Tema 13.
Espacios euclídeos. Aproximaciones de Fourier

En los temas 11 y 12 hemos dedicado nuestro esfuerzo, primero, a resolver y, luego, a enfocar con nitidez el problema de los sistemas lineales. Hemos recibido ayuda de los espacios vectoriales y de las aplicaciones lineales.

Ahora bien, no podemos perder de vista que la Geometría, la fuente inagotable de nuestra intuición, es algo más que vectores que se combinan de acuerdo con unas leyes. Piense el lector por un momento que etimológicamente la palabra Geometría significa «la ciencia de medir la tierra», y nosotros hasta ahora no hemos medido nada. Parece, pues, que hay que saber medir distancias. Pero no sólo queremos medir distancias, queremos medirlas en las mejores condiciones posibles, y eso se consigue cuando además tenemos los ángulos y la perpendicularidad, es decir, cuando disponemos de un producto escalar; en definitiva, cuando nos movemos en un espacio euclídeo.

Pero todo eso no es más que una pequeña visión de por dónde vamos a dirigir nuestros pasos en este tema. La pregunta clave es: ¿cuál o cuáles son los objetivos que nos proponemos en el estudio de los espacios euclídeos? La respuesta está en el siguiente problema de optimización: en muy diferentes contextos es necesario obtener las *mejores aproximaciones*. Precisemos este problema. Por ejemplo, las ciencias experimentales necesitan resolver el problema de ajustar una nube de datos a un modelo, con frecuencia, lineal, es decir, el problema de buscar los parámetros del modelo que más se aproxima a los datos. Otro ejemplo: los fenómenos periódicos se estudian con más facilidad si se expresan como suma de armónicos o sinusoides y, por supuesto, se buscan las sumas que más se aproximen a la onda de partida.

Es preciso decir que, en este tema, vamos a manejar con la misma intensidad los espacios $\mathbf{R}^n$, de dimensión finita, que los espacios de funciones, de dimensión infinita. Trataremos de conseguir que nuestra intuición geométrica nos indique cuál es el camino adecuado sin que ello nos atenace de tal manera que nos impida comprender conceptos como el de ortogonalidad, de proyección o de aproximaciones de Fourier en el importante espacio de las funciones periódicas.

M. Dolores Lerís
Universidad de Zaragoza, España

13.1. Producto escalar. Norma, distancia y ángulo

Producto escalar en espacios vectoriales reales

Definición: Sea E un espacio vectorial sobre $\mathbf{R}$. Un *producto escalar* en E es una aplicación del producto cartesiano $E \times E$ en $\mathbf{R}$

$$\begin{aligned} E \times E &\longrightarrow \mathbf{R} \\ (\mathbf{u},\mathbf{v}) &\rightarrow \mathbf{u} \bullet \mathbf{v} \end{aligned}$$

que verifica las siguientes propiedades:

1. *Simétrica o conmutativa:* $\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}$, $\forall\, \mathbf{u},\mathbf{v} \in E$
2. *Lineal en la primera componente:* $\forall\, \mathbf{u},\mathbf{v},\mathbf{w} \in E$, $\forall t \in \mathbf{R}$ se verifica
   $(\mathbf{u}+\mathbf{v}) \bullet \mathbf{w} = \mathbf{u} \bullet \mathbf{w} + \mathbf{v} \bullet \mathbf{w}$ y $(t\,\mathbf{u}) \bullet \mathbf{v} = t(\mathbf{u} \bullet \mathbf{v})$
3. *Definida positiva:* $\mathbf{u} \bullet \mathbf{u} \geq 0$, $\forall\, \mathbf{u} \in E$;
   además, $\mathbf{u} \bullet \mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.

Como consecuencia de la conmutatividad y de la linealidad en la primera componente, también es válida la *linealidad en la segunda componente,* es decir: $\forall\, \mathbf{u},\mathbf{v},\mathbf{w} \in E$ y $\forall t \in \mathbf{R}$ se verifica

$$\mathbf{u} \bullet (\mathbf{v}+\mathbf{w}) = \mathbf{u} \bullet \mathbf{v} + \mathbf{u} \bullet \mathbf{w} \quad \text{y} \quad \mathbf{u} \bullet (t\,\mathbf{v}) = t(\mathbf{u} \bullet \mathbf{v})$$

Un *espacio vectorial euclídeo,* o simplemente *espacio euclídeo,* es un espacio vectorial sobre $\mathbf{R}$ con un producto escalar.

El ejemplo más conocido de espacio euclídeo es $\mathbf{R}^n$ con el producto escalar estándar de dos vectores $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ y $\mathbf{v} = (v_1, v_2, \ldots, v_n)$:

$$\mathbf{u} \bullet \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{u}^t \mathbf{v}.$$

Nótese que el producto escalar de los vectores $\mathbf{u}$ y $\mathbf{v}$ es, en definitiva, el producto de la matriz fila $\mathbf{u}^t = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}$ por la matriz columna

$\mathbf{v}=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$. Esperamos que el lector haya comprendido que siempre que un vector, como $\mathbf{v}$, sea interpretado como una matriz lo pondremos en columna. Como entrenamiento, le proponemos que escriba el producto escalar de los vectores $\mathbf{u}=(1,2,-1,3)$ y $\mathbf{v}=(4,-1,-1,-1)$ en forma de producto de matrices de dos formas distintas y que obtenga su valor.

Vamos a considerar ahora un ejemplo tan importante como $\mathbf{R}^n$ pero, quizás, más novedoso. Se trata del *espacio vectorial* $\mathcal{PC}(2\pi)$ *de todas las funciones de período* 2π *y continuas en* $[-\pi,\pi]$ *salvo a lo sumo en un número finito de puntos de discontinuidad de salto finito*. Recordemos que las operaciones de este espacio vectorial son la suma de funciones y el producto por números reales que ya son conocidas. El paso de los vectores de $\mathbf{R}^n$ con un número de componentes que se pueden contar a espacios vectoriales de funciones, donde los vectores son funciones y sus componentes, las imágenes, varían en un continuo obliga a sustituir, en la definición de producto escalar, sumas por integrales. Se define el ***producto escalar*** de dos funciones f y g de $\mathcal{PC}(2\pi)$ mediante

$$f\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt.$$

La comprobación de que efectivamente se ha definido un producto escalar consiste en comprobar que se verifican las propiedades exigidas a un producto escalar en la definición. Así pues,

1. Simétrica: $f\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx=\left(\begin{array}{c}\text{conmutatividad}\\ \text{del producto de}\\ \text{números reales}\end{array}\right)=\frac{1}{\pi}\int_{-\pi}^{\pi} g(x)f(x)dx=g\bullet f.$

2. Lineal en la primera componente:

$$(f_1+f_2)\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi}(f_1(x)+f_2(x))g(x)dx=$$

$$=\frac{1}{\pi}\int_{-\pi}^{\pi} f_1(x)g(x)dx+\frac{1}{\pi}\int_{-\pi}^{\pi} f_2(x)g(x)dx=f_1\bullet g+f_2\bullet g$$

y

$$(kf)\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi} kf(x)g(x)dx=k\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx=k(f\bullet g).$$

3. La linealidad en la segunda componente, como ya hemos mencionado, es consecuencia de la linealidad en la primera y de la simetría o conmutatividad del producto escalar.
4. Definida positiva:

$$f \bullet f = \frac{1}{\pi}\int_{-\pi}^{\pi} [f(x)]^2 dx \geq 0 \quad \text{y}$$

$$f \bullet f = \frac{1}{\pi}\int_{-\pi}^{\pi} [f(x)]^2 dx = 0 \Leftrightarrow f = 0 \text{ en casi todos los puntos.}$$

Hay que observar que si en un mismo espacio vectorial se definen dos productos escalares distintos, resultan dos espacios euclídeos diferentes. Por ello al hablar de un espacio euclídeo ha de estar claro cuál es el producto escalar definido.

Longitud o norma de un vector

De la idea de producto escalar se deduce el siguiente concepto: la norma, longitud o magnitud de un vector $\mathbf{u}$.

Definición: Sea E un espacio euclídeo. Se define la *norma o longitud de un vector* $\mathbf{u} \in E$ como la raíz cuadrada positiva de $\mathbf{u} \bullet \mathbf{u}$.

Sabemos que $\mathbf{u} \bullet \mathbf{u}$ es siempre un número real positivo y, por tanto, existe la raíz cuadrada positiva $\sqrt{\mathbf{u} \bullet \mathbf{u}}$, que es la norma construida a partir del producto escalar:

$$\begin{aligned} \| \ \| : E &\longrightarrow \mathbf{R} \\ \mathbf{u} &\longrightarrow \|\mathbf{u}\| = \sqrt{\mathbf{u} \bullet \mathbf{u}} \end{aligned}$$

Todo el mundo, al menos el mundo científico, acepta que la denominación de norma o de longitud de un vector se reserva a aquellas aplicaciones $\| \ \|$ que verifiquen tres propiedades básicas:

1. $\|\mathbf{u}\| \geq 0$ para todo $\mathbf{u} \in E$; $\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$;
2. $\|k\,\mathbf{u}\| = |k|\,\|\mathbf{u}\|$, donde $|k|$ indica el valor absoluto del número real k;
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (desigualdad del triángulo), la lectura geométrica de esta desigualdad dice: «la longitud de cada lado de un triángulo es menor que la suma de las longitudes de los otros dos lados».

