4.3. La proyección ortogonal

En vez de plantearse si dos vectores son ortogonales, cabe preguntarse si dos subespacios vectoriales lo son. En algunas aplicaciones la ortogonalidad de dos subespacios indica que no tienen absolutamente nada que ver[6]. En relación con esto, en un espacio vectorial V con producto escalar, dado un subespacio $W \subset V$ se define el *complemento ortogonal* de W como

$$W^\perp = \{\vec{v} \in V \ : \ \langle \vec{v}, \vec{w} \rangle = 0 \quad \text{para todo } \vec{w} \in W\}. \tag{4.3}$$

Es decir, $W^\perp$ está constituido por todos los vectores que son ortogonales a los vectores de W. Es muy fácil comprobar que $W^\perp$ es un subespacio vectorial de V. Aunque la notación no lo refleja, naturalmente $W^\perp$ depende del espacio V en el que consideremos que W está incluido. Dada una base $\{\vec{w}_1, \dots, \vec{w}_k\}$ de W, las ecuaciones $\langle \vec{x}, \vec{w}_j \rangle = 0$ definen $W^\perp$ y resolviendo el sistema para $\vec{x} \in V$ obtenemos una base suya.

Un subespacio y su complemento ortogonal determinan una descomposición del espacio total, lo que resulta clave para introducir el concepto que da nombre a la sección.

Proposición 4.3.1. *Sea V un espacio vectorial de dimensión finita con producto escalar y sea W un subespacio suyo. Entonces* $\dim W^\perp = \dim V - \dim W$ *y para cada $\vec{v} \in V$ existe un único $\vec{w} \in W$ y un único $\vec{u} \in W^\perp$ tales que $\vec{v} = \vec{w} + \vec{u}$.*

La situación descrita en la proposición se resume en la jerga del álgebra lineal diciendo que V es la *suma directa*[7] de W y $W^\perp$ y se escribe $V = W \oplus W^\perp$.

Al vector $\vec{w}$ asociado a $\vec{v}$ se le llama *proyección ortogonal* de $\vec{v}$ sobre W. En otros términos, la proyección ortogonal sobre W es la función $P_W : V \longrightarrow W$ tal que $P_W(\vec{v}) = \vec{w}$. No es difícil ver que P_W es una aplicación lineal. En términos ópticos y geométricos, la proyección ortogonal es la sombra de un vector sobre una recta o un plano o sus generalizaciones n-dimensionales cuando los rayos de luz son perpendiculares a ellos. Hay otras proyecciones que sirven para representar sombras oblicuas pero resultan matemáticamente menos importantes y no forman parte de este curso.

La demostración de la Proposición 4.3.1 da un método para calcular la proyección ortogonal aunque también es conveniente tener en cuenta la primera de las siguientes propiedades básicas porque, a veces, simplifica dicho cálculo.

Proposición 4.3.2. *Con la notación anterior se cumple*

$$1)\ \vec{v} = P_W(\vec{v}) + P_{W^\perp}(\vec{v}) \qquad y \qquad 2)\ \|\vec{v} - P_W(\vec{v})\| \le \|\vec{v} - \vec{w}\| \quad \textit{para todo } \vec{w} \in W.$$

[6] Por ejemplo, en física cuántica significa que la probabilidad de transición es nula entre los estados correspondientes a cada uno de ellos.

[7] En general se dice que V es suma directa de dos subespacios W_1 y W_2, y se escribe $V = W_1 \oplus W_2$, cuando $V = W_1 + W_2$ (cada vector de V es suma de vectores de los subespacios) y $W_1 \cap W_2 = \{\vec{0}\}$.

