

Aplicaciones entre espacios euclídeos y unitarios

Fernando Chamizo

25 de febrero de 2019

1. La proyección ortogonal

En \mathbb{R}^2 la aplicación $\vec{x} = (x, y)^t \mapsto (x, 0)^t$ da la “parte” de \vec{x} que está en el eje X , lo que queda al anular la segunda coordenada. De la misma forma $(x, y, z)^t \mapsto (x, y, 0)^t$ “aplata” el vector sobre el plano XY . Este proceso de quedarse con algunas coordenadas es lo que se llama una *proyección* aunque la definición abstracta no recuerda demasiado a esta idea. Concretamente, dado un espacio vectorial V en matemáticas se llama *proyección* a cualquier aplicación lineal $P : V \rightarrow V$ tal que $P^2 = P$. En esta sección nos ocuparemos de un caso especial de proyecciones que como en los ejemplos anteriores corresponden a anular algunas coordenadas en una base ortonormal.

Si V es un espacio vectorial euclídeo o hermítico (unitario) y $W \subset V$ es un subespacio, en el caso de dimensión finita siempre se tiene una descomposición $V = W \oplus W^\perp$, lo cual significa que cada $\vec{v} \in V$ se escribe de forma única como $\vec{v} = \vec{a} + \vec{b}$ con $\vec{a} \in W$ y $\vec{b} \in W^\perp$. Se llama *proyección ortogonal* sobre W a la aplicación lineal P_W que asigna a cada $\vec{v} \in V$ el vector \vec{a} .

Observación. El caso de dimensión infinita aparece en física pero aquí no lo podemos tratar con generalidad porque $V = W \oplus W^\perp$ puede fallar dependiendo del caso. En la práctica es común que $\dim W < \infty$ y entonces V se pueda restringir a un subespacio de dimensión finita que contenga los vectores de interés y que incluya a W .

En gran medida la importancia de la proyección ortogonal radica en que da la mejor aproximación de cada vector de V por un vector de W . Concretamente, dado $\vec{v} \in V$

$$(1) \quad \|\vec{v} - P_W(\vec{v})\| \leq \|\vec{v} - \vec{w}\| \quad \text{para todo } \vec{w} \in W.$$

La prueba es simplemente observar que $\vec{x} = \vec{v} - P_W(\vec{v}) \in W^\perp$ y como $\vec{y} = P_W(\vec{v}) - \vec{w} \in W$ el teorema de Pitágoras asegura $\|\vec{v} - \vec{w}\|^2 = \|\vec{x} + \vec{y}\|^2 \geq \|\vec{x}\|^2$.

De la definición de la proyección ortogonal se deduce

$$(2) \quad \vec{v} = P_W(\vec{v}) + P_{W^\perp}(\vec{v}).$$

El cálculo efectivo de P_W se basa en imponer que $\vec{v} - P_W(\vec{v})$ pertenezca a W^\perp . Gracias a la relación (2) puede ser más conveniente intercambiar el papel que desempeñan W y W^\perp .

Ejemplo. Si $\dim W = 1$ entonces W está formado por los múltiplos de un vector cualesquiera $\vec{w}_0 \in W - \{\vec{0}\}$. Llamando $\lambda\vec{w}_0$ a $P_W(\vec{v})$, se tiene $\vec{v} - \lambda\vec{w}_0 \in W^\perp$ y tomando el producto escalar con \vec{w}_0 se tiene $\langle \vec{v}, \vec{w}_0 \rangle - \lambda \langle \vec{w}_0, \vec{w}_0 \rangle$. Es decir la proyección ortogonal sobre un espacio de dimensión 1 viene dada por la fórmula $\vec{v} \mapsto \vec{w}_0 \langle \vec{v}, \vec{w}_0 \rangle / \|\vec{w}_0\|^2$. Si $V = \mathbb{R}^2$ (con el producto escalar usual) y W es el eje X , tomando $\vec{w}_0 = (1, 0)$ se sigue que la proyección de $(x, y)^t$ es $(x, 0)^t$, en la línea de lo dicho al principio.

