6

Espacio afín

En este capítulo...

6.1 Definición de espació afín

■ **Definición** Dado un espacio vectorial V, decimos que un conjunto $\mathcal{A}$ es un espacio afín sobre V si existe una función

$$\begin{array}{rcl} \vec{\ } : \mathcal{A} \times \mathcal{A} & \to & V \\ (A, B) & \mapsto & \overrightarrow{AB} \end{array}$$

que satisface las dos siguientes propiedades:

A_1) para todo $A \in \mathcal{A}$ y $v \in V$ existe un único $B \in \mathcal{A}$ tal que $\overrightarrow{AB} = v$, al cual llamaremos *trasladado de A por v*, y

A_2) para cualesquiera $A, B, C \in \mathcal{A}$ se tiene que $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

A los elementos de $\mathcal{A}$ se les llama *puntos*. Si $v = \overrightarrow{AB}$ entonces decimos que v es el vector de *origen* A y *extremo* B, y también decimos que B es el *trasladado* de A por v y lo denotaremos por $A + v$. Por definición

$$\dim(\mathcal{A}) = \dim(V).$$

Si además V es un espacio vectorial Euclídeo, es decir, tiene definido un producto escalar $\langle\ ,\ \rangle$, entonces decimos que $\mathcal{A}$ es un espacio afín Euclídeo y definimos la distancia entre dos puntos A y B como

$$\mathrm{d}(A, B) = \|\overrightarrow{AB}\|.$$

Ejemplo 1 El ejemplo que debemos tener en la cabeza es $\mathbb{R}^n$. Efectivamente, el conjunto de puntos $\mathbb{R}^n$ es un espacio afín sobre el espacio vectorial $(\mathbb{R}^n, +, \cdot)$ mediante la aplicación

$$\begin{array}{rcl} \mathbb{R}^n \times \mathbb{R}^n & \to & (\mathbb{R}^n, +, \cdot) \\ \big(\underbrace{(a_1, \dots, a_n)}_{A}, \underbrace{(b_1, \dots, b_n)}_{B}\big) & \mapsto & \overrightarrow{AB} = (b_1 - a_1, \dots, b_n - a_n) \end{array}$$

De hecho $\mathbb{R}^n$ es un espacio afín Euclídeo ya que en el espacio vectorial $(\mathbb{R}^n, +, \cdot)$ podemos considerar el producto escalar usual. En este caso, la distancia entre los dos puntos A y B viene dada por

$$\mathrm{d}(A, B) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$$

Probemos a continuación algunas propiedades básicas de los espacios afines (las cuales resultan naturales si pensamos en $\mathbb{R}^n$)

■ **Proposición** *Para cualesquiera puntos A, B, C, D de un espacio afín $\mathcal{A}$ se tiene que*

1) $\overrightarrow{AB} = 0$ si y solo si $A = B$,

2) $\overrightarrow{BA} = -\overrightarrow{AB}$,

3) si $\overrightarrow{AB} = \overrightarrow{CD}$ entonces $\overrightarrow{AC} = \overrightarrow{BD}$.

Demostración. 1) Por A_2 se tiene que $\overrightarrow{AA} + \overrightarrow{AA} = \overrightarrow{AA}$, es decir, $\overrightarrow{AA} = 0$. Por A_1 deducimos que $A = B$.

2) Por A_2 se tiene que $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = 0$.

3) Si $\overrightarrow{AB} = \overrightarrow{CD}$, por A_2 tenemos que $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{CD} + \overrightarrow{BC}$ y en particular $\overrightarrow{AC} = \overrightarrow{BD}$. □

■ **Proposición** *Para cualesquiera puntos A, B, C de un espacio afín euclídeo $\mathcal{A}$ se tiene que*

1) $d(A,B) = 0 \quad \Longleftrightarrow \quad A = B,$

2) $d(A,B) = d(B,A),$

3) $d(A,C) \leq d(A,B) + d(B,C).$

Demostración. 1) $\mathrm{d}(A,B) = 0 \quad \Longleftrightarrow \|\overrightarrow{AB}\| = 0 \Longleftrightarrow \overrightarrow{AB} = 0 \Longleftrightarrow A = B.$

2) $\mathrm{d}(A,B) = \|\overrightarrow{AB}\| = \|\overrightarrow{BA}\| = \mathrm{d}(B,A).$

3) Por la desigualdad triangular que probamos en el Tema 4, sabemos que

$$\|\overrightarrow{AB} + \overrightarrow{BC}\| \leq \|\overrightarrow{AB}\| + \|\overrightarrow{BC}\|$$

y por tanto deducimos que $\mathrm{d}(A,C) \leq \mathrm{d}(A,B) + \mathrm{d}(B,C)$. □

A continuación, queremos definir lo que sería el equivalente en espacios afines de los subespacios vectoriales de un espacio vectorial. La idea es que van a ser "un punto más un subespacio vectorial".

■ **Definición** Sea $\mathcal{A}$ un espacio afín sobre el espacio vectorial V. Dado un punto $A \in \mathcal{A}$ y un subespacio vectorial W de V, llamamos *variedad afín que pasa por el punto A y tiene espacio de dirección W* al conjunto de puntos

$$A + W := \{A + w : w \in W\}.$$

O dicho de otra forma $A + W = \{B \in \mathcal{A} : \overrightarrow{AB} \in W\}$. Obsérvese que el punto A pertenece a $A + W$ ya que $A + 0 = A$.