Demostración de la Proposición 4.3.1. Si tenemos una base $\{\vec{u}_1, \dots, \vec{u}_m\}$ de W siempre la podemos completar a una base $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_m, \vec{u}_{m+1}, \dots, \vec{u}_n\}$ de V, simplemente añadiendo sucesivamente vectores linealmente independientes hasta que esto no sea posible. Aplicando el proceso de Gram-Schmidt, que deja los primeros vectores invariantes, podemos suponer que $\mathcal{B}$ es ortonormal. Automáticamente $\vec{u}_j \in W^\perp$ para $j > m$ y por ser una base ortonormal, recordando la Proposición 4.2.1, se tiene

$$\vec{v} = \big(\langle \vec{v}, \vec{u}_1\rangle \vec{u}_1 + \dots + \langle \vec{v}, \vec{u}_m\rangle \vec{u}_m\big) + \big(\langle \vec{v}, \vec{u}_{m+1}\rangle \vec{u}_{m+1} + \dots + \langle \vec{v}, \vec{u}_n\rangle \vec{u}_n\big).$$

El primer paréntesis es $\vec{w}$ y el segundo $\vec{u}$. Con nuestra notación $n = \dim V$, $m = \dim W$ y, considerando el caso $\vec{v} \in W^\perp$, deducimos que todos los vectores de $W^\perp$ son combinación lineal de $\vec{u}_{m+1}, \dots, \vec{u}_n$, que son $n - m$ vectores linealmente independientes, de donde $\dim W^\perp = \dim V - \dim W$.

La unicidad de $\vec{w}$ y $\vec{u}$ se sigue porque si $\vec{w}_1 + \vec{u}_1 = \vec{w}_2 + \vec{u}_2$ entonces $\vec{w}_1 - \vec{w}_2 = \vec{u}_2 - \vec{u}_1$ por tanto $\vec{w}_1 - \vec{w}_2$ estaría a la vez en W y en $W^\perp$ y el único vector ortogonal a sí mismo es el $\vec{0}$ porque el producto escalar es definido positivo. □

Demostración de la Proposición 4.3.2. Por la simetría del producto escalar, los vectores de W están en el ortogonal de $W^\perp$. Eso significa que en la igualdad $\vec{v} = \vec{u} + \vec{w}$ se tiene $\vec{u} \in W^\perp$ y $\vec{w} \in (W^\perp)^\perp$, esto es, $P_{W^\perp}(\vec{v}) = \vec{u}$ y se sigue la primera propiedad.

Para la segunda, nótese que $\vec{x} = \vec{v} - P_W(\vec{v}) \in W^\perp$ e $\vec{y} = P_W(\vec{v}) - \vec{w} \in W$. El teorema de Pitágoras (Proposición 4.1.3) asegura $\|\vec{v} - \vec{w}\|^2 = \|\vec{x} + \vec{y}\|^2 \ge \|\vec{x}\|^2$. □

La primera propiedad es útil para los cálculos cuando es más fácil calcular la proyección ortogonal sobre $W^\perp$ que sobre W.

La segunda propiedad es muy importante y útil en diversas aplicaciones porque dice que la proyección ortogonal es la mejor aproximación si no queremos salirnos de un subespacio. Por ejemplo, en cálculo numérico un problema natural es aproximar funciones relativamente complicadas por elementos de un subespacio de polinomios y, una vez definido un producto escalar adecuado, la proyección ortogonal sirve para obtener buenas aproximaciones de manera eficiente. Algo similar aparece cuando uno quiere sintetizar lo más fielmente posible una señal a partir de cierto subespacio de señales posibles que es capaz de generar el *hardware* y, con este fin, se proyecta ortogonalmente sobre dicho subespacio. En cierta manera, la compresión que se produce en los ficheros JPEG se debe también a que la mayor parte de las imágenes admiten una aproximación muy buena por la proyección ortogonal en un espacio de dimensión comparativamente pequeña. Genéricamente se llama método de mínimos cuadrados a la aproximación de datos por medio de aplicaciones lineales utilizando la proyección ortogonal y la famosa *recta de regresión* es solo el ejemplo más sencillo de ello.

