \cdots + a_{m-1}x^{m-1} = 0 \Rightarrow a_i = 0 : \forall i$$

$$\begin{aligned} g \in K[x] &= [g(x)(\phi)](v) = [h(x)q_{\phi,v}(x) + a_0 + a_1x + \cdots + a_{m-1}x^{m-1}](\phi)(v) = \\ &= [(a_0 + a_1x + \cdots + a_{m-1}x^{m-1})(\phi)](v) = a_0v + a_1\phi(v) + \cdots + a_{m-1}\phi^{m-1}(v) \end{aligned}$$

- Trivial por el tema de polinomios (o incluso haciendo las imágenes a mano).
- No procede.

Teorema (de Estructura)

Sea $\phi \in \text{End}(V)$ y $f \in K[x]$ irreducible, si los polinomios mínimo y característico son $p_\phi = \pm f_\phi^e$ y $q_\phi = f^a$, entonces:

$$\exists v_1, \dots, v_s \in V : V = L_\phi(v_1) \oplus \cdots \oplus L_\phi(v_s)$$

de manera que $a = a_1 \geq \cdots \geq a_s$, entendiendo dichos índices como⁷ $q_{\phi,v_i} = f^{a_i}$.

⁶La segunda igualdad de la definición del conjunto es válida porque cualquier potencia mayor que $gr(q_{\phi,v})$ será una de las potencias de ϕ menores que dicho grado, puesto que será combinación lineal de las mismas por lo razonado en el apartado del polinomio mínimo de un endomorfismo y un vector.

⁷Luego el cálculo de dichos v_i es muy sencillo, puesto que basta con calcular el subespacio $f^{a_i}(\phi)(v_i) = 0$ y descartar los vectores que hayamos encontrado en $a_i - 1$ o anteriores.

Demostración:

Vamos a hacer inducción sobre e :

- $e = 1$

$$e = 1 \Rightarrow a = 1 \Rightarrow \pm p_\phi = q_\phi = f \Rightarrow \exists B \text{ base de } V : M_B(\phi) = C_f$$

- H.I.: supongamos que $\bar{\phi} : \bar{V} \rightarrow \bar{V}$ y tenemos que $\pm p_{\bar{\phi}} = f^{\bar{e}}$ y $q_{\bar{\phi}} = f^{\bar{a}}$ de manera que $1 \leq \bar{a} \leq \bar{e} < e$ entonces $\bar{V} = L_{\bar{\phi}}(\bar{w}_2) \oplus \cdots \oplus L_{\bar{\phi}}(\bar{w}_s)$

Hemos visto que $\exists v_1 \in V : q_{\phi, v_1} = f^a$, luego $B_{v_1} = \{v_1, \dots, \phi^{m-1}(v_1)\}$ es base de $L_\phi(v_1)$ que es un subespacio invariante y también que $M_{B_{v_1}}(\phi|_{L_\phi(v_1)}) = C_{f^a}$.

Siempre podemos escribir $V = L_\phi(v_1) \oplus W$ donde W es una extensión de $L_\phi(v_1)$ hasta cubrir todo V . Luego su matriz queda como $M_B(\phi) = \begin{pmatrix} C_{f^a} & M \\ 0 & N \end{pmatrix}$, pero no está claro que pueda extender usando un W que sea ϕ -invariante y así quedara $M = 0$ y poder aplicar la hipótesis de inducción. Así que vamos a hacerlo al revés, cocientar V por $L_\phi(v_1)$ para hallar el W isomorfo que buscamos:

Definimos $\bar{V} = V/L_\phi(v_1)$ los elementos por tanto son clases de \bar{v} , es decir, $\bar{v} = \bar{w} \Rightarrow v - w \in L_\phi(v_1)$ y que un vector pertenezca a dicho subespacio quiere decir que son de la forma $g_1(\phi)(v_1)$. Veamos como se define $\bar{\phi}$ en este nuevo espacio vectorial $\bar{\phi} : \bar{V} \rightarrow \bar{V}$ de forma que $\bar{\phi}(\bar{v}) = \overline{\phi(v)}$. Vemos si $\bar{v} = \bar{w} \stackrel{?}{\Rightarrow} \overline{\phi(v)} = \overline{\phi(w)}$:

$$\overline{\phi(v)} = \overline{\phi(w)} \Leftrightarrow \phi(v) - \phi(w) \in L_\phi(v_1) \Leftrightarrow \phi(v - w) \in L_\phi(v_1) \stackrel{v-w \in L_\phi(v_1)}{\Leftrightarrow}$$

$$\Leftrightarrow \phi(g_1(\phi)(v_1)) \in L_\phi(v_1) \Leftrightarrow (xg_1(x))(\phi)(v_1) \in L_\phi(v_1) \Leftrightarrow h(x)(\phi)(v_1) \in L_{\bar{\phi}}(v_1)$$

Vemos que es lineal:

$$\bar{\phi}(\bar{v} + \bar{w}) = \overline{\phi(v + w)} = \overline{\phi(v) + \phi(w)} = \overline{\phi(v)} + \overline{\phi(w)} = \bar{\phi}(\bar{v}) + \bar{\phi}(\bar{w})$$

Del mismo modo, el producto por un escalar se verifica con normalidad.

Ahora vemos quienes son el polinomio mínimo y característico de $\bar{\phi}$, habíamos elegido que $q_{\phi, v_1} = f^a = q_\phi$, vamos a ver que $f^a(\bar{\phi})(\bar{v}) = 0$:

$$f^a(\bar{\phi})(\bar{v}) = \overline{f^a(\phi)(v)} = \overline{0(v)} = \bar{0} \Rightarrow f^a(\bar{\phi}) = 0 \Rightarrow q_{\bar{\phi}} \mid f^a \Rightarrow q_{\bar{\phi}} = f^{\bar{a}} : \bar{a} \leq a$$

Con el característico ocurre de forma igual, como tiene las mismas raíces que su mínimo nos queda que:

$$p_{\bar{\phi}} = f^{\bar{e}} : \bar{e} < e$$

Y es estrictamente menor porque: $gr(p_\phi) = n \dim V > \underbrace{\dim \bar{V}}_{n - \dim L_\phi}(v_1) = gr(f^{\bar{e}})$

Considerando que $q_{\bar{\phi}} = f^{\bar{a}}$, entonces $q_{\bar{\phi}, \bar{w}_i} = f^{b_i} \mid f^{\bar{a}} : b_i \leq \bar{a}$. Vamos a ver que significa esto último:

$$f^{b_i}(\bar{\phi})(\bar{w}_i) = 0 \Rightarrow \overline{f^{b_i}(\phi)(w_i)} \Rightarrow f^{b_i}(\phi)(w_i) \in L_\phi(v_1) \Rightarrow f^{b_i}(\phi)(w_i) = g_i(\phi)(v_1)$$

Ahora vamos a ver que:

$$f^{a-b_i}(\phi)(g_i(\phi)(v_1)) = \underbrace{f^{a-b_i}(\phi)(f^{b_i}(\phi)(w_i))}_{=f^a(\phi)(w_i)=0} = 0$$

Es decir, a $\phi(v_1)$ lo anula cierto polinomio y por lo tanto:

$$\underbrace{q_{\phi, v_1}}_{=f^a} \mid f^{a-b_i} g_i \Rightarrow f^{b_i} \mid g_i \Rightarrow g_i = f^{b_i} h_i$$

Ahora volviendo a tener en cuenta la igualdad anterior ocurre que:

$$f^{b_i}(\phi)(w_i) = f^{b_i}(\phi)(h_i(\phi)(v_1)) = f^{b_i}(\phi)[w_i - h_i(\phi)(v_1)] = 0$$

Ahora definimos el siguiente vector $v_i = w_i - h_i(\phi)(v_1)$. ¿Qué sabemos de dicho vector?, pues: $f^{b_i}(\phi)(v_i) = 0$, $\bar{v}_i = \bar{w}_i$ por como está definido v_i y $\underbrace{q_{\phi, v_i}}_{=f^{a_i}} \mid f^{b_i} : a_i \leq b_i$.

Es decir, en la descomposición inicial de $V = L_\phi(\bar{w}_2) \oplus \cdots \oplus L_\phi(\bar{w}_s)$ hemos cambiado de representante a v_i y hemos conseguido una sucesión de menor o iguales: $a_i \leq b_i \leq \bar{a} < a$.

Vamos a probar finalmente que $V = L_\phi(v_1) \oplus \cdots \oplus L_\phi(v_s)$, para ello vamos a demostrar dos cosas: que todo vector se descompone en suma de uno de esos vectores y que la suma de todos igualada a 0 obliga a que todos sean el vector nulo.

$$\begin{aligned} v \in V &\Rightarrow \bar{v} = g_2(\bar{\phi})(\bar{w}_2) + \cdots + g_s(\bar{\phi})(\bar{w}_s) = g_2(\bar{\phi})(\bar{v}_2) + \cdots + g_s(\bar{\phi})(\bar{v}_s) \Rightarrow \\ &\Rightarrow v - (g_1(\phi)(v_2) + \cdots + g_s(\phi)(v_s)) = g_1(\phi)(v_1) \Rightarrow v = g_1(\phi)(v_1) + \cdots + g_s(\phi)(v_s) \end{aligned}$$

Para ver la segunda condición:

$$\begin{aligned} 0 = g_1(\phi)(v_1) + \cdots + g_s(\phi)(v_s) &\Rightarrow -g_1(\phi)(v_1) = g_2(\phi)(v_2) + \cdots + g_s(\phi)(v_s) \Rightarrow \bar{0} = g_2(\bar{\phi})(\bar{v}_2) + \cdots + g_s(\bar{\phi})(\bar{v}_s) \Rightarrow \\ g_2(\bar{\phi})(\bar{v}_2) = \cdots = g_s(\bar{\phi})(\bar{v}_s) &= \bar{0} \Rightarrow g_2(\bar{\phi})(\bar{w}_2) = \cdots = g_s(\bar{\phi})(\bar{w}_s) = \bar{0} \end{aligned}$$

Porque como la suma es directa, necesariamente cada uno de ellos es 0. Ahora vemos que debido a esto:

$$f^{b_i} \mid g_i \Rightarrow g_i = f^{b_i} \cdot h_i \Rightarrow g_i(\phi)(v_i) = h_i(\phi)(\underbrace{f^{b_i}(\phi)(v_i)}_{=0}) = 0 \Rightarrow g_i(\phi)(v_i) = 0$$

Observación:

Este resultado es muy importante porque, si ya podíamos separar⁸ la matriz inicial en suma directa de los núcleos de las restricciones a cada factor del polinomio mínimo, ahora podemos a su vez dividir dichas restricciones en sumas directas de subespacios cíclicos. Esto provoca que la matriz de la restricción del endomorfismo sea una matriz diagonal por cajas y además que la matriz de ϕ restringido a $L_\phi(v_i)$ sea la compañera del polinomio mínimo de $\phi(v_i)$, es decir, $C_{f^{a_i}}$.

Teorema (de clasificación: forma canónica racional)

Sea $\phi \in \text{End}(V)$ y $f \in K[x]$ irreducible, si los polinomios mínimo y característico son $\pm p_\phi = f^e$ y $q_\phi = f^a$, entonces:

$$\exists B \text{ base de } V : M_B(\phi) = C_{f^{a_1}} \oplus \cdots \oplus C_{f^{a_s}} \text{ y } a = a_1 \geq a_2 \geq \cdots \geq a_s \geq 1$$

Esta descomposición es única y se la conoce como **matriz canónica racional** de ϕ . A los polinomios de las sucesivas potencias como **divisores elementales**:

$$f^{a_s} \mid f^{a_s+1} \mid \cdots \mid f^{a_2} \mid f^{a_1}$$

y caracterizan a los polinomios mínimo y característico porque:

$$\begin{cases} p_\phi = f^{a_s} \cdot \cdots \cdot \underbrace{f^{a_1}}_{=q_\phi} \\ q_\phi = f^{a_1} = f^a \end{cases}$$

Demostración:

Está clara la existencia, porque descomponemos $V = L_\phi(v_1) \oplus \cdots \oplus L_\phi(v_s)$, elegimos bases B_i de cada uno donde $B_i = \{v_i, \phi(v_i), \phi^2(v_i), \dots, \phi^{m_i-1}(v_i)\}$ tal que $m_i = \text{gr}(f^{a_i}) : f^{a_i} = q_{\phi, v_i}$ y la base resultante de la unión de todas hemos demostrado que cumple las premisas.

⁸Es destacable notar que esta descomposición no es única puesto que a pesar de que el resto de subespacios si están determinados por el 1º, el 1º podemos elegirlo libremente.

Para ver que se cumple la unicidad vamos a ver primero que se verifica la siguiente propiedad (de hecho para una cantidad numerable de sumandos):

$$\begin{aligned} g(A_1 \oplus A_2) &= \sum (b_i(A_1 \oplus A_2)^i) = \sum (b_i(A_1^i \oplus A_2^i)) = \sum ((b_i A_1^i) \oplus (b_i A_2^i)) = \\ &= \left(\sum b_i x^i \right) (A_1) \oplus \left(\sum b_i x^i \right) (A_2) = g(A_1) \oplus g(A_2) \end{aligned}$$

Es decir, que aplicar un polinomio a una matriz por cajas es lo mismo que aplicárselo a cada una de las cajas.

Veamos ahora pues:

$$g(\phi) = \begin{cases} g(C_{f^{a_1}} \oplus \cdots \oplus C_{f^{a_s}}) \\ g(C_{f^{a'_1}} \oplus \cdots \oplus C_{f^{a'_s}}) \end{cases}$$

En concreto vamos a tomar $g = f^{a-1}$ y utilizaremos el símbolo $C_{p(x)}^{(k)}$ para denotar que esa matriz compañera aparece k veces en la suma directa (porque habíamos dicho que los índices se pueden repetir):

$$\begin{cases} f^{a-1} \left(C_{f^a}^{(k_a) \geq 1} \oplus C_{f^{a-1}}^{(k_a-1) \geq 0} \oplus \cdots \oplus C_f^{(k_1)} \right) = [f^{a-1}(C_{f^a})]^{(k_a)} \\ f^{a-1}(\cdots) = [f^{a-1}(C_{f^a})]^{k'_a} \end{cases}$$

Entonces si ambas expresiones son representaciones matriciales del endomorfismo $f^{a-1}(\phi)$, entonces han de tener el mismo rango:

$$\begin{aligned} rg \left([f^{a-1}(C_{f^a})]^{k'_a} \right) &= rg(f^{a-1}(\phi)) = rg \left([f^{a-1}(C_{f^a})]^{(k_a)} \right) \Leftrightarrow \\ &\Leftrightarrow k_a \cdot rg(f^{a-1}(C_{f^a})) = k'_a \cdot rg(f^{a-1}(C_{f^a})) \xrightarrow{rg \neq 0} k'_a = k_a \end{aligned}$$

Si hacemos la misma demostración usando f^{a-2}, f^{a-3}, \dots llegamos a la misma conclusión con cada uno de los exponentes, por lo tanto la representación es única.

Observación:

Por eso, que dos matrices sean semejantes implica que tengan los mismo polinomios mínimos y característicos, pero que **dos endomorfismos tengan el mismo polinomio característico y mínimo no implica que sean semejantes**, porque deben coincidir en sus divisores elementales

Endomorfismos Nilpotentes

Hemos visto buenas propiedades para simplificar el estudio del endomorfismo inicial al estudio de endomorfismos más pequeños caracterizados por poseer un único factor en su polinomio mínimo y característico. En este subapartado vamos a sacarle provecho a todos estos enunciados válidos para el estudio adecuado de estos endomorfismos más pequeños, de manera que podamos utilizar los teoremas anteriormente citados para encontrar bases sencillas cuando el polinomio mínimo y característico sólo tiene un factor.

Definición (Endomorfismos nilpotentes)

Sea $\xi \in \text{End}(V)$, decimos que es **nilpotente de índice a** siendo $1 \leq a \leq n$ si $q_\xi = x^a$.

Observación:

Consecuentemente, como el grado del característico es la dimensión del espacio, entonces $\pm p_x i = x^n$.

Por el Teorema de Estructura y el Teorema de la Forma Canónica Racional, existe una base conveniente para la cual la matriz del endomorfismo nilpotente quedaría como:

$$M_B(\xi) = C_{x^a}^{(k_a)} \oplus \cdots \oplus C_x^{(k_1)}$$

donde el exponente⁹ (k_i) denota el número de veces que aparece en la suma directa esa caja. Además, como dichos teoremas probaban que la matriz asociada es la compañera del polinomio, entonces la caja C_{x^a} es:

$$C_{x^a} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} = J_{0,a}$$

De este modo, podemos dar la siguiente definición que será fundamental para la descomposición canónica de matrices en su base de Jordan.

Definición (Matriz Elemental de Jordan)

Se define una **matriz elemental de Jordan** como una matriz del tipo:

$$J_{a,b} = \begin{pmatrix} a & 0 & \cdots & 0 & 0 \\ 1 & a & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a \end{pmatrix}_{b \times b}$$

De este modo, podemos reescribir la matriz del endomorfismo nilpotente como:

$$M_B(\xi) = J_{0,a}^{(k_a)} \oplus \cdots \oplus J_{0,1}^{(k_1)}$$

Observación:

Esto supone la culminación del proceso de simplificación del endomorfismo, pues podemos separar la matriz inicial como suma directa de la matriz de cada factor lineal y, posteriormente, tratar cada factor lineal como si de un endomorfismo nilpotente se tratase.

Ejemplo:

- Dimensión 1:

$$J_{0,1} = (0)$$

- Dimensión 2:

$$\begin{cases} x^2 \Rightarrow a = 2 \Rightarrow & J_{0,2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ x \mid x \Rightarrow a = 1 \Rightarrow & J_{0,1}^{(2)} = 0 \end{cases}$$

- Dimensión 3:

$$\begin{cases} x^3 \Rightarrow a = 3 \Rightarrow & J_{0,3} \\ x \mid x^2 \Rightarrow a = 2 \Rightarrow & J_{0,2} \oplus J_{0,1} \\ x \mid x \mid x \Rightarrow a = 1 \Rightarrow & J_{0,1}^{(3)} = 0 \end{cases}$$

- Dimensión 4:

$$\begin{cases} x^4 \Rightarrow a = 4 \Rightarrow & J_{0,4} \\ x \mid x^3 \Rightarrow a = 3 \Rightarrow & J_{0,3} \oplus J_{0,1} \\ x^2 \mid x^2 \Rightarrow a = 2 \Rightarrow & J_{0,2}^{(2)} = 0 \\ x \mid x \mid x \mid x \Rightarrow a = 1 \Rightarrow & J_{0,1}^{(4)} = 0 \end{cases}$$

⁹Los teoremas anteriores también prueban que $1 \leq (k_a)$ y los demás ≥ 0 , es decir, que por lo menos el mínimo sale en la descomposición al menos una vez.

Cálculo de Matrices y Bases de Jordan

Una vez desarrolladas todas las herramientas teóricas anteriores para poder justificar los procedimientos que vamos a realizar, trataremos de describir un proceso más o menos unificado para calcular la matriz de Jordan de un endomorfismo y definir la base sobre la cual está expresada dicha matriz.

Conversión de trozos a endomorfismos nilpotentes

El caso general que nos encontraremos a la hora de trabajar con estos endomorfismos no será de la forma de los endomorfismos nilpotentes sino que será algo de la forma: $(x - \lambda)^e$.

Para trabajar con estos elementos como si fuesen nilpotentes, definimos $\xi = \phi - \lambda id$ que posibilita que los polinomios sean de la forma $p_\xi = x^e$ y $q_\xi = x^a$ y así podemos tratarlos como si de un nilpotente se tratase.

Una vez hallada y construida la matriz en bases de Jordan del endomorfismo nilpotente que hemos construido, regresar a nuestro ejemplo es muy sencillo, puesto que se trata de despejar de la siguiente forma:

$$\phi = \lambda id + \xi \Rightarrow M_B(\phi) = \lambda I + M_B(\xi)$$

Lo que en definitiva es rellenar con λ los 0 que dejan las bases de Jordan en las diagonales de los endomorfismos nilpotentes.

Por lo tanto, **un endomorfismo admite una matriz de Jordan si y sólo si su polinomio característico descompone completamente.**

Cálculo de la Matriz de Jordan

Supongamos que tenemos un endomorfismo $\phi \in \text{End}(V)$ sobre un espacio V de dimensión n , cuyos polinomios son:

$$\pm p_\phi(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_s)^{e_s} \quad q_\phi(x) = (x - \lambda_1)^{a_1} \cdots (x - \lambda_s)^{a_s}$$

Lo primero que sabemos es que $V =_\phi \ker((\phi - \lambda_1 I)^{e_1}) \oplus \cdots \oplus \ker((\phi - \lambda_s I)^{e_s}) = \ker((\phi - \lambda_1 I)^{a_1}) \oplus \cdots \oplus \ker((\phi - \lambda_s I)^{a_s})$, luego si calculamos las matrices de Jordan de cada sumando, luego basta componerlas para que queden como una matriz diagonal por cajas.

Centrándonos en un sumando en concreto, sabemos primero que es $\phi - \text{invariante}$ por ser el \ker de una aplicación lineal y también que el polinomio mínimo y característico de la restricción a ese sumando en el que nos hemos fijado son $p_{\phi_i}(x) = (x - \lambda_i)^{e_i}$ y $q_{\phi_i}(x) = (x - \lambda_i)^{a_i}$. Precisamente por esto, le podemos aplicar el Teorema de Estructura, luego sabemos que existe una base que descompone el espacio como:

$$\ker((x - \lambda_i)^{a_i}) =_{\phi_i} L_{\phi}(v_1) \oplus \cdots \oplus L_{\phi_i}(v_{e_i}) \Rightarrow M_B(\phi_i) = C_{(x-\lambda_i)^{a_i}} \oplus \cdots \oplus C_{(x-\lambda_i)^{a_r}}$$

donde los $a_i \geq a_{i+1} \geq \cdots \geq a_r$ y son los divisores elementales del polinomio. Sin embargo, aún sabiendo los divisores elementales la matriz resultante no es muy limpia puesto que la columna final correspondiente a los coeficientes del polinomio del que es compañera la matriz queda muy sucia. En este caso, es más provechoso transformar el endomorfismo a uno nilpotente como se ha explicado en el apartado [Conversión de trozos a endomorfismos nilpotentes](#), es decir, tomar $\xi = \phi - \lambda_i$. Una vez tenemos esto, basta con aplicar el [Teorema de Estructura](#) al nuevo endomorfismo:

$$\ker(\xi^{a_i}) =_{\xi} L_{\xi}(v_1) \oplus \cdots \oplus L_{\xi}(v_{e_i}) \Rightarrow M_B(\xi) = C_{\xi^{a_i}} \oplus \cdots \oplus C_{\xi^{a_r}}$$

Luego aquí basta tener en cuenta que $C_{\xi_j^{a_j}}$ es una Matriz Elemental de Jordan de las que se vieron en [Endomorfismos Nilpotentes](#) y recuperar la matriz de ϕ basta con tener en cuenta la siguiente igualdad:

$$M_B(\phi) = M_B(\xi) + \lambda_i I$$

Sin embargo, la verdadera pregunta está en saber cuáles son los exponentes a_1, \dots, a_r de los divisores elementales en los que se descompone el subespacio invariante. Pues la respuesta es sencilla. Nosotros somos capaces de calcular, para una dimensión concreta, todas las posibles matrices elementales de Jordan que hay en función de qué posibles divisores elementales haya. Pues bien, la matriz canónica de Jordan no deja de ser una matriz equivalente a la inicial de endomorfismo, luego deben tener el mismo rango. Por tanto, basta con elegir como matriz de Jordan (es decir, realmente para esto no nos importan los v_i ni los $L_{\phi_i}(v_i)$) que tenga el mismo rango que la restricción ϕ_i de ϕ al trozo que estamos examinando.

Ejemplo:

FALTA EL EJEMPLO

Cálculo de la Base de Jordan

Aunque tengamos la matriz de Jordan, la base es necesaria en multitud de ocasiones y muy conveniente de conocer en general. Para poder encontrarla nos situamos en la misma situación que en el apartado [Cálculo de la Matriz de Jordan](#), es decir, tenemos

$$\ker((x - \lambda_i)^{a_i}) =_{\phi_i} L_{\phi}(v_1) \oplus \dots \oplus L_{\phi_i}(v_{e_i}) \Rightarrow M_B(\phi_i) = C_{(x - \lambda_i)^{a_i}} \oplus \dots \oplus C_{(x - \lambda_i)^{a_r}}$$

Como estos a_j son los grados de q_{ϕ_i, v_j} , estamos buscando un vector de modo que $v \in \ker(\phi - \lambda_i \text{id})^{a_j}$ pero no pertenezca a los $a_m \leq a_j$. Luego, para cada exponente a_j hay que calcular un vector que pertenezca a ese núcleo pero que no pertenezca a los anteriores (para que así sea realmente ese su polinomio mínimo) y de esta forma construir la base parcial de cada uno de los sumandos, luego basta con juntarlos todos.

Ejemplo:

FALTA EL EJEMPLO

Factores invariantes

Existe una versión alternativa para clasificar los endomorfismos en clases de equivalencia. Del mismo modo que se ha visto la aproximación desde las perspectiva de los divisores elementales, ésta puede verse a través de una nueva: los factores invariantes.

Proposición

Sea $\phi : V \rightarrow V$ un endomorfismo, si denotamos por $g = q_{\phi, v}$ y por $h = q_{\phi, w}$ de modo que $\text{mcd}(g, h) = 1$, entonces ocurre que:

$$q_{\phi, v+w} = gh \text{ y } C_g \oplus C_h \sim C_{gh}$$

Demostración:

Bastaría con demostrar que tenemos un endomorfismo en K^{m+n} de manera que en una base es la suma directa de esas cajas y en la otra es la otra matriz compañera.

Sea $\phi : K^{m+n} \rightarrow K^{m+n}$ donde $M_B(\phi) = C_g \oplus C_h$ por lo que sabemos que:

$$B = \{v, \phi(v), \dots, \phi^{m+n-1}(v), w, \phi(w), \dots, \phi^{m+n-1}(w)\}$$

Luego entonces como $q_{\phi, v+w} = gh$ si escogemos el vector $v' = v+w$ vemos que $\{v', \phi(v'), \dots, \phi^{m+n-1}(v')\}$ forma una base, puesto que son linealmente independientes y hay tantos como la dimensión del espacio sobre el que están definidos y la matriz respecto de esa base es justamente la compañera.

En consecuencia, veamos que si tenemos $\phi : V \rightarrow V$ tal que $\pm p_{\phi} = f_1^{a_1} \cdot f_2^{a_2} = q_{\phi}$ donde cada f_i es un irreducible distinto, entonces sabemos que $V =_{\phi} V_1 \oplus V_2$ por el lema de estructura y, además, que $V_i = \ker(f_i^{a_i}(\phi))$ y $\exists B_i : M_{B_i}(\phi|_{V_i}) = C_{f_i^{a_i}}$. Luego volviendo a la matriz general:

$$M_{B_1 \cup B_2}(\phi) = C_{f_1^{a_1}} \oplus C_{f_2^{a_2}} \sim C_{f_1^{a_1} f_2^{a_2}}$$

Por lo tanto, se puede escoger siempre una base respecto de la cual, la matriz de ϕ es la compa era del producto de ambos trozos.

