4. Espacios con producto interno

Nuestro objetivo ahora es concentrarnos en los espacios vectoriales en los cuales tiene sentido pensar en una longitud de un vector, de distancia y ángulo entre vectores. Para ello introduciremos una función que a cada par de vectores les asigna un escalar. Esta función, llamada ***producto interno***, será una generalización del producto escalar que vieron en $\mathbb{R}^2$.

4.1. Producto interno

Definición 4.1 *Sea V un espacio vectorial sobre el cuerpo $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Un* ***producto interno*** *sobre V es una función que a cada par de vectores en V, (u, v) le asigna un escalar que se nota: $u \cdot v$, $u \times v$, $\langle u, v\rangle$ o $(u|v)$, de modo tal que para todos $uv, w \in V$ y $\alpha \in \mathbb{K}$ se verifica:*

1. $\langle u, v\rangle = \overline{\langle v, u\rangle}$.
2. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.
3. $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$.
4. $\langle u, u\rangle \geqslant 0$ *y* $\langle u, u\rangle = 0$ *si y solo si* $u = 0$.

Ejemplo 4.1

1. *Sea $V = \mathbb{R}^n$, un vector $v \in V$ lo pensamos como $v = (v_1, \cdots, v_n)^t$. Definiendo*

$$\langle u, v\rangle = u^t v = \sum_{i=1}^{n} u_i v_i.$$

Resulta $\langle u, v\rangle$ así definido un producto interno en V.
Sea $V = \mathbb{C}^n$. Definiendo

$$\langle u, v\rangle = u^t \overline{v} = \sum_{i=1}^{n} u_i \overline{v_i}.$$

donde $\overline{v}$ es el vector de $\mathbb{C}^n$ cuyas componentes son los conjugados de las componentes del vector v. Resulta $\langle u, v\rangle$ así definido un producto interno en V.
Este producto interno definido en $\mathbb{K}^n$ se conoce como ***producto interno canónico****.*

2. *En $V = \mathbb{R}^2$ podemos considerar para $u = (u_1, u_2)$ y $v = (v_1, v_2)$*

$$\langle u, v\rangle = u_1 v_1 - u_2 v_1 - u_1 v_2 + 4u_2 v_2.$$

Puede verificarse fácilmente que esto define un producto interno en V.

3. *Sea $V = \mathcal{M}_{n\times n}(\mathbb{K})$, para $A = (a_{ij})$ y $B = (b_{ij})$, definamos*

$$\langle A, B\rangle = \sum_{i,j} a_{ij}\overline{b_{ij}},$$

esto constituye un producto interno en V.
Si dada una matriz A consideramos su ***matriz adjunta*** *dada por su matriz transpuesta conjugada $B^* = \overline{B^t}$, es decir que $b^*_{ij} = \overline{b_{ji}}$, puede expresarse este producto interno en función de la traza de una matriz*

$$\langle A, B\rangle = tr(AB^*) = tr(B^* A).$$

4. *Sean $t_0, \cdots, t_n$ escalares distintos. Sea V el espacio vectorial de los polinomios sobre $\mathbb{K}$ de grado menor o igual a n. Para $p, q \in V$ definimos*

$$\langle p, q\rangle = p(t_0)\overline{q(t_0)} + \cdots + p(t_n)\overline{q(t_n)}.$$

Se comprueba fácilmente que los axiomas 1), 2) *y* 3) *se verifican. Para ver el axioma* 4) *observemos que*

$$\langle p, p\rangle = p(t_0)\overline{p(t_0)} + \cdots + p(t_n)\overline{p(t_n)} = |p(t_0)|^2 + \cdots + |p(t_n)|^2 \geqslant 0.$$

Además, también se da $\langle 0, 0\rangle = 0$. *Por último si* $\langle p, p\rangle = 0$ *entonces* p *es un polinomio que se anula en* $n+1$ *puntos distintos y esto sólo es posible si* p *es el polinomio nulo, pues* $gr(p) \leqslant n$. *Con esto concluimos que estamos en presencia de un producto interno.*