La *distancia* entre dos subconjuntos de $\mathbb{R}^n$ (o de $\mathbb{C}^n$) se define como la mínima distancia entre los puntos que las integran. Identificando las flechas que representan los vectores con los puntos de sus extremos, lo que está diciendo la segunda propiedad es que la distancia del punto que corresponde a $\vec{v}$ a los puntos del subespacio W es justamente $\|\vec{v} - P_W(\vec{v})\| = \|P_{W^\perp}(\vec{v})\|$. De esta forma, se generalizan a dimensiones superiores las fórmulas de distancia de punto a recta y punto a plano estudiadas en cursos anteriores.

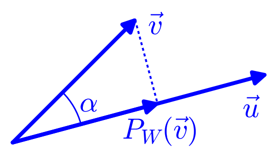
Si leemos la demostración de la Proposición 4.3.1 veremos que calcular $P_W(\vec{v})$ cuando tenemos una base ortonormal $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_n\}$ es tan sencillo como calcular productos escalares. Si $\mathcal{B}$ es ortogonal, el análogo cambiando $\vec{u}_j$ por $\vec{u}_j/\|\vec{u}_j\|$ es

$$P_W(\vec{v}) = \frac{\langle \vec{v}, \vec{u}_1\rangle}{\|\vec{u}_1\|^2}\vec{u}_1 + \frac{\langle \vec{v}, \vec{u}_2\rangle}{\|\vec{u}_2\|^2}\vec{u}_2 + \dots + \frac{\langle \vec{v}, \vec{u}_n\rangle}{\|\vec{u}_n\|^2}\vec{u}_n. \tag{4.4}$$

Esto no es otra cosa que hallar las coordenadas de $\vec{w}$ en $\mathcal{B}$ usando la Proposición 4.2.1, lo que se reducía a tomar productos escalares con $\vec{u}_j$. Por tanto es una fórmula que se puede improvisar y además explica por qué el proceso de Gram-Schmidt (4.2) funciona: en cada paso estamos calculando $\vec{u}_j = \vec{v}_j - P_{W_j}(\vec{v}_j)$ donde $W_j = \mathcal{L}(\{\vec{u}_1, \dots, \vec{u}_{j-1}\})$ y, como $\vec{u}_j \in W_j^\perp$, necesariamente es ortogonal a $\vec{u}_1, \dots, \vec{u}_{j-1}$. En definitiva, no hace falta memorizar la fórmula del proceso de Gram-Schmidt si entendemos que lo único que se está haciendo es quitar a un vector la proyección ortogonal sobre los anteriores.

Un caso particular es la proyección ortogonal sobre el espacio unidimensional $W = \mathcal{L}(\{\vec{u}\})$. Al tener un solo vector no hay que ortogonalizar nada. La fórmula resultante está de acuerdo con lo que se sigue en el plano usando simple trigonometría recordando (4.1).

$$P_W(\vec{v}) = \frac{\langle \vec{v}, \vec{u}\rangle}{\|\vec{u}\|^2}\vec{u}$$



$$P_W(\vec{v}) = \|\vec{v}\|\cos\alpha \frac{\vec{u}}{\|\vec{u}\|}.$$

Para ilustrar el uso de la primera propiedad de la Proposición 4.3.2 consideremos el problema de calcular la proyección ortogonal de $\vec{v} = (4,3,-2)^t \in \mathbb{R}^3$ sobre

$$W = \{\vec{x} \in \mathbb{R}^3 \ : \ 3x - y - z = 0\}.$$

La manera "larga" de hacerlo sería buscar una base de W, que tendría dos vectores, y seguir el procedimiento de la prueba de la Proposición 4.3.1 o (4.4). Al ser $W^\perp$ un espacio unidimensional porque $\dim W^\perp = \dim V - \dim W = 3 - 2$, es más ventajoso calcular $P_{W^\perp}(\vec{v})$. Simplemente hay que fijarse en que $W = \{\vec{x} \in \mathbb{R}^3 \ : \ \vec{x}\cdot\vec{u} = 0\}$ con $\vec{u} = (3,-1,-1)^t$. Así pues $\{\vec{u}\}$ es una base de $W^\perp$ y se tiene

$$P_{W^\perp}(\vec{v}) = \frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|}\vec{u} = (3,-1,-1)^t \quad \Rightarrow \quad P_W(\vec{v}) = \vec{v} - (3,-1,-1)^t = (1,4,-1)^t$$

donde se ha usado la primera propiedad de la Proposición 4.3.2.