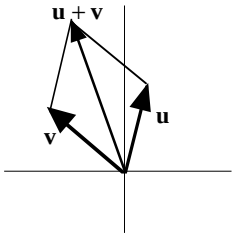


FIGURA 1: Relación entre las longitudes de los lados de un triángulo.

Los dos productos escalares que hemos definido en los ejemplos anteriores conducen a sus respectivas normas:

1. La primera de ellas es la norma estándar en el espacio euclídeo $\mathbf{R}^n$:
$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \bullet \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}\,.$$
2. También tenemos la norma $\|f\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi}[f(x)]^2 dx}$ en $\mathcal{PC}(2\pi)$ deducida del producto escalar $f \bullet g = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx$.

Destacamos por su importancia un tipo de vectores del espacio euclídeo: los vectores unitarios. Un vector $\mathbf{u}$ del espacio euclídeo E se llama *unitario* si $\|\mathbf{u}\|=1$. Por ejemplo, los vectores $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ de $\mathbf{R}^3$ son los vectores unitarios a lo largo de los ejes coordenados del espacio euclídeo tridimensional.

Para construir un vector unitario $\mathbf{u}$ en la dirección de un vector dado $\mathbf{v} \neq \mathbf{0}$ basta multiplicarlo por el inverso de su norma: $\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$.

Ejercicio 1: En el espacio vectorial $\mathcal{PC}(2\pi)$ de la funciones periódicas de período 2π y continuas a trozos, considera las funciones resultado de extender periódicamente $f(x) = x$ y $g(x) = x^2$, $x \in (-\pi, \pi]$. Calcula su producto escalar, así como la norma de cada una de ellas.

En primer lugar, su producto escalar es
$$f \bullet g = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx = \frac{1}{\pi}\int_{-\pi}^{\pi} x^3 dx = 0\,.$$

En cuanto a sus normas, tenemos

$$\|f\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} x^2 dx} = \sqrt{\frac{1}{\pi}\left[\frac{x^3}{3}\right]_{-\pi}^{\pi}} = \pi\sqrt{\frac{2}{3}} \approx 2.565;$$

$$\|g\| = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi} x^4 dx} = \sqrt{\frac{1}{\pi}\left[\frac{x^5}{5}\right]_{-\pi}^{\pi}} = \pi^2\sqrt{\frac{2}{5}} \approx 6.242.$$

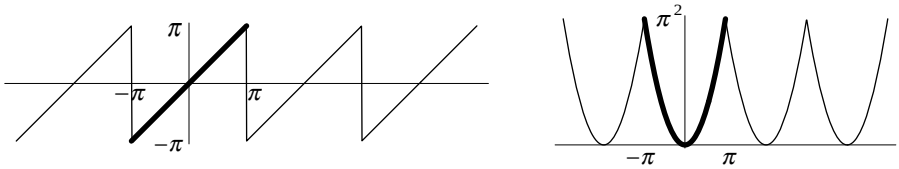


FIGURA 2: Gráfica de las extensiones periódicas de un segmento y de un arco de parábola.

Con el fin de practicar un poco más con la idea de que el producto escalar en el espacio euclídeo $\mathbf{R}^n$ se puede expresar en forma de producto de matrices, resolvemos el siguiente ejercicio:

Ejercicio 2: A es una matriz cuadrada de orden n y $\mathbf{u} \in \mathbf{R}^n$, comprobamos que la longitud de $A\mathbf{u}$ es igual a la longitud de $A^t\mathbf{u}$ siempre que $AA^t = A^tA$.

Por un lado, tenemos que

$$\|A\mathbf{u}\|^2 = (A\mathbf{u})\bullet(A\mathbf{u}) = (A\mathbf{u})^t(A\mathbf{u}) = \mathbf{u}^t\left(A^tA\mathbf{u}\right) = \mathbf{u}\bullet\left(A^tA\mathbf{u}\right);$$

y por otro,

$$\left\|A^t\mathbf{u}\right\|^2 = \left(A^t\mathbf{u}\right)\bullet\left(A^t\mathbf{u}\right) = \left(A^t\mathbf{u}\right)^t\left(A^t\mathbf{u}\right) = \mathbf{u}^t\left(AA^t\mathbf{u}\right) = \mathbf{u}\bullet\left(AA^t\mathbf{u}\right).$$

Por tanto, si $AA^t = A^tA$, entonces $\|A\mathbf{u}\|^2 = \left\|A^t\mathbf{u}\right\|^2$.

El objetivo del siguiente ejercicio es adquirir la habilidad en la utilización de las propiedades básicas tanto del producto escalar como de la norma o longitud.

Ejercicio 3: Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores de un espacio euclídeo de los cuales sabemos: $\mathbf{u}\bullet\mathbf{v} = 2$, $\mathbf{v}\bullet\mathbf{w} = -3$, $\mathbf{u}\bullet\mathbf{w} = 5$, $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$ y $\|\mathbf{w}\| = 7$.

a. Calculamos $(\mathbf{u}+\mathbf{v})\bullet(\mathbf{v}+\mathbf{w})$ aplicando las propiedades del producto escalar:

$$(\mathbf{u}+\mathbf{v})\bullet(\mathbf{v}+\mathbf{w}) = (\mathbf{u}\bullet\mathbf{v}) + (\mathbf{u}\bullet\mathbf{w}) + (\mathbf{v}\bullet\mathbf{v}) + (\mathbf{v}\bullet\mathbf{w}) = 2 + 5 + 2^2 + (-3) = 8.$$

De modo análogo se obtiene el valor de la expresión $(3\mathbf{u}+2\mathbf{w})\bullet(2\mathbf{v}-\mathbf{w})$, cuyo cálculo dejamos propuesto.

b. Obtengamos las siguientes normas: $\|\mathbf{u}+\mathbf{v}\|$, $\|3\mathbf{w}-\mathbf{u}\|$ y $\|2\mathbf{u}-\mathbf{v}+4\mathbf{w}\|$. Tan solo calculamos la primera, dejando al lector el cálculo de las otras dos.

$$\|\mathbf{u}+\mathbf{v}\|^2=(\mathbf{u}+\mathbf{v})\bullet(\mathbf{u}+\mathbf{v})=\|\mathbf{u}\|^2+2(\mathbf{u}\bullet\mathbf{v})+\|\mathbf{v}\|^2=1^2+4+2^2=9 \quad \Leftrightarrow \quad \|\mathbf{u}+\mathbf{v}\|=3$$

Una vez conocida la idea de norma o longitud de un vector disponemos de la herramienta que permite medir distancias; de hecho, la *distancia* entre $\mathbf{u}$ y $\mathbf{v}$ se define como el número real $\|\mathbf{u}-\mathbf{v}\|$ y se escribe $d(\mathbf{u},\mathbf{v})=\|\mathbf{u}-\mathbf{v}\|$.

Claro que el lector puede preguntarse si ésta es una definición «razonable» del concepto de distancia, es decir, la pregunta se puede precisar en los siguientes términos: ¿Qué propiedades se requieren para que el número que proponemos sea de verdad una distancia? La respuesta consiste, como siempre, en la observación de aquellas características mínimas comunes a lo que entendemos intuitivamente como distancia en la geometría ordinaria. Concretamente son: la distancia debe ser siempre un número real positivo, que es cero solamente cuando los puntos coinciden; la distancia debe ser independiente del orden en que éstos se consideren y, finalmente, debe cumplirse la desigualdad del triángulo, que dice: $d(u,v)\leq d(u,w)+d(w,v)$.

Ángulo entre dos vectores de un espacio euclídeo

Para introducir la noción de ángulo entre dos vectores de un espacio vectorial euclídeo necesitamos una poderosa herramienta, nos referimos a una propiedad muy importante (equivalente a la desigualdad del triángulo) que relaciona el producto escalar de dos vectores con sus longitudes. Dice así:

Desigualdad de Cauchy-Schwarz: Para cualesquiera dos vectores $\mathbf{u}$ y $\mathbf{v}$ de un espacio euclídeo E se verifica:

$$|\mathbf{u}\bullet\mathbf{v}|\leq\|\mathbf{u}\|\|\mathbf{v}\|;$$

además la igualdad sólo es válida cuando $\mathbf{u}$ es múltiplo de $\mathbf{v}$.