5. *Sea* V *el espacio vectorial de las funciones reales continuas en el intervalo* $[0, 1]$. *Para* $f, g \in V$ *sea*

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt,$$

esto define un producto interno. Los tres primeros axiomas son de inmediata verificación por las propiedades de la integración. Además

$$\langle f, f\rangle = \int_0^1 f^2(t)dt \geqslant 0.$$

Si $\langle f, f\rangle = \int_0^1 f^2(t)dt = 0$, *suponiendo que* $f \neq 0$ *y por el teorema de conservación del signo para una función continua se llegaría a que* $\langle f, f\rangle > 0$ *lo que es una contradicción.*

Definición 4.2 *Un* ***espacio producto interno*** *, es un espacio vectorial sobre* $\mathbb{K}$ *(*$\mathbb{R}$ *o* $\mathbb{C}$*) munido de un producto interno.*
Un espacio producto interno real de dimensión finita es frecuentemente llamado ***espacio euclidiano*** *.*
Un espacio producto interno complejo de dimensión finita es frecuentemente llamado ***espacio unitario*** *.*

Definición 4.3 *Sea* V *un espacio vectorial con producto interno* $\langle \cdot, \cdot\rangle$. *Se llama* ***norma*** *de un vector* v *a* $\|v\| = \sqrt{\langle v, v\rangle}$.
Resulta equivalente $\|v\|^2 = \langle v, v\rangle$.

Teorema 4.1 *Si* V *es un espacio producto interno, entonces para* $u, v \in V$ *y* $\alpha \in \mathbb{K}$ *cualesquiera se tiene*
i) $\|\alpha v\| = |\alpha|\|v\|$.
ii) $\|v\| > 0$ *para* $v \neq 0$.
iii) $|\langle u, v\rangle| \leqslant \|u\|\|v\|$. ***Desigualdad de Cauchy-Schwarz.***
iv) $\|u + v\| \leqslant \|u\| + \|v\|$. ***Desigualdad triangular*** .

Demos: Las afirmaciones *i)* y *ii)* surgen inmediatamente de las definiciones dadas.
La desigualdad *iii)* es claramente válida para $u = 0$. Supongamos entonces que $u \neq 0$ considerar

$$w = v - \frac{\langle v, u\rangle}{\|u\|^2}u.$$

Entonces $\langle w, u\rangle = 0$ y

$$0 \leqslant \|w\|^2 = \langle w, v - \frac{\langle u, u\rangle}{\|u\|^2}u\rangle = \langle w, v\rangle - \frac{\langle v, u\rangle\langle w, u\rangle}{\|u\|^2} = \langle w, v\rangle = \|v\|^2 - \frac{|\langle v, u\rangle|^2}{\|u\|^2}.$$

De aquí surge que $|\langle v, u\rangle| = |\langle u, v\rangle| \leqslant \|u\|\|v\|$.
Para probar *iv)*, veamos que

$$\begin{array}{rcr} \|u + v\|^2 & = & \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2 \\ & = & \|u\|^2 + 2\Re\langle u, v\rangle + \|v\|^2 \leqslant \\ & \leqslant & \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \leqslant \\ & \leqslant & \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = \\ & = & (\|u\| + \|v\|)^2. \end{array}$$

De aquí se tiene $\|u + v\| \leqslant \|u\| + \|v\|$. ■

Definición 4.4 *Sea V un espacio producto interno, luego se define la* **distancia** *entre dos vectores $u, v \in V$ como*

$$dist(u, v) = \|u - v\|.$$

4.2. Ortogonalidad

Definición 4.5 *Sean $u, v \in V$ espacio producto interno. Entonces diremos que* ***u y v son ortogonales*** *si $\langle u, v\rangle = 0$.*
Si un vector u es ortogonal a todo vector en un subespacio W de V se dice que u es ortogonal a W.
El conjunto de todos los vectores u ortogonales a W se llama ***complemento ortogonal de W*** *y se nota $W^{\perp}$.*