Teorema

Sea $\phi \in \text{End}(V)$, si $p_\phi = q_\phi$ entonces:

$$\phi : V \rightarrow V : \pm p_\phi = q_\phi \Rightarrow \exists B : M_B(\phi) = C_{q_\phi}$$

Por ejemplo, supongamos $\phi : R^4 \rightarrow R^4 : p_\phi = (x^2 + x + 1)(x^2 - x + 1) = q_\phi$, entonces podemos ver las dos clasificaciones simult neas:

$$\begin{cases} M_B(\phi) = C_{x^2+x+1} \oplus C_{x^2-x+1} & \text{por divisores elementales} \\ M_{B'}(\phi) = C_{x^4+x^2+1} & \text{por factores invariantes} \end{cases}$$

Paso de divisores elementales a factores invariantes

Supongamos que $p_\phi = f_1^{e_1} \cdots f_t^{e_t}$ y $q_\phi = f_1^{a_1} \cdots f_t^{a_t}$, entonces:

$$M_B(\phi) = (C_{f_1^{a_1}} \oplus C_1) \oplus \cdots \oplus (C_{f_t^{a_t}} \oplus C_t)$$

Donde $C_i = C_{f_i^{b_1}} \oplus C_{f_i^{c_i}} \oplus \cdots$, es decir, es la agrupaci n en suma directa de todas las matrices compa eras de ese divisor elevado a exponentes menores que el que poseen en el polinomio m nimo. Por lo tanto y tras reordenar la base, es semejante a esta matriz:

$$C_{f_1^{a_1}} \oplus \cdots \oplus C_{f_t^{a_t}} \oplus C \sim C_{f_1^{a_1} f_t^{a_t}} \oplus C$$

Es decir, en definitiva es reordenar la base para agrupar las ‘‘cajas’’ de forma que todos los trozos del polinomio m nimo queden juntos en una matriz compa era del m nimo y los restos que hemos llamado C_i en otra matriz a parte que llamaremos C .

En el paso siguiente hacemos lo mismo con las cajas de exponente menor que conforman la matriz C , es decir, nos va quedando una expresi n de la forma:

$$C_{f_1^{a_1} \cdots f_t^{a_t}} \oplus C_{f_1^{b_1} \cdots f_t^{b_t}} \oplus \cdots \oplus C_{f_1 \cdots f_t}$$

Cada uno de estos productos de trozos agrupados por los exponentes se llaman **factores invariantes**.

Vamos a ver un ejemplo, supongamos que $\phi : \mathbb{R}^9 \rightarrow \mathbb{R}^9$ y su $p_\phi = (x-1)^3(x^2+x+1)(x^2-x+1)^2$. De cada trozo tenemos:

$$\begin{cases} (x-1)^3 \Rightarrow \begin{cases} (x-1)^3 \\ (x-1)^2, (x-1) \\ (x-1), (x-1), (x-1) \end{cases} \\ x^2+x+1 \Rightarrow x^2+x+1 \\ (x^2-x+1)^2 \Rightarrow \begin{cases} (x^2-x+1)^2 \\ x^2-x+1, x^2-x+1 \end{cases} \end{cases}$$

Por tanto habr a 6 posibles clases de semejanza, veamos una de ellas para ver el ejemplo:

$$M_B(\phi) = C_{(x-1)^2} \oplus C_{x-1} \oplus C_{x^2+x+1} \oplus C_{(x^2-x+1)^2} \text{ por divisores elementales}$$

Pero tambi n lo podemos clasificar por factores invariantes,  Quienes son estos? pues primero escogemos los divisores elementales de mayor grado:

$$h_1(x) = (x-1)^2(x^2+x+1)(x^2-x+1)^2 \Rightarrow C_{h_1(x)}$$

$$h_2(x) = x-1 \Rightarrow C_{h_2(x)}$$

Luego por factores invariantes, la matriz queda de la siguiente forma:

$$M_{B'}(\phi) = C_{h_1(x)} \oplus C_{h_2(x)}$$

FORMAS LINEALES Y BILINEALES

Durante este capítulo vamos a desarrollar las propiedades y características de dos objetos matemáticos muy presentes: las formas lineales y bilineales. La relevancia de estos objetos va desde la definición del concepto de “ortogonalidad” básico en ramas como la geometría hasta la definición de normas a través de productos escalares.

FORMAS LINEALES

Las formas lineales son fundamentalmente aplicaciones lineales del espacio vectorial en el cuerpo sobre el que está estructurado. Es decir, son aplicaciones lineales que a cada vector les asigna un número y, por la similitud con las funciones reales escalares, toma gran relevancia analizar las propiedades que tienen.

Espacio Dual

El espacio dual es un concepto abstracto y vago a priori pero con propiedades esenciales más tarde. Sobre todo, el principio de dualidad nos permite hacer extensibles enunciados sobre V (más intuitivos para nosotros) a V^* , su dual. Del mismo modo, enunciados sorprendentes en V vienen de traducir dichas hipótesis de un enunciado muy intuitivo en V^* .

Definición (Formas Lineales y Espacio Dual)

Sea V un espacio vectorial de dimensión n sobre un cuerpo K , se define el **espacio dual** como el espacio vectorial de dimensión n definido por:

$$V^* = \text{Hom}(V, K)$$

A los elementos de dicho espacio, es decir, a las aplicaciones lineales $\omega : V \rightarrow K$ se las conoce como **formas lineales**.

Como las formas lineales se tratan de aplicaciones lineales, podemos caracterizarlas a través de la matriz que la representa, es decir, que suponiendo $B = \{v_1, \dots, v_n\}$ base de V y $\{1_k\}$ de K , entonces¹:

$$\text{Mat}_{B, \{1_k\}}(\omega) \in \text{Mat}_{1 \times n}(K) = (a_1 \ \cdots \ a_n) \text{ donde } \forall i \in \mathbb{N} : a_i = \omega(v_i)$$

¹Realmente estamos considerando el escalar a_i como un vector en un espacio de dimensión 1 como es K , luego sería $\omega(v_i) = a_i \cdot 1_k$ para indicar que está expresado en dicha base, pero se omite para no sobrecargar la notación.

Observación:

Para tener una idea más intuitiva de lo que representan estos nuevos objetos matemáticos, estamos definiendo lo que serían las “líneas” de un sistema de ecuaciones:

$$\omega(v) = \omega(x_1 v_1 + \cdots x_n v_n) = a_1 x_1 + \cdots a_n x_n$$

Es decir, que en definitiva no son más que aplicaciones que a cada vector les asignan un escalar (o elemento de K) y que evaluar un vector sobre ella no es más que sustituir las coordenadas del vector en la ecuación de la forma lineal.

Ejemplo:

La aplicación lineal $\omega : \text{Mat}_n(K) \rightarrow K$ de forma que $\omega(A) = \text{tr}(A)$ es una forma lineal que a cada matriz (elemento de $V = \text{Mat}_n(K)$) le asigna su traza (elemento de $K = \mathbb{R}$).

La aplicación lineal $\omega : C(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}$ de forma que $\omega(f) = \int_0^1 f(t)dt$ es una forma lineal que a cada función continua le asigna el valor de su integral entre 0 y 1.

Proposición (Base Dual)

Sea V un espacio vectorial de dimensión n y V^* su dual, dada una base $B = \{v_1, \dots, v_n\}$ de V , entonces el conjunto:

$$B^* = \{v_1^*, \dots, v_n^*\} : \forall i \in \mathbb{N}, v_i^* = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

es una base del espacio dual.

Demostración:

En primer lugar, trivialmente $V^* \simeq V$ porque tienen la misma dimensión y la dimensión de V es n . En segundo lugar, los elementos de V^* son aplicaciones lineales, luego para poder tenerlos definidos tendremos que conocer su valor sobre cada uno de los vectores de una base B de V , por lo tanto, la clave está en ver que la definición que se ha dado de los v_i^* genera un conjunto linealmente independiente:

$$a_1 v_1^* + \cdots a_n v_n^* = 0 \Rightarrow \forall j \in \mathbb{N} : 0 = (a_1 v_1^* + \cdots a_n v_n^*)(v_j) = a_j v_j^*(v_j) = a_j$$

Ejemplo:

Sea $B = \{e_1 - e_2, e_1 + 3e_2\}$ una base de K^2 , vamos a ver como definimos v_1^* :

$$\begin{cases} 1 = v_1^*(e_1 - e_2) = v_1^*(e_1) - v_1^*(e_2) \\ 0 = v_1^*(e_1 - 1 - 3e_2) = v_1^*(e_1) - 3v_1^*(e_2) \end{cases} \Rightarrow \begin{cases} v_1^*(e_2) = \frac{-1}{4} \\ v_1^*(e_1) = \frac{3}{4} \end{cases}$$

Entonces $\text{Mat}_{B_c, \{1_k\}}(v_1^*) = \begin{pmatrix} \frac{3}{4} & \frac{-1}{4} \end{pmatrix}$, para calcular v_2^* se hace igual.

Pero ahora si cambiamos la base a $B' = \{e_1 - e_2, e_1\}$ base de K^2 entonces volviendo a calcular todo:

$$\begin{cases} 1 = v_1^*(e_1 - e_2) \\ 0 = v_1^*(e_1) \end{cases} \Rightarrow \begin{cases} v_1^*(e_1) = 0 \\ v_1^*(e_2) = -1 \end{cases}$$

Observación:

Es cierto que existe un isomorfismo entre un espacio y su dual y entonces debe existir una aplicación que a cada vector que le asigne su dual, pero es importante resaltar que ésta aplicación no es canónica entre ambos ya que incluso la demostración de este resultado depende de las bases escogidas al estar una determinada por la otra.

Con esto queremos decir que, si $B = \{v_1, \dots, v_n\}$ y $B' = \{v'_1, \dots, v'_n\}$ son dos bases con algunos vectores comunes, por ejemplo $v_1 = v'_1$, los vectores v_1^* y v'_1^* no tienen por qué coincidir, pues están condicionados por el resto de la base.

Proposición

Sea $B = \{v_1, \dots, v_n\}$ una base de V y $B^* = \{v_1^*, \dots, v_n^*\}$ una del dual V^* , entonces las coordenadas de una forma lineal $\omega \in V^*$ son precisamente $\omega = (\omega(v_1), \dots, \omega(v_n))$.

Definición (Aplicación dual)

Sean V y W dos espacios vectoriales y $\phi \in \text{Hom}(V, W)$ una aplicación entre ellos, se define la **aplicación dual** a ϕ como:

$$V \xrightarrow{\phi} W \xrightarrow{\omega} K$$

$$\begin{aligned} \phi^* : W^* &\longrightarrow V^* \\ \omega &\longmapsto \omega \circ \phi \end{aligned}$$

Además, se verifican las siguientes propiedades:

- $\phi^*(\omega_1 + \omega_2) = (\omega_1 + \omega_2) \circ \phi = \omega_1 \circ \phi + \omega_2 \circ \phi = \phi^*(\omega_1) + \phi^*(\omega_2)$
- $\phi^*(\lambda\omega) = (\lambda\omega) \circ \phi = \lambda(\omega \circ \phi) = \lambda\phi^*(\omega)$
- $(id_V)^* = id_{V^*}$
- $(\phi_2 \circ \phi_1)^* = \phi_1^* \circ \phi_2^*$

Observación:

Por tanto, acabamos de ver que la dualidad asigna a cada espacio otro espacio vectorial y a cada aplicación su dual. Conserva la identidad y cambia de orden las composiciones por lo que se llama (esto se ve en otras asignaturas) un **functor contravariante**.

Proposición

Sea ϕ una aplicación entre dos espacios vectoriales V y W y ϕ^* su dual correspondiente, las matrices de ambas aplicaciones guardan la siguiente relación:

$$A = M_{B_v, B_w}(\phi) \Rightarrow M_{B_w^*, B_v^*}(\phi^*) = A^t$$

Espacio Bidual

Consideremos la aplicación $\langle \bullet, \bullet \rangle : V^* \times V \rightarrow K$ como aquella que evalúa en la formal el vector seleccionado, es decir, $(\omega, v) \mapsto \omega(v)$. Si fijamos v obtenemos la aplicación $\langle \bullet, v \rangle : V^* \rightarrow K$ que llamaremos ξ_v y que mapea $\omega \mapsto \langle \omega, v \rangle$

Proposición

La aplicación $\xi_v \in \text{Hom}(V^*, K)$ definida como $\xi_v(\omega) = \omega(v)$ es una aplicación lineal y biyectiva.

Demostración:

- Está bien definida, es decir, ocurre que $\forall v \in V : \eta_v(v) \in V^{**} = \text{Hom}_K(V^*, K)$

$$[\eta(v)](\omega_1 + \omega_2) = (\omega_1 + \omega_2)(v) = \omega_1(v) + \omega_2(v) = [\eta(v)](\omega_1) + [\eta(v)](\omega_2)$$

$$[\eta(v)](\lambda\omega) = (\lambda\omega)(v) = \lambda\omega(v) = \lambda[\eta(v)](\omega)$$

- Es lineal:

$$[\eta(v_1 + v_2)](\omega) = \omega(v_1 + v_2) = \omega(v_1) + \omega(v_2) = [\eta(v_1)](\omega) + [\eta(v_2)](\omega) = [\eta(v_1) + \eta(v_2)](\omega)$$

$$[\eta(\lambda v)](\omega) = \omega(\lambda v) = \lambda\omega(v) = \lambda[\eta(v)](\omega)$$

- Es biyectiva:

Se puede observar que $\dim V = \dim V^{**}$, por lo que solo hace falta ver si es inyectiva o suprayectiva para demostrarlo. Por lo que vamos a ver que es inyectiva, supongamos que no:

$$\ker(\eta_v) \neq \{0\} \Rightarrow v \in \ker(\eta_v) \Rightarrow \forall \omega \in V^* : [\eta(v)](\omega) = 0$$

Pues construimos una forma lineal que no verifique esto:

$$B = \{v, v_2, \dots, v_n\} \text{ de } V \Rightarrow \exists B^* \text{ de } V^* \Rightarrow [\eta(v)](v_1^*) = v_1^*(v) = 1 \neq 0 \Rightarrow \#$$

Definición (Espacio Bidual)

Sea V un espacio vectorial de dimensión n , se define el **espacio bidual** como $V^{**} = \text{Hom}(V^*, K)$ conformado por todas las aplicaciones ξ_v .

Proposición

Sea V un espacio vectorial de dimensión n y V^{**} su bidual, la aplicación

$$\begin{aligned} \eta : V &\longrightarrow V^{**} \\ v &\longmapsto \xi_v \end{aligned}$$

es un isomorfismo canónico entre ambos espacios.

Ortogonalidad

Un concepto fundamental que también se desarrolla, por ejemplo, en el apartado de [Formas Bilineales](#) es la ortogonalidad. Ésta nos va a permitir, en el caso del espacio dual, caracterizar los hiperplanos vectoriales con formas lineales del espacio dual.

Definición (Ortogonalidad)

Sea V un espacio vectorial y V^* su dual, decimos que $\omega \in V^*$ es **ortogonal** a $v \in V$ si $\omega(v) = 0$.

Además, sea W un subespacio de V , se define el **ortogonal de W** como:

$$W^\perp = \{\omega \in V^* : \forall w \in W, \omega(w) = 0\}$$

Proposición

La definición de ortogonalidad verifica las siguientes propiedades:

1. $W < V \Rightarrow W^\perp < V^*$
2. $\{0\}^\perp = V^*$ y $V^\perp = \{0_{V^*}\}$
3. $W_1 \subset W_2 \Rightarrow W_2^\perp \subset W_1^\perp$
4. $n = \dim V$ y $m = \dim W \Rightarrow \dim W^\perp = n - m$
5. $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$
6. $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$
7. $W_1^{\perp\perp} = W$
8. $V = W \oplus U \Rightarrow V^* = W^* \oplus U^*$

Demostración:

1. Es subespacio:

$$\omega_1, \omega_2 \in W^\perp \Rightarrow (\omega_1 + \omega_2)(w) = \omega_1(w) + \omega_2(w) \Rightarrow \omega_1 + \omega_2 \in W^\perp$$

$$\omega \in W^\perp \text{ y } \lambda \in K \Rightarrow (\lambda\omega)(w) = \lambda\omega(w) = \lambda \cdot 0 = 0 \Rightarrow \lambda\omega \in W^\perp$$

2. La primera parte es trivial porque afirmar eso implica que todas las formas lineales se anulan sobre 0, lo cual es cierto. La segunda es análoga porque una forma lineal que se anula sobre todos los vectores de V es la aplicación lineal nula.
3. Trivial porque si una forma se anula sobre todos los vectores de W_2 entonces también se anula sobre todos los vectores de W_1 . El otro contenido no es cierto porque puede haber formas que se anulen sobre W_1 pero que no se anulen sobre W_2 .
4. Esta demostración es cómo se calcula en la práctica el ortogonal de un subespacio: Supongamos $B_W = \{w_1, \dots, w_m\}$ base de W y extendemos a una base del total $B = \{w_1, \dots, w_m, w_{m+1}, \dots, w_n\}$ base de V . Ahora tenemos que demostrar que $\{w_{m+1}^*, \dots, w_n^*\}$ es base de W^\perp :

Como son independientes porque formaban parte de una base, basta con probar que son sistema de generadores, por tanto sea $\omega \in W^\perp$:

$$\omega = \sum_{i=1}^n a_i w_i^*$$

Como está en el ortogonal se anula sobre todos los vectores de W , en particular sobre los vectores de la base elegida:

$$\underbrace{\omega(w_1)}_{=a_1} = \underbrace{\dots}_{=a_i} = \underbrace{\omega(w_m)}_{=a_m} = 0$$

De este modo veamos el caso para w_1 y los demás son iguales:

$$\omega(w_1) = \left(\sum_{i=1}^n a_i (w_i)^* \right) (w_1) = a_1 w_1^*(w_1) = a_1$$

Luego se tiene que todas las coordenadas de esta forma lineal desde a_1 hasta a_m se anulan, luego se tiene que:

$$\Rightarrow \omega = a_{m+1} w_{m+1}^* + \dots + a_n w_n^*$$

5. Análoga a la demostración del punto 6
6. Hay dos contenidos que se siguen de la condición 3, en concreto: $(W_1 + W_2)^\perp \subset W_1^\perp \cap W_2^\perp$ y por otro lado $W_1^\perp + W_2^\perp \subset (W_1 \cap W_2)^\perp$ con lo cual volviendo a lo que queremos demostrar tenemos que:

$$\underbrace{(W_1 \cap W_2)^{\perp\perp}}_{\eta_v(W_1 \cap W_2)} \subset (W_1^\perp + W_2^\perp)^\perp \subset W_1^{\perp\perp} \cap W_2^{\perp\perp} = \eta_v(W_1) \cap \eta_v(W_2) \Rightarrow$$

Como la η es biyectiva, si las imágenes coinciden, los elementos tienen que ser los mismos:

$$\eta_v(w_1) = \eta_v(w_2) \Rightarrow w_1 = w_2 \Rightarrow \eta_v(W_1) \cap \eta_v(W_2) \subset \eta_v(W_1 \cap W_2)$$

Como el primer espacio que empieza la cadena de contenidos y el último son el mismo subespacio, la cadena de contenidos se vuelve cadena de igualdades, por lo que en particular:

$$\begin{aligned} (W_1 \cap W_2)^{\perp\perp} &= (W_1^\perp + W_2^\perp)^\perp \Rightarrow (W_1 \cap W_2)^{\perp\perp\perp} = (W_1^\perp + W_2^\perp)^{\perp\perp} \Rightarrow \\ &\Rightarrow \eta_v(W_1 \cap W_2)^\perp = \eta_v(W_1^\perp + W_2^\perp) \Rightarrow (W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp \end{aligned}$$

7. Veamos que ocurre que $\eta_v(W) \subset W^{\perp\perp} = (W^\perp)^\perp$:

$$\forall \omega \in W^\perp : [\eta_v(w)](\omega) = \omega(w) = 0$$

Pero como $\dim(W^\perp)^\perp = \dim V^* - \dim W^\perp = \dim V^* - (\dim V - \dim W) = \dim W = \dim \eta_v(W)$.

Teorema (Principio de dualidad)

La correspondencia $\text{Sub}(V) \rightarrow \text{Sub}(V^*)$ que lleva cada subespacio $W \rightarrow W^*$ en su dual es un antiisomorfismo de retículos completos, es decir, lleva ínfimos a supremos y viceversa y sumas a intersecciones y viceversa.

Proposición

Sea $V \xrightarrow{\phi} W$ una aplicación lineal y $W^* \xrightarrow{\phi^*} V^*$ su dual, entonces se tienen las siguientes propiedades:

1. $\ker(\phi^*) = \text{Im}(\phi)^\perp < W^*$
2. $\text{Im}(\phi^*) = \ker(\phi)^\perp < V^*$

Demostración:

1. Basta con probar que un elemento de uno está en el otro y viceversa:

$$\omega \in \ker(\phi^*) \Rightarrow \phi^*(\omega) = 0 \Rightarrow \omega \circ \phi = 0 \Rightarrow \forall v \in V : \omega(\phi(v)) = 0 \Rightarrow \omega \in (\text{Im}(\phi))^\perp$$

2. Demostración análoga a la anterior.

Observación:

Si recordamos que $A = M_{B_v, B_w}(\phi)$ y $A^t = M_{B_w^*, B_v^*}(\phi^*)$, veamos de otra forma que el rango de ambas matrices es el mismo:

$$\begin{aligned} \text{rg}(A^t) &= \text{rg}(\phi^*) = \dim W^* - \dim \ker(\phi^*) = \dim W - \dim(\text{Im}(\phi))^\perp = \\ &= \dim W - (\dim W - \dim \text{Im}(\phi)) = \dim \text{Im}(\phi) = \text{rg}(\phi) = A \end{aligned}$$

FORMAS BILINEALES

La definición $V \times V^* \rightarrow K$ de forma que $(v, \omega) \rightarrow \omega(v) \in K$, que justamente se dio para $[\eta(v)](\omega)$ tiene buenas propiedades que podemos querer “copiar” para otros objetos matemáticos. ¿Pero qué propiedades tiene esta aplicación? Si denotamos la aplicación por $\langle v, \omega \rangle$, entonces:

- $\langle v + v', \omega \rangle = \langle v, \omega \rangle + \langle v', \omega \rangle$
- $\langle \lambda v, \omega \rangle = \lambda \langle v, \omega \rangle$
- $\langle v, \omega + \omega' \rangle = \langle v, \omega \rangle + \langle v, \omega' \rangle$
- $\langle v, \lambda \omega \rangle = \lambda \langle v, \omega \rangle$

Es decir, que considerando la función fija para ω y libre para v lo que tenemos es una forma lineal sobre V y si escogemos la función fija para v y libre para ω , entonces también es una forma lineal sobre V^* .

Características de las formas bilineales

Vamos aproximarnos primero al concepto de forma bilineal y ver que consecuencias y características tiene sobre las definiciones de espacio vectorial y endomorfismos que venimos desarrollando hasta ahora.

Definición (Forma Bilineal)

Sea $f : V \times V \rightarrow K$ una aplicación, se dice que es ***K*-bilineal** si es lineal respecto de ambas componentes de la función, es decir, verifica:

- $f(v + v', w) = f(v, w) + f(v', w)$
- $f(v, w + w') = f(v, w) + f(v, w')$
- $f(av, w) = af(v, w)$
- $f(v, bw) = bf(v, w)$

Al conjunto de todas las formas bilineales se le denota como $Bil(V)$

Ejemplo:

Supongamos que tenemos $\omega_1, \omega_2 \in V^*$ y definimos $f(v, w) = \omega_1(v)\omega_2(w)$ es trivialmente una forma bilineal porque verifica dichas condiciones.

Ejemplo:

Supongamos que $V = Mat_n(K)$ y definimos $f(A, C) = tr(A \cdot C)$ y se puede comprobar fácilmente que se verifican las propiedades de bilinealidad.

Definición (Tipos de formas bilineales)

Sea f una forma K -bilineal, decimos que f es:

- **Degenerada:** $rg(f) < n$
- **Definida positiva:** $\forall x, y \in V, x, y \neq 0 : f(x, y) > 0$
- **Semidefinida positiva:** $\forall x, y \in V : f(x, y) \geq 0$
- **Semidefinida negativa:** $\forall x, y \in V : f(x, y) \leq 0$
- **Definida negativa:** $\forall x, y, x, y \neq 0 \in V : f(x, y) < 0$

Observación:

Obsérvese que $\forall x \in V : f(x, 0) = 0$ y que, por tanto, una forma que es definida positiva o definida negativa tiene por único elemento que cumple $f(v, v) = 0$ al $v = 0$.

Proposición (Operaciones con formas bilineales)

El conjunto $Bil(V)$ tiene estructura de espacio vectorial con las operaciones:

$$(f + g)(v, w) = f(v, w) + g(v, w) \qquad f \in Bil(V) \Rightarrow (\lambda f)(v, w) = \lambda f(v, w)$$

Observación:

Una forma bilineal no es una aplicación lineal porque no ocurre lo siguiente:

$$f((v, w) + (v', w')) \neq f(v, w) + f(v', w')$$

Matriz asociada y Cambio de Base

Si definimos una base de V , los vectores operandos quedarán determinados de la forma:

$$\begin{cases} v = x_1 v_1 + \cdots x_n v_n \\ w = y_1 v_1 + \cdots y_n v_n \end{cases}$$

De este modo, podemos expresar la imagen de dicha forma bilineal como:

$$f(v, w) = \sum_{i,j} (x_i y_j f(v_i, v_j)) = (x_1 \cdots x_n) (a_{ij}) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ donde } a_{ij} = f(v_i, v_j)$$

La matriz A formada por los a_{ij} es la asociada a la forma bilineal y está compuesta por las imágenes de la forma bilineal en todos los posibles pares de vectores de la base:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Teorema (de la Matriz Asociada)

Sea $f : V \times V \longrightarrow K$ una forma K -bilineal y $B = \{v_1, \dots, v_n\}$ una base de V , podemos expresar la imagen por f de los pares de vectores de V como:

$$\forall v, w \in V : f(v, w) = v \cdot M_{B,B'}(\phi) \cdot w^t$$

donde la matriz de la ecuación es la matriz:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \text{ donde } a_{ij} = f(v_i, v_j)$$

Si suponemos ahora que B' es otra base, entonces la función cambiará a:

$$f(v, w) = (x'_1 \cdots x'_n) (a'_{ij}) \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix}$$

Por tanto, llamando $P = M(B', B) \in GL_n(K)$ a la matriz de paso de B' a B , entonces:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$
$$(y_1 \cdots y_n) = (y'_1 \cdots y'_n) P^t$$

luego si reescribimos la matriz A en estos nuevos términos, vemos que se relacionan como:

$$f(v, w) = (x_1 \cdots x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (x'_1 \cdots x'_n) \cdot \underbrace{P^t \cdot A \cdot P}_{A'} \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} \Rightarrow A' = P^t A P$$

Definición (Congruencia de Matrices)

Sean $A, A' \in Mat_n(K)$, decimos que ambas matrices son **congruentes** si están relacionadas como:

$$\exists P \in GL_n(K) : A' = P^t A P$$

esta relación es de equivalencia en el espacio $Mat_n(K)$.