Demostración:

Está claro que si $\mathbf{v}$ es el vector nulo la igualdad es cierta:

$$|\mathbf{u}\bullet\mathbf{0}|=0=\|\mathbf{u}\|\|\mathbf{0}\|).$$

Supongamos, pues, que $\mathbf{v}\neq 0$ y consideremos el vector $\mathbf{u}-k\mathbf{v}$, donde el escalar es $k=\dfrac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}$.

Precisamente ese valor de k hace que el vector $k\mathbf{v}=\dfrac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}\mathbf{v}$ sea el resultado de proyectar $\mathbf{u}$ sobre la recta de dirección $\mathbf{v}$, como se observa en la figura 3. Más adelante, se justificará esta afirmación.

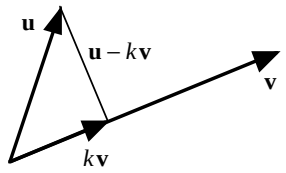


FIGURA 3: La idea geométrica de la elección de $k\mathbf{v}$.

Para el número real positivo $(\mathbf{u}-k\mathbf{v})\bullet(\mathbf{u}-k\mathbf{v})$ se tiene lo siguiente:

$$0\le(\mathbf{u}-k\mathbf{v})\bullet(\mathbf{u}-k\mathbf{v})=\mathbf{u}\bullet\mathbf{u}-k(\mathbf{u}\bullet\mathbf{v})-k(\mathbf{v}\bullet\mathbf{u})+k^2(\mathbf{v}\bullet\mathbf{v})=$$

$$=\mathbf{u}\bullet\mathbf{u}-2\frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}(\mathbf{u}\bullet\mathbf{v})+\left(\frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}\right)^2(\mathbf{v}\bullet\mathbf{v})=\mathbf{u}\bullet\mathbf{u}-\frac{(\mathbf{u}\bullet\mathbf{v})^2}{\mathbf{v}\bullet\mathbf{v}}$$

de donde se deduce que

$$|\mathbf{u}\bullet\mathbf{v}|^2\le(\mathbf{u}\bullet\mathbf{u})(\mathbf{v}\bullet\mathbf{v}),$$

que es lo que queríamos demostrar. □

De la desigualdad de Cauchy-Schwarz se deduce que podemos medir ángulos entre vectores de un espacio euclídeo E. En efecto, de la desigualdad $|\mathbf{u}\bullet\mathbf{v}|\le\|\mathbf{u}\|\|\mathbf{v}\|$ se deduce que $\dfrac{|\mathbf{u}\bullet\mathbf{v}|}{\|\mathbf{u}\|\|\mathbf{v}\|}\le 1$, es decir, $-1\le\dfrac{\mathbf{u}\bullet\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\le 1$. Por definición, el *ángulo entre dos vectores* $\mathbf{u}$ y $\mathbf{v}$ no nulos es el único número θ que verifica:

$$\cos\theta=\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|},\quad 0\le\theta\le\pi.$$

A partir de ahí podemos escribir

$$\mathbf{u}\cdot\mathbf{v}=\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta,$$

expresión que concuerda con la forma en la que se define el producto escalar de los vectores geométricos del plano y del espacio.

Recordamos cómo se calculan ángulos entre vectores en el plano euclídeo. Si, por ejemplo, $\mathbf{a}=(2,5)$ y $\mathbf{b}=(4,-1)$ son vectores del plano euclídeo, respondemos a las siguientes preguntas: ¿cuál es el coseno del ángulo entre $\mathbf{a}$ y $\mathbf{b}$?, y ¿cuál es el ángulo entre $\mathbf{a}$ y el vector $\mathbf{i}=(1,0)$?

— Si α es el ángulo entre $\mathbf{a}$ y $\mathbf{b}$, entonces

$$\cos\alpha = \frac{\mathbf{a}\bullet\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|} = \frac{3}{\sqrt{29}\,\sqrt{17}} \approx 0.135.$$

— En cuanto al ángulo β entre $\mathbf{a}$ y $\mathbf{i}$, una vez calculado

$$\cos\beta = \frac{2}{\sqrt{29}} \approx 0.3714,$$

obtenemos que

$$\beta = \arccos\frac{2}{\sqrt{29}} \approx 1.2 \text{ radianes.}$$

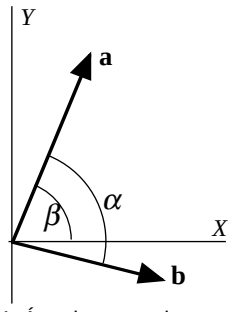


FIGURA 4: Ángulos entre los vectores del plano euclídeo del ejemplo.

13.2. Ortogonalidad

Vectores ortogonales

En el espacio ordinario se dice que dos vectores geométricos son ortogonales o perpendiculares cuando el ángulo α entre ellos es igual a $\frac{\pi}{2}$ radianes, es decir, cuando su producto escalar es cero. Esta sencilla idea se puede generalizar a un espacio euclídeo cualquiera y definir el concepto de vectores perpendiculares sin necesidad de mencionar el concepto de ángulo, usando tan solo el producto escalar.

Definición: Dos vectores $\mathbf{u}$ y $\mathbf{v}$, no nulos, de un espacio euclídeo E se dicen *perpendiculares* u *ortogonales* si $\mathbf{u}\bullet\mathbf{v}=0$. Escribiremos $\mathbf{u}\perp\mathbf{v}$.

Ejercicio 4: En el espacio euclídeo $\mathbf{R}^4$ nos preguntamos qué parejas de vectores ortogonales entre sí hay entre los siguientes vectores: $\mathbf{v}_1=(1,2,2,-1)$, $\mathbf{v}_2=(4,0,4,0)$ y $\mathbf{v}_3=(1,-1,-1,-1)$.

Esperamos que el lector verifique sin vacilar que los vectores $\mathbf{v}_2$ y $\mathbf{v}_3$ forman la única pareja de vectores perpendiculares, pues son los dos únicos vectores cuyo producto escalar es cero.

En el ejercicio 2. habíamos averiguado que el producto escalar de las

funciones resultado de extender periódicamente $f(x)=x$ y $g(x)=x^2$, $x\in[-\pi,\pi]$ es cero. En consecuencia, ambas funciones son ortogonales para el producto escalar $f\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$ definido en $\mathcal{PC}(2\pi)$. La gráfica de ambas funciones en el intervalo $[-\pi,\pi]$ se muestra en la figura 5, pero no hay ningún elemento que sugiera su ortogonalidad en el sentido geométrico.

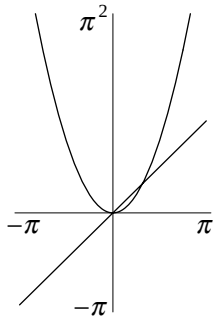


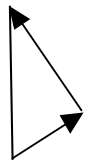
FIGURA 5: Las funciones $y=x$ e $y=x^2$.

Este ejemplo sirve de advertencia para que el lector se dé cuenta de los prejuicios que tenemos para interpretar el concepto de ortogonalidad en un espacio euclídeo arbitrario. La intuición geométrica de la noción de ortogonalidad nos es útil en los espacios $\mathbf{R}^2$ y $\mathbf{R}^3$, pero se pierde al considerar espacios euclídeos generales.

El famoso *teorema de Pitágoras* de la geometría elemental es un criterio para decidir la perpendicularidad de dos vectores $\mathbf{u}$ y $\mathbf{v}$ de un espacio euclídeo arbitrario en términos de las longitudes de cada uno de ellos. Dice lo siguiente:

«Los vectores $\mathbf{u}$ *y* $\mathbf{v}$ *son ortogonales si y sólo si* $\|\mathbf{u}+\mathbf{v}\|^2=\|\mathbf{u}\|^2+\|\mathbf{v}\|^2$*».*

Para comprobar la veracidad de esta afirmación escribimos la siguiente cadena de igualdades:



$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\bullet(\mathbf{u}+\mathbf{v})=\\ &=\mathbf{u}\bullet\mathbf{u}+2(\mathbf{u}\bullet\mathbf{v})+\mathbf{v}\bullet\mathbf{v}=\\ &=\|\mathbf{u}\|^2+\|\mathbf{v}\|^2+2(\mathbf{u}\bullet\mathbf{v}),\end{aligned}$$

de la cual concluimos que

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad \Leftrightarrow \quad \mathbf{u}\bullet\mathbf{v} = 0,$$

es decir, $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad \Leftrightarrow \quad \mathbf{u}\perp\mathbf{v}$.

Por ejemplo, hemos afirmado que los vectores $\mathbf{v}_2 = (4,0,4,0)$ y $\mathbf{v}_3 = (1,-1,-1,-1)$ del espacio euclídeo $\mathbf{R}^4$ son ortogonales, le pedimos al lector que compruebe que verifican el teorema de Pitágoras.