Nota 4.1 *El vector nulo es ortogonal a todo vector en V y es el único vector con tal propiedad.*

Ejemplo 4.2 *Sea $V = \mathbb{R}^3$ y sea W un plano que pasa por el origen y sea L la recta que pasa por el 0 y es perpendicular a W. Luego si z, w son vectores no nulos, tales que z está sobre L y $w \in W$ se tiene que $z \cdot w = 0$. Así cada vector en L es ortogonal a cada vector w en el plano. De hecho se tiene que*

$$L = W^{\perp}, \qquad\qquad W = L^{\perp}.$$

Proposición 4.1 *$i)$ Un vector $v \in W^{\perp}$ si y solo si v es ortogonal a todo vector en un conjunto que genere a W.*
$ii)$ $W^{\perp}$ es un subespacio vectorial de V.

Demos: Ejercicio.

Teorema 4.2 *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. El complemento ortogonal del espacio fila de A es el espacio nulo de A, y el complemento ortogonal del espacio columna de A es el espacio nula de A^t.*

$$(Fil(A))^{\perp} = nul(A) \qquad\qquad (Col(A))^{\perp} = nul(A^t).$$

Demos: Surge de la forma de calcular Av que si $v \in nul(A)$ resulta v ortogonal a cada fila de A (viendo las filas como vectores de $\mathbb{K}^n$). Como las filas de A generan el espacio $Fil(A)$, resulta v ortogonal a $Fil(A)$.
De manera recíproca, si v es ortogonal a $Fil(A)$, en particular resulta ortogonal a cada fila de A y así se tiene que $Av = 0$ con lo cual $v \in nul(A)$. Con esto queda probada la primera igualdad del enunciado.
Como el enunciado vale para cualquier matriz, vale en particular para A^t, es decir que el complemento ortogonal del espacio fila de A^t es el $nul(A^t)$, de donde la segunda igualdad. ■

.

Definición 4.6 *Si $S \subset V$, se dice que S es un* **conjunto ortogonal** *si todos los pares de vectores distintos de S son ortogonales.*
Un **conjunto ortonormal** *es un conjunto ortogonal S donde $\|v\| = 1$ para todo $v \in S$.*

Ejemplo 4.3

1. *La base canónicas en $\mathbb{R}^n$ o $\mathbb{C}^n$ es un conjunto ortonormal con respecto al producto interno canónico.*
2. *Sea $V = \mathcal{M}_{n\times n}(\mathbb{K})$, para $A = (a_{ij})$ y $B = (b_{ij})$, consideremos el producto interno*

$$\langle A, B\rangle = \sum_{i,j} a_{ij}\overline{b_{ij}},$$

Sea E^{pq} la matriz que tiene un 1 en la fila p, columna q y el resto de las entradas iguales a 0. El conjunto $\{E^{pq} : p, q = 1, \cdots, n\}$ es un conjunto ortonormal pues

$$\langle E^{pq}, E^{rs}\rangle = tr(E^{pq}(E^{rs})^*) = tr(E^{pq}E^{sr})\delta_{qs}tr(E^{pr}) = \delta_{qs}\delta_{pr}.$$

3. *Sea V el espacio vectorial de las funciones reales continuas en el intervalo $[0,1]$. Para $f, g \in V$ consideremos el producto interno*
$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt.$$
Consideremos las funciones $f_n(x) = \sqrt{2}\cos(2\pi nx)$ y $g_n(x) = \sqrt{2}\operatorname{sen}(2\pi nx)$. Entonces el conjunto $\{1, f_1, g_1, f_2, g_2, \cdots\}$ es un conjunto infinito ortonormal o ***sistema ortonormal*** .