Observación 6.1 1) Usando la notación de la definición anterior, es fácil comprobar que

$$W = \{\overrightarrow{PQ} : P, Q \in A + W\}.$$

Veamos primero que $W \subset \{\overrightarrow{PQ} : P, Q \in A + W\}$. Dado $w \in W$, por definición A y $A + w$ son puntos de $A + W$ y el vector de origen A y extremo $A + w$ es precisamente w. Por tanto $w \in \{\overrightarrow{PQ} : P, Q \in A + W\}$, como queríamos probar. Veamos por último que $\{\overrightarrow{PQ} : P, Q \in A + W\} \subset W$. Dados dos puntos cualquiera $P, Q \in A + W$ se tiene que $\overrightarrow{AP} \in W$ y $\overrightarrow{AQ} \in W$ y en particular

$$\overrightarrow{PQ} = \overrightarrow{PA} + \overrightarrow{AQ} = -\overrightarrow{AP} + \overrightarrow{AQ} \in W.$$

2) En particular, si nos dan un subconjunto cualquiera de puntos L de $\mathcal{A}$, para comprobar que es una variedad afín tan solo tenemos que hacer dos cosas: primero, comprobar que el conjunto de vectores $\{\overrightarrow{PQ} : P, Q \in L\}$ es un subespacio vectorial de W, y segundo, en el caso de que lo sea, comprobar que

$$L = A + \{\overrightarrow{PQ} : P, Q \in L\}$$

para algún punto $A \in L$. De hecho, y como vamos a ver en un momento, podemos tomar un punto A de L al azar.

3) Dada una variedad afín $A + W$ y dado un punto $B \in A + W$ es fácil comprobar que $A + W = B + W$, es decir, que podemos considerar cualquier punto de $A + W$ para definirla. En efecto, ya que $\overrightarrow{AB}$ es un vector del subespacio W, en particular $\overrightarrow{BA} = -\overrightarrow{AB}$ también lo es y se tiene que,

$$C \in A+W \Longleftrightarrow \overrightarrow{AC} \in W \Longleftrightarrow \overrightarrow{BA}+\overrightarrow{AC} \in W \Longleftrightarrow \overrightarrow{BC} \in W \Longleftrightarrow \overrightarrow{CB} = -\overrightarrow{BC} \in W \Longleftrightarrow C \in B+W.$$

Ejercicio 6.2 Sean A_1+W_1 y A_2+W_2 dos variedades afines de un espacio afín $\mathcal{A}$. Demuestra que, si A_1+W_1 no son disjuntos A_2+W_2, entonces

$$(A_1+W_1)\cap(A_2+W_2) = A + (W_1\cap W_2)$$

donde A es un punto cualquiera de la intersección $(A_1+W_1)\cap(A_2+W_2)$.

Ejemplo 6.3 Sea $\mathcal{A}$ un espacio afín Euclídeo y sea $A+W$ una variedad afín de $\mathcal{A}$. Dado un punto cualquiera $X\in\mathcal{A}$, podemos considerar la variedad afín ortogonal a $A+W$ que pasa por X, es decir,

$$X+W^{\perp}.$$

Con la intuición que tenemos de $\mathbb{R}^n$, la intersección de ambas debería ser un solo punto. Veamos que es así. Lo primero que debemos comprobar es que la intersección no es vacía. En efecto, sabemos que $V = W\oplus W^{\perp}$ y por tanto

$$\overrightarrow{AX} = w+v$$

donde $w\in W$ y $v\in W^{\perp}$. Veamos que el punto

$$p_{A+W}(X) := A+w,$$

al cual llamamos *proyección ortogonal* de X sobre $(A+W)$, pertenece a $(A+W)\cap(X+W^{\perp})$. Evidentemente $A+w\in A+W$, así que solo hay que comprobar que pertenece a $X+W^{\perp}$. En efecto,

$$A+w = X+\overrightarrow{XA}+w = X-\overrightarrow{AX}+w = X-w-v+w = X-v\in X+W^{\perp}.$$

Finalmente, por el Ejercicio 6.2 tenemos que la intersección no solo no es vacía sino que es un solo punto ya que

$$(A+W)\cap(X+W^{\perp}) = p_{A+W}(X)+(W\cap W^{\perp}) = p_{A+W}(X)+\{0\}.$$

Obsérvese por último que

$$p_{A+W}(X) := A+p_W(\overrightarrow{AX}),$$

donde $p_W : V\to V$ es la proyección ortogonal de V sobre W.

6.2 Sistema de referencia y coordenadas

Nuestra intención ahora es la de definir en un espacio afín algo parecido al concepto de base y de coordenadas de un espacio vectorial. Sea $\mathcal{A}$ un espacio afín sobre un espacio vectorial V de dimensión n. Fijemos un punto O cualquiera de $\mathcal{A}$ y una base cualquiera $\mathcal{B}$ de V. Observamos que dado un punto $A\in\mathcal{A}$, el vector $\overrightarrow{OA}$ tiene unas ciertas coordenadas respecto de $\mathcal{B}$,

$$\overrightarrow{OA} = (a_1,\dots,a_n)_{\mathcal{B}}.$$

Si nos dan otro punto $A'\in\mathcal{A}$ y las coordenadas de $\overrightarrow{OA}$ y $\overrightarrow{OA'}$ respecto de $\mathcal{B}$ coinciden entonces evidentemente $\overrightarrow{OA}=\overrightarrow{OA'}$ y en particular $A=A'$. Es decir, que esas coordenadas están describiendo unívocamente al punto A. Por tanto tiene sentido que digamos que

$$\mathcal{R} = \{O;\mathcal{B}\}$$

es un *sistema de referencia cartesiano* de $\mathcal{A}$ y que las coordenadas del punto A respecto de $\mathcal{R}$ son

$$A = (a_1, \ldots, a_n)_{\mathcal{R}}.$$

Además, si nos dan dos puntos $A, B \in \mathcal{A}$ cuyas coordenadas respecto de $\mathcal{R}$ son

$$A = (a_1, \ldots, a_n)_{\mathcal{R}} \qquad B = (a_1, \ldots, a_n)_{\mathcal{R}}$$

entonces tendremos que

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = -\overrightarrow{OA} + \overrightarrow{OB} = (b_1 - a_1, \ldots, b_n - a_n)_{\mathcal{R}}.$$