Recordando lo dicho antes respecto a la relación entre distancia y proyección ortogonal, la distancia del punto $Q = (4,3,-2)$ al plano determinado por W sería:

$$d(Q,W) = \|\vec{v} - P_W(\vec{v})\| = \|P_{W^\perp}(\vec{v})\| = \sqrt{3^2 + (-1)^2 + (-1)^2} = \sqrt{11}.$$

Si recuerdas la fórmula de la distancia de un punto a un plano, podrás corroborar este resultado.

Solo para practicar, vamos a ver rápidamente, sin explicitar todas las cuentas, los pasos involucrados en la manera "larga" de calcular $P_W(\vec{v})$ en este ejemplo, que consiste en no valerse de $P_{W^\perp}(\vec{v})$. En primer lugar, obtendríamos una base de W, para ello lo natural es tomar $y = \lambda$, $z = \mu$ y al sacar común, multiplicando los vectores por 3 para que no nos salgan fracciones, llegamos a la base $\mathcal{B} = \{(1,3,0)^t, (1,0,3)^t\}$. El proceso de Gram-Schmidt modifica el segundo vector para que sea ortogonal al primero, si además multiplicamos por 10 para eliminar denominadores se obtiene la base ortogonal $\mathcal{B}' = \{(1,3,0)^t, (3,-1,10)^t\}$. Finalmente, solo resta aplicar (4.4) para concluir de nuevo $P_W(\vec{v}) = (1,4,-1)^t$.

Consideremos $V = \mathbb{R}^4$ y un subespacio $W \subset V$ dado por

$$W = \{\vec{x} \in \mathbb{R}^4 \,:\, x_1 - 2x_2 + x_3 = x_1 - 3x_2 + x_3 + x_4 = 0\}.$$

Queremos calcular la proyección ortogonal de $\vec{v} = (4,8,-4,12)^t$ sobre W.

Si deseamos utilizar (4.4), lo que debemos hacer es hallar una base ortogonal de W como en el esquema del ejemplo anterior. En primer lugar, calculamos una base de W resolviendo el sistema por eliminación de Gauss:

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -3 & 1 & 1 \end{pmatrix} \underset{f_2 \mapsto f_2 - f_1}{\longrightarrow} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$$

Con $x_3 = \lambda$, $x_4 = \mu$ se sigue $x_2 = \mu$ y $x_1 = 2\mu - \lambda$, por tanto

$$\vec{x} \in W \quad \Leftrightarrow \quad \vec{x} = \lambda\vec{v}_1 + \mu\vec{v}_2 \quad \text{donde } \vec{v}_1 = (-1,0,1,0)^t,\ \vec{v}_2 = (2,1,0,1)^t.$$

Se tiene entonces que $\{\vec{v}_1, \vec{v}_2\}$ es base de W. El proceso de Gram-Schmidt afirma que reemplazando $\vec{v}_2$ por $\vec{v}_2 - \|\vec{v}_1\|^{-2}(\vec{v}_2 \cdot \vec{v}_1)\vec{v}_1$, se convierte en una base ortogonal, concretamente en

$$\{\vec{u}_1 = (-1,0,1,0)^t,\ \vec{u}_2 = (1,1,1,1)^t\}.$$

Aplicando (4.4), la proyección es

$$P_W(\vec{v}) = \frac{\vec{v} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 + \frac{\vec{v} \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 = (9,5,1,5)^t.$$