Teorema 4.3 *Un conjunto ortogonal de vectores no nulos es linealmente independiente.*

Demos: Sea S un conjunto ortonormal (finito o infinito) de vectores no nulos en un espacio producto interno V. Supongamos que $\{v_1, \cdots, v_m\}$ son vectores distintos en S y consideremos el vector
$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m.$$
Entonces
$$\langle w, v_k\rangle = \langle \sum_{j=1}^m \alpha_j v_j, v_k\rangle = \sum_{j=1}^m \alpha_j \langle v_j, v_k\rangle = \alpha_k \langle v_k, v_k\rangle.$$
Como $\langle v_k, v_k\rangle \neq 0$ resulta que $\alpha_k = \dfrac{\langle w, v_k\rangle}{\|v_k\|^2}$, para $k = 1, \cdots, m$.
Luego si $w = 0$ cada $\alpha_k = 0$, lo que prueba que S es linealmente independiente. ■

Corolario 4.1 *Si un vector w es combinación lineal de una colección ortogonal de vectores no nulos $v_1, \cdots, v_m$, entonces w es igual a la combinación lineal particular dada por*
$$w = \sum_{j=1}^m \frac{\langle w, v_k\rangle}{\|v_k\|^2} v_k.$$

Observación 4.1 *Se desprende trivialmente del teorema que si $\{v_1, \cdots, v_m\}$ es un conjunto ortogonal de vectores no nulos en un espacio con producto interno de dimensión finita V, luego $m \leqslant \dim(V)$. Es decir que la cantidad de direcciones mutuamente perpendiculares en V no puede exceder la dimensión algebraica de V. La cantidad de direcciones mutuamente perpendiculares en V es lo que intuitivamente se considera la dimensión geométrica de$V$. Es decir que estamos concluyendo que la dimensión geométrica de V no es mayor que la dimensión algebraica de V. De hecho dichas dimensiones son iguales, esto se obtiene a partir del siguiente teorema.*

Proceso de ortogonalidad de Gram-Schmidt

Teorema 4.4 ***Proceso de ortogonalización de Gram-Schmidt***
Sea V un espacio con producto interno y sea $\{v_1, v_2, \cdots, v_n\} \subset V$ linealmente independiente. Entonces pueden construirse vectores ortogonales $w_1, w_2, \cdots, w_n \in V$ tales que para cada $k = 1, \cdot, n$ resulte el conjunto $\{w_1, \cdots, w_k\}$ una base para el espacio generado por $v_1, \cdots, v_k$.

Demos: El proceso de construcción de estos vectores w es lo que se conoce como el proceso de ortogonalización de Gram-Schmidt.
Sea $w_1 = v_1$, construimos los otros vectores de manera inductiva como sigue. Supongamos que hayan sido elegidos $w_1, \cdot, w_m$ con $(1 \leqslant m \leqslant n)$ de manera que para cada $k = 1, \cdots, m$ $\{w_1, \cdots, w_k\}$ una base para $< v_1, \cdots, v_k >$.
Construyamos el siguiente vector como
$$w_{m+1} = v_{m+1} - \sum_{k=1}^m \frac{\langle v_{m+1}, w_k\rangle}{\|w_k\|^2} w_k.$$

Entonces $w_k \neq 0$, pues de serlo resultaría v_{m+1} ser una combinación lineal de $w_1, \cdots, w_m$ es decir que $v_{m+1} \in < v_1, \cdots, v_m >$ lo que contradice la hipótesis de que los vectores $v_1, \cdots, v_n$ son *l.i.*.
Además

$$\langle w_{m+1}, w_j \rangle = \langle v_{m+1}, w_j \rangle - \sum_{k=1}^{m} \frac{\langle v_{m+1}, w_k \rangle}{\|w_k\|^2} \langle w_k, w_j \rangle = \langle v_{m+1}, w_j \rangle - \langle v_{m+1}, w_j \rangle = 0.$$