De lo anterior deducimos de forma inmediata que si $A = (a_1, \ldots, a_n)_{\mathcal{R}} \in \mathcal{A}$ y $v = (v_1, \ldots, v_n)_{\mathcal{B}} \in V$ entonces el trasladado B de A por v tiene coordenadas

$$B = (a_1 + v_1, \ldots, a_n + v_n)_{\mathcal{R}}.$$

Si el espacio vectorial V es de hecho un espacio vectorial Euclídeo y la base $\mathcal{B}$ es ortonormal, entonces diremos que $\mathcal{R}$ es un *sistema de referencia rectangular.* En este caso, dados dos puntos $A = (a_1, \ldots, a_n)_{\mathcal{R}}$ y $B = (a_1, \ldots, a_n)_{\mathcal{R}}$ se tiene que

$$\mathrm{d}(A, B) = \|\overrightarrow{AB}\| = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.$$

Ejemplo 6.4 En $\mathbb{R}^n$ como espacio afín euclídeo, tenemos el sistema de referencia rectangular canónico

$$\mathcal{R}_c = \{(0, \ldots, 0) : \mathcal{B}_c\}$$

donde $\mathcal{B}_c$ es la base canónica de $\mathbb{R}^n$.

Y si nos dan otro sistema de referencia $\mathcal{R}' = \{O', \mathcal{B}'\}$, ¿cómo se realiza el cambio de coordenadas de $\mathcal{R}'$ a $\mathcal{R}$? Al igual que ocurría con las bases en los espacios vectoriales, necesitamos que $\mathcal{R}$ tenga información de $\mathcal{R}'$. En concreto, necesitamos dos cosas:

1) La matriz de cambio de $\mathcal{B}'$ a $\mathcal{B}$, y

2) Las coordenadas $O' = (b_1, \ldots, b_n)_{\mathcal{R}}$ del punto O' respecto de $\mathcal{R}$, es decir, que

$$\overrightarrow{OO'} = (b_1, \ldots, b_n)_{\mathcal{B}}.$$

Veamos que con estos datos es suficiente. Consideremos un punto cualquiera $A \in \mathcal{A}$ del cual conocemos sus coordenadas respecto de $\mathcal{R}'$,

$$A = (x'_1, \ldots, x'_n)_{\mathcal{R}'}$$

es decir, que $\overrightarrow{O'A} = (x'_1, \ldots, x'_n)_{\mathcal{B}'}$. Queremos saber cuales son las coordenadas

$$A = (x_1, \ldots, x_n)_{\mathcal{R}}$$

de A respecto de $\mathcal{R}$. Por un lado tenemos que las coordenadas de $\overrightarrow{O'A}$ respecto de $\mathcal{B}$ son

$$P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

En particular, ya que $\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A}$, obtenemos que

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

La fórmula anterior se suele escribir en forma reducida

$$X = O' + PX'$$

donde X denota la matriz columna de las coordenadas de A respecto de $\mathcal{R}$, X' denota la matriz columna de las coordenadas de A respecto de $\mathcal{R}'$ y O' denota la matriz columna de las coordenadas de O' respecto de $\mathcal{R}$.

De hecho, haciendo un pequeño artificio, incluso podemos capturar la fórmula anterior en forma matricial, lo cual siempre es más cómodo:

$$\left(\begin{array}{c} 1 \\ \hline X \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline O' & P \end{array}\right) \left(\begin{array}{c} 1 \\ \hline X' \end{array}\right)$$

A la matriz cuadrada de orden $n+1$ que aparece en la expresión anterior

$$Q = \left(\begin{array}{c|c} 1 & 0 \\ \hline O' & P \end{array}\right)$$

la llamamos matriz del cambio de referencia de $\mathcal{R}'$ a $\mathcal{R}$. Al igual que pasaba en los cambios de base, la inversa de la matriz anterior es la matriz del cambio de referencia de $\mathcal{R}$ a $\mathcal{R}'$.

Para terminar la sección, veamos que dado un sistema de referencia $\mathcal{R} = \{O, \mathcal{B}\}$ de $\mathcal{A}$, es posible describir una variedad afín $A + W$ de $\mathcal{A}$ mediante ecuaciones paramétricas e implícitas. En efecto, si $A = (a_1, \ldots, a_n)$ y $W = L[w_1, \ldots, w_\ell]$ donde los vectores $w_1, \ldots, w_\ell$ son una base de W cuyas coordenadas respecto de $\mathcal{B}$ son

$$\left\{\begin{array}{l} w_1 = (w_{11}, \ldots, w_{\ell 1})_{\mathcal{B}} \\ \vdots \\ w_\ell = (w_{1\ell}, \ldots, w_{\ell\ell})_{\mathcal{B}} \end{array}\right.$$

entonces tendremos que un punto $X = (x_1, \ldots, x_n)_{\mathcal{R}}$ pertenece a $A+W$ si y solo si $\overrightarrow{AX} \in W$, es decir, si existen parámetros $\lambda_1, \ldots, \lambda_\ell \in \mathbb{R}$ tales que

$$\left\{\begin{array}{l} x_1 = a_1 + \lambda_1 w_{11} + \ldots + \lambda_\ell w_{1\ell} \\ \vdots \\ x_n = a_n + \lambda_1 w_{\ell 1} + \ldots + \lambda_\ell w_{\ell\ell} \end{array}\right.$$

las cuales son las *ecuaciones paramétricas* de $A + W$ en el sistema de referencia $\mathcal{R}$. Si eliminamos los parańetros de estas ecuaciones obtendremos unas *ecuaciones implícitas* de $A + W$.