Ejemplo:

Si escogemos $V = \text{Mat}_2(K)$ y la base formada por:

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Si definimos la forma bilineal como $f(A, C) = \text{tr}(A \cdot C)$ vamos a calcular los coeficientes de su matriz:

$$\begin{cases} a_{11} = \text{tr}(E_{11}, E_{11}) = \text{tr} \cdots = 1 \\ a_{12} = \text{tr}(E_{11}, E_{12}) = \text{tr} \cdots = 0 \\ a_{13} = \text{tr}(E_{11}, E_{21}) = \text{tr} \cdots = 0 \\ a_{14} = \text{tr}(E_{11}, E_{22}) = \text{tr} \cdots = 0 \\ \vdots \end{cases}$$

Observación:

Del mismo modo que ocurría con las aplicaciones lineales, ocurre que la aplicación que asocia a cada forma su matriz es un isomorfismo de espacios y, en consecuencia, la dimensión del espacio $\text{Bil}(V)$ es:

$$f \xrightarrow{\Phi} M_B(f) \Rightarrow \text{Bil}(V) \simeq \text{Mat}_n(K) \Rightarrow \dim \text{Bil}(V) = n^2$$

Simetría y antisimetría

Como nos encaminamos a una clasificación general de las formas bilineales, es necesario observar las características de dos tipos fundamentales de formas: las simétricas y las antisimétricas, pues más adelante veremos que son las piezas fundamentales de cualquier forma bilineal.

Definición (Simetría y Antisimetría)

Sea $f : V \times V \rightarrow K$ una forma K -bilineal, decimos que f es **simétrica** si:

$$f(v, w) = f(w, v) \Rightarrow A = M_B(f) \in \text{Mat}_n(K) \text{ simétricas}$$

y decimos que f es **antisimétrica** si:

$$f(v, w) = -f(w, v) \Rightarrow A = M_B(f) \in \text{Mat}_n(K) \text{ antisimétricas}$$

Recordemos por un instante que cualquier matriz se podía escribir como combinación de una matriz simétrica y antisimétrica:

$$A = \frac{A + A^t}{2} + \frac{A - A^t}{2}$$

así que es razonable pensar que con las formas bilineales suceda lo mismo.

Teorema

Sea $f \in \text{Bil}(V)$, entonces se puede² escribir como combinación lineal de una forma simétrica y otra antisimétrica:

$$f = f_s + f_a : \begin{cases} f_s(v, w) = \frac{f(v, w) + f(w, v)}{2} & \Rightarrow A_{f_s} = \frac{A + A^t}{2} \\ f_a(v, w) = \frac{f(v, w) - f(w, v)}{2} & \Rightarrow A_{f_a} = \frac{A - A^t}{2} \end{cases}$$

Observación:

En consecuencia, podemos calcular la dimensión de estos dos subespacios de $\text{Bil}(V)$:

$$\begin{cases} \text{Bil}_s(V) = \{f_s : V \times V \rightarrow K\} & \Rightarrow \dim \text{Bil}_s(V) = \frac{n(n+1)}{2} \\ \text{Bil}_a(V) = \{f_a : V \times V \rightarrow K\} & \Rightarrow \dim \text{Bil}_a(V) = \frac{n(n-1)}{2} \end{cases}$$

²De este modo, es fácil comprobar que toda la teoría desarrollada no tiene sentido a menos que $\chi(K) \neq 2$ puesto que si no la antisimetría no tendría sentido.

Definición (Vectores Isótopos)

Sea V un espacio vectorial y $f : V \times V \rightarrow K$ una forma bilineal, un vector v se dice **isótopo respecto a f** si verifica $f(v, v) = 0$.

Proposición

Sea V un espacio vectorial y f una forma bilineal sobre dicho espacio, entonces:

$$f \text{ antisimétrica} \Leftrightarrow \forall v \in V : f(v, v) = 0$$

Demostración:

■ \Rightarrow :

Trivial por la definición de antisimetría

■ \Leftarrow :

$$0 = f(v + w, v + w) \Rightarrow 0 = \underbrace{f(v, v)}_{=0} + f(v, w) + f(w, v) + \underbrace{f(w, w)}_{=0} \Rightarrow f(v, w) = -f(w, v)$$

Corolario

Si f es simétrica donde $f \neq 0$, entonces $\exists v \in V : f(v, v) \neq 0$

Forma cuadrática asociada

Reviste gran importancia, sobre todo en los capítulos posteriores, el valor que toma la función bilineal sobre el mismo vector como ambas componentes de la función. Hemos visto que podemos caracterizar las formas antisimétricas como aquellas en las que todos los vectores son isótopos. Por ejemplo, para que una forma no sea degenerada, el único vector isótopo debe ser el 0. Por todo esto, tiene sentido definir una nueva función que a cada vector $v \in V$ le asigne su imagen por f como ambas componentes.

Definición (Forma Cuadrática)

Sea f una forma bilineal, se define su **forma cuadrática asociada** como la aplicación

$$\begin{aligned} q : V &\longrightarrow K \\ v &\longmapsto f(v, v) \end{aligned}$$

Proposición

Sea $q : V \rightarrow K$ una forma cuadrática, existe una única forma bilineal f simétrica de la cual es asociada y viene dada por la expresión:

$$f(v, w) = \frac{1}{2} (q(v + w) - q(v) - q(w))$$

Es decir, que existe una correspondencia biyectiva entre las formas bilineales simétricas y las formas cuadráticas asociadas.

Demostración:

$$f \in \text{Bil}_s(V) \Rightarrow f(v+w, v+w) = \begin{cases} q(v+w) \\ q(v) + q(w) + 2f(v, w) \end{cases} \Rightarrow f(v, w) = \frac{1}{2} (q(v + w) - q(v) - q(w))$$

Observación:

Es notable destacar que la proposición anterior no se verifica para una forma bilineal en general, por ejemplo, la forma cuadrática nula proviene de cualquier antisimétrica:

$$f \in \text{Bil}_a(V) \Rightarrow q(v) = 0 \text{ pero } \Leftarrow \text{ no es cierto}$$

Observación:

Cuando hablamos de la matriz de la forma cuadrática no estamos hablando más que de la matriz de la forma bilineal de la que procede, ya que el cálculo sería el mismo que cuando vimos la matriz de la forma bilineal, pero cambiando los y_i por x_i .

Ortogonalidad

En este apartado, vamos a tratar de relacionar la ortogonalidad tal y como se definió en el capítulo [Ortogonalidad](#) con una nueva definición a través de una forma bilineal. De esta manera, el objetivo fundamental es intentar sacar provecho de los resultados que ésta tenía en el espacio dual para utilizarlos en la que se va a ver ahora.

Definición (Formas Izquierda y Derecha)

Sea $f : V \times V \rightarrow \mathbb{K}$ una forma bilineal, definimos la **forma izquierda** $f_{iz} \in \text{Hom}(V, V^*)$ y la **forma derecha** $f_d \in \text{Hom}(V, V^*)$ de la forma bilineal como aquellas que verifican que:

$$f_{iz} : V \rightarrow V^* \text{ donde } f_{iz}(v)(w) = f(v, w)$$

$$f_d : V \rightarrow V^* \text{ donde } f_d(v)(w) = f(w, v)$$

es decir, que f_{iz} asocia a cada vector v la forma de V^* que al aplicarse sobre w tiene como resultado la aplicación bilineal f con la derecha sucede de modo análogo.

Proposición

Sea $f \in \text{Bil}(V)$ y f_{iz} y f_d sus formas izquierda y derecha respectivamente, dadas unas bases B y B^* de V y V^* sus matrices son:

$$M_{B, B^*}(f_{iz}) = A^t$$

$$M_{B, B^*}(f_d) = A$$

Demostración:

$$M_{B, B^*}(f_{iz}) = (c_{ij}) \text{ donde } f_{iz}(v_i) = \sum c_{ji} v_j^* \Rightarrow f_{iz}(v_i)(v_k) = \left(\sum_{j=1}^n c_{ji} v_j^* \right) (v_k) = c_{ki}$$

Pero por otro lado tenemos que:

$$f_{iz}(v_i)(v_k) = f(v_i, v_k) = a_{ik} \Rightarrow c_{ki} = a_{ik}$$

Con lo cual se tiene que:

$$M_{B, B^*}(f_{iz}) = M_B(f)^t = A^t$$

Del mismo modo y con una demostración análoga tenemos que:

$$M_{B, B^*}(f_d) = A$$

Observación:

En consecuencia, tenemos³ que:

$$\begin{cases} \text{rg}(f) = \text{rg}(f_{iz}) = n - \dim \ker(f_{iz}) \\ \text{rg}(f) = \text{rg}(f_d) = n - \dim \ker(f_d) \end{cases} \Rightarrow \dim \ker f_{iz} = \dim \ker f_d = n - \text{rg}(f)$$

³En una forma bilineal no hay un núcleo concreto porque no se puede considerar subespacio de nada. Lo único que podemos llamar núcleo es la intersección de los núcleos izquierdo y derecho y considerarlo un conjunto de pares que se anulan y que estudiaremos más tarde.

Esto quiere decir pues, que los núcleos de ambas funciones no tienen por qué coincidir, pero siempre tendrán la misma dimensión.

Proposición

Sea f una forma bilineal, entonces:

$$f \text{ no degenerada} \Leftrightarrow f_{iz} \text{ isomorfismo o } f_d \text{ isomorfismo}$$

Definición (Ortogonalidad respecto de una forma bilineal)

Sea $W < V$, $f \in \text{Bil}_a(V)$ y $v, w \in V$, decimos que v y w son **ortogonales respecto de f** si $f(v, w) = 0$. Además, se define el **ortogonal a un subespacio** como:

$$W^\perp = \{v \in V : f(v, w) = 0 : \forall w \in W\} < V$$

Ejemplo:

- El conjunto W y su ortogonal W^\perp no tienen porque ser disjuntos, por ejemplo considerando:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = M_B(f) \text{ donde } f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

los vectores isótropos de esta forma bilineal son:

$$v = (x_1, x_2) : f(v, v) = 0 \Rightarrow f((x_1, x_2), (y_1, y_2)) = x_1 y_2 - x_2 y_1$$

Si ocurre que $f(v, v) = 0$ quiere decir que $v \in L(v)^\perp$ con lo cual estos vectores quedarían definidos por:

$$x_1 x_2 - x_2 x_1 = 0 \Leftrightarrow 0 = 0$$

Lo que implica que todos los vectores son ortogonales a sí mismos⁴.

- Del mismo modo, para cualquier forma bilineal antisimétrica:

$$\{0\}^\perp = \mathbb{R}^2$$

porque $f(v, 0) = f(v, 0) + f(v, 0) \Rightarrow f(v, 0) = 0$.

$$(\mathbb{R}^2)^\perp = 0$$

Por que $f(v, w) = 0 : \forall w \in \mathbb{R}^2 \Rightarrow f_{iz}(v)(w) \Rightarrow v \in \ker(f_{iz})$, pero como el $rg(f) = 2 = rg(f_{iz})$ entonces $\ker(f_{iz}) = \{0\} \Rightarrow v = 0 \Rightarrow (\mathbb{R}^2)^\perp = \{0\}$.

- Consideramos $W : x_1 - x_2 = 0$, para que un vector sea ortogonal a esa recta, tiene que serlo a todos los vectores de esa recta, por lo que basta que sea ortogonal al generador de esa recta:

$$0 = f(v, e_1 + e_2) = x_1 - x_2$$

Luego el ortogonal de W es W^\perp la propia recta.

Ortogonalidad en productos escalares

En primer lugar, vamos a definir la noción de producto escalar, que va a ser fundamental para la comprensión de otros capítulos o áreas como la geometría analítica o el espacio vectorial euclídeo.

Definición (Producto Escalar)

Sea V un espacio vectorial sobre \mathbb{R} de dimensión $\dim_{\mathbb{R}} V = n$ y $f \in \text{Bil}_s(V)$ una forma bilineal simétrica, decimos que esta es **producto escalar** si es bilineal simétrica y definida positiva.

⁴Pero no hemos demostrado que sean ortogonales a todos los vectores, que no es cierto.

Vista la definición, si consideramos f un producto escalar y la siguiente aplicación: para cada $v \in V$ le asignamos la aplicación $v^* : V \rightarrow K$ que mapea $w \mapsto f(v, w)$, es decir, la forma bilineal f fija para el vector v o, lo que es lo mismo, la aplicación $f_{iz}(v)$. Entonces la aplicación que a cada v le asigna su correspondiente v^* es un isomorfismo y además canónico por la proposición sobre f no degeneradas.

Por otro lado, tenemos que la correspondencia entre v y la aplicación η_v que vimos en el apartado de [Espacio Bidual](#) es un isomorfismo canónico entre V y V^{**} .

Por tanto, dado $v \in V$ podemos considerar, ya no $L(v)^\perp < V^*$, sino $L(v)^\perp < V$ debido al siguiente razonamiento:

- Consideramos $v^* \in V^*$ y calculamos $L(v^*)^\perp$, de este modo, estamos considerando todos los $\eta_v \in V^{**}$ de modo que $\eta_v(v^*) = 0$, esto es, que $v^*(v) = 0$.
- Como, por ser f un producto escalar, tenemos que $v \mapsto v^*$ es un isomorfismo, podemos considerar $L(v)^\perp = L(v^*)^\perp < V^{**}$ por ser $L(v) \simeq L(v^*)$.
- Del mismo modo, como también hemos visto que $w \mapsto \eta_w$ es un isomorfismo, entonces $L(v^*)^\perp \simeq W < V$.
- Por tanto, cuando escogemos un vector v y decimos que otro vector w es ortogonal a él, estamos diciendo que $\eta_w(v^*) = v^*(w) = f(v, w) = 0$.

Surge ahora la pregunta de ¿y para qué todo este razonamiento si me podía limitar a considerar la definición de ortogonalidad previa sin relacionarla con la ortogonalidad del espacio dual? Pues muy sencillo: porque como una es equivalente a la otra a través de isomorfismos todos los enunciados válidos para una también lo son para la otra. De este modo, todas las propiedades deducidas para la ortogonalidad que se vio para formas lineales también se tiene para la ortogonalidad entendida desde el punto de vista bilineal.

Clasificación de formas bilineales

Del mismo modo que en el tema de clasificación de endomorfismos, tratábamos de agrupar las aplicaciones lineales en clases de equivalencia y escoger el representante más sencillo posible dentro de cada clase, en este apartado vamos a tratar de hacer lo mismo pero con las formas bilineales.

Teorema (de Clasificación de Formas Antisimétricas)

Sea $f \in \text{Bil}_a(V)$, entonces $\text{rg}(f) = 2 \cdot s$ y $\exists B = \{v_1, \dots, v_n\}$ base de V tal que la matriz de f queda como:

$$M_B(f) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{(s)} \oplus (0)^{(t)} \text{ donde } 2s + t = \dim V$$

Demostración:

Supongamos que $f = 0$, está trivialmente probado por lo que podemos suponer f distinto de 0.

Por tanto, supongamos $f \neq 0 \Rightarrow \exists v, w : v \neq w : f(v, w) \neq 0$, deben ser distintos porque todos los vectores son isótopos al ser la forma antisimétrica. Vamos a probar primero que son linealmente independientes denotando por $a = f(v, w)$:

$$\lambda v + \mu w = 0 \Rightarrow f(v, \lambda v + \mu w) = 0 \Rightarrow \mu a = 0 \Rightarrow \mu = 0 \Rightarrow \lambda = 0$$

Ahora para elegir la base escogemos como primeros vectores $v_1 = a^{-1}v$, $v_2 = w$ viendo primero que:

$$f(v_1, v_2) = a^{-1}f(v, w) = a^{-1}a = 1$$

Y los escogemos de este modo porque ocurre con ambos que:

$$\Pi = L(v, w) = L(v_1, v_2) \Rightarrow V = \Pi \oplus \Pi^\perp$$

Vamos a probar esto último:

$$v = x_1 v_1 + x_2 v_2 + x_3 w_3 + \cdots + x_n w_n \in \Pi^\perp \Leftrightarrow \begin{cases} 0 = f(v_1, v) = x_1 \cdot 0 + x_2 \cdot 1 + \cdots = x_2 + \cdots \\ 0 = f(v_2, v) = x_1 \cdot -1 + x_2 \cdot 0 + \cdots = -x_1 + \cdots \end{cases}$$

Luego como posee dos ecuaciones independientes ocurre que $\dim \Pi^\perp = n - 2$ así que falta solo ver que $\Pi \cap \Pi^\perp = \{0\}$ para poder afirmar que son suma directa, por lo tanto, sabiendo que si $v \in \Pi \Rightarrow \lambda_1 v_1 + \lambda_2 v_2$:

$$v \in \Pi \cap \Pi^\perp \Rightarrow \begin{cases} 0 = f(\lambda_1 v_1 + \lambda_2 v_2, v_1) = -\lambda_2 \\ 0 = f(\lambda_1 v_1 + \lambda_2 v_2, v_2) = \lambda_1 \end{cases} \Rightarrow v = 0 \Rightarrow \Pi \cap \Pi^\perp = \{0\}$$

Para terminar la demostración tomamos una base de V de forma que $B = \{ \underbrace{v_1, v_2}_{B \text{ base de } \Pi} \underbrace{v_3, \dots, v_n}_{\text{base de } \Pi^\perp} \}$,

con lo cual la matriz de f en la base elegida queda como:

$$M_B(f) = M_{\{v_1, v_2\}}(f|_{\Pi \times \Pi}) \oplus M_{\{v_3, \dots, v_n\}}(f|_{\Pi^\perp \times \Pi^\perp})$$

Como la segunda matriz se puede considerar como una forma restringida a una dimensión más pequeña, aplicamos inducción y terminamos.

Observación:

La demostración nos propone un método para calcular la base B que genera esa representación matricial:

- Como la matriz es no nula, $\exists v, w \in V : f(v, w) \neq 0$ que serán los primeros vectores de la base.
- Calculamos el ortogonal del plano formado por ambos vectores, es decir, $\Pi = L(v_1, v_2) \Rightarrow \Pi^\perp = L(v_1, v_2)^\perp$.
- La restricción de la forma bilineal f al subespacio Π^\perp es de nuevo una forma bilineal que podemos simplificar con este método, es decir, calcularíamos dos vectores tal que $f(v, w) \neq 0$, etc.

Ejemplo:

Si estamos en dimensión 2 tenemos que:

$$A \in A_2(K) \Rightarrow \begin{cases} A = 0 \\ A \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{cases} \Rightarrow f \in \text{Bil}_a(V) \Rightarrow \begin{cases} f = 0 \\ f(v, w) = x_1 y_2 - x_2 y_1 \end{cases}$$

Si estamos en dimensión 3 tenemos que:

$$A \in A_3(K) \Rightarrow \begin{cases} A = 0 \\ A \equiv \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{cases} \Rightarrow f \in \text{Bil}_a(V) \Rightarrow \begin{cases} f = 0 \\ f(v, w) = x_1 y_2 - x_2 y_1 \end{cases}$$

Por ejemplo, supongamos que tenemos la siguiente matriz:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \in A_3(K)$$

Como $A \neq 0$, si consideremos $A = M_{B_c}(f)$ de forma que $f : K^3 \times K^3 \rightarrow K$, entonces:

$$A \equiv \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Para calcular la base B en la que la matriz de la forma bilineal es esa tenemos en cuenta lo siguiente:

- Como la matriz es antisimétrica todos los vectores son isótropos, es decir, que $f(v, v) = 0 : \forall v \in V$. De este modo, la diagonal de la matriz siempre debe ser 0.
- Como la matriz es no nula, existe una pareja de vectores de forma que $\exists v, w \in V : f(v, w) \neq 0$ que en este caso se verifica con e_1 y e_2 ya que $f(e_1, e_2) = 1$. De este modo tomamos como primeros vectores de la base $v_1 = e_1 \cdot 1$ y $v_2 = e_2 \cdot 1$.
- Calculamos el ortogonal del plano formado por ambos vectores, es decir:

$$\Pi^\perp = \begin{cases} 0 = f(e_1, v) = x_2 + x_3 \\ 0 = f(e_2, v) = -x_1 + x_3 \end{cases} \Rightarrow B = \{(1, -1, 1)\} \text{ es base de } \Pi^\perp$$

- Tomamos la base $B = \{e_1, e_2, e_1 - e_2 + e_3\}$ y solo queda calcular $P = M(B, B_c)$.

Teorema (de Clasificación de Formas Simétricas)

Sea $f \in \text{Bil}_s(V)$, entonces $\exists B = \{v_1, \dots, v_n\}$ base de V respecto de la cual $M_B(f)$ es diagonal.

Demostración:

Si $f = 0$ no hay nada que probar puesto que se cumple trivialmente.

Suponiendo que $f \neq 0 \Rightarrow \exists v \in V : f(v, v) \neq 0$, es decir, existe algún vector no isótropo. En la práctica, sean cuales sean los vectores escogidos se debe cumplir que o bien v , o bien w o $v + w$ sean NO isótropos. De este modo llamamos v_1 a ese v que no es isótropo.

En primer lugar vamos a ver que $L(v_1)^\perp$ es un hiperplano:

Extendemos a una base del total de forma que $B = \{v_1, w_2, \dots, w_n\}$ de forma que ahora calculamos el subespacio ortogonal de la siguiente forma:

$$v = x_1 v_1 + x_2 w_2 + \dots + x_n w_n \in L(v)^\perp \Leftrightarrow 0 = f(v_1, v) = \underbrace{f(v_1, v_1)}_{\neq 0} x_1 + \dots$$

Por lo que ya vemos que es un hiperplano porque posee una única ecuación con lo cual $\dim L(v)^\perp = n - 1$. Por lo que basta ahora con comprobar que la intersección de ambos subespacios es 0:

$$v \in L(v_1) \cap L(v_1)^\perp \Rightarrow \begin{cases} 0 = f(\lambda v_1, v_1) = \lambda f(v_1, v_1) \Rightarrow \lambda = 0 \\ v = \lambda v_1 \end{cases} \Rightarrow v = 0$$

Ahora para escoger la base B que hace diagonal la matriz tomamos $B = \{ \underbrace{v_1}_{\text{base de } L(v_1)^\perp}, \underbrace{v_2, \dots, v_n}_{\text{base de } L(v_1)^\perp} \}$

luego la matriz de la forma bilineal en esta base queda de la siguiente forma:

$$M_B(f) = \begin{pmatrix} f(v_1, v_1) & \dots & 0 \\ \vdots & M_B(f|_{L(v_1)^\perp \times L(v_1)^\perp}) & \\ 0 & & \end{pmatrix}$$

Es decir, que nos quedan todos los ceros menos en la posición a_{11} y después una caja más pequeña que es la forma bilineal restringida a al subespacio ortogonal hallado. Por inducción se termina teniendo que ocurre de nuevo en los sucesivos trozos más pequeños luego la matriz queda diagonal.

Métodos de diagonalización de formas simétricas

Fundamentalmente tenemos tres métodos plausibles:

- Por vectores isotropos:

Hemos visto en la demostración de que existe una base respecto de la cual es diagonal que para obtener dicha matriz diagonal se siguen los siguientes pasos:

1. Buscamos un vector no isotropo que llamaremos v
2. Calculamos el subespacio ortogonal que siempre será un hiperplano y además $L(v) \oplus L(v)^\perp = V$
3. Como v es ortogonal a todo lo que queda de V , la fila y la columna que contiene a v son 0 menos la posición de v .
4. La restricción de f al subespacio ortogonal hallado vuelve a ser una forma bilineal y repetimos el proceso.

Repetiendo este proceso un número finito de veces, cada v es un vector de la base respecto de la cual es diagonal.

Ejemplo:

FALTA EJEMPLO

- Sumando filas con filas o columnas con columnas (que es como multiplicar a izquierda o derecha por matrices elementales) y aplicar el mismo cambio sobre las filas o las columnas (las que no se hayan hecho) para que de esta forma sea como multiplicar por una matriz elemental y su traspuesta por el otro lado, hasta tener una matriz diagonal. El producto de todas esas matrices sobre la identidad es la matriz P de cambio de base.

Ejemplo:

FALTA EJEMPLO

- El último método consiste en obtener la forma cuadrática asociada y llegar mediante métodos algebraicos como completar cuadrados a una forma cuadrática asociada de la forma $q(v) = x^2 + y^2 + z^2 + \dots$ porque esta forma es la asociada a una matriz diagonal.

Ejemplo:

$$q(v) = -x^2 + 4xy + 3y^2 + 2z^2 \Rightarrow q(v) = -(x-2y)^2 + 4y^2 + 3y^2 + 2z^2 = -\underbrace{(x-2y)^2}_{x'} + 7\underbrace{y^2}_{y'} + 2\underbrace{z^2}_{z'}$$

Luego tenemos que:

$$\begin{cases} x' = x + 2y \\ y' = y \\ z' = z \end{cases} \Rightarrow \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=M(B,B')} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Clasificación de formas bilineales simétricas sobre \mathbb{C}

Teorema

Sea $f \in \text{Bil}_s(V)$ cuyo rango es $\text{rg}(f) = r$ y V un \mathbb{C} -espacio vectorial, entonces:

$$\exists B = \{v_1, \dots, v_n\} \text{ base de } V : M_B(f) = I_r \oplus 0_{n-r}$$

Demostración:

Sabemos por el teorema general que $M_B(f)$ es diagonal respecto de alguna base, es decir:

$$M_B(f) = \begin{pmatrix} d_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & d_r & \vdots & \vdots \\ \vdots & \dots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

Si cambiamos los primeros r vectores por los que siguen a continuación y los demás los dejamos igual

$$v'_i = \begin{cases} \frac{1}{\sqrt{d_i}} v_i & i = 1, \dots, r \\ v_i & i > r \end{cases}$$

conseguiamos que con este cambio la forma bilineal valga:

$$f(v'_i, v'_i) = f\left(\frac{1}{\sqrt{d_i}} v_i, \frac{1}{\sqrt{d_i}} v_i\right) = \left(\frac{1}{\sqrt{d_i}}\right)^2 f(v_i, v_i) = \left(\frac{1}{\sqrt{d_i}}\right)^2 d_i = 1$$

es decir, que quede la matriz que habíamos enunciado.