Ejemplo. Uno de los postulados más controvertidos de la física cuántica es que las mediciones están asociadas a proyecciones cuya norma al cuadrado da la probabilidad. Por ejemplo, medir el espín arriba o abajo corresponde a proyectar sobre \vec{e}_1 o \vec{e}_2 de la base canónica de \mathbb{C}^2 (sobre los espacios lineales generados) y el espín en el eje X (a derecha o izquierda) a proyectar sobre $\vec{u}_1 = (1, 1)^t / \sqrt{2}$ o $\vec{u}_2 = (1, -1)^t / \sqrt{2}$. La proyección de \vec{e}_1 sobre \vec{u}_1 y después sobre \vec{e}_1 es $(1/2, 0)^t$ con lo que tiene probabilidad 1/4, lo cual físicamente da lugar a la paradoja de que si entre las partículas en la que hemos medido espín arriba seleccionamos las que tienen espín derecho, al medir de nuevo el espín solo la cuarta parte de ellas lo conservarán. La medición ha cambiado el estado.

Ejemplo. Hallemos la proyección ortogonal de $\vec{v} = (1, 2, 3)^t \in \mathbb{R}^3$ en $W = \{(x, y, z) \in \mathbb{R}^3 : x + y - 2z = 0\}$. Se tiene $\dim W^\perp = 1$, claramente W^\perp está generado por $\vec{w}_0 = (1, 1, -2)^t$, el vector normal al plano W . Según la fórmula del ejemplo anterior, $P_{W^\perp}(\vec{v})$ es $-\vec{w}_0/2$ y (2) implica que $P_W(\vec{v})$ es $\vec{v} + \vec{w}_0/2 = (3/2, 5/2, 2)^t$. Obviamente este vector está en W .

Vamos a derivar una fórmula general para hallar la proyección ortogonal en coordenadas para referencia posterior pero lo importante es el método, no hace falta memorizarla. Consideremos primero que V es un espacio euclídeo ($\mathbb{K} = \mathbb{R}$) y que hemos fijado una base ortonormal \mathcal{B} (por ejemplo la canónica en el caso de \mathbb{R}^m). En dicha base el producto escalar es $\langle \vec{x}, \vec{y} \rangle = \vec{x}^t \vec{y}$ donde identificamos en el segundo miembro vectores y sus coordenadas en \mathcal{B} . Si $\{\vec{w}_1, \dots, \vec{w}_n\}$ es una base de $W \subset V$ sea S la matriz cuyas columnas son las coordenadas de estos vectores en la base \mathcal{B} y sea G la matriz con elementos $g_{ij} = \langle \vec{w}_i, \vec{w}_j \rangle$. Esta matriz coincide con $S^t S$. Escribiendo $P_W(\vec{v}) = \sum_j \lambda_j \vec{w}_j$, se tiene $\langle \vec{v} - \sum_j \lambda_j \vec{w}_j, \vec{w}_i \rangle = 0$ que implica

$$(3) \quad \sum_j \langle \vec{w}_i, \vec{w}_j \rangle \lambda_j = \langle \vec{v}, \vec{w}_i \rangle$$

o en notación matricial $G\vec{\lambda} = S^t \vec{v}$ donde $\vec{\lambda} = (\lambda_1, \dots, \lambda_n)^t$. Entonces

$$(4) \quad P_W(\vec{v}) = S G^{-1} S^t \vec{v} \quad \text{con} \quad G = S^t S$$

porque $\sum_j \lambda_j \vec{w}_j = S\vec{\lambda}$.

En un par de ocasiones de este argumento hemos usado la simetría del producto escalar. En el caso en que V es hermítico hay que poner en esos pasos un conjugado. El resultado en ese caso, con la misma notación, es:

$$(5) \quad P_W(\vec{v}) = S\overline{G}^{-1}\overline{S}^t\vec{v} \quad \text{con} \quad G = S^t\overline{S}.$$

En ambos casos no es posible en general usar en G^{-1} la inversa de un producto y simplificar porque la matriz S no es cuadrada si $W \subsetneq V$.

Ejemplo. Aunque sea menos eficiente, vamos a rehacer el último ejemplo con lo aprendido en (3) y (4). Una base de W está formada por $\vec{w}_1 = (2, 0, 1)^t$ y $\vec{w}_2 = (0, 2, 1)^t$. Si escribimos $\vec{v} = \lambda\vec{w}_1 + \mu\vec{w}_2 + \vec{u}$ con $\vec{u} \in W^\perp$ nos libramos de \vec{u} multiplicando por los \vec{w}_j de donde $5 = 5\lambda + \mu$ y $7 = \lambda + 5\mu$, que es el sistema (3). La solución es $\lambda = 3/4$, $\mu = 5/4$ que conduce a $P_W(\vec{v}) = \lambda\vec{w}_1 + \mu\vec{w}_2 = (3/2, 5/2, 2)^t$ como antes.