6.3 Aplicaciones afines

Con el espíritu de las secciones anteriores, queremos definir una aplicación afín como "un punto más una aplicación lineal".

■ **Definición** Sean $\mathcal{A}$ y $\mathcal{A}'$ sobre espacios vectoriales V y V'. Dada una aplicación lineal

$$\vec{f} : V \to V'$$

y dos puntos $A \in \mathcal{A}$ y $A' \in \mathcal{A}'$, decimos que la aplicación

$$\begin{array}{rcl} f : \mathcal{A} & \to & \mathcal{A}' \\ X & \mapsto & A' + \vec{f}(\overrightarrow{AX}) \end{array}$$

es una *aplicación afín* con aplicación lineal asociada $\vec{f}$.

Observación 6.5 1) Obsérvese que por definición $f(A) = A' + \vec{f}(\overrightarrow{AA}) = A'$.

2) Si nos dan un punto $B \in \mathcal{A}$ cualquiera, entonces se tiene que

$$f(X) = f(B) + \vec{f}(\overrightarrow{BX}),$$

ya que

$$f(X) = A' + \vec{f}(\overrightarrow{AX}) = A' + \vec{f}(\overrightarrow{AB} + \overrightarrow{BX}) =$$
$$= A' + \vec{f}(\overrightarrow{AB}) + \vec{f}(\overrightarrow{BX}) = f(B) + \vec{f}(\overrightarrow{BX}).$$

Esta igualdad nos está asegurando que podemos usar un punto cualquiera para definir la aplicación.

3) La identidad $I : \mathcal{A} \mapsto \mathcal{A} : X \mapsto X$ es una aplicación afín cuya aplicación lineal asociada es también la identidad: en efecto, si fijamos un punto cualquiera $A \in \mathcal{A}$, podemos escribir

$$I(X) = X = A + \overrightarrow{AX}.$$

4) Si sabemos que $f : \mathcal{A} \to \mathcal{A}'$ es una aplicación afín, podemos recuperar la aplicación lineal asociada $\vec{f}$ de la siguiente forma:

$$\begin{array}{rcl} \vec{f} : V & \to & V' \\ \overrightarrow{XY} & \mapsto & \overrightarrow{f(X)f(Y)} \end{array}$$

5) Si nos dan una aplicación

$$f : \mathcal{A} \to \mathcal{A}'$$

¿cómo podemos saber si es una aplicación afín? Fácil, tan solo tenemos que fijar un punto cualquiera $A \in \mathcal{A}$ y comprobar que la correspondencia

$$\begin{array}{rcl} \vec{f} : V & \to & V' \\ \overrightarrow{AX} & \mapsto & \vec{f}(\overrightarrow{AX}) := \overrightarrow{f(A)f(X)} \end{array}$$

es una aplicación lineal (la cual será evidentemente la aplicación lineal asociada a f). En caso de que lo sea, tendremos que

$$\begin{array}{rcl} f : \mathcal{A} & \to & \mathcal{A}' \\ X & \mapsto & f(A) + \vec{f}(\overrightarrow{AX}) \end{array}$$

es una aplicación afín.

Ejercicio 6.6 Demostrar que la composición de dos aplicaciones afines es una aplicación afín.

Ejercicio 6.7 Demostrar que si $f, g : \mathcal{A} \to \mathcal{A}'$ son aplicaciones afines, entonces

$$f + g : \mathcal{A} \to \mathcal{A}' : X \mapsto f(X) + g(X)$$

es una aplicación afín.

Ejercicio 6.8 Demuestra que una aplicación afín $f : \mathcal{A} \to \mathcal{A}'$ es una biyección si y solo si $\vec{f}$ es un isomorfismo. En este caso, decimos que f es una *isomorfismo afín.*

Ejemplo 6.9 La aplicación

$$\begin{array}{rcl} f : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x + y + 1, x - z, x + y + z - 1) \end{array}$$

es una aplicación afín.

Ejemplo 6.10 Sea $\mathcal{A}$ un espacio afín sobre V y sea $v \in V$. La *traslación* de vector v es la aplicación

$$\begin{array}{rcl} t_v : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & X + v \end{array}$$

la cual, si fijamos un punto A, podemos escribir como

$$t_v(X) = (A + v) + \overrightarrow{AX}$$

y por tanto es evidentemente una aplicación afín cuya aplicación lineal asociada es la identidad. Por ejemplo, la traslación en $\mathbb{R}^3$ de vector $v = (1, 1, 2)$ es la aplicación

$$\begin{array}{rcl} t_v : \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (x, y, z) + (1, 1, 2) = (x + 1, y + 1, z + 2) \end{array}$$

Ejemplo 6.11 Sea $A + W$ una variedad afín de un espacio afín $\mathcal{A}$ sobre V. Recuérdese que en el Ejemplo 6.3 vimos que dado un punto $X \in \mathcal{A}$ podemos considerar la proyección ortogonal de X sobre $A + W$. Evidentemente, esto define una aplicación afín

$$\begin{array}{rcl} p_{A+W} : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & p_{A+W}(X) = A + p_W(\overrightarrow{AX}) \end{array}$$

que llamamos *proyección ortogonal* de $\mathcal{A}$ sobre $A + W$.