Una alternativa que no requiere hallar una base ortogonal es usar que $\vec{v} - P_W(\vec{v})$ debe pertenecer a $W^\perp$, por definición, por tanto su producto escalar con $\vec{v}_1$ y $\vec{v}_2$ es nulo. Escribiendo $P_W(\vec{v}) = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2$, lo cual es lícito porque $P_W(\vec{v}) \in W$, llegamos a un sistema de dos ecuaciones con dos incógnitas:

$$\begin{cases} (\vec{v} - \lambda_1\vec{v}_1 - \lambda_2\vec{v}_2) \cdot \vec{v}_1 = 0, \\ (\vec{v} - \lambda_1\vec{v}_1 - \lambda_2\vec{v}_2) \cdot \vec{v}_2 = 0, \end{cases} \quad \Rightarrow \quad \begin{cases} -8 - 2\lambda_1 + 2\lambda_2 = 0, \\ 28 + 2\lambda_1 - 6\lambda_2 = 0. \end{cases}$$

La solución $\lambda_1 = 1$, $\lambda_2 = 5$ implica $P_W(\vec{v}) = \vec{v}_1 + 5\vec{v}_2 = (9,5,1,5)^t$ que coincide con el resultado anterior.

Supongamos que el ejemplo anterior nos pidieran una base de $W^\perp$. Esto es inmediato si reescribimos la definición de W como

$$W = \{\vec{x} \in \mathbb{R}^4 \,:\, \vec{u}_1 \cdot \vec{x} = \vec{u}_2 \cdot \vec{x} = 0\} \qquad \text{con} \quad \vec{u}_1 = (1, -2, 1, 0)^t,\ \vec{u}_2 = (1, -3, 1, 1)^t.$$

Vista así, nos dice que los vectores de W son justamente los que tienen a $\vec{u}_1$ y a $\vec{u}_2$ como ortogonales, por tanto estos generan $W^\perp$ y conforman una base suya ya que son linealmente independientes.

Veamos ahora un ejemplo más abstracto en nuestro espacio favorito de polinomios $V = \mathbb{R}_2[x]$ en el que definiremos el producto escalar extraño dado por

$$\langle P, Q\rangle = P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1).$$

Siguiendo los requerimientos de un examen anterior, vamos a calcular la proyección ortogonal de x^2 sobre el subespacio

$$W = \mathcal{L}\big(\{P_1,\, P_2\}\big) \qquad \text{con} \quad P_1 = (x+1)^2, \quad P_2 = x^2 + 1.$$

Si $Q = \lambda_1 P_1 + \lambda_2 P_2 = P_W(x^2)$ es la proyección ortogonal buscada, entonces $x^2 - Q \in W^\perp$ o lo que es lo mismo $\langle P_1, x^2 - Q\rangle = 0$ y $\langle P_2, x^2 - Q\rangle = 0$. Unos cálculos sencillos muestran $\langle P_1, x^2\rangle = \langle P_2, x^2\rangle = 4$ con lo que las ecuaciones anteriores son:

$$\begin{cases} 4 - \lambda_1\langle P_1, P_1\rangle - \lambda_2\langle P_1, P_2\rangle = 0, \\ 4 - \lambda_1\langle P_2, P_1\rangle - \lambda_2\langle P_2, P_2\rangle = 0, \end{cases} \quad \Rightarrow \quad \begin{cases} 4 = 17\lambda_1 + 9\lambda_2, \\ 4 = 9\lambda_1 + 9\lambda_2 \end{cases}$$

donde se ha sustituido $\langle P_1, P_1\rangle = 17$ y $\langle P_1, P_2\rangle = \langle P_2, P_2\rangle = 9$. Resolviendo el sistema, $\lambda_1 = 0$ y $\lambda_2 = 4/9$ por tanto la proyección ortogonal es $Q = 4(x^2+1)/9$. Este sería el polinomio que mejor aproxima a x^2 dentro de nuestro subespacio cuando usamos las distancias que derivan de nuestro producto escalar.