Por lo tanto $\{w_1, \cdots, w_{m+1}\}$ es un conjunto ortogonal que consta de $m+1$ vectores no nulos en el subespacio generado por $v_1, \cdots, v_{m+1}$ y resulta una base para este subespacio. ■

Observación 4.2 *Por ejemplo para para $n = 4$ los vectores toman la forma*

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1 \\
w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2 \\
w_4 &= v_4 - \frac{\langle v_4, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_4, w_2 \rangle}{\|w_2\|^2} w_2 - \frac{\langle v_4, w_3 \rangle}{\|w_3\|^2} w_3.
\end{aligned}$$

Corolario 4.2 *Todo espacio con producto interno de dimensión finita tiene una base ortonormal.*

Demos: Sea V un espacio con producto interno de dimensión finita y $\{v_1, \cdots, v_n\}$ una base para V. Aplicando el proceso de Gram-Schmidt se construye una base ortogonal $\{w_1, \cdots, w_n\}$. Luego para obtener una base ortonormal, se reemplaza cada vector w_k por $w_k/\|w_k\|$. ■

Observación 4.3 *Una de las ventajas principales de usar base ortonormales sobre usar bases ordenadas cualesquiera radica en el hecho que son mucho más simples los cálculos donde se hacen intervenir a las coordenadas.*
Sea V un espacio vectorial de de dimensión finita con producto interno. Sea $\mathcal{B} = \{v_1, \cdots, v_n\}$ una base ordenada de V. Luego podemos definir una matriz asociada al producto interno del siguiente modo

$$G = (g_{ij}) = (\langle v_i, v_j \rangle),$$

podemos usar esta matriz para escribir el producto interno de dos vectores cualesquiera de V en función de las coordenadas del siguiente modo: $[u]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $[w]_{\mathcal{B}} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

$$\langle u, w \rangle = \langle \sum_{j=1}^{n} x_j v_j, \sum_{j=1}^{n} y_j v_j \rangle = [u]_{\mathcal{B}}^t G [\overline{w}]_{\mathcal{B}}.$$

En el caso que la base sea ortonormal, la matriz G es la identidad y puede expresarse el producto interno en función sólo de las coordenadas

$$\langle u, w \rangle = \sum_{j=1}^{n} x_j \overline{y}_j.$$

Ejemplo 4.4 *Consideremos los vectores $v_1 = (3,0,4)^t, v_2 = (-1,0,7)^t, v_3 = (2,9,11)^t$ en $\mathbb{R}^3$ con el producto interno canónico. Aplicando el proceso de Gram-Schmidt se obtienen los vectores $w_1 = (3,0,4)^t$, $w_2 = (-4,0,3)$ y $w_3 = (0,9,0)$ que constituyen una base ortogonal. Si queremos una base ortonormal, basta dividir cada vector por su norma y obtenemos $u_1 = (\frac{3}{5}, 0, \frac{4}{5})$, $u_2 = (\frac{-4}{5}, 0, \frac{3}{5})$ y $u_3 = (0,1,0)$.*

4.3. Proyección ortogonal. Mejor aproximación

Sea V un espacio vectorial con producto interno sobre $\mathbb{K}$. Sea $0 \neq u \in V$. Queremos descomponer un vector $y \in V$ en la suma de dos vectores, uno que sea múltiplo de v y otro que sea ortogonal a u. Es decir queremos escribir

$$y = \widehat{y} + z, \tag{10}$$

con $\widehat{y} = \alpha u$ para algún escalar $\alpha \in \mathbb{K}$, y z tal que $\langle u, z \rangle = 0$.
Dado un escalar cualquiera $\alpha \in \mathbb{K}$, llamemos $z = y - \alpha u$, de modo que se cumpla la igualdad (10). Entonces $y - \widehat{y}$ es ortogonal a u si y sólo si

$$0 = \langle y - \alpha u, u \rangle = \langle y, u \rangle - \alpha \langle u, u \rangle,$$

es decir sólo si

$$\alpha = \frac{\langle y, u \rangle}{\|u\|^2}, \qquad \widehat{y} = \frac{\langle y, u \rangle}{\|u\|^2} u.$$

El vector $\widehat{y}$ es la ***proyección ortogonal de y sobre u*** y z es la componente de y ortogonal a u.