Ejemplo 6.12 Con la notación del ejemplo anterior, podemos definir también la simetría de $\mathcal{A}$ respecto de $A + W$, la cual viene dada por

$$\begin{array}{rcl} s_{A+W} : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & A + s_W(\vec{AX}) \end{array}$$

donde s_w denota la simetría ortogonal de V respecto de W. Es fácil probar (ejercicio) que $s_{A+W}(X) = X - 2p_W(X)$ usando el hecho de que $s_W = I - 2p_W$.

Al igual que hicimos con la aplicaciones lineales, nos gustaría trabajar con las aplicaciones afines de forma matricial. Sean $\mathcal{A}$ y $\mathcal{A}'$ espacios afines sobre V y V' respectivamente. Sean $\mathcal{R} = \{O; \mathcal{B}\}$ y $\mathcal{R}' = \{O'; \mathcal{B}'\}$ sistemas de referencia de $\mathcal{A}$ y $\mathcal{A}'$ respectivamente. Consideremos una aplicación afín

$$f : \mathcal{A} \to \mathcal{A}$$

con aplicación lineal asociada $\vec{f} : V \to V'$. Sabemos entonces que para cualquier punto $X \in \mathcal{A}$ se tiene que

$$f(X) = f(O) + \vec{f}(\overrightarrow{OX}).$$

Denotemos por $X = (x_1, \dots, x_n)_{\mathcal{R}}$ y $f(X) = (y_1, \dots, y_m)_{\mathcal{R}'}$ a las coordenadas de X y $f(X)$ respecto de $\mathcal{R}$ y $\mathcal{R}'$. Entonces,

1) si M es la matriz de $\vec{f}$ respecto de $\mathcal{B}$ y $\mathcal{B}'$, y

2) $f(O) = (c_1, \dots, c_m)_{\mathcal{R}'}$,

se tiene que

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} + M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Esta expresión podemos reescribirla de la siguiente forma

$$\left(\begin{array}{c} 1 \\ \hline y_1 \\ \vdots \\ y_m \end{array}\right) = \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_1 & & & \\ \vdots & & M & \\ c_m & & & \end{array}\right)}_{N} \left(\begin{array}{c} 1 \\ \hline x_1 \\ \vdots \\ x_n \end{array}\right)$$

donde a la matriz N la denominamos *matriz de f asociada a los sistemas de referencia $\mathcal{R}$ y $\mathcal{R}'$*.

Ejemplo 6.13 La aplicación

$$\begin{array}{rcl} f:\mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (x+y+1, x-z, x+y+z-1) \end{array}$$

es una aplicación afín cuya matriz respecto del sistema de referencia canónico $\mathcal{R}_c$ es

$$N = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{array}\right)$$

La traslación de vector $v = (1,1,2)$,

$$\begin{array}{rcl} t_v:\mathbb{R}^3 & \to & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (x,y,z)+(1,1,2) = (x+1, y+1, z+2) \end{array}$$

tiene matriz asociada respecto del sistema de referencia canónico $\mathcal{R}_c$

$$N' = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right)$$

En particular, la matriz de $f \circ t_v$ respecto del sistema de referencia canónico $\mathcal{R}_c$ es

$$N \cdot N' = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & 1 \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 3 & 1 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ 3 & 1 & 1 & 1 \end{array}\right)$$

Ejercicio 6.14 Sea $f:\mathcal{A} \to \mathcal{A}'$ una aplicación afín. Demostrar que

$$\mathrm{Im}(f) = \{f(A) : A \in \mathcal{A}\}$$

es una variedad afín de $\mathcal{A}'$. Dado $B \in \mathcal{A}'$, demostrar que $f^{-1}(B)$ es una variedad afín de $\mathcal{A}$.

6.4 Movimientos

En esta sección $\mathcal{A}$ denotará siempre un espacio afín euclídeo. Dado que tenemos un concepto de distancia en $\mathcal{A}$, resulta natural estudiar las aplicaciones afines que respetan dicha distancia.

■ **Definición** Decimos que una aplicación afín $f : \mathcal{A} \to \mathcal{A}$ es un *movimiento* si conserva la distancia entre puntos, es decir,

$$\mathrm{d}(A, B) = \mathrm{d}(f(A), f(B)) \qquad \text{para cualesquiera } A, B \in \mathcal{A}$$

■ **Proposición** *Una aplicación afín $f : \mathcal{A} \to \mathcal{A}$ es un movimiento si y solo si su aplicación lineal asociada es un endomorfismo ortogonal.*

Demostración. Para cualesquiera $A, B \in \mathcal{A}$ se tiene que

$$\mathrm{d}(A, B) = \mathrm{d}(f(A), f(B)) \iff \|\overrightarrow{AB}\| = \|\overrightarrow{f(A)f(B)}\| = \|\vec{f}(\overrightarrow{AB})\|$$

es decir, que f respeta la distancia si y solo si $\vec{f}$ respeta la norma. □

■ **Corolario** *Sea $\mathcal{R} = \{0; \mathcal{B}\}$ un sistema de referencia rectangular de $\mathcal{A}$. Sea $f : \mathcal{A} \to \mathcal{A}$ una aplicación afín cuya matriz respecto de $\mathcal{R}$ es*

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline C & M \end{array}\right).$$

Entonces f es un movimiento si y solo si M es ortogonal.