Vector ortogonal a un subespacio

Definición: Consideramos un subespacio W del espacio euclídeo E. Se dice que un *vector* $\mathbf{u} \in E$ es *ortogonal a* W si $\mathbf{u}$ es perpendicular a todos los vectores de W, es decir, si $\mathbf{u}\bullet\mathbf{w} = 0$ para todo $\mathbf{w} \in W$.

En la práctica, para comprobar que un vector $\mathbf{u}$ es ortogonal a un subespacio W no es necesario realizar infinitos productos escalares, uno para cada uno de los infinitos vectores de W, sino que basta comprobar que los productos escalares de $\mathbf{u}$ con cada uno de los vectores de una base o de un sistema generador de W son nulos. Enunciamos esta propiedad a continuación:

Propiedad: El vector $\mathbf{u}$ es ortogonal al subespacio W generado por el sistema $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$ si y sólo si $\mathbf{u}\bullet\mathbf{w}_i = 0$ para $i = 1,\ldots,k$.

Demostración:

Está claro que si $\mathbf{u}$ es ortogonal al subespacio W, entonces el producto escalar de $\mathbf{u}$ por cualquier vector de W es 0, en particular, $\mathbf{u}\bullet\mathbf{w}_i = 0$ para $i = 1,\ldots,k$.

Para comprobar el recíproco, tomamos un vector cualquiera $\mathbf{w} \in W$ y lo escribimos como combinación lineal del sistema generador $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$: $\mathbf{w} = t_1\mathbf{w}_1 + \ldots + t_k\mathbf{w}_k$; y finalmente calculamos el producto escalar de $\mathbf{u}$ por $\mathbf{w}$:

$$\mathbf{u}\bullet\mathbf{w} = t_1(\mathbf{u}\bullet\mathbf{w}_1) + \ldots + t_k(\mathbf{u}\bullet\mathbf{w}_k) = 0,$$

que vale cero, como queríamos comprobar. □

Ejercicio 5: Calculamos todos los vectores que sean perpendiculares al subespacio W de $\mathbf{R}^4$ generado por los vectores $\mathbf{u} = (1,4,4,1)$ y $\mathbf{v} = (2,9,8,2)$.

Se requiere determinar todos los vectores $\mathbf{x} = (x_1,x_2,x_3,x_4)$ tales que

$\mathbf{x}\bullet\mathbf{w}=0$ para todo $\mathbf{w}\in W$. Por la propiedad anterior, el problema se reduce a calcular los vectores $\mathbf{x}$ que sean perpendiculares a los dos del sistema generador de W, es decir, a resolver el sistema $\begin{cases}\mathbf{x}\bullet\mathbf{u}=0\\ \mathbf{x}\bullet\mathbf{v}=0\end{cases}$, que escrito en forma matricial es $\underbrace{\begin{pmatrix}1&4&4&1\\2&9&8&2\end{pmatrix}}_{A}\underbrace{\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}}_{\mathbf{x}}=\begin{pmatrix}0\\0\end{pmatrix}$. Las soluciones son precisamente el espacio nulo de la matriz A. Una vez reducida esta matriz A a la forma escalonada: $U=\begin{pmatrix}1&4&4&1\\0&1&0&0\end{pmatrix}$ se deduce que el espacio nulo de A está generado por el sistema $\left\{\begin{pmatrix}-1\\0\\0\\1\end{pmatrix},\begin{pmatrix}-4\\0\\1\\0\end{pmatrix}\right\}$.

Por tanto, el conjunto de vectores ortogonales a W, o espacio nulo de la matriz A, está formado por los vectores $s(-1,0,0,1)+t(-4,0,1,0)$, donde s y t son números reales. Nótese que este conjunto de vectores es un subespacio de $\mathbf{R}^4$.

Ejercicio 6: Encuentra un vector ortogonal al espacio fila y un vector ortogonal al espacio columna de $A=\begin{pmatrix}1&2&1\\2&4&3\\3&6&4\end{pmatrix}$.

a. Para localizar un vector (x,y,z) ortogonal al espacio fila basta exigir que ese vector sea perpendicular a cada una de las filas de A, es decir, hay que determinar una solución del sistema homogéneo:

$$A\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}, \text{ es decir, } \begin{pmatrix}1&2&1\\2&4&3\\3&6&4\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

En otras palabras, el espacio nulo de A es el subespacio ortogonal al espacio fila de A.

La eliminación gaussiana conduce al sistema equivalente $\begin{pmatrix}1&2&1\\0&0&1\\0&0&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$, luego, el espacio nulo de la matriz A está generado

por el vector $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$.

En resumen, cuando el subespacio es $W = \mathcal{F}ila(A)$, con una base formada por $(1,2,1)$ y $(0,0,1)$, entonces todos los vectores ortogonales a W son los del subespacio nulo $\mathcal{N}(A)$, generado por $(-2,1,0)$.

b. Ahora miramos el espacio columna de A: los vectores ortogonales al espacio columna de A son los vectores (x,y,z) tales que $\begin{pmatrix} x & y & z \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$, que al trasponerlo, da lugar al sistema lineal $A^t\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, es decir, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$; cuya expresión es la habitual de los que ya hemos resuelto en otros temas y que nos informa de que el espacio nulo de A^t es el formado por todos los vectores ortogonales al espacio columna de A. Al aplicar de nuevo la eliminación gaussiana se llega al sistema equivalente: $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ y, por tanto, el espacio nulo de A^t está generado por el vector: $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

En resumen, si $W = \mathcal{C}ol(A)$, con una base formada por $(1,2,3)$ y $(1,3,4)$, entonces su subespacio ortogonal es $\mathcal{N}(A^t)$, generado por $(-1,-1,1)$.

Ejercicio 7: En el espacio euclídeo $\mathbf{R}^3$ determina el subespacio formado por todos los vectores perpendiculares al plano de ecuación $2x + y - z = 0$. Lo mismo para la recta de ecuaciones paramétricas $\begin{cases} x = 2t \\ y = -t \\ z = 3t \end{cases}$.

a. La ecuación que verifican las coordenadas de los vectores del plano $\Pi = \{(x,y,z):\ 2x + y - z = 0\}$ puede expresarse en la forma equivalente

$\begin{pmatrix} 2 & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$, en la que es fácil leer que los vectores del plano Π son perpendiculares al subespacio generado por $(2,1,-1)$.

Entonces el subespacio ortogonal a Π es la recta $\{t(2,1,-1): \quad t \in \mathbf{R}\}$, véase la figura 7.

b. En cuanto a los vectores perpendiculares a los de la recta $L: \begin{cases} x = 2t \\ y = -t \\ z = 3t \end{cases}$, nuestra intuición geométrica nos dice que ha de ser un plano y efectivamente así es. Como los vectores de la recta L están generados por $\mathbf{v} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$, al escribir la condición de perpendicularidad: $\underbrace{\begin{pmatrix} 2 & -1 & 3 \end{pmatrix}}_{\mathbf{v}^t}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$ hemos averiguado la ecuación general $2x - y + 3z = 0$ del plano perpendicular a la recta L. Si resolvemos esa ecuación lineal, conseguimos la base $\left\{\begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}\right\}$ de ese plano, es decir, del subespacio nulo de la matriz fila $(2,-1,3)$.

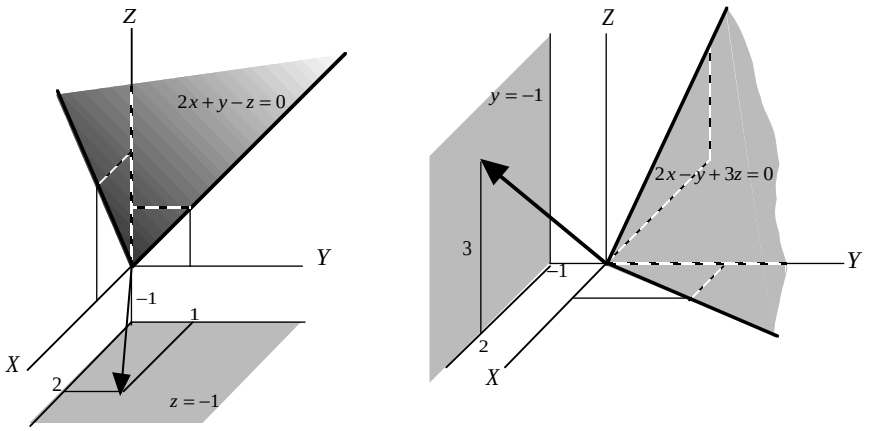


FIGURA 7: Representación de las figuras del ejercicio 7.