Clasificación de formas bilineales simétricas sobre \mathbb{R} : Ley de inercia

Teorema

Sea $f \in \text{Bil}_s(V)$ cuyo rango es $\text{rg}(f) = r$ y V un \mathbb{R} -espacio vectorial, entonces:

$$\exists B = \{v_1, \dots, v_n\} \text{ base de } V : M_B(f) = I_s \oplus (-I_t) \oplus 0_{n-r} \text{ donde } r = s + t$$

Además, s y t son únicos y forman lo que llamamos **signatura** de la forma bilineal:

$$\varepsilon(f) = (s, t)$$

Demostración:

Sabemos por el teorema general que $M_B(f)$ es diagonal respecto de alguna base, es decir:

$$M_B(f) = \begin{pmatrix} d_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & d_r & \vdots & \vdots \\ \vdots & \dots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

En este caso procedemos de la misma manera que con los complejos, pero distinguiendo entre positivos (desde d_1 hasta d_s) y negativos (desde d_{s+1} hasta d_r) porque aquí no existen ciertas raíces. Se puede definir una base B' de forma que se cambian los positivos por 1 y los negativos por menos 1:

$$v'_i = \begin{cases} \frac{1}{e_i} v_i & 1 \leq i \leq s : d_i = e_i^2 \\ \frac{1}{e_i} v_i & s+1 \leq i \leq r : d_i = -e_i^2 \\ v_i & r < i \leq n \end{cases}$$

Entonces tenemos que:

$$f(v'_i, v'_i) = \begin{cases} \frac{1}{e_i^2} f(v_i, v_i) = \frac{1}{e_i^2} d_i = 1 \\ \frac{1}{e_i^2} f(v_i, v_i) = \frac{1}{e_i^2} d_i = -1 \\ f(v_i, v_i) = 0 \end{cases}$$

Por tanto:

$$M_{B'}(f) = I_s \oplus (-I_t) \oplus 0_{n-r}$$

A priori parece que esta clasificación no clasifica por rango porque para el mismo rango podemos tener distintos para el número de 1 y -1. Veamos pues la unicidad, supongamos que tenemos dos matrices de esta forma:

$$M_B(f) = I_s \oplus (-I_t) \oplus 0_{n-r}$$

$$M_{B'}(f) = I_{s'} \oplus (-I_{t'}) \oplus 0_{n-r}$$

A pesar de que estos valores de la signatura pueden ser distintos, el rango no puede serlo, por lo tanto se tiene que:

$$\begin{cases} rg(f) = r = s + t \\ rg(f) = r = s' + t' \end{cases}$$

Si nos fijamos en el subespacio generado por los vectores positivos de la primera matriz y nos fijamos en el generado por los vectores no positivos de la segunda matriz vamos a ver que son suma directa y generan el espacio V :

$$W_1 = L(v_1, \dots, v_s) \Rightarrow \dim W_1 = s$$

$$W_2 = L(v'_{s'+1}, \dots, v'_n) \Rightarrow \dim W_2 = n - s'$$

Veamos que la intersección es nula:

$$\sum_{i=1}^s a_i v_i = \sum_{j=s'+1}^n a'_j v'_j \in W_1 \cap W_2 \Rightarrow \begin{cases} f(v, v) = \sum_{i=1}^s a_i^2 f(v_i, v_i) = \sum_{i=1}^s a_i^2 \geq 0 & \text{por } \in W_1 \\ f(v, v) = \sum_{j=s'+1}^n (a'_j)^2 f(v'_j, v'_j) = \sum_{j=s'+1}^n -(a'_j)^2 \leq 0 & \text{por } \in W_2 \end{cases}$$

Por lo tanto, como deben ser iguales, entonces necesariamente por ser todos los cuadrados positivos, se tiene que:

$$a_1^2 + \dots + a_s^2 = 0 \Rightarrow a_1 = \dots = a_s = 0 \Rightarrow v = 0$$

Por tanto, usando la fórmula de Grassmann:

$$n \geq \dim(W_1 + W_2) = \dim W_1 + \dim W_2 = s + n - s' \Rightarrow s' \geq s$$

Por un razonamiento análogo cambiando la base llegaríamos a $s \geq s'$, en consecuencia $s = s'$, implicando $t = t'$.

Teorema

Sea $f \in \text{Bil}_s(V)$ y $\dim V = n$, entonces verifica:

- Si $s + t < n$, entonces es **degenerada**.
- Si $\varepsilon(m, 0)$ con $m < n$, entonces es **semidefinida positiva**.
- Si $\varepsilon(0, m)$ con $m < n$, entonces es **semidefinida negativa**.
- Si $\varepsilon(n, 0)$, entonces es **definida positiva**.
- Si $\varepsilon(0, n)$, entonces es **definida negativa**.

Ejemplo:

Sobre $n = 2$ y $rg(f) = 2$ hay 3 clases de equivalencia por congruencia de matrices, una por cada posible signatura:

$$\varepsilon(f) = (2, 0) \Rightarrow f((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2 \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\varepsilon(f) = (1, 1) \Rightarrow f((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2 \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\varepsilon(f) = (0, 2) \Rightarrow f((x_1, x_2), (y_1, y_2)) = -x_1 y_1 - x_2 y_2 \Rightarrow \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Es notable destacar que la tercera es menos la primera pero a pesar de ello, NO son congruentes.

Viendo las formas cuadráticas asociadas:

$$q(x_1, x_2) = x_1^2 + x_2^2$$

$$q(x_1, x_2) = x_1^2 - x_2^2$$

$$q(x_1, x_2) = -(x_1^2 + x_2^2)$$

Si ahora $rg(f) = 1$, entonces:

$$\varepsilon(f) = (1, 0) \Rightarrow f((x_1, x_2), (y_1, y_2)) = x_1 y_1$$

$$\varepsilon(f) = (0, 1) \Rightarrow f((x_1, x_2), (y_1, y_2)) = -x_1 y_1$$

Y viendo las formas cuadráticas asociadas entonces:

$$q(x_1, x_2) = x_1^2$$

$$q(x_1, x_2) = -x_1^2$$

ESPACIO VECTORIAL EUCLÍDEO

El espacio vectorial euclídeo es un concepto fundamental que servirá de antesala del espacio afín, la geometría euclídea, la física, etc. Nos permite modelar el espacio n dimensional de \mathbb{R} y estudiar analíticamente las propiedades de dicho espacio. Sin ir más lejos, permite modelar el espacio tridimensional real (o al menos sirve como base) para poder hacer un estudio exhaustivo de la geometría en tres dimensiones, para poder hacer modelos físicos sobre el comportamiento de cierto objeto, etc.

ESPACIO VECTORIAL EUCLÍDEO

Definición (Base Ortonormal)

Sea f un producto escalar y V un espacio vectorial euclídeo, entonces la base B para la cual la forma bilineal es la identidad se dice que es una base **ortonormal**.

Definición (Espacio Vectorial Euclídeo)

Se define un **espacio vectorial euclídeo** como una dupla (V, f) formada por un espacio vectorial sobre \mathbb{R} de dimensión finita y un producto escalar.

De este modo, decimos que **dos espacios vectoriales son isomorfos** cuando:

$$(V, f) \simeq (V', f') \Leftrightarrow \begin{cases} \exists \phi : V \rightarrow V' \text{ isomorfismo} \\ f(\phi(v), \phi(w)) = f'(v, w) : \forall v, w \in V \end{cases}$$

Teorema

Sean V y V' espacios vectoriales euclídeos, éstos son isomorfos si y solo si la dimensión de V es la misma que la de V' , es decir:

$$(V, f) \simeq (V', f') \Leftrightarrow \dim V = \dim V'$$

Demostración:

La implicación en el sentido derecho está clara por la definición de isomorfismo de espacios euclídeos y por la condición de isomorfismos de espacios vectoriales.

Sea $B = \{v_1, \dots, v_n\}$ base ortonormal de V respecto de f y $B' = \{v'_1, \dots, v'_n\}$ base ortonormal de V' respecto de f' , entonces definimos:

$$\phi \left(\sum_{i=1}^n x_i v_i \right) = \sum_{i=1}^n x_i v'_i$$

Esta es la única aplicación lineal que lleva los vectores de la base B a los de la base V' ya tenemos completamente definida la aplicación. Tal y como la hemos definido se tiene que $f(v_i, v_j) = \delta_{ij}$ donde cada v_i, v_j son vectores de la base B . De este modo, $f(v'_i, v'_j) = \delta_{ij}$ también porque hemos dicho que B' es ortonormal respecto de f' luego esto ya prueba la biyectividad, etc.

Observación:

De este razonamiento extraemos que lo que clasifica los espacios euclídeos es la dimensión del espacio vectorial porque si tenemos dos productos escalares distintos basta con definir la aplicación biyectiva que transforma uno en otro y ya son isomorfos.

Teorema (del Criterio de Silvester)

Sea $f \in \text{Bil}_s(V)$ y $A = (a_{ij}) = M_B(f) \in \text{Mat}_n(\mathbb{R})$ donde $a_{ij} = f(v_i, v_j)$, entonces:

$$f \text{ es producto escalar} \Leftrightarrow \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix} > 0 : \forall i = 1, 2, \dots, n$$

es decir, basta con ver que todos los menores superiores izquierdos son positivos.

Demostración:

■ \Rightarrow :

Como f es producto escalar, f es definida positiva y entonces f es definida positiva sobre $L(v_1, \dots, v_i) = V_i$ vectores de la base de V porque si lo es sobre V , sobre una restricción también. Por lo que la matriz de dicha forma bilineal quedaría como:

$$M_{\{v_1, \dots, v_i\}}(f|_{V_i}) = \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix} \Rightarrow |M| > 0$$

Esto ocurre porque si g es definida positiva y la matriz en alguna base $M_B(g) = I_n \Rightarrow |M_B(g)| = 1$, como para cambiar a la base que queramos se usa una matriz P inversible, se tiene que:

$$|M_B(f)| = |P^t I_n P| = |P|^2 > 0$$

■ \Leftarrow

Supongamos que $B = \{v_1, \dots, v_n\}$, vamos a definir una base $B' = \{v'_1, \dots, v'_n\}$ de forma que $L(v_1, \dots, v_i) = L(v'_1, \dots, v'_i)$ y que $f(v'_i, v'_j) = 0$ si $i \neq j$ y $f(v'_i, v'_i) > 0$ por inducción.

En el primer paso, $v'_1 = v_1$ es trivial la demostración solo falta probar que $f(v'_1, v'_1) > 0$:

$$f(v'_1, v'_1) = a_{11} \Rightarrow a_{11} > 0$$

Por la condición de partida de que el determinante es positivo.

En el paso inductivo, considerando construido hasta el paso i y vamos a verlo para el paso $i+1$. Consideramos $W = L(v_1, \dots, v_{i+1}) = L(v'_1, \dots, v'_i, v_{i+1})$ que se cumplen porque hemos demostrado que generan el mismo subespacio hasta el paso i . Sabemos que $f|_{W \times W}$ es no degenerada porque el rango es máximo:

$$A = M_{\{v_1, \dots, v_{i+1}\}}(f|_{W \times W}) = \begin{pmatrix} a_{11} & \cdots & a_{1,i+1} \\ \vdots & \ddots & \vdots \\ a_{i+1,1} & \cdots & a_{i+1,i+1} \end{pmatrix} \Rightarrow |A| > 0 \Rightarrow A \neq 0 \Rightarrow f \in \text{Bil}_s(\mathbb{R}) \text{ no degenerada}$$

Si ahora consideramos el subespacio $L(v'_1, \dots, v'_i)^\perp$, entonces $L(v'_1, \dots, v'_i) = L(v_1, \dots, v_i) < W$ y como f es no degenerada entonces se cumple que¹ $\dim U^\perp = \text{co dim } U$, por tanto:

$$\dim L(v'_1, \dots, v'_i) = \text{co dim } L(v'_1, \dots, v'_i) = \dim W - \dim L = i+1 - i = 1$$

¹Este resultado se puede ver que es cierto en las dos siguientes proposiciones

Es decir, que puedo elegir un vector en W de forma que sea ortogonal a todos los vectores $L(v_1, \dots, v_i)$, por tanto:

$$\exists v'_{i+1} \in W : v'_{i+1} \in L(v'_1, \dots, v'_i)^\perp \Rightarrow W$$

Para completar la demostración falta comprobar que $f(v'_{i+1}, v'_{i+1}) > 0$, veamos que:

$$\begin{aligned} M_{\{v_1, \dots, v_{i+1}\}}(f|_{W \times W}) &= \begin{pmatrix} a_{11} & \cdots & a_{1i+1} \\ \vdots & \ddots & \vdots \\ a_{i+1,1} & \cdots & a_{i+1,i+1} \end{pmatrix} \Rightarrow B = M_{\{v'_1, \dots, v'_{i+1}\}}(f|_{W \times W}) = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{i+1} \end{pmatrix} \Rightarrow \\ \Rightarrow B &= C_1 \oplus C_2 : C_2 = d_{i+1} \Rightarrow |B| = |C_1| \cdot |C_2| \Rightarrow |C_2| > 0 \end{aligned}$$

Proposición

Sea f una forma bilineal no degenerada en un espacio vectorial euclideo, la dimensión de un subespacio ortogonal es la codimensión del subespacio del que es ortogonal, es decir:

$$\dim W^\perp = n - m : \dim W = m$$

Además, si f es producto escalar:

$$W^\perp \oplus W = V$$

Demostración:

1. Supongamos que la dimensión es m , pues sea $\{w_1, \dots, w_m\}$ base de W , entonces:

$$v \in W^\perp \Leftrightarrow 0 = f(v, w_i) = f_d(w_i)(v) : V \rightarrow V^*$$

De este modo, podemos denominar a $f_d(w_i) = \omega_i \in V^*$ puesto que es una forma lineal. Es decir, que para poder definir que vectores están en W^\perp podemos entenderlos como m formas lineales que se anulan sobre v , luego tenemos m ecuaciones cartesianas igualadas a 0. Estas ecuaciones lineales son independientes porque como se trata de una forma no degenerada, el rango es máximo y en consecuencia la función f_d es un isomorfismo. En consecuencia, m ecuaciones con n incógnitas implican que el espacio de soluciones (W^\perp) tiene dimensión $n - m$, es decir, la codimensión.

2. Vemos que:

$$v \in W \cap W^\perp \Rightarrow f(v, v) = 0 \Rightarrow v = 0$$

Luego como la intersección es nula y la suma de las dimensiones es todo n entonces queda probado:

$$\dim W + \dim W^\perp = \dim W + \dim W^\perp - \dim W \cap W^\perp = n = \dim V$$

Proposición

Sea f definida positiva, los vectores ortogonales son independientes.

$$f(w_i, w_j) = 0 \Rightarrow \{w_i, w_j\} \text{ independientes}$$

Demostración:

$$\lambda_1 w_1 + \dots + \lambda_n w_n = 0 \Rightarrow 0 = f(w_i, \lambda_1 w_1 + \dots + \lambda_n w_n) = \lambda_i \underbrace{f(w_i, w_i)}_{\neq 0} \Rightarrow \lambda_i = 0 : \forall i$$

Teorema (Método de Gram-Schmidt)

Sea (V, f) un espacio vectorial euclídeo y $B = \{v_1, \dots, v_n\}$ una base de V , entonces

$$B' = \begin{cases} v'_1 = v_1 \\ v'_2 = v_2 + \lambda_{2,1}v'_1 \\ v'_3 = v_3 + \lambda_{3,2}v'_2 + \lambda_{3,1}v'_1 \\ \vdots \\ v'_n = v_n + \lambda_{n,n-1}v'_{n-1} + \dots + \lambda_{n,1}v'_1 \end{cases}$$

forma una base ortogonal de V , es decir, podemos encontrar $\lambda_{i,j} \in \mathbb{R}$ únicos de forma que la base formada por esos vectores sea ortogonal respecto de f .

Demostración:

Primero vamos a ver como encontrar estos $\lambda_{i,j}$ **únicos** conociendo que la base formada debe ser ortogonal:

$$0 = f(v'_2, v'_1) = f(v_2 + \lambda_{2,1}v_1, v_1) = f(v_2, v_1) + \lambda_{2,1}f(v_1, v_1) \Rightarrow \lambda_{2,1} = -\frac{f(v_2, v_1)}{f(v_1, v_1)}$$

Para el caso v'_3 se imponen de nuevo las condiciones que sean necesarias:

$$\begin{cases} 0 = f(v'_3, v'_2) = f(v_3, v'_2) + \lambda_{3,2}f(v'_2, v'_2) + \lambda_{3,1}\underbrace{f(v'_1, v'_2)}_{=0} \\ 0 = f(v'_3, v'_1) = f(v_3, v'_1) + \lambda_{3,2}\underbrace{f(v'_2, v'_1)}_{=0} + \lambda_{3,1}f(v'_1, v'_1) \end{cases}$$

De esta manera, por inducción siempre queda un sistema de ecuaciones cuyas incógnitas son los $\lambda_{i,j}$ y que siempre posee una solución única. Además cabe destacar (la demostración es trivial) que:

$$L(v_1, \dots, v_i) = L(v'_1, \dots, v'_i)$$

Del mismo modo, a partir de aquí se puede construir una base ortonormal, modificando los vectores hallados en el paso anterior para que su módulo sea unitario, es decir:

$$v''_i = \frac{v'_i}{\sqrt{f(v'_i, v'_i)}}$$

Vamos a comprobarlo:

$$f(v''_i, v''_i) = f\left(\frac{v'_i}{\sqrt{f(v'_i, v'_i)}}, \frac{v'_i}{\sqrt{f(v'_i, v'_i)}}\right) = \frac{1}{f(v'_i, v'_i)}f(v'_i, v'_i) = 1$$

Es notable destacar que la matriz de el producto escalar en esta base es la matriz identidad, porque todos los vectores son ortogonales dos a dos y ellos mismos son 1.

Definición (Norma de un vector)

Sea f producto escalar en un espacio vectorial euclídeo, se define la **norma de un vector** como:

$$\|v\| = \sqrt{f(v, v)} = \sqrt{q(v)}$$

Proposición

La función norma definida anteriormente $\|\cdot\| : V \rightarrow \mathbb{R}$, cumple las siguientes propiedades:

$$\blacksquare \|v\| \geq 0, \|v\| = 0 \Leftrightarrow v = 0$$

- $\|\lambda v\| = |\lambda| \cdot \|v\|$
- $\|v + w\| \leq \|v\| + \|w\|$

Demostración:

- Trivial
- Por bilinealidad
- Primero vamos a probar un resultado conocido como la desigualdad de Schwarz:

$$f \text{ producto escalar} \Rightarrow f(v, w)^2 \leq q(v)q(w)$$

Vamos a considerar $q(v + \lambda w)$. Es sencillo ver que $0 \leq q(v + \lambda w) = q(v) + \lambda^2 q(w) + 2\lambda f(v, w)$. Este resultado lo podemos entender como un polinomio con coeficientes reales con grado dos sobre λ que es la indeterminada. Para que ese polinomio se anule, por venir de la forma cuadrática que viene, tiene que ocurrir que:

$$q(v + \lambda w) = 0 \Rightarrow v + \lambda w = 0 \Rightarrow \begin{cases} v = -\lambda w \\ \nexists \lambda \in \mathbb{R} \end{cases}$$

Luego el polinomio o no tiene ninguna raíz real o tiene solo una que sería la que hace v y w proporcionales, luego el discriminante tiene que ser menor igual que 0:

$$\Delta \leq 0 \Rightarrow (v, w)^2 - 4q(w)q(v) \leq 0 \Rightarrow f(v, w)^2 \leq q(v)q(w)$$

Conociendo este resultado, vamos a probar ahora la desigualdad triangular y con probar que $\|v + w\|^2 \leq (\|v\| + \|w\|)^2$ es suficiente:

$$\|v + w\|^2 = q(v + w) = f(v + w, v + w) = q(v) + 2f(v, w) + q(w) \leq q(v) + 2\sqrt{q(v)}\sqrt{q(w)} + q(w) =$$

Porque de la desigualdad que hemos probado antes tenemos que $|f(v, w)| \leq \sqrt{q(v)}\sqrt{q(w)}$. Y ahora para terminar tenemos que:

$$= (\sqrt{q(v)} + \sqrt{q(w)})^2 = (\|v\| + \|w\|)^2$$

Además, recuperando la desigualdad de Schwarz podemos dar otra interpretación a la definición que se ha dado de norma:

$$\frac{|f(v, w)|}{\|v\| \cdot \|w\|} \leq 1 \Rightarrow -1 \leq \underbrace{\frac{f(v, w)}{\|v\| \cdot \|w\|}}_{\cos(\varphi)} \leq 1 : 0 \leq \varphi \leq \pi$$

Es decir, que podemos entender el producto escalar como una forma de relacionar los módulos de los vectores y el ángulo que los relaciona.

Definición (Endomorfismos Autoadjuntos)

Sea (V, f) un espacio vectorial euclídeo y $\phi \in \text{End}(V)$, decimos que es un **endomorfismo autoadjunto respecto de f** si cumple que:

$$f(\phi(v), w) = f(v, \phi(w)) : \forall v, w \in V$$

Proposición

Sea (V, f) un espacio vectorial euclídeo, $\phi \in \text{End}(V)$ un endomorfismo y $C = M_B(f)$, entonces:

$$\phi \text{ autoadjunto} \Leftrightarrow A^t C = C A$$

Además, cómo podemos elegir una base conveniente sobre la que $M_{B'}(f) = I$, entonces se tiene también:

$$\phi \text{ autoadjunto} \Leftrightarrow A^t = A$$

Demostración:

Si definimos una base B de V y llamamos $A = M_B(\phi)$ y $C = M_B(f)$, entonces siendo $v = (x_1, \dots, x_n)$ y $w = (y_1, \dots, y_n)$ se tiene:

$$(x_1, \dots, x_n)A^t C \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (x_1, \dots, x_n)CA \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Por tanto, ϕ es autoadjunto si y solo si:

$$A^t C = CA$$

Del mismo modo, si elegimos B convenientemente hemos visto que la matriz de f es I luego el endomorfismo ϕ en esa base cumple que:

$$A = A^t$$

Teorema (Espectral)

Sea (V, f) un espacio vectorial euclídeo, si $\phi \in \text{End}(V)$ es autoadjunto entonces existe una base B de V ortonormal respecto de f formada por vectores propios respecto de ϕ .

Demostración:

Vamos a ver primero que q_ϕ descompone completamente, para ello supongamos que $f = x^2 + a_1x + a_0 \in R[x]$ irreducible de forma que $f \nmid q_\phi$, entonces $\exists v \in V : v \neq 0 : q_{\phi,v} = f$. Para buscarlo, en primer lugar vemos que podemos escribir q_ϕ de esta manera:

$$q_\phi = f^a \cdot g_a \geq 1 \wedge f \nmid g$$

De este modo, vimos en el Teorema de Caley Hamilton que siempre existe un $w \in V : q_{\phi,w} = f^a$ y ahora para elegir el v basta con seleccionarlo de esta forma:

$$v = f(\phi)^{a-1}(w)$$

Puesto que $f(\phi)(v) = f(\phi)^a(w) = 0 \Rightarrow f(x) = q_{\phi,v}$ por ser f irreducible y en consecuencia el polinomio más pequeño que se anula en v .

Como el polinomio mínimo sobre v es de grado 2, entonces el plano generado por $\pi = L(v, \phi(v))$ tiene $\dim \pi = 2$ porque si tuviese dimensión 1 existiría un polinomio mínimo de grado 1 más pequeño que dividiría a f y eso no es posible. Por definición $L(v, \phi(v))$ es invariante. Si consideramos la restricción del producto escalar sobre ese plano:

$$f|_{\pi \times \pi} : \pi \times \pi \rightarrow \mathbb{R}$$

De manera que es sencillo ver que esta sigue siendo un producto escalar restringido a ese subespacio, vamos a denotar $g = f|_{\pi \times \pi}$.

Como es invariante $\phi|_{\pi \times \pi} \in \text{End}(\pi)$ y, del mismo modo, este ϕ restringido con el producto escalar restringido sigue siendo autoadjunto. ¿Cual es la matriz de la restricción $\psi = \phi|_{\pi \times \pi}$?

$$M_{\{v, \phi(v)\}}(\psi) = \begin{pmatrix} 0 & -a_0 \\ 1 & -a_1 \end{pmatrix}$$

Por otro lado, como tenemos un producto escalar, existirán bases ortonormales respecto de las cuales la matriz de ψ será simétrica por ser autoadjunta así que:

$$\exists B = \{w_1, w_2\} : M_B(\psi) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Por tanto, aunque las matrices sean distintas el polinomio mínimo sigue siendo el mismo, con lo cual, debemos llegar al mismo resultado a partir de las dos matrices al calcular el polinomio característico:

$$x^2 + a_1x + a_0 = p_\psi = x^2 - (a + c)x + ac - b^2$$

Como f es irreducible de grado 2, el discriminante es menor que 0 luego:

$$0 > a_1^2 - 4a_0 = (a+c)^2 - 4(ac-b^2) = (a-c)^2 + 4b^2 \geq 0 \Rightarrow \#$$

Osea que ya podemos suponer que el polinomio mínimo descompone completamente:

$$q_\phi = (x - \lambda_1)^{a_1} + \dots + (x - \lambda_t)^{a_t}$$

Así que solo falta ver que todos los a_i son 1, por lo que vamos a ver que:

$$\ker(\phi - \lambda id)^2 \subset \ker(\phi - \lambda id)$$

Puesto que si estabiliza la sucesión de núcleos directamente en el primer \ker ya tenemos que es diagonalizable.

Supongamos que $v \in \ker(\phi - \lambda id)^2$, aplicando el producto escalar tenemos que:

$$f((\phi - \lambda id)(v), (\phi - \lambda id)(v)) = f(v, (\phi - \lambda id)^2(v)) = f(v, 0) = 0 \Rightarrow (\phi - \lambda id)(v) = 0 \Rightarrow v \in \ker(\phi - \lambda id)$$

Cabe destacar que la suma de endomorfismos autoadjuntos es de nuevo un endomorfismo autoadjunto y en este caso ϕ y $\lambda \cdot id$ lo son.

Sin embargo, aunque hemos demostrado que en esa cierta base ϕ es diagonal, puede ocurrir que la base no sea ortonormal, cosa de la que nos tenemos que cerciorar. Sabemos que $B = B_1 \cup \dots \cup B_t$ donde B_i es base de V_{ϕ, λ_i} . En primer lugar, sí podemos garantizar que vectores propios de autovalores distintos son ortogonales:

$$\begin{aligned} f(\phi(v_i), v_j) &= f(v_i, \phi(v_j)) \Rightarrow f(\lambda_i v_i, v_j) = f(v_i, \lambda_j v_j) \Rightarrow \lambda_i f(v_i, v_j) = \lambda_j f(v_i, v_j) \Rightarrow \\ &\Rightarrow \underbrace{(\lambda_i - \lambda_j)}_{\neq 0} f(v_i, v_j) = 0 \Rightarrow f(v_i, v_j) = 0 \end{aligned}$$

Así que sabemos que todos los vectores entre B_i distintas ya son ortogonales. Para poder terminar, si restrinjo a cada subespacio propio el producto escalar, este sigue siendo un producto escalar y, en consecuencia, existe una base ortonormal para cada V_{ϕ, λ_i} formada además por vectores propios así que ya he terminado.

Observación:

Supongamos que tenemos (\mathbb{R}^n, \cdot) con un producto escalar y un endomorfismo autoadjunto ϕ . Si escogemos una base ortonormal B de manera $M_B(\phi) = A$ no sea diagonal y, aplicando el teorema, cambiamos a una base ortonormal donde $M_{B'}(\phi) = A'$ sea diagonal, vemos que:

$$M_{B'}(f) = P^t M_B(f) P \Rightarrow I_n = P^t I_n P \Rightarrow P^t = P^{-1}$$

es decir, que, en consecuencia, ϕ se diagonaliza simultáneamente por congruencia y por semejanza, ya que:

$$A' = P^{-1} A P = P^t A P$$

APLICACIONES ORTOGONALES

Las aplicaciones ortogonales son aplicaciones de gran importancia pues conservan el producto escalar, esto quiere decir, en términos geométricos que respetan el “ángulo” (puesto que aún no hemos definido la noción de ángulo) que forman varios vectores entre sí. Cabe destacar que hemos dicho que únicamente respetan la posición relativa entre ellos, no que no deformen dichos vectores aumentando su módulo o disminuyéndolo.

Características de las aplicaciones ortogonales

Primero vamos aproximarnos al concepto de aplicación ortogonal y ver que consecuencias tiene con las definiciones de espacio vectorial euclídeo que se han dado. Una vez conozcamos los detalles técnicos que las hacen tan especiales, estaremos listos para poder clasificar en su conjunto los tipos de aplicaciones en cada dimensión.

Definición (Aplicación Conservativa)

Sea (V, f) un espacio vectorial euclídeo y $\phi \in \text{End}(V)$, decimos que ϕ **conserva** f si verifica que:

$$f(\phi(v), \phi(w)) = f(v, w) : \forall v, w \in V$$

Proposición

Si ϕ conserva f , entonces ϕ es \mathbb{R} -lineal.