Demostración. Es inmediato, ya que la matriz M es la matriz de $\vec{f}$ respecto de la base ortonormal $\mathcal{B}$, la cual es ortogonal si y solo si $\vec{f}$ es un endomorfismo orotogonal. □

Como ya sabemos cómo son los endomorfismos ortogonales de los espacios vectoriales euclídeos $\mathbb{R}^2$ y $\mathbb{R}^3$, podemos clasificar los movimientos de $\mathbb{R}^2$ y $\mathbb{R}^3$ como espacios afines euclídeos. La mejor forma de hacerlo es a través de los posibles puntos fijos de un movimiento $f : \mathcal{A} \to \mathcal{A}$. Observemos primero que los puntos fijos de f, si no es un conjunto vacío, es una variedad afín de $\mathcal{A}$. En efecto, si existe un punto $A \in \mathcal{A}$ que es fijo entonces $f(A) = A$ y por tanto se tiene que

$$X = f(X) = f(A) + \vec{f}(\overrightarrow{AX}) = A + \vec{f}(\overrightarrow{AX}) \iff \overrightarrow{AX} = \vec{f}(\overrightarrow{AX}) \iff$$

$$\iff \overrightarrow{AX} \in \mathrm{Ker}(\vec{f} - I) \iff X \in A + \mathrm{Ker}(\vec{f} - I).$$

Es decir, los puntos fijos de f son

$$A + \mathrm{Ker}(\vec{f} - I). \tag{$*$}$$

Para clasificar un movimiento $f : \mathbb{R}^2 \to \mathbb{R}^2$ debemos estudiar sus puntos fijos:

- Si todo $\mathbb{R}^2$ queda fijo por medio de f, entonces se trata claramente de la *identidad*.
- Si f tiene una recta de puntos fijos, entonces por $(*)$ deducimos que $\lambda = 1$ es un autovalor de $\vec{f}$ y la dimensión de su subespacio propio $V_1 = \mathrm{Ker}(\vec{f} - I)$ es 1. Por tanto $\vec{f}$ debe ser la simetría axial s_{V_1} respecto de V_1. De hecho, si tomamos un punto fijo concreto A entonces podemos escribir

  $$f(X) = A + s_{V_1}(\overrightarrow{AX}),$$

  es decir, que f es la *simetría ortogonal* respecto de la variedad afín $A + V_1$ (véase el Ejemplo 6.12).
- Si f tiene solo un punto fijo $A \in \mathbb{R}^2$, de nuevo por $(*)$ se tiene que $\mathrm{Ker}(\vec{f} - I) = \{0\}$ y por tanto $\vec{f}$ debe ser un giro de ángulo α (lo cual incluye a la simetría central, es decir, el giro de ángulo π). En particular observamos que f es simplemente un *giro* de ángulo α alrededor del punto fijo A.

- Si f no tiene puntos fijos, pueden ocurrir dos cosas. Observemos primero que, denotando $O = (0,0) \in \mathbb{R}^2$, tenemos que $X \in \mathbb{R}^2$ es un punto fijo si

$$X = f(X) = f(O) + \vec{f}(\overrightarrow{OX}) \Longleftrightarrow \overrightarrow{f(O)X} = \vec{f}(\overrightarrow{OX}) \Longleftrightarrow \overrightarrow{f(O)O} + \overrightarrow{OX} = \vec{f}(\overrightarrow{OX})$$

$$\Longleftrightarrow \overrightarrow{f(O)O} = \vec{f}(\overrightarrow{OX}) - \overrightarrow{OX} \Longleftrightarrow \overrightarrow{f(O)O} = (\vec{f} - I)(\overrightarrow{OX})$$

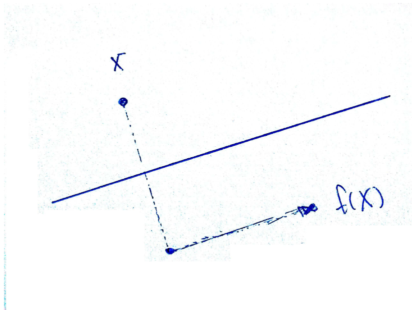
Como f no tiene puntos fijos, entonces no existen puntos $X \in \mathbb{R}^2$ que satisfagan la última condición, es decir, tales que $\overrightarrow{f(O)O} = (\vec{f} - I)(\overrightarrow{OX})$. Deducimos entonces que $\lambda = 1$ debe ser autovalor de $\vec{f}$. Si no lo fuese, entonces $\vec{f} - I$ sería un isomorfismo (su determinante no es cero) y por lo tanto el punto

$$X = O + (\vec{f} - I)^{-1}(\overrightarrow{f(O)O})$$

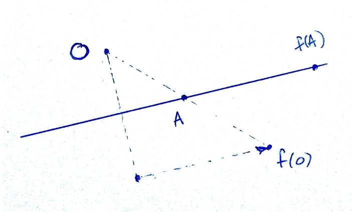
sería un punto fijo, contradicción. En particular $\vec{f}$ debe ser la identidad o una simetría axial respecto de $V_1 = \mathrm{Ker}(\vec{f} - I)$. Si $\vec{f}$ es la identidad, entonces f es una *traslación* cuyo vector de traslación es $\overrightarrow{Of(O)}$ ya que

$$f(X) = f(O) + \overrightarrow{OX} = O + \overrightarrow{Of(O)} + \overrightarrow{OX} = O + \overrightarrow{OX} + \overrightarrow{Of(O)} = X + \overrightarrow{Of(O)}.$$

Si $\vec{f}$ es una simetría respecto de $V_1 = \mathrm{Ker}(\vec{f} - I)$, es decir, $\vec{f} = s_{V_1}$, veamos que f es una *simetría deslizante*, es decir, es la composición de una simetría axial con una traslación cuyo vector de traslación es paralelo al eje de la simetría.