Si queremos proceder con la fórmula (4), las columnas de S serían las coordenadas de \vec{w}_1 y \vec{w}_2 . El resultado sería

$$(6) \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \frac{1}{24} \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{24} \begin{pmatrix} 10 & -2 \\ -2 & 10 \\ 4 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 5/2 \\ 2 \end{pmatrix}$$

que coincide de nuevo con la solución anterior.

2. El método de mínimos cuadrados

En estadística descriptiva la *recta de regresión* que ajusta datos a una recta se obtiene con el llamado *método de mínimos cuadrados*. Hay unas fórmulas ya hechas para obtener la recta que nosotros evitaremos para explicar cómo se deducen de la proyección ortogonal (lo habitual es usar técnicas básicas de análisis). Este enfoque tiene sus ventajas pues permite tratar de la misma forma problemas relacionados.

Comencemos primero con el caso de la recta de regresión. Tenemos una lista de datos formando un vector $\vec{y} \in \mathbb{R}^n$ que queremos ajustar linealmente con unos datos de partida \vec{x} . Es decir, buscamos un a y un b tales que y_i sea aproximadamente $a + bx_i$, la recta $y = a + bx$ es la *recta de regresión*. Introduciendo la matriz X de dimensiones $n \times 2$ que tienen como primera columna unos y como segunda columna \vec{x} , lo que buscamos es un $\vec{\beta} \in \mathbb{R}^2$ tal que

$$(7) \quad \vec{y} \approx X\vec{\beta} \quad \text{con} \quad \vec{\beta} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Si tomamos $V = \mathbb{R}^n$ y $W = \text{Im } X$ (el subespacio generado por las columnas de la matriz), entonces (1) dice que la mejor aproximación está dada por la proyección ortogonal. Es la mejor en el sentido de que minimiza la norma usual en \mathbb{R}^n . Esta viene dada por sumas de cuadrados, lo que justifica el nombre del método.

Ajustando la notación de (4) a la de (7) se tiene $P_W(\vec{y}) = XG^{-1}X^t\vec{y}$. Recordando que $G = X^tX$ y que nosotros estamos interesados en $\vec{\beta}$ en lugar de en $X\vec{\beta}$, la fórmula que resulta para la mejor aproximación en el sentido indicado es:

$$(8) \quad \vec{\beta} = (X^tX)^{-1}X^t\vec{y}.$$

La distancia $\|\vec{y} - X\vec{\beta}\|$ es mínima. En vez de esta cantidad se suele utilizar como medida de la bondad del ajuste el *error cuadrático medio*, que no es más que $\|\vec{y} - X\vec{\beta}\|^2/n$; el cuadrado del error en relación con el tamaño de la muestra.

Una vez más, en un caso real la matriz X será muy alta y con solo dos columnas por lo que no tiene sentido invertir $(X^tX)^{-1}$ como si fuera un producto de matrices cuadradas. Si comparamos (8) y (7), vemos que $(X^tX)^{-1}X^t$ funciona como una especie de “inversa aproximada” de la matriz no cuadrada X . A veces se la conoce como la *inversa de Moore-Penrose*.

Ejemplo. Digamos que queremos aplicar el método de mínimos cuadrados para encontrar una regla que ajuste linealmente la calificación en un examen en términos del número de horas de estudio de acuerdo con la siguiente tabla:

n° horas	0	1	2	3
calificación	1	3	7	9

Con la notación anterior tenemos

$$(9) \quad X^t = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}, \quad \vec{y} = (1, 3, 7, 9)^t.$$

El cálculo es entonces

$$(10) \quad (X^tX)^{-1}X^t\vec{y} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}^{-1} \begin{pmatrix} 20 \\ 44 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 14 & -6 \\ -6 & 4 \end{pmatrix} \begin{pmatrix} 20 \\ 44 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 4 \\ 14 \end{pmatrix}.$$