En la línea de lo visto anteriormente, otra forma de llegar a la solución es ortogonalizar $\{P_1, P_2\}$. Aplicando el proceso de Gram-Schmidt y quitando denominadores, se obtiene la base ortogonal $\{Q_1, Q_2\}$ con $Q_1 = (x+1)^2$ y $Q_2 = 4x^2 - 9x + 4$. Las cuentas, tras tomar $Q_1 = P_1$, son

$$P_2 - \frac{\langle P_2, Q_1\rangle}{\|Q_1\|^2} Q_1 = x^2 + 1 - \frac{2\cdot 0 + 1\cdot 1 + 4\cdot 2}{0^2 + 1^2 + 4^2}(x+1)^2 = \frac{8x^2 - 18x + 8}{17}$$

y multiplicando por $17/2$ para simplificar, se obtiene Q_2. Con ello la proyección ortogonal es:

$$P_W(x^2) = \frac{\langle x^2, Q_1\rangle}{\|Q_1\|^2} Q_1 + \frac{\langle x^2, Q_2\rangle}{\|Q_2\|^2} Q_2 = \frac{4Q_1}{0^2 + 1^2 + 4^2} + \frac{16Q_2}{17^2 + 4^2 + (-1)^2} = \frac{4}{9}(x^2+1).$$

Para no olvidarnos de los números complejos, calculemos la proyección ortogonal de $\vec{u} = (2+i, 1+7i)^t$ sobre la “recta compleja” dada por el subespacio

$$W = \{\vec{x} \in \mathbb{C}^2 \,:\, (1+i)x_1 - 3ix_2 = 0\}.$$

Tomando $x_2 = \lambda$ se deduce $x_1 = 3(1+i)\lambda/2$ por tanto $W = \mathcal{L}(\{\vec{u}\})$ con $\vec{u} = \big(\frac{3}{2}+\frac{3}{2}i, 1\big)^t$. Basta entonces usar (4.4) para $n = 1$. Los cálculos con mucho detalle son:

$$P_W(\vec{v}) = \frac{(2+i)(\frac{3}{2}-\frac{3}{2}i)+(1+7i)\cdot 1}{|\frac{3}{2}+\frac{3}{2}i|^2+1^2}\begin{pmatrix}\frac{3}{2}+\frac{3}{2}i\\ 1\end{pmatrix} = (1+i)\begin{pmatrix}\frac{3}{2}+\frac{3}{2}i\\ 1\end{pmatrix} = \begin{pmatrix}3i\\ 1+i\end{pmatrix}.$$

Ligeramente más sencillo desde el punto de vista computacional hubiera sido tomar $x_2 = 2\lambda$ para eliminar los denominadores. Otra posibilidad es $x_2 = (1+i)\lambda$ para no dividir por números complejos.

Es posible automatizar el cálculo de proyecciones ortogonales con una fórmula aunque esto no es tan conveniente en la práctica si no somos una máquina, porque las cuentas se vuelven opacas y la fórmula no conlleva una reducción en los cálculos cuando se llevan a cabo de la forma habitual. Inspeccionando la demostración, está claro que se reduce a lo que ya habíamos hecho en algunos de los ejemplos salvo que resolvemos un sistema con la matriz inversa.

Proposición 4.3.3. *Dado un subespacio $W \subset K^n$ con $K = \mathbb{R}$ o $\mathbb{C}$, sea S una matriz cuyas columnas forman una base de W. Entonces*

$$P_W(\vec{v}) = S(S^\dagger S)^{-1}S^\dagger\vec{v} \qquad \textit{para todo} \quad \vec{v}\in K^n.$$

Si alguien está tentado a escribir $(S^\dagger S)^{-1} = S^{-1}(S^\dagger)^{-1}$ y simplificar, que se contenga porque S genéricamente no es cuadrada y no tiene sentido invertirla. Solo es cuadrada cuando $W = K^n$ en cuyo caso es evidente de partida que $P_W(\vec{v}) = \vec{v}$.