Observación 4.4 *La proyección ortogonal de y sobre u es igual a la proyección ortogonal de y sobre βu para cualquier $\beta \in \mathbb{K}$. Es por esto que la proyección está determinada por el subespacio L generado por u, y se la suele llamar* ***proyección ortogonal de y sobre L***

$$\widehat{y} = \mathrm{proy}_L y = \frac{\langle y, u \rangle}{\|u\|^2} u.$$

Nos planteamos ahora el siguiente problema, sea W un subespacio de un espacio con producto interno V y sea v un vector arbitrario en V. Nos preguntamos si existe una mejor aproximación a v por vectores en W. Esto quiere decir si existe en W algún vector $\overline{w}$ tal que sea de todos los vectores en W el que se encuentra a menor distancia de v. Se dice que $\overline{w}$ es una ***mejor aproximación a v en W*** si

$$\|v - \overline{w}\| \leqslant \|v - w\|, \forall w \in W.$$

Si pensamos este problema en $\mathbb{R}^2$ o $\mathbb{R}^3$, se ve intuitivamente que lo que buscamos es el pie de la perpendicular sobre W del vector v y en tal caso $\overline{w}$ resulta único. Estas ideas intuitivas son correctas para los subespacios de dimensión finita y para algunos (no todos) los subespacios de dimensión infinita. Veamos el siguiente resultado

Teorema 4.5 *de la descomposición ortogonal*
Sea W un subespacio vectorial de V espacio vectorial sobre $\mathbb{K}$ con producto interno y sea $v \in V$.
i) El vector $\overline{w} \in W$ es una mejor aproximación de v, por vectores de W si y sólo si $v - \overline{w}$ es ortogonal a W.
ii) Si existe una mejor aproximación a v por vectores de W, la misma es única.
iii) Si W es de dimensión finita y $\{w_1, \cdots, w_n\}$ es una base ortonormal de W, entonces el vector

$$\overline{w} = \sum_{k=1}^{n} \frac{\langle v, w_k \rangle}{\|w_k\|^2} w_k$$

es la única mejor aproximación a v por vectores de W.

Demos: *i*) Observemos que si u es cualquier vector en V se tiene que $v - u = (v - \overline{w}) + (\overline{w} - u)$ y

$$\|v - u\|^2 = \|v - \overline{w})\|^2 + 2\Re\langle v - \overline{w}, \overline{w} - u \rangle + \|(\overline{w} - u)\|^2.$$

Supongamos ahora que $v - \overline{w}$ es ortogonal a todo vector en W, $\overline{w} \in W$ y que $u \neq \overline{w}$. Entonces, como $\overline{w} - u \in W$, se tiene que

$$\|v - u\|^2 = \|v - \overline{w})\|^2 + \|(\overline{w} - u)\|^2 > \|v - \overline{w})\|^2.$$

Recíprocamente, supongamos que $\|v - u\| > \|v - \overline{w})\|$ para todo $u \in W$, entonces de la primera ecuación anterior se tiene que

$$2\Re\langle v - \overline{w}, \overline{w} - u\rangle + \|(\overline{w} - u)\|^2 \geqslant 0, \forall u \in W.$$

Como todo vector de $\tau \in W$ puede expresarse como $\overline{w} - u$ con $u \in W$ se tiene que

$$2\Re\langle v - \overline{w}, \tau\rangle + \|\tau\|^2 \geqslant 0, \forall \tau \in W.$$