Por tanto la recta de regresión es $y = 4/5 + 14x/5$. Aproximadamente multiplicando el número de horas por $14/5$ y sumando $4/5$ da la calificación. El error cuadrático medio será

$$(11) \quad \frac{1}{4} \|\vec{y} - X\vec{\beta}\|^2 = \frac{1}{4} \|(1, 3, 7, 9)^t - \frac{1}{5}(1, -3, 3, -1)^t\|^2 = \frac{1}{4} \cdot \frac{1+9+9+1}{25} = \frac{1}{5}.$$

Si cambiamos el enunciado inicial y ponemos 5 y 7 en lugar de las dos últimas calificaciones, la recta $y = 1 + 2x$ da un ajuste perfecto y eso se traduce en un error cuadrático medio nulo que en términos de álgebra lineal significa que estamos proyectando sobre un subespacio que ya contiene al vector.

Observación. Diga lo que diga el álgebra lineal de la mejor aproximación, las necesidades de la estadística a veces son distintas. Por ejemplo, el método de mínimos cuadrados es muy sensible a los valores atípicos (a veces llamados *outliers*) que pueden venir de un error en la toma de los datos. En ese caso, es conveniente hacer un tratamiento previo que “cocine” los datos tratando de identificar los que posiblemente se deban a errores.

La misma fórmula (8) permite tratar situaciones más generales. Por ejemplo, digamos que buscamos una aproximación por medio de una parábola $y = a + bx + cx^2$. El problema se escribe en la forma (7) pero ahora con $\vec{\beta} = (a, b, c)^t$ y X con una tercera columna que tenga los cuadrados de las coordenadas de \vec{x} . Para grados superiores se procedería de forma análoga.

Ejemplo. Si hacemos un ajuste parabólico de los datos del ejemplo anterior resulta que $c = 0$, lo que significa que la parábola se convierte en la recta que ya conocemos. Esto es una casualidad. Para evitarla, supongamos que la última calificación es 8 en lugar de 9. Entonces

$$(12) \quad X^t = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix}, \quad \vec{y} = (1, 3, 7, 8)^t.$$

Los cálculos serían ahora

$$(13) \quad (X^t X)^{-1} X^t \vec{y} = \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix}^{-1} \begin{pmatrix} 19 \\ 41 \\ 103 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 19 & -21 & 5 \\ -21 & 49 & -15 \\ 5 & -15 & 5 \end{pmatrix} \begin{pmatrix} 19 \\ 41 \\ 103 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 3 \\ 13 \\ -1 \end{pmatrix}.$$

De este modo la mejor parábola en el sentido de minimizar distancias es $y = (3 + 13x - x^2)/4$.

3. Aplicaciones ortogonales, unitarias y autoadjuntas

Una vez que tenemos un producto escalar, una manera de medir, es natural estudiar las transformaciones que lo dejan fijo. En \mathbb{R}^2 con el producto euclídeo habitual parece que todo lo que podemos hacer para preservar distancias y ángulos es trasladar, girar y hacer simetrías. Más adelante en el curso clasificaremos todas las posibilidades en \mathbb{R}^2 y \mathbb{R}^3 . Por ahora daremos una definición genérica que comprende el caso lineal (una traslación no define una aplicación lineal).

Se dice que un endomorfismo biyectivo $L : V \rightarrow V$ en un espacio euclídeo V es una *aplicación ortogonal* si

$$(14) \quad \langle L(\vec{x}), L(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle \quad \text{para todo } \vec{x}, \vec{y} \in V.$$

Si el espacio V es unitario, se dice que L es una *aplicación unitaria*.

En realidad $\langle L(\vec{x}), L(\vec{y}) \rangle = \langle \vec{x}, \vec{y} \rangle$ implica $\text{Nuc } L = \{\vec{0}\}$ y entonces no es necesario pedir la biyectividad en el caso de dimensión finita. Siguiendo con este caso, si con el abuso de notación habitual identificamos L con su matriz en una base ortonormal $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ y cada vector con su coordenadas, entonces el producto escalar será el usual $\vec{x}^t \vec{y}$ o $\vec{x}^t \overline{\vec{y}}$ y la condición (14) se traduce en que L es ortogonal o unitaria si y solo si su matriz en una base ortonormal satisface, respectivamente,