Definamos el punto $A = O + \frac{1}{2}\overrightarrow{Of(O)}$ y veamos que $\overrightarrow{Af(A)} \in V_1$.



En efecto,

$$f(A) = f(O) + \vec{f}(\overrightarrow{OA}) = f(O) + \vec{f}(\frac{1}{2}\overrightarrow{Of(O)}) = f(O) + \frac{1}{2}\vec{f}(\overrightarrow{Of(O)})$$

y por tanto

$$\overrightarrow{Af(A)} = \overrightarrow{AO} + \overrightarrow{Of(O)} + \overrightarrow{f(O)f(A)} = -\frac{1}{2}\overrightarrow{Of(O)} + \overrightarrow{Of(O)} + \frac{1}{2}\vec{f}(\overrightarrow{Of(O)}) =$$

$$= \frac{1}{2}\overrightarrow{Of(O)} + \frac{1}{2}\vec{f}(\overrightarrow{Of(O)})$$

y en particular, puesto que $\vec{f}^2 = I$ por ser $\vec{f}$ una simetría,

$$\vec{f}(\overrightarrow{Af(A)}) = \frac{1}{2}\vec{f}(\overrightarrow{Of(O)}) + \frac{1}{2}\vec{f}^2(\overrightarrow{Of(O)}) =$$

$$= \frac{1}{2}\vec{f}(\overrightarrow{Of(O)}) + \frac{1}{2}(\overrightarrow{Of(O)}) = \overrightarrow{Af(A)},$$

como queríamos probar. Finalmente, podemos escribir f como $t \circ \vec{s}_{A+V_1}$ donde

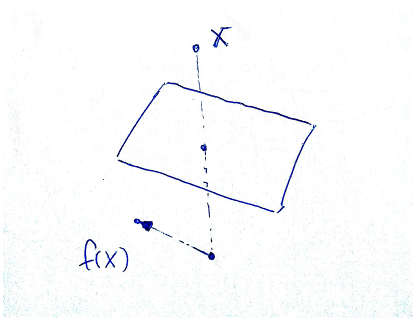
$$\begin{array}{rcl} \vec{s}_{A+V_1} : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & A + \vec{s}_{V_1}(\overrightarrow{AX}) \end{array} \qquad \begin{array}{rcl} t : \mathcal{A} & \to & \mathcal{A} \\ X & \mapsto & X + \overrightarrow{Af(A)} \end{array}$$

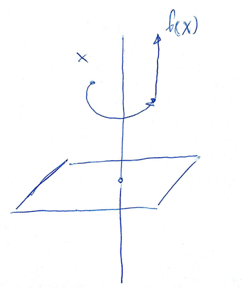
donde $\vec{s}_{A+V_1}$ es la simetría axial respecto de $A+V_1$ y t es una traslación de vector $\overrightarrow{Af(A)}$ el cual es paralelo a la recta $A+V_1$ puesto que $\overrightarrow{Af(A)} \in V_1$. Observamos que aunque f no tiene puntos fijos, la recta $A+V_1$ es invariante, es decir, $f(A+V_1) = A+V_1$.

Para clasificar los movimientos $f : \mathbb{R}^3 \to \mathbb{R}^3$ podemos actuar de forma parecida, clasificando en función de la variedad afín de puntos fijos, pero seremos más breves con los detalles:

- Si todo $\mathbb{R}^3$ queda fijo entonces f es la *identidad*.
- Si f tiene un plano de puntos fijos entonces f es una *simetría ortogonal* respecto del plano de puntos fijos.
- Si f tiene una recta de puntos fijos entonces f es una *rotación axial* cuyo eje es la recta de puntos fijos.
- Si f solo deja un punto fijo entonces f es la *composición de una rotación axial y una simetría* respecto al plano perpendicular del eje de la rotación que se intersecan en el punto fijo de f.
- Si f no tiene puntos fijos, entonces f es una *traslación*, o una *simetría deslizante* o un *movimiento helicoidal*.

La simetría deslizante es la composición de una simetría respecto de un plano con una traslación cuyo vector de traslación es paralelo al plano de la simetría (y aunque no tiene puntos fijos, sí que tiene un plano invariante: el propio plano de la simetría). El movimiento helicoidal es la composición de una rotación axial con una traslación cuyo vector de traslación es paralelo al eje de la rotación (y aunque no tiene puntos fijos, sí que tiene una recta invariante: el eje de la rotación axial).





6.5 Cónicas

Antes de comenzar esta sección, recordemos que en el siguiente enlace se pueden refrescar los conceptos sobre cónicas que necesitaremos

http://www.mat.ucm.es/ eliasbaro/WebMatBas/Hojas/conicas.pdf

Recordemos que las ecuaciones en $\mathbb{R}^2$ de las cónicas según nos las mostraron en cursos anteriores son:

$$\begin{array}{rcl} \text{Elipse} & \rightsquigarrow & \frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \\ \text{Parábola} & \rightsquigarrow & y^2=2px \\ \text{Hipérbola} & \rightsquigarrow & \frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \end{array}$$

Estas son las ecuaciones de las cónicas cuando los elementos geométricos que la definen están bien situados respecto del sistema de referencia canónico $\mathcal{R}_c$, y las conocemos como *ecuaciones reducidas*. Por ejemplo, si fijamos los puntos $F_1 = (-d, 0)$ y $F_2 = (d, 0)$ y consideramos la elipse con focos F_1 y F_2 y semieje $a > d$, es decir, los puntos $X \in \mathbb{R}^2$ tales que

$$\mathrm{d}(X, F_1) + \mathrm{d}(X, F_2) = 2a$$

entonces la ecuación que deben satisfacer las coordenadas (x, y) de un punto para pertenecer a la elipse es precisamente $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ donde $b = \sqrt{a^2 - d^2}$.