13.3. Sistemas ortogonales: el sistema trigonométrico

Sistemas de vectores mutuamente ortogonales

Un conjunto $\mathcal{S}=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ de vectores no nulos de un espacio euclídeo E se dice *sistema* o *conjunto ortogonal* si los vectores que lo integran son ortogonales dos a dos, es decir, verifican:

$$\mathbf{v}_i\cdot\mathbf{v}_j=0\text{, si } i\neq j.$$

Si, además, los vectores $\mathbf{v}_1,\ldots,\ \mathbf{v}_k$ son unitarios, el sistema $\mathcal{S}$ se llama *conjunto ortonormal.*

Por ejemplo, $\{\mathbf{i}=(1,0)\,,\mathbf{j}=(0,1)\}$ es un sistema ortogonal en $\mathbf{R}^2$ y también lo es el sistema $\{\mathbf{u}=(1,0,2,0)\,,\mathbf{v}=(2,3,-1,0)\,,\mathbf{w}=(0,0,0,4)\}$ en $\mathbf{R}^4$.

Los conjuntos ortogonales verifican la siguiente propiedad sobre la forma de expresar los elementos que están en el espacio por ellos generado.

Propiedad: Si $\mathcal{S}=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ es un conjunto ortogonal, de vectores no nulos, y $\mathbf{u}$ es un vector perteneciente al subespacio generado por $\mathcal{S}$, entonces $\mathbf{u}$ se expresa como combinación lineal de los vectores de $\mathcal{S}$ de la siguiente forma:

$$\mathbf{u}=t_1\mathbf{v}_1+\ldots+t_k\mathbf{v}_k=\frac{\mathbf{u}\bullet\mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1+\ldots+\frac{\mathbf{u}\bullet\mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k.$$

Demostración: Para determinar el primer coeficiente t_1, se calcula el producto escalar de cada miembro de la igualdad $\mathbf{u}=t_1\mathbf{v}_1+\ldots+t_k\mathbf{v}_k$ por el vector $\mathbf{v}_1$:

$$\mathbf{u}\bullet\mathbf{v}_1=(t_1\mathbf{v}_1+\ldots+t_k\mathbf{v}_k)\bullet\mathbf{v}_1=t_1(\mathbf{v}_1\bullet\mathbf{v}_1)+\ldots+t_k(\mathbf{v}_k\bullet\mathbf{v}_1)=t_1(\mathbf{v}_1\bullet\mathbf{v}_1),$$

donde la última igualdad es cierta debido a la ortogonalidad de los vectores del sistema $\mathcal{S}$. Como consecuencia de lo anterior y de que $\mathbf{v}_1\neq\mathbf{0}$, deducimos que

$$t_1=\frac{\mathbf{u}\bullet\mathbf{v}_1}{\|\mathbf{v}_1\|^2}.$$

Análogo razonamiento conduce a probar que los restantes coeficientes son los que se han enunciado en la propiedad. □

Los números $t_1 = \dfrac{\mathbf{u} \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}, \ldots,\ t_k = \dfrac{\mathbf{u} \bullet \mathbf{v}_k}{\|\mathbf{v}_k\|^2}$ se denominan *coeficientes de Fourier* de $\mathbf{u}$ respecto al conjunto ortogonal $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

No es ningún descubrimiento que el vector nulo está en cualquier subespacio; en particular, está en el subespacio generado por el sistema ortogonal $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, y claro está que ello obliga a que el vector nulo se escriba como combinación lineal de los elementos de $\mathcal{S}$ en la forma descrita en la anterior propiedad, es decir,

$$\mathbf{0} = \frac{\mathbf{0} \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \ldots + \frac{\mathbf{0} \bullet \mathbf{v}_k}{\|\mathbf{v}_k\|^2}\mathbf{v}_k = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_k .$$

En resumen, la única forma de combinar los vectores de un sistema ortogonal para producir el vector nulo es aquella en la que los coeficientes son nulos.

Las líneas precedentes contienen la demostración de la siguiente propiedad.

Propiedad: En un espacio euclídeo E, todo sistema ortogonal es linealmente independiente. En particular, en un espacio euclídeo E de dimensión n, todo sistema ortogonal que conste de n elementos es una base de E.

Definición: Sea E un espacio euclídeo de dimensión n. Un sistema ortonormal formado por n vectores se llama *base ortonormal* de E.

Sin duda alguna, las bases ortonormales más conocidas son las que usamos habitualmente en la geometría del plano y del espacio: las bases estándar. En ambos casos, estamos habituados a representar las coordenadas de un vector respecto a estas bases dibujando las proyecciones ortogonales sobre cada eje. Merece nuestra atención esa característica geométrica que recordamos en la figura 8, pues en breve estudiaremos las proyecciones ortogonales en cualquier espacio euclídeo y veremos que las coordenadas calculadas en la propiedad anterior son precisamente los coeficientes de las proyecciones sobre cada vector de la base ortogonal.

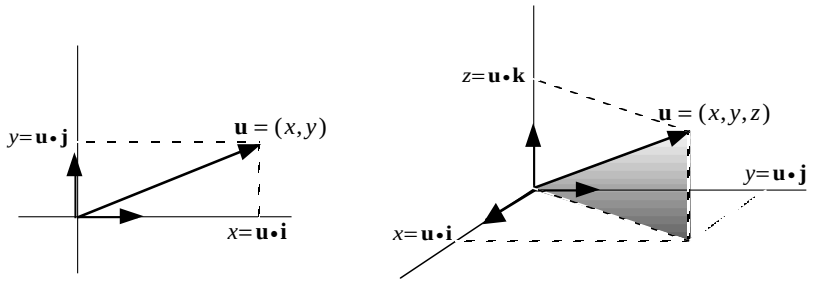


FIGURA 8: Las coordenadas respecto a una base ortogonal son las proyecciones sobre cada eje.

Ejercicio 8: Considera la base ortonormal (compruébese) de $\mathbf{R}^3$ formada por los vectores: $\mathbf{u}=\left(\frac{2}{3},\ \frac{2}{3},\ -\frac{1}{3}\right)$, $\mathbf{v}=\left(-\frac{1}{3},\ \frac{2}{3},\ \frac{2}{3}\right)$ y $\mathbf{w}=\left(\frac{2}{3},\ -\frac{1}{3},\ \frac{2}{3}\right)$. Escribe el vector $\mathbf{a}=2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$ como una combinación lineal de los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ (trata de usar dos métodos distintos).

El problema consiste en averiguar los valores de los coeficientes x, y y z que cumplen $\mathbf{a}=2\mathbf{i}+3\mathbf{j}+4\mathbf{k}=x\mathbf{u}+y\mathbf{v}+z\mathbf{w}$.

a. La igualdad $\mathbf{a}=2\mathbf{i}+3\mathbf{j}+4\mathbf{k}=x\mathbf{u}+y\mathbf{v}+z\mathbf{w}$ conduce al sistema lineal:

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix};$$

cuya solución se puede obtener por el método de eliminación de Gauss.

b. Mucho más fácil es utilizar el hecho de que la base dada es ortonormal y, por tanto, los coeficientes se calculan directamente por las fórmulas:

$$x=\mathbf{a}\bullet\mathbf{u},\ y=\mathbf{a}\bullet\mathbf{v},\ z=\mathbf{a}\bullet\mathbf{w}.$$

Se obtienen los siguientes valores: $x=2$, $y=4$ y $z=3$, luego

$$\mathbf{a}=2\mathbf{u}+4\mathbf{v}+3\mathbf{w}.$$

El sistema trigonométrico

Vamos a considerar en este apartado el espacio euclídeo $\mathcal{PC}(2\pi)$ de todas las funciones de período 2π, continuas en $[-\pi,\pi]$ salvo a lo sumo en un número finito de puntos de discontinuidad de salto finito. El producto escalar de este espacio euclídeo es el que ya definimos anteriormente:

$f \bullet g = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$ y la norma deducida: $\| f \| = \sqrt{f \bullet f} = \sqrt{\frac{1}{\pi}\int_{-\pi}^{\pi}[f(t)]^2\,dt}$.

Puesto que el producto escalar definido en $\mathcal{PC}(2\pi)$ exige el cálculo de integrales, recordamos algunas propiedades del Cálculo integral de gran utilidad en este contexto:

a. Si f es una función periódica de período 2π, entonces el valor de su integral es el mismo en cualquier intervalo de amplitud 2π:

$$\int_{-\pi}^{\pi} f(t)\,dt = \int_{a}^{a+2\pi} f(t)\,dt \ , \ \forall\, a \in \mathbf{R}.$$

b. Si f es una función par, entonces: $\int_{-a}^{a} f(t)\,dt = 2\int_{0}^{a} f(t)\,dt, \ \forall\, a \in \mathbf{R}$.

c. Si f es una función impar, entonces: $\int_{-a}^{a} f(t)\,dt = 0 \ , \ \forall\, a \in \mathbf{R}$.