En particular si $u \in W$, $u \neq \overline{w}$, puede expresarse

$$\tau = -\frac{\langle v - \overline{w}, \overline{w} - u\rangle}{\|\overline{w} - u\|^2}(\overline{w} - u),$$

entonces la desigualdad se reduce a

$$-2\frac{|\langle v - \overline{w}, \overline{w} - u\rangle|^2}{\|\overline{w} - u\|^2} + \frac{|\langle v - \overline{w}, \overline{w} - u\rangle|^2}{\|\overline{w} - u\|^2} \geqslant 0$$

Lo cual se cumple sólo si $\langle v - \overline{w}, \overline{w} - u\rangle = 0$, por lo tanto $v - \overline{w}$ es ortonormal a todo vector en W.
ii) Supongamos para $v \in V$ existen dos proyecciones ortogonales en W, sean $\overline{w}$ y $\tilde{w}$. Luego podemos escribir a $v = z_1 + \overline{w} = z_2 + \tilde{w}$ donde $z_1 = (v - \overline{w}), z_2 = (v - \tilde{w}) \in W^{\perp}$ y $\overline{w}, \tilde{w} \in W$. Luego resulta que

$$W^{\perp} \ni z_1 - z_2 = \overline{w} - \tilde{w} \in W.$$

Luego el vector $\overline{w} - \tilde{w}$ está en W y el $W^{\perp}$, por lo tanto $\langle \overline{w} - \tilde{w}, \overline{w} - \tilde{w}\rangle = 0$, lo que muestra que $\overline{w} = \tilde{w}$.
iii) Supongamos que W es un subespacio de dimensión finita de V, luego sabemos que W tiene una base ortogonal (por corolario de Gram-Schmidt). Sea $\{w_1, \cdots, w_n\}$ cualquier base ortogonal de W y sea

$$\overline{w} = \sum_{k=1}^{n} \frac{\langle v, w_k\rangle}{\|w_k\|^2} w_k.$$

Luego se ve fácilmente que $v - \overline{w}$ resulta ortogonal a cada uno de los elementos de la base, luego es ortogonal a cualquier combinación lineal de dichos vectores y así a todo vector en W. Si $u \in W$ y $u \neq \overline{w}$ se tiene que $\|v - \overline{w}\| \leqslant \|v - u\|$. Por lo tanto $\overline{w}$ es la mejor aproximación a v en W. ■

Definición 4.7 *Siempre que exista el vector $\overline{w}$ del teorema anterior se lo llama* ***proyección ortogonal****. Si todo vector de V tiene proyección ortogonal sobre W, la aplicación que a cada vector de V le asigna su proyección ortogonal sobre W se llama* ***proyección ortogonal de V sobre W****.*

Por el teorema anterior sabemos que siempre existe la proyección ortogonal de un espacio con producto interno sobre un subespacio de dimensión finita. Pero podemos deducir otro resultado del teorema anterior.

Corolario 4.3 *Sea V un espacio con producto interno, W un subespacio de dimensión finita y P_W la proyección ortogonal de V sobre W. Entonces la aplicación $v \to v - P_W v$ es la proyección ortogonal de V sobre $W^{\perp}$.*

Demos: Sea $v \in V$, como $v - P_W v \in W^{\perp}$ y para cualquier $u \in W^{\perp}$ se tiene que $v - u = P_W v + (v - P_W v - u)$, como $P_W v \in W$ y $(v - P_W v - u) \in W^{\perp}$ resulta que

$$\|v - u\|^2 = \|P_W v\|^2 + \|v - P_W v - u\|^2 \geqslant \|v - (v - P_W v)\|^2.$$

Así vemos que $v - P_W v$ es la mejor aproximación a v en $W^{\perp}$. ■

Ejemplo 4.5 *Consideremos $\mathbb{R}^3$ con el producto interno canónico, sea $W = s.e.\{(3,12,-1)\}$, luego calculemos la proyección ortogonal del vector $(-10,2,8)$ sobre W*

$$\overline{w} = \frac{\langle(-10,2,8),(3,12,-1)\rangle}{9+144+1}(3,12,-1) = \frac{-14}{154}(3,12,-1).$$