Pero evidentemente los elementos geométricos de las cónicas no tienen por qué estar tan bien colocados con respecto del sistema de referencia canónico. Sin embargo, es fácil imaginar que respecto de un cierto sistema de referencia rectangular $\mathcal{R}'$ de $\mathbb{R}^2$ los elementos geométricos sí están bien colocados y por tanto la ecuación respecto de $\mathcal{R}'$ es reducida, por ejemplo, la de una elipse

$$\frac{[x']^2}{a^2}+\frac{[y']^2}{a^2}=1.$$

donde $(x', y')_{\mathcal{R}'}$ son la coordenadas respecto de $\mathcal{R}'$. Tan solo se trata de una cónica de las de siempre, solo que la hemos trasladado y girado. ¿Cómo será la ecuación de esta cónica respecto de $\mathcal{R}_c$? Las ecuaciones de paso de $\mathcal{R}_c$ a $\mathcal{R}'$ serán de la forma

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

y por tanto la ecuación de la cónica usando las coordenadas respecto de $\mathcal{R}$ será

$$\frac{(\alpha_1 x + \alpha_2 y + \alpha_3)^2}{a^2} + \frac{(\beta_1 x + \beta_2 y + \beta_3)^2}{b^2} = 1.$$

Si desarrollamos la ecuación anterior, seguro que nos quedará algo de la forma:

$$a_{11}x^2 + a_{22}y^2 + 2a_{21}xy + a_{01}x + a_{02}y + a_{00} = 0. \tag{$*$}$$

Así que es natural preguntarse lo siguiente: si nos dan una ecuación del tipo $(*)$, ¿cómo podemos saber de qué cónica se trata? ¿cómo podemos calcular un sistema de referencia respecto del cual la cónica tiene una ecuación reducida?

Dada una ecuación del tipo $(*)$, la podemos escribir de la forma

$$X^t AX + BX + a_{00} = 0$$

donde

$$A = \begin{pmatrix} a_{11} & a_{22} \\ a_{12} & a_{22} \end{pmatrix} \qquad B = \begin{pmatrix} a_{01} & a_{02} \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

Nuestro objetivo es encontrar un sistema de referencia rectangular $\mathcal{R}'$ de forma que cuando escribimos la ecuación anterior respecto de las coordenadas respecto de $\mathcal{R}'$ nos quede reducida. Para ello hay que seguir dos pasos:

1) calcular una matriz P ortogonal tal que $P^{-1}AP$ es una matriz diagonal, y

2) completar cuadrados.

Lo mostramos con un par de ejemplos.

Ejemplo 6.15 Consideremos la ecuación

$$x^2 + 6x + 5y + 14 = 0$$

la cual podemos escribir como

$$X^t AX + BX + 14 = 0$$

donde

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & \frac{5}{2} \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

La matriz A ya es diagonal, así que tan solo queda completar cuadrados:

$$0 \quad = \quad x^2 + 6x + 5y + 14 = (x+3)^2 - 9 + 5y + 14 = (x+3)^2 + 5y - 5 = (x+3)^2 + 5(y-1)$$

Así que si hacemos la sustitución

$$\begin{cases} x' = x + 3 \\ y' = y - 1 \end{cases}$$

entonces la ecuación quedaría

$$[x']^2 + 5y' = 0$$

es decir, es una parábola. O dicho de otra manera, la ecuación con la que empezamos, respecto del sistema de referencia $\mathcal{R}' = \{(-3, 1); (1, 0), (0, 1)\}$ es $[x']^2 + 5y' = 0$.

Ejemplo 6.16 Consideremos la ecuación

$$2x^2 - 4xy - y^2 - 4x - 8y + 14 = 0$$

la cual podemos escribir como

$$X^t AX + BX + 14 = 0$$

donde

$$A = \begin{pmatrix} 2 & -2 \\ -2 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} -4 & -8 \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

Haciendo los cálculos necesarios, podemos diagonalizar la matriz A de forma que

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

En particular, si hacemos el cambio de variable

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

entonces la ecuación quedará de la forma

$$[X']^t P^t A P X' + B P X' + 14 = 0$$

es decir,

$$-2[x']^2 + 3[y']^2 - \frac{20}{\sqrt{5}} x' + 14 = 0$$

y completando cuadrados queda

$$-2[x' + \frac{5}{\sqrt{5}}]^2 + 3[y']^2 + 24 = 0$$

así que si hacemos la sustitución

$$\left\{ \begin{array}{l} x'' = x' + \frac{5}{\sqrt{5}} \\ y'' = y' \end{array} \right.$$

la ecuación queda

$$\frac{[x'']^2}{12} - \frac{[y'']^2}{4} = 1$$

es decir, es una hipérbola. Recapitulando, el cambio de variable que hemos realizado es

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{5}} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \quad \leftrightsquigarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\frac{5}{\sqrt{5}} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Es decir, que la hipérbola tiene ecuación reducida respecto del sistema de referencia rectangular:

$$\mathcal{R}' = \{(-\frac{5}{\sqrt{5}}, 0); (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}), (-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})\}$$