Los vectores más destacados del espacio vectorial $\mathcal{PC}(2\pi)$ son las funciones $c_1(t) = \cos t$, $s_1(t) = \sin t$ y todas las de la forma $c_n(t) = \cos(nt)$ y $s_n(t) = \sin(nt)$ para $n = 1,2,3,\ldots$

Afirmamos que, con el producto escalar $f \bullet g = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$, ***el conjunto de funciones*** $\{1, \cos t, \sin t, \cos 2t, \sin 2t, \ldots, \cos nt, \sin nt\}$ ***es un sistema ortogonal y, además, tienen norma uno salvo la primera.***

Vamos a comprobar la ortogonalidad de algunas de ellas:

a. La función constante 1 es ortogonal a las demás:

$$1 \bullet c_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos nt\,dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} \cos nt\,dt = \frac{2}{\pi}\left[-\frac{\sin nt}{n}\right]_0^{\pi} = 0;$$

$$1 \bullet s_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin nt\,dt = \text{(función impar)} = 0.$$

b. La función $c_1(t) = \cos t$ es ortogonal a las demás:

— para cualquier valor de n: $c_1 \bullet s_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \sin nt\,dt = \text{(función impar)} = 0;$

— y siempre que $n \neq 1$:

$$c_1 \bullet c_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos t \cos nt\, dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} \cos t \cos nt\, dt$$

$$= \frac{1}{\pi}\int_{0}^{\pi} \left[\cos(n+1)t + \cos(n-1)t\right] dt = \frac{1}{\pi}\left[-\frac{\sin(n+1)t}{n+1} - \frac{\sin(n-1)t}{n-1}\right]_0^{\pi} = 0.$$

c. La función $s_1(t) = \sin t$ es ortogonal a las demás:

— para cualquier valor de n: $s_1 \bullet c_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin t \cos nt\, dt = \text{(función impar)} = 0$;

— y siempre que $n \neq 1$:

$$s_1 \bullet s_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin t \sin nt\, dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} \sin t \sin nt\, dt$$

$$= \frac{1}{\pi}\int_{0}^{\pi} \left[\cos(n+1)t - \cos(n-1)t\right] dt = \frac{1}{\pi}\left[-\frac{\sin(n+1)t}{n+1} + \frac{\sin(n-1)t}{n-1}\right]_0^{\pi} = 0.$$

Y así sucesivamente.

Calculamos ahora las longitudes o normas de las funciones del sistema trigonométrico $\{1, \cos t, \sin t, \cos 2t, \sin 2t, \ldots, \cos nt, \sin nt\}$:

a. La norma de la función constante $f(t) = 1$, $\forall\, t$, es $\sqrt{2}$, puesto que

$$\| f \|^2 = f \bullet f = \frac{1}{\pi}\int_{-\pi}^{\pi} [f(t)]^2\, dt = \frac{1}{\pi}\int_{-\pi}^{\pi} dt = 2.$$

b. Para cualquier número natural n, la norma de las funciones $c_n(t) = \cos(nt)$ es

$$\| c_n \|^2 = c_n \bullet c_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2 nt\, dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} \cos^2 nt\, dt =$$

$$= \frac{1}{\pi}\int_{0}^{\pi} (1 + \cos 2nt)\, dt = \frac{1}{\pi}\left[t + \frac{\sin 2nt}{2n}\right]_0^{\pi} = 1.$$

c. Para cualquier valor natural de n, la norma de las funciones $s_n(t) = \sin(nt)$ es

$$\| s_n \|^2 = s_n \bullet s_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2 nt\, dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} \sin^2 nt\, dt =$$

$$= \frac{1}{\pi}\int_{0}^{\pi} (1 - \cos 2nt)\, dt = \frac{1}{\pi}\left[t - \frac{\sin 2nt}{2n}\right]_0^{\pi} = 1.$$

13.4. Matrices ortogonales

Fijamos nuestra atención en el espacio euclídeo $\mathbf{R}^m$ y en las matrices A, de tamaño $m \times n$, cuyas columnas forman sistemas ortogonales en $\mathbf{R}^m$. No se trata de una idea caprichosa, todo lo contrario, a partir de ella conseguiremos las matrices ortogonales, que juegan un papel importante en la resolución de sistemas lineales.

Si las columnas de una matriz A, de tamaño $m \times n$, son ortogonales, entonces el producto $A^t A$ es una matriz diagonal de orden n, siendo los elementos diagonales el cuadrado de las longitudes de las columnas de A. Es decir,

$$A = \begin{pmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{pmatrix}, \text{ entonces } A^t A = \begin{pmatrix} \|\mathbf{c}_1\|^2 & 0 & \cdots & 0 \\ 0 & \|\mathbf{c}_2\|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|\mathbf{c}_n\|^2 \end{pmatrix}.$$

Como consecuencia, si las columnas de A son ortonormales, entonces

$$A^t A = I_n.$$

Es muy importante que no pase desapercibido el siguiente hecho: «Una matriz A, de tamaño $m \times n$, cuyas columnas son ortogonales, debe tener un número de columnas menor o igual que el de filas: $n \leq m$». ¿Sabe el lector por qué hacemos esta afirmación tan contundente? Si no consigue averiguarlo, conviene que repase el apartado sobre sistemas ortogonales.

Definimos ya las matrices ortogonales.

Definición: Una matriz A se dice *ortogonal* si es cuadrada y sus columnas son ortonormales.

Propiedad: Una matriz cuadrada A es ortogonal si y sólo si $A^t A = I_n$. En este caso, no sólo $A^t A = I_n$, sino que además $AA^t = I_n$.

Demostración: Sólo necesita demostración la igualdad $AA^t = I_n$.

Al ser n el rango de la matriz cuadrada A, ya que las n columnas de A son linealmente independientes, también han de ser linealmente independientes las n filas de A. En consecuencia, A tiene inversa y, como la inversa es única, esa inversa debe ser A^t, la inversa a izquierda de A. □

Observaciones: Toda matriz ortogonal A tiene las siguientes características:

1. A es inversible y la inversa A^{-1} es su traspuesta: $A^{-1} = A^t$.
2. Las filas de A también son vectores ortonormales, pues también $AA^t = I_n$.

Como consecuencia de que las matrices ortogonales son inversibles y su inversa es la traspuesta, los sistemas lineales $Q\mathbf{x} = \mathbf{b}$ cuya matriz de coeficientes sea una matriz ortogonal son bien sencillos de resolver, pues la solución $\mathbf{x} = Q^t\mathbf{b}$ se consigue mediante una sencilla multiplicación de la matriz Q^t por el vector de los términos independientes $\mathbf{b}$.

Ejercicio 9: Encuentra una tercera columna de modo que la matriz $Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & ? \\ \frac{1}{\sqrt{3}} & 0 & ? \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & ? \end{pmatrix}$ sea ortogonal, ¿cuánta libertad hay para elegir esa tercera columna? Verifica que las filas también son ortonormales.

a. La tercera columna $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ de Q debe cumplir las condiciones de ortogonalidad y de longitud que se escriben en el siguiente sistema (no lineal):

$$\begin{cases} \frac{1}{\sqrt{3}}a + \frac{1}{\sqrt{3}}b + \frac{1}{\sqrt{3}}c = 0 \\ \frac{1}{\sqrt{2}}a \qquad - \frac{1}{\sqrt{2}}c = 0 \\ a^2 + \quad b^2 + \quad c^2 = 1 \end{cases}.$$

Hay dos soluciones: $\begin{cases} a = \frac{1}{\sqrt{6}} & b = -\frac{2}{\sqrt{6}} & c = \frac{1}{\sqrt{6}} \\ a = -\frac{1}{\sqrt{6}} & b = \frac{2}{\sqrt{6}} & c = -\frac{1}{\sqrt{6}} \end{cases}$.

b. Es fácil comprobar que las filas de $Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \pm\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \mp\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \pm\frac{1}{\sqrt{6}} \end{pmatrix}$ son ortonormales.

Ejercicio 10: Considera el sistema lineal $\begin{cases} -\frac{3}{\sqrt{14}}x+\frac{1}{\sqrt{6}}y-\frac{2}{\sqrt{21}}z=1 \\ \frac{2}{\sqrt{14}}x+\frac{2}{\sqrt{6}}y-\frac{1}{\sqrt{21}}z=2 \\ -\frac{1}{\sqrt{14}}x+\frac{1}{\sqrt{6}}y+\frac{4}{\sqrt{21}}z=3 \end{cases}$.

a. La matriz de coeficientes Q es ortogonal. La comprobación es directa, sin dificultades.

b. Resuelve el sistema: la única solución del sistema es

$$Q\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}=Q^t\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}=\begin{pmatrix} -\frac{3}{\sqrt{14}} & \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{14}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{21}} & -\frac{1}{\sqrt{21}} & \frac{4}{\sqrt{21}} \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Rightarrow \begin{cases} x=-\frac{2}{\sqrt{14}} \\ y=\frac{8}{\sqrt{6}} \\ z=\frac{8}{\sqrt{21}} \end{cases}.$$

Ejercicio 11: Comprueba que si A es una matriz ortogonal de orden n, entonces

1. El producto escalar de dos vectores cualesquiera $\mathbf{u},\mathbf{v}\in\mathbf{R}^n$ no varía al transformarlos en $A\mathbf{u}$ y $A\mathbf{v}$, es decir, $A\mathbf{u}\bullet A\mathbf{v}=\mathbf{u}\bullet\mathbf{v}$.
2. La longitud de un vector $\mathbf{u}\in\mathbf{R}^n$ no cambia al ser multiplicado por A, es decir, $\|A\mathbf{u}\|=\|\mathbf{u}\|$.