Demostración:

Es lineal ya que: $\begin{cases} \phi(v+w) = \phi(v) + \phi(w) \\ \phi(\lambda v) = \lambda\phi(v) \end{cases}$. Para comprobar, por ejemplo, la segunda:

$$f(\phi(\lambda v) - \lambda\phi(v), \phi(\lambda v) - \lambda\phi(v)) = f(\phi(\lambda v), \phi(\lambda v)) + \lambda^2 f(\phi(v), \phi(v)) - 2\lambda f(\phi(\lambda v), \phi(v))$$

Como f es producto escalar

$$= f(\lambda v, \lambda v) + \lambda^2 f(v, v) - 2\lambda f(\lambda v, v) = \lambda^2 f(v, v) + \lambda^2 f(v, v) - 2\lambda^2 f(v, v) = 0$$

Para comprobar la primera de las condiciones habría que ver que:

$$f(\phi(v+w) - \phi(v) - \phi(w), \phi(v+w) - \phi(v) - \phi(w))$$

Definición (Aplicación Ortogonal)

Sea (V, f) un espacio vectorial euclídeo y $\phi \in \text{End}(V)$, decimos que ϕ es una **aplicación ortogonal** si ϕ es lineal y conserva f . Denotamos por $O(V) = \{\phi \in \text{End}(V) : \phi \text{ ortogonal}\}$ al conjunto de aplicaciones ortogonales.

Proposición (Grupo Ortogonal)

Sea (V, f) un espacio vectorial euclídeo, las aplicaciones ortogonales verifican las siguientes propiedades:

- $\phi \in O(V) \Rightarrow$ biyectiva
- $\phi \in O(V) \Rightarrow \phi^{-1} \in O(V)$
- $\phi, \psi \in O(V) \Rightarrow \psi \circ \phi \in O(V)$

Por tanto, el conjunto de aplicaciones ortogonales con operación de composición forman lo que se conoce como **Grupo Ortogonal** $(O(V), \circ)$.

Demostración:

1. Por ser un endomorfismo, basta con probar que es inyectiva, así que:

$$v \in \ker \phi \Rightarrow \phi(v) = 0 \Rightarrow f(\phi(v), \phi(v)) = f(v, v) = 0 \Rightarrow v = 0$$

2. Como es biyectiva, existe la inversa y entonces es trivial comprobar que es ortogonal.

3. Trivial

Observación:

El conjunto de isomorfismos se denotaba por $GL(V)$, de modo que $(GL(V), \circ)$ tenía estructura de grupo. Es sencillo ver que $O(V) < GL(V)$ y como $GL(V) \simeq Mat_n(K)$ es sencillo demostrar que $O(V) \simeq \{A \in Mat_n(K) : A^t = A^{-1}\}$.

Definición (Orientación de una Base)

Sean B, B' bases² de V y $P = M(B', B) \in GL_n(\mathbb{R})$ la matriz de cambio entre ambas, decimos que:

$$\begin{cases} |P| > 0 & \Rightarrow B \text{ y } B' \text{ tienen la } \textbf{misma orientación} \\ |P| < 0 & \Rightarrow B \text{ y } B' \text{ tienen } \textbf{distinta orientación} \end{cases}$$

De este modo, si tenemos $\phi \in GL(V)$ y tomamos una base cualquiera $B = \{v_1, \dots, v_n\}$ y consideramos la base resultante de las imágenes de los vectores a través de ϕ que denotamos por $B' = \{\phi(v_1), \dots, \phi(v_n)\}$, entonces se verifica que:

$$M(B', B) = M_B(\phi) = A \Rightarrow \begin{cases} |A| > 0 \\ |A| < 0 \end{cases}$$

Por tanto, podemos distinguir en $GL(V)$ (y por extensión en $O(V)$) dos grupos de aplicaciones, las que conservan la orientación de las bases y las que invierten esta orientación.

Definición (Grupo Conservativo e Inversivo)

Sea $GL(V)$ el conjunto³ de isomorfismos de un espacio vectorial en sí mismo, definimos $GL^+(V)$ como el grupo de isomorfismos que conservan la orientación de las bases y $GL^-(V)$ como el que las invierte. Además, la unión de ambos es el total:

$$GL(V) = GL^+(V) \sqcup GL^-(V)$$

Ejemplo:

Podemos descomponer $\mathbb{R}^n = W \oplus W^\perp$ porque en este contexto f es un producto escalar y entonces la suma es directa.

Si consideramos $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ la simetría ortogonal respecto de W (simetría con base W y dirección W^\perp). Para ver que es ortogonal vamos a tomar dos vectores cuales quiera y vamos a ver si se verifican las premisas de ortogonalidad:

$$\begin{cases} v_1 \in \mathbb{R}^n : v_1 = w_1 + w'_1 \\ v_2 \in \mathbb{R}^n : v_2 = w_2 + w'_2 \end{cases}$$

Sabiendo que las simetrías ya son aplicaciones lineales, vemos si conserva el producto escalar definido:

$$\begin{cases} f(\phi(v_1), \phi(v_2)) = f(w_1 - w'_1, w_2 - w'_2) = f(w_1, w_2) + f(w'_1, w'_2) \\ f(v_1, v_2) = f(w_1 + w'_1, w_2 + w'_2) = f(w_1, w_2) + f(w'_1, w'_2) \end{cases} \Rightarrow f(\phi(v_1), \phi(v_2)) = f(v_1, v_2)$$

Luego es una aplicación ortogonal.

Por el contrario, hay que tener cuidado porque no siempre podemos verificar esto. Ocurre que sea $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ la proyección ortogonal respecto de W [proyección con base W y dirección W^\perp] no es una aplicación ortogonal porque a pesar de que descompone como suma directa de subespacios, no es biyectiva.

²La base canónica B_c de \mathbb{R}^n es positiva por convenio.

³La definición es completamente válida y análoga para el grupo $O(V)$.

Proposición (Propiedades de las aplicaciones ortogonales)

Sea $\phi \in \text{End}(V)$ con matriz asociada $A = M_B(\phi)$ y f un producto escalar con matriz $C = M_B(f)$, se verifican:

- $\phi \in O(V) \Leftrightarrow A^t C A = C$
- $\phi \in O(V) \stackrel{B \text{ ortonormal}}{\Leftrightarrow} A^t A = I_n$
- $\phi \in O(V) \Rightarrow |A| = \pm 1 : \begin{cases} \phi \in O^+(V) & \Rightarrow |\phi| = 1 \\ \phi \in O^-(V) & \Rightarrow |\phi| = -1 \end{cases}$
- Los únicos autovalores posibles son $\sigma(\phi) \subset \{-1, 1\}$
- $\phi \in O(V)$ y $W <_\phi V \Rightarrow W^\perp <_\phi V$. Además, $V = W \oplus W^\perp$.

Demostración:

- Trivialmente por la definición que se ha dado de aplicación ortogonal: $f(\phi(v), \phi(w)) = f(v, w)$
- De la primera se deduce fácilmente por ser $C = I_n$ en base ortonormal.
- Si la base es ortonormal, entonces $A^t A = I \Rightarrow |A^t A| = |I| \Rightarrow |A^t| |A| = |I| \Rightarrow |A| = \pm 1$. Además, como $|A|$ es un invariante para cualquier matriz que represente a A , tenemos que se cumple a pesar de que no esté respecto de una base ortonormal.
- Supongamos que $\phi(v) = \lambda v \Rightarrow 0 < \|v\|^2 = f(v, v) = f(\phi(v), \phi(v)) = f(\lambda v, \lambda v) = \lambda^2 \|v\|^2 \Rightarrow \lambda^2 = 1 \Rightarrow \lambda = \pm 1$
- La segunda parte ya la tenemos demostrada, queda demostrar que el ortogonal es ϕ -invariante. Tenemos que ver que:

$$f(v, w) = 0 : \forall w \in W \Rightarrow f(\phi(v), w) = 0 : \forall w \in W$$

Para ello aplicamos el siguiente razonamiento:

$$\phi \in O(V) \text{ y la dimensión es finita} \Rightarrow \phi : V \rightarrow V \text{ biyectiva} \Rightarrow \phi|_W : W \rightarrow W \text{ biyectiva}$$

De este modo, puedo considerar w como $\phi(w')$ para algún $w' \in W$ luego se tiene que:

$$f(\phi(v), w) = f(\phi(v), \phi(w')) = f(v, w') = 0$$

Clasificación de aplicaciones ortogonales

Nuestro objetivo ahora es el que viene siendo en los capítulos anteriores, poder llegar a un conjunto de clases de equivalencia en las que poder meter todas las aplicaciones de cierto tipo y elegir un representante adecuado para dicha clase.

Dimensión 1

Si tomamos $\dim V = 1$, entonces tenemos que:

$$O(V) = \{id_v, -id_v\}$$

Dimensión 2

Para $\dim V = 2$, consideramos $\phi \in O(V) \Rightarrow M_B(\phi) = A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ en una base $B = \{v_1, v_2\}$ ortonormal. Vamos a analizar su matriz atendiendo a las condiciones de ortogonalidad que debe cumplir:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A^t A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}$$

Cuando vemos esta relación rápidamente surge la idea de considerar estos valores como el coseno y el seno de un único ángulo, quedando:

$$\begin{cases} a = \cos \theta \\ b = \sin \theta \\ c = \cos \theta' \\ d = \sin \theta' \end{cases} : 0 \leq \theta < 2\pi$$

Las condiciones de que sean igual a 1, de este modo se verifican, pero para que se verifiquen las de la izquierda se tiene que:

$$0 = ac + bd = \cos \theta \cos \theta' + \sin \theta \sin \theta' = \cos(\theta - \theta') \Leftrightarrow \theta' = \begin{cases} \theta + \frac{\pi}{2} \\ \theta + \frac{3\pi}{2} \end{cases}$$

Por tanto, ahora podemos distinguir dos casos para clasificar las matrices, suponiendo que $\theta' = \theta + \frac{\pi}{2}$ o $\theta' = \theta + \frac{3\pi}{2}$. De este modo, la primera matriz posible que tenemos es:

$$\theta' = \theta + \frac{\pi}{2} \Rightarrow \begin{cases} c = \cos \theta' = -\sin \theta = -b \\ d = \sin \theta' = \cos \theta = a \end{cases} \Rightarrow A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in O^+(V)$$

$$\theta' = \theta + \frac{3\pi}{2} \Rightarrow \begin{cases} c = \cos \theta' = \sin \theta = b \\ d = \sin \theta' = -\cos \theta = -a \end{cases} \Rightarrow A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \in O^-(V)$$

Por tanto, si dividimos en ambas clasificaciones vemos que:

$$\phi \in O^+(V) \Rightarrow M_B(\phi) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad 0 \leq \theta < 2\pi$$

Que corresponde a una rotación de los vectores del plano de ángulo θ , además como:

$$p_\phi = x^2 - 2x \cos \theta + 1 \Rightarrow x = \frac{2 \cos \theta \pm \sqrt{4 \cos^2 \theta - 4}}{2} = \cos \theta \pm \sqrt{\cos^2 \theta - 1} \Rightarrow \cos \theta = 1 \Rightarrow x = \pm 1$$

No existen autovalores en general, excepto en el caso extremo de la rotación de ángulo 0 y la de ángulo 180° (porque en el fondo con como simetrías).

Y, por otro lado, considerando las negativas:

$$\phi \in O^-(V) \Rightarrow M_B(\phi) = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} : a^2 + b^2 = 1$$

Estas matrices pueden ser consideradas simetrías del plano porque podemos considerar que:

$$p_\phi = x^2 - 1 = (x + 1)(x - 1) = q_\phi \Rightarrow M_{B'}(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dimensión 3

Si tomamos $\dim V = 3$, entonces el polinomio característico tiene como mínimo una raíz real que ya hemos visto que solo puede ser ± 1 . La caja restante corresponde a una rotación o a una simetría:

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{array} \right) \in O^+(V)$$

$$\left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{array} \right) \in O^-(V)$$

Cabe destacar que este posible caso teórico en el fondo está incluido en el primero por ser un caso de rotación especial:

$$\left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & -1 & 0 \\ 0 & 0 & 1 \end{array} \right) \in O^+(V)$$

Clasificación en dimensión arbitraria

Para dimensión arbitraria tenemos el siguiente teorema:

$$\phi \in O(V) \Rightarrow \exists B \text{ ortonormal de } V :$$

$$M_B(\phi) = I_s \oplus -I_t \oplus \begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix} \oplus \dots \oplus \begin{pmatrix} \cos \theta_m & \sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{pmatrix} : s, t, m \geq 0 \wedge \theta_i \in (0, 2\pi) : n = s+t+m$$

Demostración:

Como el caso de dimensión 1 es trivial y además ya lo hemos visto en los ejemplos anteriores, así que suponemos cierto para dimensiones menores que n y vamos a verlo para n .

- Supongamos primero que $\lambda = \pm 1 \in \sigma(\phi)$. Entonces existen vectores propios para ese autovalor, consideramos $v_1 \neq 0 \in V_{\phi, \lambda}$ y el subespacio generado por él $W = L(v_1) <_{\phi} V$. Por ser producto escalar tenemos que:

$$V =_{\phi} W \oplus W^{\perp}$$

Si restringimos la aplicación a este subespacio invariante, vuelve a ser un endomorfismo y además ortogonal, luego si consideramos $B = \underbrace{\{v_1\}}_{\in W}, \underbrace{\{v_2, v_3, \dots, v_n\}}_{\in W^{\perp}}$:

$$\phi|_{W^{\perp}} \in O(W^{\perp}) \Rightarrow M_{\{v_1, v_2, \dots, v_n\}}(\phi) = \left(\begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & M(\phi|_{W^{\perp}}) & \\ 0 & & & \end{array} \right) = \left(\begin{array}{c|ccc} \pm 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & M(\phi|_{W^{\perp}}) & \\ 0 & & & \end{array} \right)$$

Y como esas restricciones tienen dimensión menor que n , entonces se cumple por hipótesis de inducción.

- Supongamos, por otro lado, que $\sigma(\phi)$ es vacío, entonces sabemos que:

$$q_{\phi} = f_1^{a_1} \dots f_t^{a_t} : f_1, \dots, f_t \in \mathbb{R}[x] \text{ irreducibles de grado } 2$$

Ahora puedo fijarme en f_1 y encontrar un elemento que tenga ese polinomio mínimo, es decir:

$$\exists v_1 \in V : q_{\phi, v_1} = f_1 = x^2 + a_1 x + a_0$$

De este modo:

$$\pi = L(v_1, \phi(v_1)) <_{\phi} V : \dim \pi = 2$$

Y, por tanto, por ser producto escalar se tiene que:

$$V = \pi \oplus \pi^{\perp}$$

Restringimos de nuevo $\phi|_{\pi^{\perp}}$ que sigue siendo ortogonal y en ese trozo la matriz que conforma ya cumple las hipótesis. Los vectores de π son ortogonales a los del perpendicular y podemos seleccionarlos de forma que conformen una base ortonormal en π por lo que compondrán una de las cajas que habíamos definido como rotaciones.

ESPACIO AFÍN

Durante este capítulo vamos a desarrollar toda la teoría necesaria para introducir los conceptos de espacio y aplicaciones afines. Esta nueva estructura junto con las aplicaciones asociadas a la misma nos van a permitir modelar el espacio real y la geometría tal y como la conocemos de forma analítica, completando la definición de espacio vectorial euclídeo que se dio.

ESPACIO AFÍN

Con esta nueva estructura buscamos enriquecer la noción de espacio vectorial euclídeo. Con éste último habíamos sido capaces de introducir las nociones ortogonalidad, producto escalar, etc. Sin embargo, dicha definición nos dejaba anclados a subespacios fijos al origen, pues todos los subespacios contienen al 0. Supuesto el plano $z = 0$, si quisiéramos describir los paralelos a distintas alturas la definición de espacio vectorial euclídeo se nos quedaría corta. De esta necesidad nace el espacio afín.

Definición (Espacio Afín)

Un **espacio afín** sobre un cuerpo K , es una terna (A, V, φ) que verifica:

- $A \neq \emptyset$.
- V es un \mathbb{K} -espacio vectorial de dimensión finita.
- φ es una función que a cada par de puntos le asigna un vector de V , es decir:

$$\begin{aligned}\varphi : A \times A &\longrightarrow V \\ (p, q) &\longmapsto \vec{pq} = \varphi(p, q)\end{aligned}$$

- $\forall p, q, r \in A : \varphi(p, q) = \varphi(p, r) + \varphi(r, q)$
- $\forall p \in A$ la siguiente función es una biyección:

$$\begin{aligned}\varphi_p : A &\longrightarrow V \\ q &\longmapsto \vec{pq} = \varphi(p, q)\end{aligned}$$

Ejemplo:

Por ejemplo, definimos el espacio afín usual como:

$$\begin{cases} A = \mathbb{K}^n \\ V = \mathbb{K}^n \\ \varphi(p, q) \stackrel{def}{=} (b_1 - a_1, \dots, b_n - a_n)_{B_c} \in \mathbb{K}^n : p = (a_1, \dots, a_n), q = (b_1, \dots, b_n) \end{cases}$$

Las primeras 3 propiedades son trivialmente ciertas, comprobamos la 4:

$$\begin{aligned} r = (c_1, \dots, c_n) \Rightarrow \varphi(p, r) + \varphi(r, q) &= (c_1 - a_1, \dots, c_n - a_n)_{B_c} + (b_1 - c_1, \dots, b_n - c_n)_{B_c} = \\ &= (c_1 - a_1 + b_1 - c_1, \dots, c_n - a_n + b_n - c_n)_{B_c} = (b_1 - a_1, \dots, b_n - a_n)_{B_c} = \varphi(p, q) \end{aligned}$$

Comprobamos que φ_p cumple la 5, es decir, es biyectiva:

■ Inyectiva.

$$\begin{aligned} (b_1 - a_1, \dots, b_n - a_n)_{B_c} = \varphi_p(q) = \varphi_p(r) &= (c_1 - a_1, \dots, c_n - a_n)_{B_c} \Rightarrow \\ \Rightarrow b_i - a_i = c_i - a_i \Rightarrow b_i &= c_i : \forall i \Rightarrow b = c \end{aligned}$$

■ Suprayectiva

Sea $v = (x_1, \dots, x_n)_{B_c}$, ¿tenemos un $\varphi_p(q) = v : q \in A$?

$$\varphi_p(q) = (b_1 - a_1, \dots, b_n - a_n)_{B_c} \stackrel{?}{=} (x_1, \dots, x_n)_{B_c} \Rightarrow b_i - a_i = x_i$$

Luego ya tenemos la definición de cada b_i y, por lo tanto, ya está demostrado.

Observación:

En general, se puede considerar cualquier espacio vectorial como un espacio afín si lo definimos de la siguiente manera:

$$\begin{cases} V \\ A = V \\ \varphi(v, w) = w - v \in V \end{cases} \Rightarrow (V, V, \varphi) \text{ es espacio afín}$$

Aunque esto pone de manifiesto que todo espacio vectorial puede considerarse como un espacio afín, esta forma no es única porque podemos dar una definición distinta para φ o incluso para algún otro elemento en casos concretos.

Proposición

Sea (A, V, φ) un espacio afín, se verifican las siguientes propiedades:

1. $p\vec{p} = 0$
2. $q\vec{p} = -p\vec{q}$
3. Propiedad del Paralelogramo: $p\vec{q} = r\vec{s} \Leftrightarrow p\vec{r} = q\vec{s}$

Demostración:

1. $p\vec{p} + p\vec{p} = p\vec{p} \Rightarrow p\vec{p} = 0$
2. $p\vec{q} + q\vec{p} = p\vec{p} = 0 \Rightarrow q\vec{p} = -p\vec{q}$

$$\varphi_p(q) = p\vec{q} = 0 = p\vec{p} = \varphi_p(p) \xrightarrow{\varphi_p \text{ biyectiva}} q = p$$

3. Hacemos " \Rightarrow " porque la otra dirección es completamente análoga:

$$p\vec{r} = p\vec{q} + q\vec{r} = r\vec{s} + q\vec{r} = q\vec{s}$$

Definición (Notación de traslación de vector v)

Hemos definido la φ como la asignación de un vector concreto para cada par de puntos de A que escojamos, es decir, $\vec{pq} = v \in V$, entonces podemos emplear la siguiente notación:

$$p + v \stackrel{\text{def}}{=} q$$

En la que el operador “suma” no tiene por qué ser la suma¹ usual, sino aquella que cumpla la igualdad de la notación, es decir, será la aplicación:

$$\begin{aligned} + : A \times V &\longrightarrow A \\ (p, v) &\longmapsto q \end{aligned}$$

donde $q \in A$ es el único punto tal que $v = \vec{pq} = \varphi_p(q)$. Por tanto, realmente estamos diciendo que:

$$\boxed{p + v = \varphi_p^{-1}(v) = q}$$

que está bien definida por ser una biyección φ_p .

Proposición

Sea φ la aplicación de un espacio afín y la notación para la aplicación φ_p^{-1} , entonces esta cumple las siguientes propiedades:

1. $p + 0 = p$
2. $p + (v + w) = (p + v) + w$
3. $\forall p, q \in A : \exists! v \in V : p + v = q$

Demostración:

1. Porque si consideramos que $\vec{pp} = 0$, entonces sumarle a p el vector nulo implica que su destino es el mismo.
2. Trivial, porque basta con considerar
$$\begin{cases} v = \vec{pq} \\ w = \vec{qr} \\ v + w = \vec{pr} \end{cases} \Rightarrow r = q + w = q + \vec{qr} = r$$
3. Trivial por ser φ_p^{-1} biyectiva.

Definición (Dimensión de un Espacio Afín)

Sea (A, V, φ) un espacio afín de forma que $\dim V = n$, entonces definimos la **dimensión de A** como la dimensión de V , esto es, $\dim A = \dim V = n$.

Variedades Afines

Del mismo modo que en la estructura de espacio vectorial teníamos los subespacios, en la estructura de espacio afín tenemos el término variedad para referirnos a su homólogo. De esta manera, la restricción a dichas variedades componen de nuevo una estructura de espacio afín.

Definición (Variedad Afín)

Dado un espacio afín (A, V, φ) , un punto $p \in A$ y un subespacio $W < V$ se define la **variedad afín que pasa por p con dirección W** como el conjunto:

$$p + W = \{p + w : w \in W\} = \{q \in A : \vec{pq} \in W\}$$

¹En el fondo lo que estamos haciendo es dado un punto y un vector, la aplicación que hemos definido asocia a ese par donde acabaría el punto tras trasladarlo lo que manda el vector.

Observación:

Es notable destacar que $X = p + W$, $W < V$ y $\varphi|_{X \times X}$ tienen de nuevo estructura de espacio afín porque:

1. $X \neq \emptyset : p = p + 0 \in X$
2. $W < V$
3. $\varphi|_{X \times X} : X \times X \rightarrow W$

$$\varphi(q, r) = \varphi(p + w_1, p + w_2) = \varphi(p + w_1, p + w_1 + (w_2 - w_1)) = w_2 - w_1 \in W$$

4. Trivial
5. Inyectiva es inmediato porque es la restricción de una que sí es inyectiva.
6. Es suprayectiva si se verifica $\forall w \in W : \exists ! r \in X : \varphi_q(r) = w$, es decir, sea $w = \varphi_q(r) = \varphi(q, r) = \vec{qr} : r \in A$ único. $r = q + \vec{qr} = p + (\vec{pq} + \vec{qr})$

$$\begin{cases} \vec{pq} = \varphi(p, q) [p, q \in X] \in W \\ \vec{qr} = \varphi(q, r) = \varphi_q(r) = w \in W \end{cases} \Rightarrow r = q + \vec{qr} = p + (\vec{pq} + \vec{qr}) \in p + W = X$$

Proposición

Sea $X = p + W$ una variedad afín, entonces:

- El punto de definición de la variedad se puede cambiar por cualquier otro de la misma:

$$\forall q \in X : X = q + W$$

- La dirección de la variedad, que denotaremos por \vec{X} queda determinada por todos los vectores formados por el punto de definición y otro punto de X o todos los vectores generados por dos puntos:

$$W = \underbrace{\{\vec{pq} : q \in X\}}_{p\vec{X}} = \underbrace{\{\vec{qr} : q, r \in X\}}_{\vec{X}}$$

Demostración:

- Vamos a demostrarlo por el doble contenido, es decir, vamos a ver que siendo p el punto de definición de la variedad afín, siempre puede escribirse cualquier elemento de la variedad en términos de otro punto q .

- $p + W \subseteq q + W$:

$$p + w = q + (\vec{qp} + w) \in q + W$$

Y esto ya demuestra el contenido porque $w \in W$ y $\vec{qp} \in W$ también debido a que $q \in X \Rightarrow q = p + w' \Rightarrow \vec{pq} = w' \Rightarrow \vec{qp} = -w' \in W$

- \supseteq es análogo

- Ahora vamos a ver que cualquier vector de W se puede definir como el vector definido por dos puntos de la variedad: En primer lugar, es trivial ver que $p\vec{X} \subseteq \vec{X}$, luego lo realmente interesante es ver $\vec{X} \subseteq W$?

$$\vec{qr}_{p, r \in X} = \vec{qp} + \vec{pr} = \vec{pr} - \vec{pq}$$

Como $q, r \in X$, entonces ambos pueden escribirse de esta manera:

$$\begin{cases} r = p + \vec{pr} = p + w \\ q = p + \vec{pq} = p + w' \end{cases}$$

Luego entonces:

$$p\vec{r} - p\vec{q} = w - w' \in W$$

Para terminar vamos a ver $W \subseteq p\vec{X}$:

$$w \in W \Rightarrow q = p + w \in X \Rightarrow w = p\vec{q} \in p\vec{X}$$

Teorema (Operaciones con Variedades)

■ **Intersección**

Si la intersección entre dos variedades no es nula, es decir, se cortan, entonces ésta define de nuevo una variedad, es decir:

$$X_1, X_2 \text{ variedades} \Rightarrow X_1 \cap X_2 \text{ es variedad afin}$$

Además esta nueva variedad queda definida por cualquier punto que está en la intersección y la intersección de las direcciones:

$$X_1 \cap X_2 = p + (\vec{X}_1 + \vec{X}_2) : \begin{cases} p \in X_1 \cap X_2 \\ X_1 \cap \vec{X}_2 = \vec{X}_1 \cap \vec{X}_2 \end{cases}$$

■ **Unión:**

Del mismo modo que ocurría con los subespacios vectoriales, la unión de variedades no es en general una variedad:

$$X_1, X_2 \text{ variedades} \Rightarrow X_1 \cup X_2 \text{ no es variedad}$$

■ **Suma:**

La suma de variedades afines es la variedad más pequeña que contiene a la unión de ambas variedades, queda definida como:

$$X_1 + X_2 = p_1 + \vec{X}_1 + \vec{X}_2 + L(p_1\vec{p}_2)$$

donde $L(p_1\vec{p}_2)$ es el subespacio generado por el vector entre dos puntos cualesquiera de ambas variedades.