$$(15) \quad L^t L = I \quad \text{o} \quad L^t \overline{L} = I.$$

Se habla correspondientemente de *matrices ortogonales* o *unitarias*. Equivalentemente, estas matrices satisfacen que su inversa es su traspuesta conjugada (entendiendo que en el caso real la conjugación no tiene efecto). La columna j -ésima de la matriz es $L(\vec{e}_j)$ y $\langle L(\vec{e}_j), L(\vec{e}_k) \rangle = \langle \vec{e}_j, \vec{e}_k \rangle$ permite probar que una matriz es ortogonal o unitaria si y solo si sus columnas forman una base ortonormal.

Ejemplo. Girar la base canónica de \mathbb{R}^3 alrededor del eje OZ un ángulo α , da lugar a los vectores $(\cos \alpha, -\sin \alpha, 0)^t$, $(\sin \alpha, \cos \alpha, 0)^t$ y $(0, 0, 1)^t$. Es fácil ver que conforman una matriz A que cumple $AA^t = I$. Esta es la matriz de la aplicación lineal correspondiente al giro.

Si trabajamos con una base que no es ortonormal las caracterizaciones matriciales (15) dejan de ser ciertas.

Ejemplo. Aunque $\langle \vec{x}, \vec{y} \rangle = x_1y_1 - x_2y_2$ no define estrictamente un producto escalar en \mathbb{R}^2 porque no es definido positivo, en relatividad especial es conveniente trabajar con él como si lo fuera. La base canónica usual no es ortonormal y ni siquiera podemos definir la norma del segundo vector. Para cualquier $v \in (-1, 1)$ las *transformaciones de Lorentz* en unidades relativistas $(x_1, x_2)^t \mapsto (1 - v^2)^{-1/2}(x_1 - vx_2, -vx_1 + x_2)^t$ se comportan como una aplicación ortogonal, es decir, satisfacen (14). Para ello basta comprobar la identidad matricial

$$(16) \quad (1 - v^2)^{-1} \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

que se reduce a un cálculo sencillo.

Observación. Las aplicaciones unitarias preservan las normas al cuadrado las cuales en física cuántica corresponden a probabilidades. El hecho de que la probabilidad total debe ser 1 se manifiesta en que los *operadores de evolución* que rigen cómo cambian con el tiempo los sistemas aislados, son aplicaciones unitarias.

Ejemplo. El operador de evolución de la precesión del espín de un electrón alrededor de $(1, 0, 1)^t$ con frecuencia angular 2ω tiene como matriz $U = U(t)$

$$(17) \quad U = \begin{pmatrix} \cos(\omega t) + \frac{i}{\sqrt{2}} \operatorname{sen}(\omega t) & \frac{i}{\sqrt{2}} \operatorname{sen}(\omega t) \\ \frac{i}{\sqrt{2}} \operatorname{sen}(\omega t) & \cos(\omega t) - \frac{i}{\sqrt{2}} \operatorname{sen}(\omega t) \end{pmatrix}.$$

Es una matriz unitaria para cualquier t porque $U^t \bar{U} = I$ o alternativamente porque las columnas son vectores ortogonales de norma 1 en \mathbb{C}^2 .

Como curiosidad, la interpretación física es que $|u_{11}|^2$ es la probabilidad de que un electrón inicialmente con espín arriba sea detectado en un tiempo t con espín arriba. El resto de los elementos de U corresponden a otras combinaciones del espín inicial y el detectado.

Un concepto que tiene una motivación menos intuitiva pero todavía una importancia capital es la de *aplicación autoadjunta*. Nos limitamos de nuevo al caso de dimensión finita (por sutilezas aludidas en una observación posterior). En este caso lo que se pide es que

$$(18) \quad \langle L(\vec{x}), \vec{y} \rangle = \langle \vec{x}, L(\vec{y}) \rangle \quad \text{para todo } \vec{x}, \vec{y} \in V.$$