En efecto,

— Primero el producto escalar:

$$A\mathbf{u}\bullet A\mathbf{v}=(A\mathbf{u})^t A\mathbf{v}=\mathbf{u}^t\left(A^t A\mathbf{v}\right)=(A \text{ es ortogonal})=\mathbf{u}^t\mathbf{v}=\mathbf{u}\bullet\mathbf{v}.$$

— Ahora las longitudes:

$$\|A\mathbf{u}\|^2=A\mathbf{u}\bullet A\mathbf{u}=(\text{apartado anterior})=\mathbf{u}\bullet\mathbf{u}=\|\mathbf{u}\|^2.$$

Las matrices ortogonales de orden dos, como las de orden tres, son especialmente interesantes, pues describen las aplicaciones lineales del plano, o del espacio, en sí mismo que conservan longitudes y ángulos (véase el ejercicio 11). En el siguiente ejercicio echamos una ojeada a las matrices ortogonales de orden dos.

Ejercicio 12: Comprobemos que toda matriz ortogonal de orden 2 es necesariamente de la forma $G=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, siendo α un número real, o de la forma $S=\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ con $a^2+b^2=1$.

Si exigimos que la matriz $A=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ sea ortogonal, debemos imponer que sus entradas verifiquen las siguientes condiciones: $\begin{cases} ac+bd=0 \\ a^2+b^2=1 \\ c^2+d^2=1 \end{cases}$. De ellas, se deduce que $\begin{cases} d=\pm a \\ c=\mp b \\ a^2+b^2=1 \end{cases}$ y, en definitiva, que sólo hay dos tipos de matrices ortogonales de orden dos:

a. La primera es $G=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, con α número real cualquiera, que es la matriz asociada a la rotación del plano de ángulo α.

b. La segunda es $S=\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$, con $a^2+b^2=1$, y describe una simetría ortogonal. El eje de la simetría es la recta: $\begin{cases} x=bt \\ y=(1-a)t \end{cases}$.

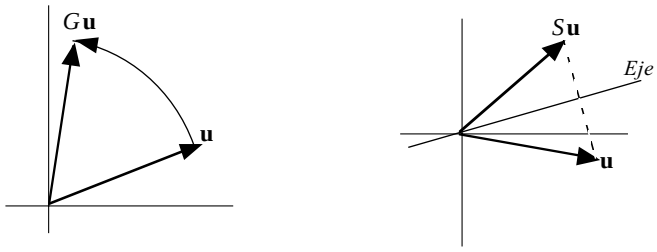


FIGURA 9: Giro y simetría en el plano.

13.5. Proyecciones ortogonales

Proyección ortogonal sobre un vector

Sea E un espacio vectorial euclídeo y $\mathbf{u}$ y $\mathbf{v}$ dos vectores de E distintos de $\mathbf{0}$. La *proyección ortogonal* de $\mathbf{u}$ sobre $\mathbf{v}$ (a lo largo de $\mathbf{v}$) es un vector $k\mathbf{v}$ con la misma dirección que $\mathbf{v}$ de modo que $\mathbf{u}-k\mathbf{v}$ sea perpendicular a $\mathbf{v}$.

Calculemos el valor del coeficiente k de acuerdo con la condición de perpendicularidad exigida. De la siguiente cadena de equivalencias:

$$(\mathbf{u}-k\mathbf{v})\perp\mathbf{v} \quad \Leftrightarrow \quad (\mathbf{u}-k\mathbf{v})\bullet\mathbf{v}=0 \quad \Leftrightarrow \quad \mathbf{u}\bullet\mathbf{v}-k(\mathbf{v}\bullet\mathbf{v})=0 \quad \Leftrightarrow \quad k=\frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}$$

se deduce que el vector proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$, que se denota por $\mathbf{p}=\mathrm{Proy}_{\mathbf{v}}\mathbf{u}$, es $\mathbf{p}=k\mathbf{v}=\dfrac{\mathbf{u}\bullet\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v}$.

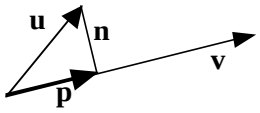


FIGURA 10: Proyección del vector $\mathbf{u}$ en la dirección de $\mathbf{v}$.

Si ponemos $\mathbf{n}=\mathbf{u}-\mathbf{p}$, es sencillo observar que: « $\mathbf{u}$ se descompone en la suma de dos vectores ortogonales entre sí: $\mathbf{u}=\mathbf{p}+\mathbf{n}$, con $\mathbf{p}$ en la dirección de $\mathbf{v}$ y $\mathbf{n}$ perpendicular a $\mathbf{v}$».

En los dos ejercicios siguientes trabajamos en el espacio euclídeo $\mathbf{R}^n$ con el producto escalar estándar.

Ejercicio 13: Considera los vectores $\mathbf{a}=(2,5)$ y $\mathbf{b}=(4,-1)$ del plano. Calcula la proyección ortogonal de $\mathbf{a}$ sobre $\mathbf{b}$.

El vector $\mathbf{p}=\mathrm{Proy}_{\mathbf{b}}\mathbf{a}$, proyección de $\mathbf{a}$ sobre $\mathbf{b}$, es

$$\mathbf{p}=\frac{\mathbf{a}\bullet\mathbf{b}}{\|\mathbf{b}\|^2}\mathbf{b}=\frac{3}{17}(4,-1)=\left(\frac{12}{17},-\frac{3}{17}\right)\approx(0.7,-0.18).$$

Esta proyección nos da la siguiente información:

— Por un lado, $\mathbf{a}$ se descompone en *suma de dos vectores perpendiculares:* $\mathbf{a}=\mathbf{p}+(\mathbf{a}-\mathbf{p})=\mathbf{p}+\mathbf{n}$, donde $\mathbf{n}=\left(\frac{22}{17},\ \frac{88}{17}\right)\approx(1.3,\ 5.18)$.

— Por otro lado, *la distancia mínima* de $\mathbf{a}$ a la recta de dirección $\mathbf{b}$ se alcanza en $\mathbf{p}$ y es $\|\mathbf{a}-\mathbf{p}\|=\|\mathbf{n}\|=\sqrt{\left(\frac{22}{17}\right)^2+\left(\frac{88}{17}\right)^2}\approx 5.34$.

Ejercicio 14: Sea $\mathbf{F}=(1,1,1)$ un vector fuerza que mueve una masa puntual a lo largo de la recta $r:\begin{cases} x=0 \\ y=t \\ z=2t \end{cases}$, donde $t\in\mathbf{R}$. Calculamos la proyección de $\mathbf{F}$ sobre la recta r.

El vector proyección de $\mathbf{F}$ sobre r lo llamamos $\mathbf{T}$, entonces

$$\mathbf{T}=\frac{\mathbf{F}\bullet\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v}=\frac{3}{5}(0,1,2)=\left(0,\ \frac{3}{5},\ \frac{6}{5}\right)=(0,\ 0.6,\ 1.2).$$

La descomposición en suma de vectores ortogonales es

$$\mathbf{F}=\mathbf{T}+\mathbf{N}=(0,\ 0.6,\ 1.2)+(1,\ 0.4,\ -0.2).$$

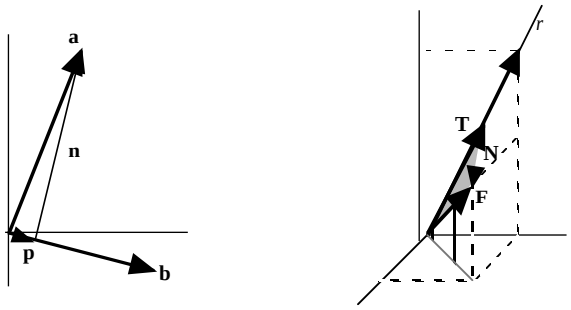


FIGURA 11: A la izquierda, la proyección del vector $\mathbf{a}=(2,5)$ en la dirección de $\mathbf{b}=(4,-1)$ y, a la derecha, la proyección de $\mathbf{F}=(1,1,1)$ sobre la recta r.

Proyección ortogonal de un vector sobre un subespacio de dimensión finita

Acabamos de ver que en un espacio euclídeo E es posible expresar cualquier vector $\mathbf{u}$, distinto de $\mathbf{0}$, como una suma de dos vectores ortogonales:

$$\mathbf{u}=\mathbf{p}+\mathbf{n},$$

donde $\mathbf{p}$ es la proyección sobre un vector $\mathbf{v}\neq\mathbf{0}$ fijado de antemano. Además averiguamos que: $\mathbf{p}=\dfrac{\mathbf{u}\bullet\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v}$ y $\mathbf{n}=\mathbf{u}-\mathbf{p}$.