Demostración. Sea $\{\vec{v}_1, \dots, \vec{v}_n\}$ la base de W representada por las columnas de S. La proyección ortogonal es de la forma $\sum_j \lambda_j\vec{v}_j$ para ciertos λ_j que es nuestro objetivo calcular. En forma matricial tenemos

$$P_W(\vec{v}) = S\vec{\lambda} \qquad \text{con} \quad \vec{\lambda} = (\lambda_1, \dots, \lambda_n)^t.$$

Lo que caracteriza los λ_j es que $\vec{v} - P_W(\vec{v})$ debe ser ortogonal a todos los $\vec{v}_i$, es decir,

$$0 = \vec{v}\cdot\vec{v}_i - \sum_j \lambda_j\vec{v}_j\cdot\vec{v}_i = \vec{v}_i^\dagger\vec{v} - \sum_j \vec{v}_i^\dagger\vec{v}_j\lambda_j$$

donde en la última expresión hay que entender que los vectores se consideran como matrices columna. Por la forma en que se multiplican las matrices, $\vec{v}_i^\dagger\vec{v}$ es la i-ésima coordenada del vector $S^\dagger\vec{v}$ y $\vec{v}_i^\dagger\vec{v}_j$ es el elemento ij de $S^\dagger S$. Se concluye $S^\dagger\vec{v} = S^\dagger S\vec{\lambda}$ y de aquí $\vec{\lambda} = (S^\dagger S)^{-1}S^\dagger\vec{v}$. Necesariamente $S^\dagger S \in \mathcal{M}_n(K)$ es invertible porque si no tuviera rango máximo $\vec{v} = \vec{0}$, $\vec{\lambda} \neq \vec{0}$ sería una posibilidad absurda. □

Una ventaja de la Proposición 4.3.3 es que produce explícitamente la matriz de la proyección ortogonal extendida a un endomorfismo, lo que puede tener cierto interés si vamos a proyectar muchos vectores. Con la extensión nos referimos a que en principio, en la situación de la proposición, $P_W : K^n \longrightarrow W$ pero como $W \subset K^n$ en realidad el resultado de la proyección lo podemos entender como un elemento de K^n. Si uno quisiera ser muy riguroso, escribiría $i \circ P_W$ donde $i : W \longrightarrow K^n$ es la inclusión.

Veamos un ejemplo que podamos comprobar con cálculos anteriores.

Recordemos que habíamos hallado $P_W(\vec{v})$ para cierto $\vec{v}$ y un subespacio $W \subset \mathbb{R}^4$ con una base compuesta por $\vec{v}_1 = (-1, 0, 1, 0)^t$ y $\vec{v}_2 = (2, 1, 0, 1)^t$. Nos preguntamos cuál es la matriz de P_W vista como aplicación lineal $\mathbb{R}^4 \longrightarrow \mathbb{R}^4$. A partir de las coordenadas de los vectores de la base,

$$S = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad S^\dagger S = \begin{pmatrix} 2 & -2 \\ -2 & 6 \end{pmatrix} \quad \Rightarrow \quad (S^\dagger S)^{-1} = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$$

Por tanto la matriz del endomorfismo que da la proyección ortogonal es:

$$\frac{1}{4}\begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 3 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Si la aplicamos a $\vec{v} = (4, 8, -4, 12)^t$, obtendremos $P_W(\vec{v}) = (9, 5, 1, 5)^t$, como antes.