La proyección ortogonal de $\mathbb{R}^3$ sobre W es la transformación lineal P_W definida por

$$(x_1,x_2,x_3) \mapsto P_W(x_1,x_2,x_3) = \frac{3x_1+12x_2-x_3}{154}(3,12,-1).$$

Obviamente el $rang(P_W) = 1$, luego por el teorema del rango se tiene que $\dim(nul(P_W) = 2$. Además se tiene que

$$P_W(x_1,x_2,x_3) = 0 \Leftrightarrow 3x_1+12x_2-x_3 = 0 \Leftrightarrow (x_1,x_2,x_3) \in W^{\perp},$$

por lo tanto $W^{\perp} = nul(P_W)$ y $\dim(W^{\perp}) = 2$.
La proyección ortogonal de V sobre $W^{\perp}$ es la transformación lineal $I - P_W$.

Lo observado en este ejemplo puede generalizarse.

Definición 4.8 *Un operador lineal $T \in \mathcal{L}(V)$ se dice **idempotente** si $T^2 = T$.*

Teorema 4.6 *Sea W un subespacio de dimensión finita de un espacio con producto interno V y sea P_W la proyección ortogonal de V sobre W. Entonces P_W es una transformación lineal idempotente de V sobre W, $W^{\perp} = nul(P_W)$ y $V = W \oplus W^{\perp}$.*

Demos: Para $v \in V$, $P_W v$ es la mejor aproximación a v por un vector en W, luego si $v \in W$ se tiene que $P_W v = v$, así se tiene que $P_W^2 = P_W$, es decir es idempotente.
Sean $u, v \in V$ y $\alpha \in \mathbb{K}$, así por el Teorema 4.5 se tiene que $v - P_W v$ y $u - P_W u$ son ortogonales a todo vector en W, y así $\alpha(v - P_W v) + (u - P_W u) = (\alpha v + u) - (\alpha P_W v - P_W u) \in W^{\perp}$, por la unicidad de la proyección ortogonal sobre W se tiene que $P_W(\alpha v + u) = \alpha P_W v - P_W u$.
Sea ahora $v \in V$, así $P_W v$ es el único vector en W tal que $v - P_W v \in W^{\perp}$. Así $P_W v = 0 \Leftrightarrow v \in W^{\perp}$ y así $nul(P_W) = W^{\perp}$.
Finalmente vemos que la igualdad $v = P_W v + v - P_W v$ prueba que $V = W + W^{\perp}$ y claramente $W \cap W^{\perp} = \{0\}$, luego $V = W \oplus W^{\perp}$. ■

Corolario 4.4 *Bajo las condiciones del teorema, $I - P_W$ es la proyección ortogonal de V sobre $W^{\perp}$, es una transformación lineal idempotente de V en $W^{\perp}$ con espacio nulo igual a W.*

Corolario 4.5 ***Desigualdad de Bessel***
Sea $\{v_1, \cdots, v_n\}$ un conjunto ortogonal de vectores no nulos en un espacio con producto interno V. Sea $u \in V$, entonces

$$\sum_{k=1}^{n} \frac{|\langle u, v_k\rangle|^2}{\|v_k\|^2} \leqslant \|u\|^2,$$

la igualdad vale si y solo si

$$u = \sum_{k=1}^{n} \frac{\langle u, v_k\rangle}{\|v_k\|^2} v_k.$$

Demos: Sea $w = P_W u = \sum_{k=1}^{n} \frac{\langle u,v_k\rangle}{\|v_k\|^2} v_k$, luego $u = w + x$ con $\langle w, x\rangle = 0$. Así

$$\|u\|^2 = \|w\|^2 + \|x\|^2,$$

con lo que queda probada la desigualdad. Viendo que $\|w\|^2 = \sum_{k=1}^{n} \frac{\langle u,v_k\rangle}{\|v_k\|^2} v_k$ se completa la prueba. ■