Demostración:

■ Veamos la intersección:

Primero uno de los contenidos:

$$\begin{cases} p \in X_1 \Rightarrow X_1 = p + \vec{X}_1 \supseteq p + (\vec{X}_1 \cap \vec{X}_2) \\ p \in X_2 \Rightarrow X_2 = p + \vec{X}_2 \supseteq p + (\vec{X}_1 \cap \vec{X}_2) \end{cases} \Rightarrow X_1 \cap X_2 \supseteq p + (\vec{X}_1 \cap \vec{X}_2)$$

Ahora vemos el otro contenido:

$$q \in X_1 \cap X_2 \Rightarrow q = \begin{cases} p + w_1 = p + p\vec{q}_1 : q_1 \in X_1 \\ p + w_2 = p + p\vec{q}_2 : q_2 \in X_2 \end{cases} \Rightarrow q = q_1 = q_2 \Rightarrow w = w_1 = w_2 \Rightarrow q = p + w \in p + (\vec{X}_1 \cap \vec{X}_2)$$

■ Veamos la unión y la suma:

Se deja para mirar en el libro

Teorema (Fórmula de las Dimensiones)

Sean X_1 y X_2 dos variedades afines, entonces se verifica que:

$$\begin{cases} X_1 \cap X_2 \neq \emptyset \Rightarrow \dim(X_1 + X_2) = \dim X_1 + \dim X_2 - \dim X_1 \cap X_2 \\ X_1 \cap X_2 = \emptyset \Rightarrow \dim(X_1 + X_2) = \dim X_1 + \dim X_2 - \dim \vec{X}_1 \cap \vec{X}_2 + 1 \end{cases}$$

Demostración:

- $X_1 \cap X_2 \neq \emptyset$, entonces hay que tener en cuenta que $p_1\vec{p}_2 \in \vec{X}_1 + \vec{X}_2$:

$$\begin{aligned}\dim(X_1 + X_2) &= \dim(\overrightarrow{X_1 + X_2}) = \dim(\vec{X}_1 + \vec{X}_2 + L(p_1\vec{p}_2)) = \dim(\vec{X}_1 + \vec{X}_2) = \\ &= \dim \vec{X}_1 + \dim \vec{X}_2 - \dim(\vec{X}_1 \cap \vec{X}_2) = \dim X_1 + \dim X_2 - \dim(X_1 \cap X_2)\end{aligned}$$

- $X_1 \cap X_2 = \emptyset$, entonces hay que tener en cuenta que $p_1\vec{p}_2 \notin \vec{X}_1 + \vec{X}_2$:

$$\begin{aligned}\dim(X_1 + X_2) &= \dim(\overrightarrow{X_1 + X_2}) = \dim(\vec{X}_1 + \vec{X}_2 + L(p_1\vec{p}_2)) = \dim(\vec{X}_1 + \vec{X}_2) + \underbrace{\dim p_1\vec{p}_2}_{=1} - \underbrace{\dim(\vec{X}_1 + \vec{X}_2 \cap p_1\vec{p}_2)}_{=0} \\ &= \dim \vec{X}_1 + \dim \vec{X}_2 - \dim(\vec{X}_1 \cap \vec{X}_2) + 1 - 0 = \dim X_1 + \dim X_2 - \dim(X_1 \cap X_2) + 1\end{aligned}$$

Definición (Variedad Afín definida por puntos)

Sea (A, V, φ) un espacio afín y $P = \{p_0, p_1, \dots, p_m\}$ un conjunto de puntos de A , la **variedad afín generada por P** será:

$$A(P) = p_0 + L(p_0\vec{p}_1, \dots, p_0\vec{p}_m)$$

donde dicha definición² no depende del punto que se tome como “origen”.

Proposición

Sea (A, V, φ) un espacio afín y $P = \{p_0, p_1, \dots, p_m\}$, la variedad generada por P es la menor variedad que contiene a los puntos de P .

Demostración:

Definimos la el conjunto:

$$A(P) = p_0 + L(p_0\vec{p}_1, \dots, p_0\vec{p}_m)$$

Por definición es una variedad y además es trivial que contiene a todos los puntos, ¿pero es la más pequeña posible que contiene a todos los puntos? Sí, porque su dirección es la suma de las direcciones de cada variedad “lineal” que compondría cada vector.

Definición (Posición relativa)

Sean X_1, X_2 variedades de (A, V, φ) , decimos que son:

- **Secantes:** $X_1 \cap X_2 \neq \emptyset$.
- **Paralelas:** $\vec{X}_1 \subseteq \vec{X}_2$ ó $\vec{X}_2 \subseteq \vec{X}_1$.
- **Cruzadas:** en caso contrario.

Observación:

Es destacable decir que el hecho de no exigir que el paralelismo implique la igualdad de las direcciones permite definir correctamente el concepto de paralelismo entre variedades de distinta dimensión, es decir, de recta paralela a un plano, plano paralelo a hiperplano, etc.

²Trivialmente se verifican las condiciones de Variedad Afín.

Ecuaciones cartesianas

Sea (A, V, φ) un espacio afín tal que $\dim A = n$ y $X = q_0 + \vec{X}$ una variedad afín tal que $\dim X = m$, $\vec{X} = \{u_1, \dots, u_k\}$ y $q_0 = (a_1, \dots, a_n)$. Nos proponemos, fijado un sistema $R = \{p_0, B = \{v_1, \dots, v_n\}\}$ de referencia cartesiana, encontrar las condiciones que debe cumplir un punto $p \in X$.

Cualquier $p = (x_1, \dots, x_n) \in X$ pertenece si y sólo si $q_0 \vec{p} \in \vec{X}$, por tanto, debe ser combinación lineal de los que forman la dirección de la variedad:

$$rg \begin{pmatrix} x_1 - a_1 & \overbrace{b_1^1}^{u_1} & \dots & \overbrace{b_1^k}^{u_k} \\ \vdots & \vdots & & \vdots \\ x_n - a_n & b_n^1 & \dots & b_n^k \end{pmatrix} = rg \begin{pmatrix} \overbrace{b_1^1}^{u_1} & \dots & \overbrace{b_1^k}^{u_k} \\ \vdots & & \vdots \\ b_n^1 & \dots & b_n^k \end{pmatrix} = k$$

Escogemos de la segunda matriz un menor no nulo de orden k (debe haber alguno por tener rango k) y consideraremos por simplicidad que es el primero:

$$\begin{vmatrix} b_1^1 & \dots & b_1^k \\ \vdots & \ddots & \vdots \\ b_k^1 & \dots & b_k^k \end{vmatrix} \neq 0$$

Por tanto, lo que tiene que ocurrir para que se cumpla la afirmación anterior sobre el rango es:

$$\forall i = k+1, \dots, n : \begin{vmatrix} x_1 - a_1 & b_1^1 & \dots & b_1^k \\ \vdots & \ddots & & \vdots \\ x_k - a_k & b_k^1 & \dots & b_k^k \\ x_i - a_i & b_i^1 & \dots & b_i^k \end{vmatrix} = 0$$

Todos estos determinantes generan $n-k$ ecuaciones lineales que forman un sistema con n incógnitas.

Ecuaciones paramétricas

Supongamos que partimos del sistema no homogéneo de las ecuaciones cartesianas:

$$X = \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \dots + a_{pn}x_n = b_p \end{cases} \Rightarrow AX = B$$

En primer lugar y por el desarrollo que se ha hecho para las cartesianas, sabemos que p es la codimensión de la variedad, luego habrá algún menor cuadrado de orden p cuyo determinante no sea nulo. Por comodidad de notación y sin pérdida de generalidad³ podemos suponer que:

$$\begin{vmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{p1} & \dots & a_{pp} \end{vmatrix} \neq 0$$

De este modo, podemos parametrizar $n-p$ incógnitas y pasarlas a formar parte de la matriz B . En consecuencia, bastaría con resolver el sistema en función de los parámetros:

$$\begin{cases} x_{p+1} = \lambda_{p+1} \\ \vdots \\ x_n = \lambda_n \\ x_1 = c_{1,p+1}\lambda_{p+1} + \dots + c_{1n}\lambda_n + d_1 \\ \vdots \\ x_p = c_{p,p+1}\lambda_{p+1} + \dots + c_{pn}\lambda_n + d_p \end{cases}$$

³Puesto que permutar las variables y las ecuaciones entre sí podría convertir el menor concreto del caso en el que se ha escogido para la explicación.

En definitiva, la forma de calcular la ecuación paramétrica de dicha variedad es resolver el sistema formado por sus ecuaciones cartesianas en función de $n - p$ parámetros.

De modo análogo, si supiésemos cuál es el punto q_0 de definición de la variedad X y la dirección $\vec{X} = \{v_1, \dots, v_k\}$ que la compone para obtener las paramétricas bastaría con escribir:

$$p = q_0 + \lambda_1 v_1 + \dots + \lambda_k v_k \Leftrightarrow \begin{cases} x_1 = q_0^1 + \lambda_1 v_1 \\ \vdots \\ x_k = q_0^k + \lambda_k v_k \\ x_{k+1} = q_0^{k+1} \\ \vdots \\ x_n = q_0^n \end{cases}$$

Paso de paramétricas a cartesianas

Si tenemos por ejemplo el plano $\pi = (-1, 1, 0, 0) + L(e_1 - e_2, e_3 - e_4)$ y queremos saber sus ecuaciones paramétricas tenemos que:

$$\pi = \begin{cases} x = 1 + \lambda \\ y = -1 - \lambda \\ z = \mu \\ t = 1 - \mu \end{cases}$$

Para pasar a cartesianas nos basamos en que:

$$p = (x, y, z, t) \in \pi \Rightarrow \exists \lambda, \mu \in \mathbb{R} : \begin{cases} x = 1 + \lambda \\ y = -1 - \lambda \\ z = \mu \\ t = 1 - \mu \end{cases}$$

tiene solución única, luego si consideramos como incógnitas los parámetros y como constantes las coordenadas tenemos que:

$$rg \begin{pmatrix} x-1 & 1 & 0 \\ y+1 & -1 & 0 \\ z & 0 & 1 \\ t-1 & 0 & 1 \end{pmatrix} = rg \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

Y una vez que sabemos el rango de la matriz de coeficientes basta con escoger un menor no nulo de ese rango en la matriz ampliada y orlarlo con todos los menores de mayor rango, de modo que todos estos sean nulos:

$$\begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} \Rightarrow \begin{vmatrix} x-1 & 1 & 0 \\ y+1 & -1 & 0 \\ z & 0 & 1 \end{vmatrix} = 0 \quad \begin{vmatrix} y+1 & -1 & 0 \\ z & 0 & 1 \\ t-1 & 0 & 1 \end{vmatrix} = 0$$

Y de aquí salen las dos ecuaciones correspondientes al plano en \mathbb{R}^4

Sistema de Referencia

Sea (A, V, φ) la terna usual de espacio afín, nos gustaría poder determinar de forma única un punto $p \in A$ arbitrario del mismo modo que los vectores quedaban determinados de forma única por una base en espacios vectoriales.

Para ello, escogemos un punto $p_0 \in A$ fijo y una base $B = \{v_1, \dots, v_n\}$ de V y, de este modo, podemos describir dicho punto p como:

$$p_0 \vec{p} = v = x_1 v_1 + \dots + x_n v_n$$

Y, de hecho, estas coordenadas son únicas pues $\varphi_{p_0}(p) = v$ y es una biyección por hipótesis. Luego estamos en condiciones de dar la siguiente definición.

Definición (Sistema de Referencia Afín)

Sea (A, V, φ) un espacio afín, $p_0 \in A$ un punto cualquiera y $B = \{v_1, \dots, v_n\}$ una base de V , definimos un **sistema de referencia afín** como el conjunto $R = \{p_0, B\}$.

Dado un punto $p \in A$ cualquiera, se definen sus **coordenadas cartesianas** como los coeficientes del vector asignados por φ_{p_0} , es decir:

$$p = (x_1, \dots, x_n)_R \Leftrightarrow \varphi_{p_0}(p) = \vec{p_0 p} = (x_1, \dots, x_n)_B = x_1 v_1 + \dots + x_n v_n$$

A pesar de que la idea de tener un punto a modo de “origen” de coordenadas y una base B a modo de direcciones es muy intuitiva para dar una definición de sistema de referencia, también podemos dar una definición alternativa que prescinde de la base definida para V y se supedita en puntos. La relación entre lo que denominaremos el origen y el resto de puntos⁴ de referencia es lo que hará las veces de base B , puesto que determinan los vectores sus vectores.

Definición (Sistema de Referencia por puntos)

Sea (A, V, φ) un espacio afín, $p_0 \in A$ un punto cualquiera, entonces decimos que el conjunto $\{p_0, \dots, p_n\}$ es un **sistema de referencia afín** si y sólo si $\{p_0 \vec{p}_1, \dots, p_0 \vec{p}_n\}$ es base de V .

Observación:

Luego entonces $R = \{p_0, B = \{p_0 \vec{p}_1, \dots, p_0 \vec{p}_n\}\}$ es el sistema de referencia cartesiano equivalente.

Definición (Dependencia Lineal Afín)

Sea $\{p_0, p_1, \dots, p_m\}$ un conjunto de puntos de A con $m \leq n$, decimos que son **afínmente independientes** si $\{p_0 \vec{p}_1, \dots, p_0 \vec{p}_m\}$ son linealmente independientes.

Observación:

Además, la independencia de estos vectores no dependen de quién sea escogido como origen de los vectores.

Cambio de sistema de referencia

Definido un sistema de referencia cartesiano $R = \{p_0, B\}$ formado por un punto $p_0 \in A$ y una base $B = \{v_1, \dots, v_n\}$ de V , habíamos dicho que las coordenadas de un punto $p \in A$ cualquiera venían dadas por:

$$p = (x_1, \dots, x_n)_R \stackrel{def}{\Leftrightarrow} \vec{p_0 p} = (x_1, \dots, x_n)_B = x_1 v_1 + \dots + x_n v_n$$

La pregunta ahora es ¿cómo cambian dichas coordenadas si cambiamos de punto $p_0 \in A$ o de base B ? Supongamos que cambiamos a un nuevo sistema de referencia $R' = \{p'_0, B'\}$. Las coordenadas de p ahora son $(x'_1, \dots, x'_n)_{R'}$ y suponiendo las coordenadas de $p'_0 = (a_1, \dots, a_n)_R$ podemos relacionar ambos vectores:

$$\begin{aligned} \vec{p_0 p} &= \vec{p_0 p'_0} + \vec{p'_0 p} \Rightarrow \vec{p_0 p} - \vec{p_0 p'_0} = \vec{p'_0 p} \Rightarrow (x_1, \dots, x_n)_B - (a_1, \dots, a_n)_B = (x'_1, \dots, x'_n)_{B'} \\ &\Rightarrow (x_1 - a_1, \dots, x_n - a_n)_B = (x'_1, \dots, x'_n)_{B'} \Rightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + M(B', B) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \end{aligned}$$

Sin embargo, esta notación nos es muy incómoda pues no resulta un cambio de “base” similar al que estamos acostumbrados, por tanto, el siguiente cambio algebraico conserva la igualdad y usa una notación más cómoda.

⁴Es notable destacar que si $n = \dim A$, serán necesarios $n + 1$ puntos porque p_0 va a constituir el origen.

Teorema (Cambio de Sistema de Referencia)

Sea (A, V, φ) un espacio afín y $R = \{p_0, B\}$ y $R' = \{p'_0, B'\}$ sistemas de referencia afines, entonces las coordenadas de un punto $p \in A$ respecto de ambos sistemas se relacionan a través de la siguiente igualdad:

$$\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ a_1 & & & \\ \vdots & & M(B', B) & \\ a_n & & & \end{pmatrix}}_{M(R, R')} \begin{pmatrix} 1 \\ x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

APLICACIONES AFINES

Del mismo modo que vimos la importancia que revestían las aplicaciones lineales en los espacios vectoriales, en esta sección vamos a tratar definir y caracterizar el análogo de dichas aplicaciones lineales en espacios afines: las aplicaciones afines.

Definición (Aplicación Afín)

Sean (A, V, φ) y (A', V', φ') dos espacios afines, una aplicación $\phi : A \rightarrow A'$ entre ambos conjuntos y $\lambda : V \rightarrow V'$ una aplicación K -lineal, decimos que el par (ϕ, λ) es una **aplicación afín** si y sólo si:

$$\forall q \in A : q = p + v \Rightarrow \phi(q) = \phi(p) + \lambda(v)$$

Observación:

Por la definición que se ha dado, dada una aplicación ϕ concreta la aplicación lineal λ queda definida de manera única en función de las imágenes que otorgue ϕ a cada par de puntos, por este motivo, es común denotarla por $\vec{\phi}$:

$$v = \vec{pq} \Rightarrow \lambda(v) = \vec{\phi}(v) = \overrightarrow{\phi(p)\phi(q)}$$

Por tanto, para determinar si una aplicación ϕ es afín basta con ver que la asociada a ella $\vec{\phi}$ es lineal.

Teorema

Sean A, A' espacios afines y $R = \{p_0, B = \{v_1, \dots, v_n\}\}$ un sistema de referencia, entonces:

$$\forall p'_0 \in A' \text{ y } B' = \{v'_1, \dots, v'_n\} \text{ base de } V' : \exists ! \phi : A \rightarrow A' \text{ afín tal que } \begin{cases} \phi(p_0) = p'_0 \\ \vec{\phi}(v_i) = v'_i : \forall i = 1, \dots, n \end{cases}$$

es decir, que los valores que toma la aplicación afín sobre los elementos del sistema de referencia determinan de forma única la aplicación afín.

Demostración:

Sea $p \in A$ un punto cualquiera de A , entonces:

$$p = p_0 + \sum_{i=1}^n a_i v_i \Rightarrow \phi(p) = p'_0 + \sum_{i=1}^n a_i v'_i$$

Y se comprueba trivialmente que cumple las propiedades de afín.

Teorema

Sea $\{p_0, p_1, \dots, p_n\}$ sistema de referencia afín de A , entonces:

$$\forall p'_0, p'_1, \dots, p'_n \in A' : \exists ! \phi : A \rightarrow A' \text{ afín} : \phi(p_i) = p'_i : \forall i = 1, \dots, n$$

es decir, el teorema anterior se cumple para los sistemas dados por puntos.

Teorema

Sea $\phi : A \rightarrow A'$ una aplicación afín, entonces el carácter de ϕ depende estrictamente de $\vec{\phi}$, es decir:

- ϕ inyectiva $\Leftrightarrow \vec{\phi}$ inyectiva
- ϕ suprayectiva $\Leftrightarrow \vec{\phi}$ suprayectiva
- ϕ biyectiva $\Leftrightarrow \vec{\phi}$ biyectiva

Demostración:

Se va a realizar la demostración de la primera afirmación, dejando al lector las restantes.

■ \Rightarrow

$$\lambda(\vec{pq}) = \lambda(\vec{pr}) \Rightarrow \phi(q) = \phi(p) + \lambda(\vec{pq}) = \phi(p) + \lambda(\vec{pr}) = \phi(r) \xrightarrow{\phi \text{ inyectiva}} q = r \Rightarrow \vec{pq} = \vec{pr}$$

■ \Leftarrow

$$\phi(q) = \phi(r) \Rightarrow \lambda(\vec{qr}) = \overrightarrow{\phi(q)\phi(r)} = 0 \xrightarrow{\lambda \text{ inyectiva}} \vec{qr} = 0 \Rightarrow q = r$$

Definición (Grupo Afín)

Dado un espacio afín A , se denomina **grupo afín** de dicho espacio a :

$$GA(A) = \{ \phi : A \rightarrow A : \phi \text{ es afín y biyectiva} \}$$

que forma un grupo junto con las operaciones de composición y suma de aplicaciones.

Observación:

De este hecho anterior se deduce, aunque no se va a demostrar que:

- Si ϕ_1 y ϕ_2 son afinidades, entonces $\phi_1 \circ \phi_2$ también lo es.
- Si ϕ es afinidad con $\vec{\phi}$ como aplicación lineal asociada, entonces ϕ^{-1} también lo es con aplicación lineal asociada $\vec{\phi}^{-1}$.

De la misma manera que las aplicaciones lineales fueron caracterizadas a través de una matriz que representaba su interacción con el espacio vectorial, nos encaminamos a encontrar un objeto matemático similar que representa dicha interacción en las variedades afines.

Sabemos que la imagen de un punto cualquiera $p \in A$ está descrita en términos de A' pues $\phi(p) \in A'$. Además, conocemos la imagen $\phi(p_0) \in A'$ del origen del primer sistema de referencia con respecto al segundo, por tanto, podemos dar coordenadas a los puntos:

$$p = (x_1, \dots, x_n)_R \Rightarrow p_0 \vec{p} = (x_1, \dots, x_n)_B$$

$$\phi(p_0) = (a_1, \dots, a_m)_{R'} \Rightarrow \overrightarrow{p'_0 \phi(p_0)} = (a_1, \dots, a_m)_{B'}$$

$$\phi(p) = (x'_1, \dots, x'_m)_{R'} \Rightarrow \overrightarrow{p'_0 \phi(p)} = (x'_1, \dots, x'_m)_{B'}$$

Ahora tenemos que intentar relacionar las coordenadas de $\phi(p)$ en el sistema R' con sus coordenadas en R :

$$\begin{aligned} \overrightarrow{p'_0 \phi(p)} &= \overrightarrow{p'_0 \phi(p_0)} + \overrightarrow{\phi(p_0) \phi(p)} = \overrightarrow{p'_0 \phi(p_0)} + \vec{\phi}(\vec{p_0 p}) \Rightarrow \vec{\phi}(\vec{p_0 p}) = \overrightarrow{p'_0 \phi(p)} - \overrightarrow{p'_0 \phi(p_0)} \Rightarrow \\ &\Rightarrow \vec{\phi}((x_1, \dots, x_n)_B) = (x'_1, \dots, x'_m)_{B'} - (a_1, \dots, a_m)_{B'} = (x'_1 - a_1, \dots, x'_m - a_m)_{B'} \end{aligned}$$

Luego, expresado en forma matricial:

$$\begin{pmatrix} x'_1 - a_1 \\ \vdots \\ x'_m - a_m \end{pmatrix} = M_{B,B'}(\vec{\phi}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} + M_{B,B'}(\vec{\phi}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Pero, de nuevo, para tratar de obtener una expresión más cómoda y similar a un cambio de base podemos expresarla de la siguiente manera.

Proposición (Matriz de una aplicación afín)

Sean (A, V, φ) y (A', V', R') dos espacios afines de dimensiones n y m respectivamente y $R = \{p_0, B\}$ y $R' = \{p'_0, B'\}$ sus respectivos sistemas de referencia cartesianos. Dada una aplicación $\phi: A \rightarrow A'$ afín, ésta queda totalmente definida a través de la matriz:

$$\begin{pmatrix} 1 \\ x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ a_1 & & & \\ \vdots & & M_{B,B'}(\vec{\phi}) & \\ a_m & & & \end{pmatrix}}_{M_{R,R'}(\phi)} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

donde la primera columna son las coordenadas que tiene la imagen del origen de R expresadas en R' , es decir, $\phi(p_0) = (a_1, \dots, a_m)_{R'}$.

Variedades ϕ -invariantes

Del mismo modo que teníamos subespacios ϕ -invariantes en espacios vectoriales, podemos preguntarnos por las variedades X que se mantienen fijas bajo ϕ , esto es, para las cuales $\phi|_X$ vuelve a ser una variedad. En este apartado, vamos a desarrollar que características tienen éstas y su íntima relación con los subespacios invariantes mencionados anteriormente.

Definición (Conjunto de Puntos Fijos)

Dada una aplicación afín $\phi: A \rightarrow A$ definimos el **conjunto de puntos fijos** como:

$$A_\phi = \{p \in A : \phi(p) = p\}$$

es decir, son aquellos para los que ϕ es la identidad.

Proposición

Dada una aplicación afín ϕ , entonces:

$$A_\phi \neq \emptyset \Rightarrow A_\phi \text{ es variedad y } \overrightarrow{A_\phi} = V_{\vec{\phi},1}$$

Demostración:

Vamos a tratar de demostrar que $A_\phi \neq \emptyset \stackrel{?}{\Rightarrow} A_\phi = p_0 + V_{\vec{\phi},1}$, donde $p_0 \in A_\phi$.

En primer lugar, veamos que $p_0 + V_{\vec{\phi},1} \subseteq A_\phi$, para ello escogemos un $p = p_0 + v$ donde $v \in V_{\vec{\phi},1}$: $v = p_0 \vec{p}$ de forma que $p = p_0 + v = p_0 + p_0 \vec{p}$. Entonces ocurre que por ser ϕ afín:

$$\phi(p) = \phi(p_0) + \vec{\phi}(v) = p_0 + v = p$$

Por otro lado, veamos el otro contenido, es decir, $p_0 + V_{\vec{\phi},1} \supseteq A_\phi$; para ello queremos demostrar $p \in A_\phi \stackrel{?}{\Rightarrow} p = p_0 + v : v \in V_{\vec{\phi},1}$:

$$\vec{\phi}(p_0 \vec{p}) = \overrightarrow{\phi(p_0)\phi(p)} = p_0 \vec{p} \in V_{\vec{\phi},1} \Rightarrow p = p_0 + p_0 \vec{p} \in p_0 + V_{\vec{\phi},1}$$

Observación:

Por tanto, empezamos a ver la fuerte correlación entre los espacios invariantes y las variedades invariantes por ϕ .

Proposición

Sea $\phi : A \rightarrow A$ una aplicación afín. Si el 1 no es un autovalor de $\vec{\phi}$, entonces sólo hay un punto fijo:

$$A_\phi = \{p_0\} \Leftrightarrow 1 \notin \sigma(\vec{\phi})$$

Demostración:

$$\begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline a_1 & a_{11} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_n & a_{n1} & \dots & a_{nn} \end{array} \right) \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \Leftrightarrow (A - I_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -a_1 \\ \vdots \\ -a_n \end{pmatrix}$$

Para que la solución sea única, entonces $rg(A - I_n) = n$, luego:

$$A_\phi = \{p_0\} \Leftrightarrow rg(A - I_n) = n \Leftrightarrow p_A(1) = |A - I_n| \neq 0 \Leftrightarrow 1 \notin \sigma(\vec{\phi})$$

Observación:

Es importante dejar claro que cuando el 1 pertenece al espectro de la aplicación asociada, no podemos suponer que el conjunto de puntos fijos es no vacío porque, por ejemplo:

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 1 & 1 & 0 \\ 0 & 0 & r \end{array} \right) \Rightarrow p \in A_\phi \Rightarrow \begin{cases} x_1 = 1 + x_1 \\ x_2 = rx_2 \end{cases} \Rightarrow \#$$

Proposición

Sea $\phi : A \rightarrow A'$ una afinidad y $X = a + \vec{X} \subset A$ una variedad, entonces:

$$\phi(X) = \phi(a) + \vec{\phi}(\vec{X})$$

Del mismo modo, si $Y = b + \vec{Y} \subset A'$ es una variedad y $a \in \phi^{-1}(Y)$, entonces:

$$\phi^{-1}(Y) = a + \vec{\phi}^{-1}(\vec{Y})$$

Demostración:

La primera parte es consecuencia inmediata de la definición de afinidad. En cuanto a la segunda:

$$\phi(a + \vec{\phi}^{-1}(\vec{Y})) = \phi(a) + \vec{Y} = b + \vec{Y} \Rightarrow \phi^{-1}(b + \vec{Y}) \supset a + \vec{\phi}^{-1}(\vec{Y})$$

Por otro lado:

$$c \in \phi^{-1}(b + \vec{Y}) \Rightarrow f(c) \in b + \vec{Y} = f(a) + \vec{Y} \Rightarrow \overrightarrow{f(a)f(c)} = \vec{\phi}(\vec{ac}) \in \vec{Y} \Rightarrow \vec{ca} \in \vec{\phi}^{-1}(\vec{Y}) \Rightarrow c \in a + \vec{\phi}^{-1}(\vec{Y})$$

Corolario

Sea $\phi : A \rightarrow A'$ un afinidad, entonces:

- $Im(\phi)$ es una variedad de A' .
- Sean $X, Y \subset A$ tales que $X \parallel Y$, entonces $\phi(X) \parallel \phi(Y)$.
- Si $p, q \in A$ están alineados, entonces $\phi(p), \phi(q) \in A'$ están alineados.