Una particularidad importante de una aplicación autoadjunta L es que si actúa sobre uno de los sumandos de $W \oplus W^\perp$, entonces también actúa sobre el otro. Es decir, si $L(W) \subset W$ entonces $L(W^\perp) \subset W^\perp$. La prueba se reduce a jugar un poco con (18) y la extenderemos más adelante para conseguir una descomposición muy simple de una clase de aplicaciones que incluyen a las autoadjuntas y a las unitarias. Otra propiedad interesante es que para $\mathbb{K} = \mathbb{R}$ la *forma cuadrática* (objeto de otro tema) $\langle L(\vec{x}), \vec{x} \rangle$ determina la aplicación L si sabemos que es autoadjunta. En relación con esto, incluso fuera del caso autoadjunto, hay una inesperada rigidez del caso complejo que tiene alguna relevancia en física. Así $\langle L(\vec{x}), \vec{x} \rangle = 0$ para todo $\vec{x} \in V$ no puede darse para ninguna L no idénticamente nula si $\mathbb{K} = \mathbb{C}$. Sin embargo, hay contraejemplos para $\mathbb{K} = \mathbb{R}$ cuando L no es autoadjunta, por ejemplo un giro de 90° en \mathbb{R}^2 .

Un argumento similar al llevado a cabo con las aplicaciones ortogonales y unitarias implica que L es autoadjunta en un espacio euclídeo o unitario si y solo si su matriz en una base ortonormal satisface, respectivamente,

$$(19) \quad L = L^t \quad \text{o} \quad L = \overline{L}^t.$$

Lo primero son las *matrices simétricas* de toda la vida y lo segundo lo que habíamos llamado *matrices hermíticas* en relación con las formas bilineales.

Observación. En el caso complejo a veces, especialmente en el ámbito físico, se usa la palabra *hermítico* como sustituto de autoadjunto. Se habla por ejemplo de *aplicaciones hermíticas* u *operadores hermíticos*. En realidad en dimensión infinita se hacen sutiles diferencias en matemáticas entre operadores simétricos, hermíticos y autoadjuntos. Los tres conceptos coinciden en dimensión finita y aunque este no es el caso que habitualmente involucra la mecánica cuántica, en la práctica es raro que los físicos estén interesados por estas diferencias (hay una anécdota entre Heisenberg y Friedrichs que apoya esta afirmación).

Como cabe suponer, la terminología “autoadjunta” deriva de que hay un concepto de “adjunta”. La *aplicación adjunta* de un endomorfismo $L : V \rightarrow V$ con V euclídeo o unitario es otro endomorfismo $L^* : V \rightarrow V$ tal que

$$(20) \quad \langle L(\vec{x}), \vec{y} \rangle = \langle \vec{x}, L^*(\vec{y}) \rangle \quad \text{para todo } \vec{x}, \vec{y} \in V.$$

De nuevo nos centramos en el caso de dimensión finita para evitar ejemplos patológicos. Está claro que $L^* = L$ en (20) es lo mismo que (19), de ahí el nombre. Procediendo como en (19) se tiene que en una base ortonormal la matriz de L^* es simplemente \overline{L}^t , entendiendo que la conjugación no tiene ningún efecto en el caso real.

En sintonía con una notación mencionada en otra ocasión, en física es común emplear L^\dagger en lugar de L^* .

Ejemplo. La adjunta del giro de ángulo α alrededor del eje OZ considerado en un ejemplo anterior es el giro de ángulo $-\alpha$ porque trasponer es lo mismo que invertir, al ser ortogonal.

La definición de aplicación adjunta adquiere más sentido al introducir lo que se llama el *espacio dual*. No está en el temario aunque es posible que aparezca en algún tema de física teórica dentro del grado. En breve es el espacio vectorial compuesto por las llamadas *formas lineales*, las aplicaciones lineales $\phi : V \longrightarrow \mathbb{K}$. Tal espacio existe sin necesidad de que V tenga un producto escalar pero si lo tiene, se prueba que en dimensión finita y en algunos casos destacados de dimensión infinita se cumple que cada $\phi \in V^*$ es de la forma $\phi(\vec{x}) = \langle \vec{x}, \vec{y}_0 \rangle$ para cierto $\vec{y}_0 \in V$. La adjunta nos dice cómo pasan las aplicaciones lineales entre espacios vectoriales (incluso sin ser endomorfismos) a aplicaciones lineales entre los espacios duales.

La intuición que sirve de ayuda para imaginar el espacio dual en dimensión finita es pensar en él como en un espacio de vectores fila, así cuando multiplicamos uno de sus elementos por un vector, nos da un número.