Exprimiendo el silicio [opcional]. El siguiente código implementa en `sagemath` la función `proj_Wv` que calcula la proyección ortogonal sobre un subespacio a partir de una base suya utilizando la fórmula anterior. Su uso está ilustrado con dos de los ejemplos anteriores.

```
def proj_Wv(S,v):
        G = S.conjugate_transpose()*S
        M = S*G.inverse()*S.conjugate_transpose()
        return M*v

# S = Base de W en columna
# v = vector a proyectar

S = matrix(4,2,[-1,2,0,1,1,0,0,1])
v = vector([4,8,-4,12])
print(proj_Wv(S,v))
print('----------------')

S = matrix(2,1,[3/2+3/2*I, 1])
v = vector([2+I, 1+7*I])
print(proj_Wv(S,v))
```

La salida es `(9, 5, 1, 5)` y `(3*I, I + 1)`, en consonancia con los resultados obtenidos antes.

Puestos a dar solución automática a los ejemplos, la siguiente variante `proj_Wv2` (que llama a `proj_Wv`) permite introducir las ecuaciones del subespacio en vez de una base suya.

```
def proj_Wv2(A, v):
        W = A.right_kernel()
```

```
        S = matrix([item for item in W.basis()]).transpose()
        return proj_Wv(S,v)

# A = matrix con las ecuaciones de W
# v = vector a proyectar

A = matrix(QQ,2,4,[1,-2,1,0, 1,-3,1,1])
v = vector([4,8,-4,12])
print(proj_Wv2(A,v))
print('----------------')

A = matrix(1,2,[1+I, -3*I])
v = vector([2+I, 1+7*I])
print(proj_Wv2(A,v))
print('----------------')
```

4.4. Matrices ortogonales y unitarias

Una pregunta recurrente en matemáticas es el tipo de funciones que preservan ciertas estructuras. En este sentido, a pesar de que no ha sido exactamente nuestro enfoque, las aplicaciones lineales son las funciones que preservan la estructura de espacio vectorial.

En esta línea e incluso dentro de la geometría básica, cabe preguntarse qué funciones dejan invariantes las distancias en el plano o el espacio, lo que podría llamarse preservar la estructura métrica. Claramente los giros lo hacen: la distancia de Barcelona a Madrid no cambia a lo largo del día con la rotación de la Tierra. Un prurito matemático lleva a plantearse un enunciado que determine cuáles son exactamente. Aquí restringiremos nuestro análisis a las aplicaciones lineales; a cambio, no nos quedaremos en el plano o el espacio sino que estudiaremos la situación en $\mathbb{R}^n$ y en $\mathbb{C}^n$ para cualquier dimensión.

Proposición 4.4.1. *Sea $f : K^n \longrightarrow K^n$ con $K = \mathbb{R}$ o $\mathbb{C}$ un endomorfismo que preserva las longitudes, es decir, tal que $\|f(\vec{x})\| = \|\vec{x}\|$ para todo $\vec{x} \in K^n$, entonces f también preserva el producto escalar, es decir $f(\vec{x}) \cdot f(\vec{y}) = \vec{x} \cdot \vec{y}$ para todo $\vec{x} \in K^n$.*

Demostración. Por hipótesis y por la linealidad de f

$$\|\vec{x} + \vec{y}\|^2 = \|f(\vec{x} + \vec{y})\|^2 = \|f(\vec{x}) + f(\vec{y})\|^2.$$

Utilizando que $\|\vec{z}\|^2 = \vec{z} \cdot \vec{z}$, al desarrollar se sigue

$$\|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\vec{x} \cdot \vec{y} = \|f(\vec{x})\|^2 + \|f(\vec{y})\|^2 + 2f(\vec{x}) \cdot f(\vec{y}).$$

Los dos primeros sumandos de ambos miembros se cancelan porque f preserva las distancias y se obtiene el resultado buscado. □

En términos geométricos, la interpretación de este resultado a la luz de (4.1) es que preservar longitudes implica preservar ángulos, algo que no es tan intuitivo.

Ya sabíamos (Proposición 2.3.1) que en K^n todas las aplicaciones lineales venían dadas por multiplicar por una matriz. En nuestro caso, tratamos con endomorfismos y esta matriz es necesariamente cuadrada. El siguiente resultado resuelve totalmente el problema que nos habíamos planteado.