Definición (Variedad Invariante)

Sea $\phi : A \rightarrow A$ una variedad de un espacio afín en si mismo, decimos que una variedad $X = q + \vec{X}$ es ϕ -**invariante** si y sólo si:

$$\phi(q + \vec{X}) \subset q + \vec{X}$$

es decir, aquellas variedades cerradas bajo ϕ .

Teorema (Caracterización de la Invarianza)

Sea ϕ una afinidad y X una variedad, ésta es ϕ -invariante si y sólo si:

- $\vec{\phi}(\vec{X}) \subset \vec{X}$
- $\forall p \in A : \phi(p) \in X \Leftrightarrow \overrightarrow{p\phi(p)} \in \vec{X}$

Observación:

Esto quiere decir que las direcciones de las variedades ϕ -invariantes son necesariamente subespacios $\vec{\phi}$ -invariantes. Por tanto, acabamos de encontrar un método muy preciso para poder calcular las variedades ϕ -invariantes de la dimensión deseada. Para ello, basta con calcular los subespacios $\vec{\phi}$ -invariantes de dimensión k requerida. Dada una base de dicho subespacio $B_k = \{v_1, \dots, v_k\}$ solo hace falta verificar que:

$$rg(v_1, \dots, v_k) = rg(\overrightarrow{p\phi(p)}, v_1, \dots, v_k)$$

de este modo obtendremos las ecuaciones de todos los puntos que pueden componer una variedad ϕ -invariante con el subespacio seleccionado como dirección.

Aplicaciones notorias

En este apartado vamos a caracterizar una serie de aplicaciones recurrentes y muy importantes en distintas áreas y aplicaciones de las matemáticas. Se trata, por tanto, de hacer un estudio exhaustivo de las mismas para tener soltura con sus propiedades fundamentales.

Aplicación constante

Se trata de la aplicación que transforma todos los puntos en el mismo punto constante y fijo p_0 , es decir, concentra todo el espacio afín en un único punto. Trivialmente queda caracterizada por ser $\vec{\phi}$ la aplicación nula, pues:

$$\vec{\phi}(\vec{pq}) = \overrightarrow{\phi(p)\phi(q)} = p'_0 \vec{p}'_0 = 0$$

Definición (Aplicación constante)

Sea $\phi : A \rightarrow A'$ una afinidad, se denomina **aplicación constante** p_0 si:

- $\phi(p) = p_0$
- $\vec{\phi}(v) = 0$
- **Matriz asociada:**

Para dar la matriz más sencilla posible, tomamos $R = \{p_0, B\}$ cualquiera y $R' = \{p'_0, B'\}$ donde $p'_0 = \phi(p_0)$. De este modo, la matriz asociada queda como:

$$M_{R,R'}(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{array} \right)$$

Traslación de vector u

Se trata de una aplicación de un espacio afín en sí mismo que transforma todos los puntos en sus trasladados a través del vector u , es decir, que traslada cada punto en la dirección de v y una distancia dada por el módulo de u . Trivialmente, la aplicación ϕ es la identidad, pues se conservan las distancias entre puntos:

$$\phi(\vec{pq}) = \overrightarrow{\phi(p)\phi(q)} = \overrightarrow{\phi(p)p} + \vec{pq} + \overrightarrow{q\phi(q)} = -u + \vec{pq} + u = \vec{pq}$$

Definición (Traslación de vector u)

Sea $\phi : A \rightarrow A$ una afinidad, se denomina **traslación de vector u** si:

- $\phi(p) = \varphi_p^{-1}(u) = p + u$
- $\vec{\phi}(v) = v$
- **Matriz Asociada:**

Como el espacio de llegada y salida son el mismo, entonces tomamos $R = \{p_0, B\}$ cualquiera en ambos espacios y la matriz asociada queda como:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ a_1 & & & \\ \vdots & & & \\ a_n & & & I_n \end{array} \right)$$

donde (a_1, \dots, a_n) es el vector de la traslación en la base B .

Por otro lado, podemos simplificar aún más esta expresión eligiendo una base $B' = \{v'_1, \dots, v'_n\}$ de forma que v'_1 sea el vector de la traslación, pues la matriz quedaría como:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 1 & & & \\ 0 & & & \\ \vdots & & & \\ 0 & & & I_n \end{array} \right)$$

Observación:

Cabe destacar dos cosas, la primera es que si el vector de la traslación es el nulo, esto querría decir que todos los puntos son fijos y, por tanto, la matriz sería la identidad.

Homotecia

Todas aquellas aplicaciones afines donde $\vec{\phi} = r \cdot id_v$ donde⁵ $r \neq 0, 1$ se denominan homotecias de razón r . Éstas son aplicaciones del espacio afín en sí mismo que consisten en expansiones o contracciones radiales entorno a un punto que queda fijo c . Los demás vectores modifican su distancia con respecto a él un factor de r .

La demostración de que hay un punto fijo es sencilla, primero hay que ver que es único:

$$p, q \in A_\phi \Rightarrow \vec{pq} = \overrightarrow{\phi(p)\phi(q)} = \vec{\phi(pq)} = r \cdot \vec{pq} \Leftrightarrow r \cdot \vec{pq} = r^2 \vec{pq} \Rightarrow r(1-r)\vec{pq} = 0 \Rightarrow \vec{pq} = 0 \Rightarrow p = q$$

Para ver quién es este punto fijo:

$$\begin{aligned} c \in A_\phi \Rightarrow r \cdot \vec{pc} &= \vec{\phi(p\vec{c})} = \overrightarrow{\phi(p)\phi(c)} = \overrightarrow{\phi(p)c} = \overrightarrow{\phi(p)p} + \vec{pc} \Rightarrow \\ \Rightarrow \overrightarrow{p\phi(p)} &= (1-r)\vec{pc} \Rightarrow \vec{pc} = \frac{1}{1-r} \overrightarrow{p\phi(p)} \Rightarrow c = p + \frac{1}{1-r} \cdot \overrightarrow{p\phi(p)} \end{aligned}$$

De este modo, podemos definir el centro de homotecia como:

$$\boxed{c = p + \frac{1}{1-r} \cdot \overrightarrow{p\phi(p)}}$$

Definición (Homotecia de razón r)

Sea $\phi : A \rightarrow A$ una afinidad, se denomina **homotecia de razón r** si:

- $\phi(p) = c + r \cdot \vec{cp}$ donde $c \in A_\phi$.
- $\vec{\phi(v)} = r \cdot v$
- **Matriz asociada:**

Si tomamos $R = \{c, B\}$, siendo c el centro⁶ y B una base cualquiera, entonces su matriz asociada sería:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & rI_n & \\ 0 & & & \end{array} \right)$$

⁵Tal y como se ve en las demostraciones, consideramos ϕ_r homotecia si excluimos $r = 1$ y $r = 0$ para que no sean ni la identidad ni la aplicación nula.

⁶Para calcular el centro, basta con escribir la matriz tomando como centro un punto cualquiera y resolver $v = A \cdot v$

Proyección

La aplicación afín que verifica que $\phi^2 = \phi$ se denomina proyección. Además, esta condición es equivalente a que $\vec{\phi}^2 = \vec{\phi}$ y $A_\phi \neq \emptyset$, de hecho, $A_\phi = \text{Im}(\phi)$. Para demostrarlo observemos que:

■ "⇒":

Si tenemos en cuenta que:

$$\overrightarrow{\psi \circ \phi(p\vec{q})} = \overrightarrow{(\psi \circ \phi)(p)(\psi \circ \phi)(q)} = \overrightarrow{\psi(\phi(p))\psi(\phi(q))} = \overrightarrow{\psi(\phi(p)\phi(q))} = \overrightarrow{\psi(\vec{\phi}(p\vec{q}))}$$

Entonces se tiene:

$$\phi^2 = \phi \xrightarrow{\overrightarrow{\psi \circ \phi} = \overrightarrow{\psi} \circ \vec{\phi}} \vec{\phi}^2 = \vec{\phi} \circ \vec{\phi} = \overrightarrow{\phi \circ \phi} = \vec{\phi}^2 = \vec{\phi}$$

Veamos que $A_\phi \neq \emptyset$:

$$\forall p \in A \Rightarrow \phi(\phi(p)) = \phi(p) \Rightarrow \phi(p) \in A_\phi \Rightarrow \begin{cases} A_\phi \neq \emptyset \\ \text{Im}(\phi) \subseteq A_\phi \end{cases}$$

Pero además, $p \in A_\phi \Rightarrow p = \phi(p) \in \text{Im}(\phi) \Rightarrow \text{Im}(\phi) \supseteq A_\phi$, por lo que $\text{Im}(\phi) = A_\phi$

■ "⇐"

Sea $p_0 \in A_\phi$ un punto fijo cualquiera y $p \in A$ un punto cualquiera, entonces podemos escribir:

$$p = p_0 + v : v = p_0 \vec{p} \Rightarrow \begin{cases} \phi(p) = \phi(p_0) + \vec{\phi}(v) = p_0 + \vec{\phi}(v) \\ \phi^2(p) = \phi^2(p_0) + \vec{\phi}^2(v) = p_0 + \vec{\phi}(v) \end{cases} \Rightarrow \phi(p) = \phi^2(p)$$

Matriz asociada

Sea $R = \{p_0, B\}$, siendo $p_0 \in A_\phi$ un punto fijo y $B = \underbrace{\{v_1, \dots, v_m\}}_{V_{\vec{\phi},1}}, \underbrace{\{v_{m+1}, \dots, v_n\}}_{V_{\vec{\phi},0}}$.

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ \vdots & & & 1 \\ 0 & & & \ddots & \\ & & & & 0 \end{array} \right)$$

Escogemos el sistema de esta manera porque de este modo la matriz es lo más sencilla posible (el punto fijo hace 0 en la primera columna y la base de vectores propios hace sencilla la matriz de la aplicación lineal).

Origen del nombre proyección

$\vec{\phi} : V \rightarrow V$ es una proyección vectorial con base $V_{\vec{\phi},1}$ y dirección $V_{\vec{\phi},0}$

$$\phi \left\{ \begin{array}{l} \text{Base} = A_\phi = \begin{cases} \text{Im}(\phi) \\ p_0 + V_{\vec{\phi},1} \end{cases} \\ \text{Dirección} = V_{\vec{\phi},0} \end{array} \right.$$

Veamos que papel juegan esta base y esta dirección.

$$\{\phi(p)\} = A_\phi \cap (p + V_{\vec{\phi},0})$$

■ " \subseteq ":

Tenemos dos posibles situaciones, si p es punto fijo:

$$\phi(p) \in \text{Im}(\phi) = A_\phi$$

Si p no es punto fijo:

$$\vec{\phi}(\overrightarrow{p\phi(p)}) = \overrightarrow{\phi(p)\phi^2(p)} = \overrightarrow{\phi(p)\phi(p)} = 0 \Rightarrow \overrightarrow{p\phi(p)} \in V_{\vec{\phi},0} \Rightarrow \phi(p) = p + \overrightarrow{p\phi(p)} \in p + V_{\vec{\phi},0}$$

■ " \supseteq ":

Si $A_\phi \cap (p + V_{\vec{\phi},0}) \neq \emptyset \Rightarrow A_\phi \cap (p + V_{\vec{\phi},0})$ es variedad, por lo que por lo demostrado antes solo falta ver que la dimensión es 0:

$$\dim(A_\phi \cap (p + V_{\vec{\phi},0})) = \dim \overrightarrow{A_\phi \cap (p + V_{\vec{\phi},0})} = \dim(\overrightarrow{A_\phi} \cap \overrightarrow{p + V_{\vec{\phi},0}}) = \dim(V_{\vec{\phi},1} \cap V_{\vec{\phi},0}) = \dim\{0\} = 0$$

Simetrías

Sea $\phi : A \rightarrow A$ afín y $\chi(K) \neq 2$. La aplicación afín conocida como **simetría** cumple:

$$\phi^2 = id_A \Leftrightarrow \vec{\phi}^2 = id_V \text{ y } A_\phi \neq \emptyset$$

Demostración

■ " \Rightarrow "

Del mismo modo que lo hemos hecho antes vemos que:

$$\phi^2 = id_V \xrightarrow{\vec{\psi \circ \phi} = \vec{\psi} \circ \vec{\phi}} \vec{\phi}^2 = \vec{\phi} \circ \vec{\phi} = \overrightarrow{\phi \circ \phi} = id_A = id_V$$

Veamos que $A_\phi \neq \emptyset$:

Sea $p \in A$ un punto cualquiera y elegimos el punto $m = p + \frac{1}{2}\overrightarrow{p\phi(p)}$, de este modo ocurre que:

$$\phi(m) = \phi(p) + \frac{1}{2}\vec{\phi}(\overrightarrow{p\phi(p)}) = p + \overrightarrow{p\phi(p)} + \frac{1}{2}\overrightarrow{\phi(p)\phi^2(p)} = p + \overrightarrow{p\phi(p)} + \frac{1}{2}\overrightarrow{\phi(p)p} = p + \overrightarrow{p\phi(p)} - \frac{1}{2}\overrightarrow{p\phi(p)} = m$$

■ " \Leftarrow "

Sea $p_0 \in A_\phi$ un punto fijo cualquiera y $p = p_0 + v : v = p_0\vec{p}$ un punto cualquiera, entonces:

$$\phi(p) = \phi(p_0) + \vec{\phi}(v) \Rightarrow \phi^2(p) = \phi^2(p_0) + \vec{\phi}^2(v) = p_0 + id_V(v) = p_0 + v = p$$

Matriz Canónica de una Simetría

Sea $R = \{p_0, B\}$, siendo $p_0 \in A_\phi, B = \{\underbrace{v_1, \dots, v_m}_{\in V_{\vec{\phi}, 1}}, \underbrace{v_{m+1}, \dots, v_n}_{\in V_{\vec{\phi}, -1}}\}$

$$M_R(\phi) = \left(\begin{array}{c|cccccc} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & & & & & \\ \vdots & & \ddots & & & & \\ \vdots & & & 1 & & & \\ \vdots & & & & -1 & & \\ \vdots & & & & & \ddots & \\ \vdots & & & & & & -1 \\ 0 & & & & & & \end{array} \right)$$

Homología

Sea $\phi : A \rightarrow A$ una aplicación afín, se dice que es una **homología general** si cumple:

$$q_{\vec{\phi}} = (x-1)(x-r) : r \neq 1, 0, -1 \text{ y } A_{\phi} \neq \emptyset$$

Al autovalor r se le conoce como la **razón de la homología**, a la variedad⁷ $W = p_0 + L(v_1, \dots, v_m) = p_0 + V_{\vec{\phi},1}$ como la **base** y a la variedad $U = L(v_{m+1}, \dots, v_n) = V_{\vec{\phi},r}$ como la **dirección**.

■ Matriz asociada:

Consideramos $R = \{p_0, B\}$, siendo $p_0 \in A_{\phi}$, y $B = \{\overbrace{v_1, \dots, v_m}^{\in V_{\vec{\phi},1}}, \overbrace{v_{m+1}, \dots, v_n}^{\in V_{\vec{\phi},r}}\}$ su matriz asociada queda como:

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & rI_{n-m} \end{array} \right)$$

Definición (Homología general con desplazamiento)

Se define una **homología general con desplazamiento** como la composición de una homología general con una traslación cuyo vector de traslación es paralelo a la base de la homología.

Sea $B = \{v_1, v_2\}$ base de vectores propios, entonces la matriz asociada resultante queda como:

$$M_R(\phi) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & 0 & r \end{array} \right) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_2 & 0 & r \end{array} \right) = M_R(\phi_2 \circ \phi_1) \Rightarrow \phi = \phi_2 \circ \phi_1$$

Se observa que la matriz de la aplicación ϕ_2 es una traslación y que el vector de traslación es paralelo al vector v_2 que es el que compone la base de la homología general. La aplicación ϕ_1 es una homología no elegida en su forma canónica definida, puesto que no hemos puesto de centro a un punto fijo.

Definición (Homología Especial)

Se define una **homología especial** como una homología general, pero con el único autovalor 1 que posee multiplicidad doble, pues si no sería traslación.

Sea $R = \{p_0, B\}$, siendo $p_0 \in A_{\phi}$ un punto fijo y eligiendo una base de forma que $M_B(\vec{\phi}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ la matriz canónica de Jordan asociada a su aplicación lineal, entonces su matriz queda como:

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right)$$

Denotamos por base a $A_{\phi} = p_0 + L(v_1)$ y por dirección a $L(v_2)$.

Definición (Homología Especial con desplazamiento)

De modo análogo, definimos la homología especial con desplazamiento como una composición de una homología especial con una traslación cuyo vector es paralelo a la dirección de la homología.

Sea $R = \{p_0, B\}$, siendo $M_B(\vec{\phi}) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ a_1 & 1 & 1 \\ a_2 & 0 & 1 \end{array} \right) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_2 & 0 & 1 \end{array} \right) \left(\begin{array}{c|cc} 1 & 0 & 0 \\ a_1 & 1 & 1 \\ 0 & 0 & 1 \end{array} \right) = M_R(\phi_2 \circ \phi_1)$$

⁷Cabe destacar que sabemos que los vectores de la base y de la dirección forman una base por ser $\vec{\phi}$ diagonalizable.

Clasificación de las Afinidades del Plano

Del mismo modo que hemos estado clasificando y caracterizando los endomorfismos, las aplicaciones ortogonales, ... nuestro objetivo ahora es clasificar todas las posibles aplicaciones afines en el plano. El propósito es poder tener un conocimiento muy preciso de lo que se espera encontrar como aplicación afín en el plano para, llegado el momento, tener claras las posibilidades y propiedades en el manejo de este objeto en áreas como la geometría.

Clasificación en \mathbb{C}

Sea $\dim \mathcal{A} = 2$, $K = \mathbb{C}$ y $\phi \in GA(\mathcal{A})$, entonces:

■ $1 \in \sigma(\vec{\phi})$

$$\begin{cases} q_{\vec{\phi}} = (x-1) \Rightarrow \phi \text{ traslación} \Rightarrow \begin{cases} \phi = id & A_{\phi} = \mathcal{A} \\ \phi \neq id & A_{\phi} = \emptyset \end{cases} \\ q_{\vec{\phi}} = (x-1)(x-r) : r \neq 1, 0, -1 \Rightarrow \phi \text{ diagonalizable} \Rightarrow \begin{cases} \phi \text{ h. g.} & A_{\phi} \neq \emptyset \\ \phi \text{ h. g. con desp.} & A_{\phi} = \emptyset \end{cases} \\ q_{\vec{\phi}} = (x-1)(x+1) \Rightarrow \begin{cases} \phi \text{ simetría} & A_{\phi} \neq \emptyset \\ \phi \text{ simetría con desplazamiento} & A_{\phi} = \emptyset \end{cases} \\ q_{\vec{\phi}} = (x-1)^2 \Rightarrow \begin{cases} \phi \text{ h. e.} & A_{\phi} \neq \emptyset \\ \phi \text{ h. e. con desp.} & A_{\phi} = \emptyset \end{cases} \end{cases}$$

■ $1 \notin \sigma(\vec{\phi})$

En este caso, si conocemos el conjunto de puntos fijos, pues es $A_{\phi} = \{p_0\}$, luego tomaremos como referencia $R = \{p_0, B\}$ con una base B conveniente para cada caso.

$$\begin{cases} q_{\vec{\phi}} = (x-r) : r \neq 1, 0 \Rightarrow \vec{\phi} = r \cdot id \Rightarrow \phi \text{ homotecia} \\ q_{\vec{\phi}} = (x-r)(x-s) : r, s \neq 1, 0 \Rightarrow \text{composición de homologías generales} \\ q_{\vec{\phi}} = (x-r)^2 : r \neq 1, 0 \Rightarrow \text{homología especial compuesta con homotecia} \end{cases}$$

Composición de homologías generales

Sea la base de vectores propios $B = \{v_1, v_2\}$:

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & s \end{array} \right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix} = M_R(\phi_2 \circ \phi_1)$$

Homología especial compuesta con homotecia

Sea la base $B = \{v_1, v_2\}$ tal que $M_B(\vec{\phi}) = \begin{pmatrix} r & 1 \\ 0 & r \end{pmatrix}$ sea la forma de Jordan asociada:

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & r & 1 \\ 0 & 0 & r \end{array} \right) \Rightarrow M_{R'}(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & r & r \\ 0 & 0 & r \end{array} \right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & r \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Clasificación en \mathbb{R}

Sea $\dim \mathcal{A} = 2$, $K = \mathbb{R}$ y $\phi \in GA(\mathcal{A})$, se mantienen todos los casos anteriores, pero hemos de añadir el caso en el que el polinomio no descompone:

$$q_{\vec{\phi}} = x^2 + a_1x + a_0 \in \mathbb{R}[x] \text{ irreducible} \Rightarrow 1 \notin \sigma(\vec{\phi}) \Rightarrow A_{\phi} = \{p_0\}$$

Luego si expresamos el polinomio mínimo de esta manera

$$q_{\vec{\phi}} = (x - \alpha)(x - \bar{\alpha}) : \alpha = a + bi \in \mathbb{C} \setminus \mathbb{R}$$

la matriz asociada queda como:

$$M_B(\vec{\phi}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \Rightarrow M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{array} \right)$$

Observación:

A pesar de que la notación es parecida, esta expresión no es extrapolable a las rotaciones vistas en las aplicaciones ortogonales.

ISOMETRÍAS

Con los elementos anteriores, el espacio vectorial euclídeo y el espacio afín, podemos construir una nueva estructura que permite modelar de forma feaciente la realidad geométrica del mundo en el que vivimos: el espacio afín euclídeo. En esta sección vamos a analizar las característica que aporta esta amalgama de estructuras al conjunto y a estudiar las aplicaciones que conservan dicho espacio que denominaremos movimientos rígidos o isometrías.

Definición (Espacio Afín Euclídeo)

Sea (A, V, φ) un espacio afín, y $f : V \times V \rightarrow \mathbb{R}$, entonces definimos un **espacio afín euclídeo** como la estructura (A, V, φ, f) .

Definición (Distancia)

Sea (A, V, f, φ) un espacio afín euclídeo, definimos la **distancia entre dos puntos** como:

$$d(p, q) = \|\vec{pq}\| = \sqrt{q(\vec{pq})}$$

y dicha definición verifica las siguientes propiedades:

- $d(p, q) \geq 0 : \forall p, q \in A$
 - $d(p, q) = 0 \Rightarrow p = q$
 - $d(p, q) \leq d(p, r) + d(r, p)$
-

Teorema (de Pitágoras)

$$p_0\vec{p}_1 \perp p_0\vec{p}_2 \Rightarrow \|p_1\vec{p}_2\|^2 = \|p_0\vec{p}_1\|^2 + \|p_0\vec{p}_2\|^2$$

Demostración:

$$\|p_1\vec{p}_2\|^2 = f(p_1\vec{p}_2, p_1\vec{p}_2) = f(p_1\vec{p}_0 + p_0\vec{p}_2, p_1\vec{p}_0 + p_0\vec{p}_2) = f(p_1\vec{p}_0, p_1\vec{p}_0) + f(p_0\vec{p}_2, p_0\vec{p}_2) = \|p_1\vec{p}_0\|^2 + \|p_0\vec{p}_2\|^2$$

Definición (Distancia entre variedades)

Sea $p \in A$ un punto y $X \subseteq A$ una variedad, definimos:

$$d(p, X) = \inf\{d(p_1, q) : q \in X\} = d(p, X) = \min\{d(p_1, q) : q \in X\}$$

Sean $X, Y \subseteq A$ variedades afines, definimos la distancia entre ellas como:

$$d(X, Y) = \inf\{d(p, q) : p \in X, q \in Y\} = \min\{d(p, q) : p \in X, q \in Y\}$$

y en ambos casos el ínfimo se alcanza.

Proposición

Un punto pertenece a una variedad si y sólo si su distancia respecto a ella es 0.

$$d(p, X) = 0 \Leftrightarrow p \in X$$

Dadas dos variedades $X, Y \subseteq A$ ambas intersecan si y sólo si su distancia es nula.

$$X \cap Y \neq \emptyset \Leftrightarrow \exists p_0 \in X \cap Y : d(p_0, p_0) = 0$$

En este tema en concreto toman gran importancia lo que definiremos como referencias ortonormales, puesto que las bases ortonormales permiten dar la notación de matriz identidad, y en consecuencia la operación de producto escalar es la usual, al producto escalar y esto es preferible cuando trabajamos con vectores en estas estructuras.

Definición (Referencia Ortonormal)

Sea $R = \{p_0, B\}$ un sistema de referencia, diremos que es **ortonormal**, si lo es la base B que lo compone.

Clasificación de Movimientos Rígidos

De nuevo y como acostumbramos, vamos a clasificar todos las posibles isometrías que podemos encontrar en espacios habituales como pueden ser el plano y el espacio. Esto nos dará una ventaja significativa a la hora de trabajar en estos ambientes pues de tener una aplicación y verificar que es un movimiento rígido, podemos adquirir mucha información acerca de la misma y dar una expresión matricial más amable para trabajar con ella.

Definición (Isometrías)

Sea $\phi : A \rightarrow A$ una aplicación afín, decimos que es un **movimiento rígido o isometría de A** si su aplicación lineal asociada $\vec{\phi}$ es ortogonal, es decir, que para R ortonormal la matriz $M_B(\vec{\phi})$ es ortogonal.

Los movimientos son todos biyectivos por lo que existe la aplicación inversa que también es un movimiento. Como la composición de movimientos es de nuevo un movimiento, entonces el conjunto de estas aplicaciones junto con la composición forman un grupo.

Proposición

Una aplicación afín es un movimiento rígido si y sólo si conserva las distancias.

$$\phi \text{ movimiento} \Leftrightarrow d(p, q) = d(\phi(p), \phi(q))$$

Demostración:

$$d(\phi(p), \phi(q))^2 = \|\overrightarrow{\phi(p)\phi(q)}\|^2 = f(\overrightarrow{\phi(p)\phi(q)}, \overrightarrow{\phi(p)\phi(q)}) = f(\vec{\phi}(\vec{pq}), \vec{\phi}(\vec{pq})) \stackrel{\vec{\phi} \in O(V)}{=} f(\vec{pq}, \vec{pq}) = \|\vec{pq}\|^2 = d(p, q)^2$$

Ejemplo:

Sea $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de forma que su matriz en base canónica sea:

$$M_{R_c}(\phi) = \frac{1}{5} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 4 & 3 \\ -5 & -3 & 4 \end{pmatrix}$$

Veamos que ϕ es movimiento:

$$M_{B_c}(\vec{\phi}) = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix} \Rightarrow \det(\vec{\phi}) = 1 \Rightarrow \vec{\phi} \in O^+(2)$$

Como $1 \notin \sigma(\phi) \Rightarrow A_\phi = \{p_0\} : p_0 = \begin{cases} (\frac{4}{5} - 1)x + \frac{3}{5}y = 0 \\ -1 - \frac{3}{5}x + (\frac{4}{5} - 1)y = 0 \end{cases}$. Luego si escogemos un sistema apropiado $R = \{p_0, B = \{v_1, v_2\}\}$ su matriz quedaría como:

$$M_R(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{array} \right)$$

Ejemplo: En \mathbb{R}^3 consideramos la matriz:

$$M_{R_c}(\phi) = \frac{1}{3} = \left(\begin{array}{c|ccc} 3 & 0 & 0 & 0 \\ \hline 3 & 2 & -1 & 2 \\ 3 & -1 & 2 & 2 \\ 0 & 2 & 2 & -1 \end{array} \right) \Rightarrow M_{B_c}(\vec{\phi}) = \frac{1}{3} \begin{pmatrix} 2 & -1 & 2 \\ -1 & 2 & 2 \\ 2 & 2 & -1 \end{pmatrix} \text{ ortogonal} \Rightarrow \vec{\phi} \in O(3)$$

Como $\det(\vec{\phi}) < 0 \Rightarrow \det(\vec{\phi}) = -1$, luego existe una base respecto de la cual:

$$M_B(\vec{\phi}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}$$

Como la traza es un invariante, tenemos que:

$$\text{tr}(\vec{\phi}) = 1 = -1 + 2c \Rightarrow c = 1 \Rightarrow s = 0 \Rightarrow M_B(\vec{\phi}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Luego ϕ será una simetría especular si hay puntos fijos:

$$A_\phi : \begin{cases} 1 - x - y + 2z = 0 \\ 1 - x - y + 2z = 0 \\ 2x + 2y - 4y = 0 \end{cases} \Rightarrow A_\phi = \emptyset$$

Lo que quiere decir que $\vec{\phi}$ es simetría ortogonal, pero ϕ no es simetría (es simetría con desplazamiento).

Clasificación en el plano

Sea $\phi : A \rightarrow A$ un movimiento rígido y A un espacio afín euclídeo tal que $\dim A = 2$, entonces:

- $\vec{\phi} \in O^+(V)$
- $A_\phi \neq \emptyset$

Definición (Giro)

Elegimos como $R = \{p_0, B\}$ donde $p_0 \in A_\phi$ es fijo y B es ortonormal para que:

$$M_R(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} : 0 \leq \theta \leq \pi$$

Como casos especiales tenemos que $\begin{cases} \theta = 0 & \Rightarrow \phi = id_A \\ \theta = \pi & \Rightarrow \phi = \text{sim. central} \end{cases}$

- $A_\phi = \emptyset$

Definición (Giro con Desplazamiento)

Elegimos como $R = \{p, B\}$ donde B es ortonormal (ahora no podemos coger p conveniente) de forma que:

$$M_R(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ a_1 & \cos \theta & -\sin \theta \\ a_2 & \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & 0 & 1 \end{pmatrix}_{(a_1, a_2) \neq (0,0)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

Vemo, por tanto, que se trata de la composición de un giro con centro p_0 y ángulo θ (la segunda matriz), seguido de una traslación de vector $v = (a_1, a_2)$.

- $\vec{\phi} \in O^-(V)$

- $A_\phi \neq \emptyset$

Definición (Simetría Ortogonal Axial)

Elegimos como $R = \{p_0, B\}$ con $p_0 \in A_\phi$ fijo y B ortonormal, de forma que:

$$M_R(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

La base corresponde a la variedad $p_0 + L(v_1) = A_\phi$ y la dirección $L(v_2) = L(v_1)^\perp$.

- $A_\phi = \emptyset$

Definición (Simetría Ortogonal Axial con desplazamiento)

En este caso, no tenemos puntos que queden fijos, luego la matriz es:

$$M_R(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Se trata de la composición de una traslación de vector $v = (a_1, a_2)$ (la primera matriz) con una simetría ortogonal axial, es decir, se trata de una simetría con desplazamiento.

Clasificación en el espacio

Sea $\phi : A \rightarrow A$ un movimiento rígido y A es un espacio afín euclídeo tal que $\dim A = 3$, entonces:

- $\vec{\phi} \in O^+(V)$

- $A_\phi \neq \emptyset$

Definición (Rotación)

Elegimos un sistema de referencia $R = \{p_0, B\}$ donde $p_0 \in A_\phi$ y B sea una base ortonormal, entonces:

$$M_R(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{pmatrix} : 0 \leq \theta \leq \pi$$

Con lo cual, tal y como lo vimos el eje de rotación es $p_0 + L(v_1)$ y el ángulo θ . Como

casos especiales: $\begin{cases} \theta = 0 & \text{Identidad} \\ \theta = \pi & \text{Simetría axial} \end{cases}$

- $A_\phi = \emptyset$

Definición (Movimiento Helicoidal)

En este caso solo podemos elegir una base ortonormal para poder obtener un sistema de referencia ortonormal, pero ya el punto no es fijo:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ a_1 & 1 & 0 & 0 \\ a_2 & 0 & \cos \theta & -\sin \theta \\ a_3 & 0 & \sin \theta & \cos \theta \end{array} \right)_{(a_1, a_2, a_3) \neq (0,0,0)} = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ a_1 & & & \\ 0 & & I_3 & \\ 0 & & & \end{array} \right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a_2 & 0 & \cos \theta & -\sin \theta \\ a_3 & 0 & \sin \theta & \cos \theta \end{array} \right)$$

Se le llama movimiento helicoidal por ser la composición de una rotación y una traslación de vector $a_1 v_1$ (el movimiento que haría una hélice). Como caso especial, cuando los giros son de ángulo 0, entonces tenemos las traslaciones dentro de este grupo.

- $\vec{\phi} \in O^-(3)$

- $A_\phi \neq \emptyset$

Definición (Rotación con Simetría Especular)

Elegimos un sistema de referencia ortonormal $R = \{p_0, B\}$ donde $p_0 \in A_\phi$ y B ortonormal, entonces:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{array} \right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \theta & -\sin \theta \\ 0 & 0 & \sin \theta & \cos \theta \end{array} \right)$$

Se trata de una rotación seguida de una simetría especular, cuya base es el plano de la rotación.

- $A_\phi = \emptyset$

Definición (Simetría Especular con Traslación)

De este modo, sabemos que $\sigma(\phi) = \{-1, 1, 1\}$, luego en una referencia ortonormal tenemos:

$$M_R(\phi) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ a_1 & -1 & 0 & 0 \\ a_2 & 0 & 1 & 0 \\ a_3 & 0 & 0 & 1 \end{array} \right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ 0 & & & \\ a_2 & & I_3 & \\ a_3 & & & \end{array} \right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ a_1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

Se trata de una simetría especular compuesta con una traslación, cuyo vector de traslación está contenido en la dirección del plano que es base de la simetría.

CÓNICAS

En multitud de problemas físicos o analíticos usualmente trabajamos con este tipo de objetos matemáticos: las cónicas. El estudio profundo de su representación óptima y de los métodos para un manejo adecuado de dicho objeto nos permitirá en el futuro poder obviar la parte algebraica (y centrarnos en la sustancial) de los problemas en los que veamos involucrados este tipo de elementos.

Definición (Cónica)

Sea $A = \mathbb{R}^2$ plano afín real usual y $R = \{p_0, B\}$ un sistema de referencia, definimos una cónica como el conjunto de puntos que verifican:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{01}x + 2a_{02}y + a_{00} = 0 : a_{ij} \in \mathbb{R} \text{ y } (a_{11}, a_{12}, a_{22}) \neq (0, 0, 0)$$

La condición que se ha dado en la definición se puede expresar de forma matricial de la siguiente forma:

$$0 = \begin{pmatrix} 1 & x & y \end{pmatrix} A \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} \text{ donde } A = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{01} & a_{11} & a_{12} \\ a_{02} & a_{12} & a_{22} \end{pmatrix} = M_R(C)$$

Proposición

Sea C una cónica en el plano afín real \mathbb{R}^2 , ésta queda representada por la matriz de coeficientes:

$$M_R(C) = \begin{pmatrix} a_{00} & a_{01} & a_{02} \\ a_{01} & a_{11} & a_{12} \\ a_{02} & a_{12} & a_{22} \end{pmatrix}$$

donde a_{ij} son los coeficientes de la ecuación general de una cónica.

Como la matriz de paso $P = M(R', R)$ verifica que $\begin{pmatrix} 1 & x & y \end{pmatrix} = P \begin{pmatrix} 1 \\ x' \\ y' \end{pmatrix}$, entonces $M_R(C) = A$ guarda la siguiente relación con la matriz $M_{R'}(C) = A'$:

$$0 = \begin{pmatrix} 1 & x & y \end{pmatrix} A \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = \begin{pmatrix} 1 & x' & y' \end{pmatrix} \underbrace{P^t A P}_{A'} \begin{pmatrix} 1 \\ x' \\ y' \end{pmatrix} \Rightarrow A = P^t A P$$

Clasificación Afín

En este apartado vamos a clasificar las cónicas en 7 clases de equivalencia desde el punto de vista afín y a escoger los mejores representantes de cada una. De esta manera, trabajar con una cónica será más efectivo y sencillo.

Definición (Equivalencia Afín)

Dos cónicas C y C' son **afínmente equivalentes** si respecto de sistemas distintos ambas tienen la misma matriz, es decir:

$$C \sim C' \Leftrightarrow M_R(C) = M_{R'}(C')$$

esta relación define una relación de equivalencia cuyo conjunto cociente cuenta con exactamente 7 clases.

Teorema

Toda cónica real es **afínmente equivalente** a una y sólo una de las siguientes:

Elipse no degenerada:	$x^2 + y^2 - 1 = 0$
Hipérbola no degenerada:	$x^2 - y^2 - 1 = 0$
Parábola no degenerada:	$x^2 - y = 0$
Elipse degenerada:	$x^2 + y^2 = 0$
Hipérbola degenerada:	$x^2 - y^2 = 0$
Parábola simplemente degenerada:	$x^2 - 1 = 0$
Parábola doblemente degenerada:	$x^2 = 0$

Demostración:

Consideremos la matriz de una cónica

$$A = \left(\begin{array}{c|cc} a_{00} & a_{01} & a_{02} \\ a_{01} & & \\ \hline & A_0 & \\ a_{02} & & \end{array} \right)$$

y veamos que podemos escribir su ecuación de la siguiente manera:

$$0 = (1 \quad x \quad y) A \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} \Leftrightarrow 0 = (1 \quad x \quad y) A \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} = (x \quad y) A_0 \begin{pmatrix} x \\ y \end{pmatrix} + 2(a_{01} \quad a_{02}) \begin{pmatrix} x \\ y \end{pmatrix} + a_{00}$$

De este modo, si consideramos $R = \{p_0, B\}$ un sistema de referencia, podemos interpretar A_0 como la matriz de una forma cuadrática, puesto que es simétrica, es decir:

$$M_B(q_0) = A_0 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Como el cuerpo son los reales, aplicando la Ley de Inercia, se tiene que:

$$A_0 \equiv A'_0 = \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} : \varepsilon_i = 1, -1, 0$$

Como son congruentes, existe una matriz de paso de una a otra y por tanto podemos expresar la relación entre ambas a través de $A_0 = P^t A'_0 P$, siendo $P = M(B, B')$ (y teniendo la siguiente relación en consecuencia $\begin{pmatrix} x' \\ y' \end{pmatrix} = P \begin{pmatrix} x \\ y \end{pmatrix}$).

Ahora si sustituimos convenientemente en la expresión inicial tenemos que:

$$\begin{aligned} (x \quad y) P^t A'_0 P \begin{pmatrix} x \\ y \end{pmatrix} + 2(a_{01} \quad a_{02}) \begin{pmatrix} x \\ y \end{pmatrix} + a_{00} = \\ (x' \quad y') A'_0 \begin{pmatrix} x' \\ y' \end{pmatrix} + 2(a_{01} \quad a_{02}) \begin{pmatrix} x \\ y \end{pmatrix} + a_{00} = \\ \varepsilon_1 (x')^2 + \varepsilon_2 (y')^2 + 2(a_{01} \quad a_{02}) \begin{pmatrix} x \\ y \end{pmatrix} + a_{00} = \end{aligned}$$

$$\varepsilon_1(x')^2 + \varepsilon_2(y')^2 + 2(a_{01} \ a_{02}) P^{-1} \begin{pmatrix} x' \\ y' \end{pmatrix} + a_{00} =$$

$$\varepsilon_1(x')^2 + \varepsilon_2(y')^2 + 2a'_{01}x' + 2a'_{02}$$

Ambos $\varepsilon_1, \varepsilon_2$ no pueden ser 0 simultáneamente pues entonces $A'_0 = A' = 0$ y todos los coeficientes del término cuadrático se anularían. Siempre un ε_i es 1 o -1, pero además podemos suponer que el primero es siempre 1, puesto que si fuese -1 bastaría con multiplicar el polinomio por -1 (los puntos no cambian) y si el segundo fuera el que toma el valor no nulo entonces bastaría con reordenar la base para que dicho ε_2 pase a ser ε_1 .

A pesar del arreglo algebraico que se ha hecho en el razonamiento anterior para simplificar la explicación y la notación, en el fondo lo que estamos haciendo es cambiar de sistema de referencia $R = \{p_0, B\}$ a un sistema $R' = \{p_0, B'\}$, puesto que teniendo en cuenta que $P = M(B, B')$ se satisface que:

$$\begin{pmatrix} 1 \\ x' \\ y' \end{pmatrix} = \underbrace{\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & & \\ \hline 0 & & P \end{array} \right)}_{M(R, R')} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix}$$

Sabiendo que $\varepsilon_1 = 1$ siempre, distinguimos casos:

- Si $\varepsilon_2 = 1$:

$$(x')^2 + (y')^2 + 2a'_{01}x' + 2a'_{02}y' + a'_{00} = (x' + a'_{01})^2 + (y' + a'_{02})^2 + \underbrace{a'_{00} - a'^2_{01} - a'^2_{02}}_{a''_{00}}$$

$$\begin{cases} x''^2 + y''^2 - 1 & a''_{00} < 0 \\ x''^2 + y''^2 & a''_{00} = 0 \\ \text{No es cónica} & a''_{00} > 0 \end{cases}$$

Para poder verificar que el cambio que hemos hecho al completar cuadrados es correcto, observamos la siguiente relación:

$$\begin{cases} x'' = x' + a'_{01} \\ y'' = y' + a'_{02} \end{cases} \Rightarrow \begin{pmatrix} 1 \\ x'' \\ y'' \end{pmatrix} = \underbrace{\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline a'_{01} & 1 & 0 \\ \hline a'_{02} & 0 & 1 \end{array} \right)}_{M(R', R'')} \begin{pmatrix} 1 \\ x' \\ y' \end{pmatrix}$$

Esta relación es correcta porque la matriz cuadrada 2x2 correspondiente a la identidad representa el cambio de base entre B' y B'' que se mantienen iguales, es decir, $B' = B''$. Sin embargo, el centro a cambiado porque ahora lo marcan las coordenadas a'_{01} y a'_{02} de modo que:+

$$R'' = \{p''_0, B''\} : p''_0, \vec{p}'_0 = a'_{01} \cdot v_1 + a'_{02} \cdot v_2 \Rightarrow p''_0 = (-a'_{01}, -a'_{02})_{R'}$$

- Si $\varepsilon_2 = -1$

$$x'^2 - y'^2 + 2a'_{01}x' + 2a'_{02}y' + a'_{00} = (x' + a'_{01})^2 - (y' - a'_{02})^2 + \underbrace{a'_{00} - a'^2_{01} - a'^2_{02}}_{a''_{00}}$$

$$\begin{cases} x''^2 - y''^2 - 1 & a''_{00} < 0 \\ x''^2 - y''^2 & a''_{00} = 0 \\ x''^2 - y''^2 + 1 & a''_{00} > 0 \end{cases}$$

En este caso, el procedimiento de justificación de que los cambios algebraicos corresponden a cambios de sistemas de referencia es completamente análogo.

- Si $\varepsilon_2 = 0$

$$x'^2 + 2a'_{01}x' + 2a'_{02}y' + a'_{00} = (x' + a'_{01})^2 + 2a'_{02}y' + a'_{00} - a'^2_{01} = x''^2 + 2a'_{02}y' + a''_0 =$$

- $a'_{02} \neq 0$:

$$0 = x''^2 - (-2a'_{02}y' - a''_0) \Rightarrow \begin{cases} x'' = x' + a'_{01} \\ y'' = -2a'_{02}y' - a''_0 \end{cases} \Rightarrow x''^2 - y''$$

Comprobamos que se trata de un cambio de sistema de referencia.

$$\begin{pmatrix} 1 \\ x'' \\ y'' \end{pmatrix} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ a'_{01} & 1 & 0 \\ a'_{02} & 0 & -2a'_{02} \end{array} \right) \begin{pmatrix} 1 \\ x' \\ y' \end{pmatrix}$$

Como la matriz 2x2 es inversible, corresponde a un cambio de base por ser inversible y como las coordenadas del centro son válidas todo es correcto.

- $a'_{02} = 0$:

$$0 = x''^2 + a''_{00} \xrightarrow{1/a''_{00}} x''^2 - 1 : \begin{cases} \text{No es cónica} & a''_{00} > 0 \\ \text{Rectas Paralelas} & a''_{00} < 0 \\ \text{Recta única en el plano (doble sobre sí)} & a''_{00} = 0 \end{cases}$$

Caracterización de las Cónicas

Sea C una cónica y

$$A = M_R(C) = \left(\begin{array}{c|cc} a_{00} & a_{01} & a_{02} \\ a_{01} & & \\ \hline & A_0 & \\ a_{02} & & \end{array} \right) \quad A' = M_{R'}(C) = \left(\begin{array}{c|cc} a'_{00} & a'_{01} & a'_{02} \\ a'_{01} & & \\ \hline & A'_0 & \\ a'_{02} & & \end{array} \right)$$

sus matrices de coeficientes en dos sistemas de referencia distintos, la relación que existe entre ambas viene dada por:

$$A' = P^t A P \text{ donde } P^t = \left(\begin{array}{c|cc} 1 & c_1 & c_2 \\ 0 & & \\ \hline & P_0^t & \\ 0 & & \end{array} \right) \text{ y } P = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ c_1 & & \\ \hline & P_0 & \\ c_2 & & \end{array} \right)$$

Siendo $P = M(R', R)$, $P_0 = M(B', B)$ y c_1 y c_2 la relación entre los centros de ambas. Del mismo modo, por ser afínmente equivalentes, entonces $A \equiv A'$ y se mantienen invariantes:

- El rango: $rg(A) = rg(A')$
- La signatura no porque multiplicamos por -1, pero sí se mantiene: $(s, t) = \varepsilon(A) = \varepsilon(A') = (s', t') \Rightarrow \delta = |s - t| = |s' - t'|$
- El signo del determinante de A : $sign|A| = d$

Teorema

Sea A la matriz de coeficientes de una cónica en cierto sistema de referencia R y utilizando la siguiente notación:

$$\begin{cases} r = \text{rango} \\ \delta = |s - t| \\ d = \det A \end{cases}$$

podemos clasificar dicha cónica en una y sólo una de las siguientes clases de equivalencia afín:

- C elipse no degenerada:

$$\left(\begin{array}{c|cc} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 3, \delta = 1, d < 0 \\ r_0 = 2, \delta_0 = 2, d_0 > 0 \end{cases}$$

- C elipse degenerada:

$$\left(\begin{array}{c|cc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 2, \delta = 2, d = 0 \\ r_0 = 2, \delta_0 = 2, d_0 > 0 \end{cases}$$

- C hipérbola no degenerada:

$$\left(\begin{array}{c|cc} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 3, \delta = 1, d > 0 \\ r_0 = 2, \delta_0 = 0, d_0 < 0 \end{cases}$$

- C hipérbola degenerada:

$$\left(\begin{array}{c|cc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 2, \delta = 0, d = 0 \\ r_0 = 2, \delta_0 = 0, d_0 < 0 \end{cases}$$

- C parábola no degenerada:

$$\left(\begin{array}{c|cc} 0 & 0 & \frac{-1}{2} \\ 0 & 1 & 0 \\ \frac{-1}{2} & 0 & 0 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 3, \delta = 1, d < 0 \\ r_0 = 1, \delta_0 = 1, d_0 = 0 \end{cases}$$

- C parábola simplemente degenerada:

$$\left(\begin{array}{c|cc} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 2, \delta = 0, d = 0 \\ r_0 = 1, \delta_0 = 1, d_0 = 0 \end{cases}$$

- C parábola doblemente degenerada:

$$\left(\begin{array}{c|cc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right) \equiv A \Rightarrow \begin{cases} r = 1, \delta = 1, d = 0 \\ r_0 = 1, \delta_0 = 1, d_0 = 0 \end{cases}$$

Observación:

Además, si consideramos C cónica de forma que $A = M_R(C)$ y $A' = M_{R'}(C)$, entonces podemos considerar la matriz de paso P como la matriz de una aplicación afín en vez de una matriz de paso. Esta aplicación afín lleva los puntos de C a los de C' y los expresa respecto del propio sistema de referencia R , sin embargo, la matriz de la cónica cambia a la canónica que queríamos:

$$P = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ c_1 & M(B, B') & \\ c_2 & & \end{array} \right) = M_R(\phi) : \phi(x, y)_R = (x', y')_R$$

Ejemplo:

En la referencia Canónica consideramos:

$$0 = 5x^2 + 5y^2 + 6xy - 4x + 4y + 2$$

Tras realizar las transformaciones convenientes, se tiene:

$$0 = 5\left(x + \frac{3y}{5} - \frac{2}{5}\right)^2 - \frac{9y^2}{5} - \frac{4}{5} + \frac{12y}{5} + 5y^2 + 4y + 2 = 5x'^2 + \frac{16}{5}y^2 + \frac{32}{5}y + \frac{6}{5}$$

$$0 = 5x'^2 + \frac{16}{5}(y+1)^2 - \frac{16}{5} + \frac{6}{5} = 5x'^2 + \frac{16}{5}y'^2 - 2 =$$

Luego la relación que existe es:

$$\begin{cases} x' = x + \frac{3y}{5} - \frac{2}{5} \\ y' = y + 1 \end{cases}$$

Como $d_0 < 0, rg = 3$, se trata de una elipse no degenerada, así que su forma canónica será:

$$x''^2 + y''^2 - 1 \Rightarrow \begin{cases} x'' = \sqrt{\frac{5}{2}}x' = \sqrt{\frac{5}{2}}(x + \frac{3y}{5} - \frac{2}{5}) \\ y'' = \sqrt{\frac{8}{5}}y' = \sqrt{\frac{8}{5}}(y + 1) \end{cases}$$

Podemos definir por tanto $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ de forma que su matriz quede como:

$$M_{R_c}(\phi) = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline -\frac{2}{5}\sqrt{\frac{5}{2}} & \sqrt{\frac{5}{2}} & \frac{3}{5}\sqrt{\frac{5}{2}} \\ \sqrt{\frac{8}{5}} & 0 & \sqrt{\frac{8}{5}} \end{array} \right)$$

Clasificación Afín-Euclídea

En esta clasificación vamos añadir una restricción más a la equivalencia entre cónicas. Ya no vale cualquier cambio de referencia, pues antes podíamos deformar el espacio afín (y, en consecuencia, la cónica) modificando las distancias y los ángulos, pero ahora la restricción va a ser que dichas características se mantengan invariantes. Es decir, que las aplicaciones que transformen una cónica en otra van a ser isometrías o los cambios entre sistemas de referencia han de ser ortonormales.

Definición (Equivalencia Afín-Euclídea)

Sea \mathbb{R}^2 espacio afín euclídeo usual y la definición habitual de cónica, decimos que dos cónicas C y C' son afín-euclídeamente equivalentes si y sólo si poseen la misma matriz respecto de sistemas de referencia **ortonormales** distintos, es decir,

$$C \equiv C' \Leftrightarrow M_R(C) = M_{R'}(C')$$

esta definición constituye una relación de equivalencia con 7 clases de equivalencia distintas.

Teorema

Sean $a, b \in \mathbb{R} : 0 < a \leq b$, toda cónica C de \mathbb{R}^2 es afín-euclídeamente equivalente¹ a una y sólo una de las siguientes cónicas canónicas:

Elipse no degenerada:	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$
Hipérbola no degenerada:	$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$
Parábola no degenerada:	$x^2 - ay = 0$
Elipse degenerada:	$x^2 + y^2 = 0$
Hipérbola degenerada:	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$
Parábola simplemente degenerada:	$x^2 - a^2 = 0$
Parábola doblemente degenerada:	$x^2 = 0$

Demostración:

Sea $R = \{p_0, B\}$ donde B es una base ortonormal, entonces consideramos:

$$0 = a_{11}x^2 + \dots + a_{00} = (x \ y) A_0 \begin{pmatrix} x \\ y \end{pmatrix} + 2(a_{01} \ a_{02}) \begin{pmatrix} x \\ y \end{pmatrix} + a_{00}$$

¹La existencia de a y un b denota que todas las cónicas que son afín-euclídeamente equivalentes a esas cónicas lo son **para unos valores concretos de a y b , pero no para todos**, es decir, estamos considerando la clase de equivalencia salvo cambio en los parámetros a y b (que determinan las características como el radio de una elipse, etc.)

Ya dijimos en su momento que:

$$A_0 = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \in Mat_2(\mathbb{R}) \text{ simétrica}$$

Luego puede considerarse como la matriz de un endomorfismo autoadjunto de forma que:

$$A_0 = M_B(\vec{\psi}), \quad \vec{\psi}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

Por las características que acabamos de dar $\vec{\psi}$ es autoadjunto, luego por el Teorema Espectral:

$$\exists B' \text{ ortonormal} : A'_0 = M_{B'}(\vec{\psi}) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} : \sigma(A_0) = \{\lambda_1, \lambda_2\}$$

Podemos convertir entonces esta matriz A_0 a la A'_0 que hemos denotado a través del siguiente cambio de base:

$$A_0 = P^{-1} A'_0 P = P^t A'_0 P : P = M(B, B') \text{ ortogonal}$$

Luego retomando la ecuación inicial, tenemos que:

$$\begin{aligned} 0 &= a_{11}x^2 + \dots + a_{00} = \begin{pmatrix} x & y \end{pmatrix} A_0 \begin{pmatrix} x \\ y \end{pmatrix} + 2 \begin{pmatrix} a_{01} & a_{02} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + a_{00} = \\ &= \begin{pmatrix} x & y \end{pmatrix} P^t A'_0 P \begin{pmatrix} x \\ y \end{pmatrix} + 2 \begin{pmatrix} a_{01} & a_{02} \end{pmatrix} P^t \begin{pmatrix} x' \\ y' \end{pmatrix} + a'_{00} = \begin{pmatrix} x' & y' \end{pmatrix} A'_0 \begin{pmatrix} x' \\ y' \end{pmatrix} + 2 \begin{pmatrix} a_{01} & a_{02} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + a'_{00} = \\ &= \lambda_1 x'^2 + \lambda_2 y'^2 + 2a'_{01}x' + 2a'_{02}y' + a'_{00} \end{aligned}$$

Y ahora vamos a ver quienes son estos autovalores que hemos hallado. Desde luego debe ocurrir que $(\lambda_1, \lambda_2) \neq (0, 0)$ puesto que si no no estaríamos hablando de una cónica y, como en la demostración anterior de cónicas, podemos suponer que uno de los dos será positivo y asumir que este será λ_1 (multiplicando por -1 toda la ecuación en caso de que no se verifique).

- $\lambda_2 > 0$

Completamos cuadrados y queda lo siguiente:

$$\lambda_1(x' + \lambda_1^{-1}a'_{01})^2 + \lambda_2(y' + \lambda_2^{-1}a'_{02})^2 + a''_{00} = \lambda_1 x''^2 + \lambda_2 y''^2 + a''_{00}$$

Luego tenemos la siguiente distinción de casos:

$$\begin{cases} a''_{00} > 0 \Rightarrow C = \emptyset \\ a''_{00} = 0 \Rightarrow C \text{ es elipse degenerada} \\ a''_{00} < 0 \Rightarrow \frac{x''^2}{\frac{-a''_{00}}{\lambda_1}} + \frac{y''^2}{\frac{-a''_{00}}{\lambda_2}} - 1 = 0 \Rightarrow C \text{ es elipse no degenerada} \end{cases}$$

- Para el resto de posibles valores de λ_2 se hace completamente análogo.