Ahora, en lugar de proyectar un vector sobre otro, veamos cómo podemos proyectar un vector sobre un subespacio de dimensión finita. Conseguiremos ampliar fácilmente las ideas estudiadas antes acerca de proyecciones sobre un

vector si suponemos que podemos elegir una base ortogonal en el subespacio de proyección. Esta hipótesis no supone ninguna pérdida de generalidad porque siempre existe una base ortogonal en un espacio euclídeo de dimensión finita, aunque quizás sea necesario trabajar un poco para conseguir una.

Suponemos que E es un espacio euclídeo y que W es un subespacio de dimensión finita y $\mathcal{S}=\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$ una de sus bases ortogonales.

La *proyección ortogonal* del vector $\mathbf{u}$ *sobre el subespacio* W es el único vector $\mathbf{p}\in W$ que verifica la condición $(\mathbf{u}-\mathbf{p})\perp W$. Esta condición significa que ha de verificarse $(\mathbf{u}-\mathbf{p})\bullet\mathbf{w}_i=0$, para $i=1,\ldots,k$. Es decir,

$$0=[\mathbf{u}-(t_1\mathbf{w}_1+\ldots+t_k\mathbf{w}_k)]\bullet\mathbf{w}_i=\mathbf{u}\bullet\mathbf{w}_i-t_i(\mathbf{w}_i\bullet\mathbf{w}_i),$$

de donde se deduce que los coeficientes o coordenadas de $\mathbf{p}$ respecto a la base $\mathcal{S}$ son $t_i=\dfrac{\mathbf{u}\bullet\mathbf{w}_i}{\mathbf{w}_i\bullet\mathbf{w}_i}=\dfrac{\mathbf{u}\bullet\mathbf{w}_i}{\|\mathbf{w}_i\|^2}$, para $i=1,\ldots,k$. En resumen, el vector proyección de $\mathbf{u}$ sobre W que denotamos $\mathbf{p}=\mathrm{Proy}_W\mathbf{u}$ es

$$\mathbf{p}=\frac{\mathbf{u}\bullet\mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1+\ldots+\frac{\mathbf{u}\bullet\mathbf{w}_k}{\|\mathbf{w}_k\|^2}\mathbf{w}_k;$$

el vector diferencia $\mathbf{n}=\mathbf{u}-\mathbf{p}$ se llama *componente de* $\mathbf{u}$ *ortogonal a* W.

Los números $t_i=\dfrac{\mathbf{u}\bullet\mathbf{w}_i}{\|\mathbf{w}_i\|^2}$ son los coeficientes de Fourier de $\mathbf{u}$ respecto de la base ortogonal $\mathcal{S}$ del subespacio W.

De nuevo, hemos expresado un vector $\mathbf{u}$ cualquiera como suma de dos vectores ortogonales: $\mathbf{u}=\mathbf{p}+\mathbf{n}$, donde $\mathbf{p}\in W$ y $\mathbf{n}=\mathbf{u}-\mathbf{p}\perp W$.

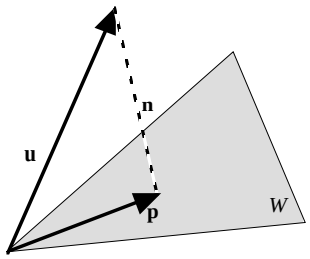


FIGURA 12: Proyección de un vector geométrico sobre un subespacio de dimensión finita.

Ejercicio 15: Proyectamos el vector $\mathbf{b}=(0,0,3)$ sobre cada uno de los vectores

$$\mathbf{a}_1=\left(\frac{2}{3},\ \frac{2}{3},\ -\frac{1}{3}\right),\quad \mathbf{a}_2=\left(-\frac{1}{3},\ \frac{2}{3},\ \frac{2}{3}\right) \text{ y } \mathbf{a}_3=\left(\frac{2}{3},\ -\frac{1}{3},\ \frac{2}{3}\right).$$

Como los tres vectores son de longitud uno:

$$\text{Proy}_{\mathbf{a}_1}\mathbf{b}=\mathbf{p}_1=(\mathbf{b}\bullet\mathbf{a}_1)\mathbf{a}_1=\left(\mathbf{b}^t\mathbf{a}_1\right)\mathbf{a}_1=(-1)\mathbf{a}_1,$$
$$\text{Proy}_{\mathbf{a}_2}\mathbf{b}=\mathbf{p}_2=(\mathbf{b}\bullet\mathbf{a}_2)\mathbf{a}_2=\left(\mathbf{b}^t\mathbf{a}_2\right)\mathbf{a}_2=2\mathbf{a}_2,$$
$$\text{Proy}_{\mathbf{a}_3}\mathbf{b}=\mathbf{p}_3=(\mathbf{b}\bullet\mathbf{a}_3)\mathbf{a}_3=\left(\mathbf{b}^t\mathbf{a}_3\right)\mathbf{a}_3=2\mathbf{a}_3.$$

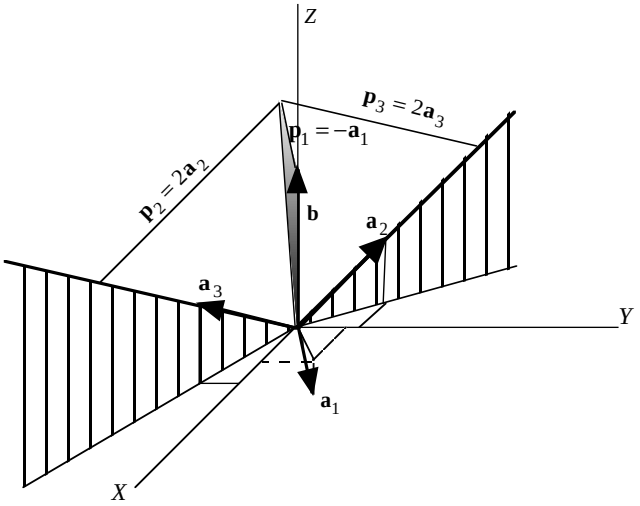


FIGURA 13: Proyección del vector geométrico $\mathbf{b}$ sobre cada uno de los vectores dados.

a. ¿Cuál es la proyección de $\mathbf{b}$ sobre el plano Π generado por $\mathbf{a}_1$ y $\mathbf{a}_2$?

Al ser $\mathbf{a}_1$ y $\mathbf{a}_2$ ortogonales: $\text{Proy}_{\Pi}\mathbf{b}=\mathbf{p}_1+\mathbf{p}_2=(-1)\mathbf{a}_1+2\mathbf{a}_2=\left(-\frac{4}{3},\ \frac{2}{3},\ \frac{5}{3}\right).$

b. ¿Cuál es la suma de las tres proyecciones unidimensionales?

La suma de las tres proyecciones ha de ser el vector $\mathbf{b}$, pues $\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\}$ es una base de $\mathbf{R}^3$. Comprobémoslo:

$$\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3=-\mathbf{a}_1+2\mathbf{a}_2+2\mathbf{a}_3=(0,0,3)=\mathbf{b}.$$

c. ¿Por qué $P = \mathbf{a}_1\mathbf{a}_1^t + \mathbf{a}_2\mathbf{a}_2^t + \mathbf{a}_3\mathbf{a}_3^t$ es igual a la matriz identidad de orden tres?

Observemos que $\left(\mathbf{a}_1\mathbf{a}_1^t\right)\mathbf{b} = \mathbf{a}_1\left(\mathbf{a}_1^t\mathbf{b}\right) = \mathbf{a}_1\left(\mathbf{a}_1 \bullet \mathbf{b}\right) = \mathbf{p}_1 = \text{Proy}_{\mathbf{a}_1}\mathbf{b}$, y del mismo modo los otros dos sumandos:

$$\left(\mathbf{a}_2\mathbf{a}_2^t\right)\mathbf{b} = \mathbf{p}_2 = \text{Proy}_{\mathbf{a}_2}\mathbf{b} \quad \text{y} \quad \left(\mathbf{a}_3\mathbf{a}_3^t\right)\mathbf{b} = \mathbf{p}_3 = \text{Proy}_{\mathbf{a}_3}\mathbf{b}.$$

Por tanto,

$$P\mathbf{b} = \left(\mathbf{a}_1\mathbf{a}_1^t + \mathbf{a}_2\mathbf{a}_2^t + \mathbf{a}_3\mathbf{a}_3^t\right)\mathbf{b} = \mathbf{a}_1\mathbf{a}_1^t\mathbf{b} + \mathbf{a}_2\mathbf{a}_2^t\mathbf{b} + \mathbf{a}_3\mathbf{a}_3^t\mathbf{b} = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = \mathbf{b},$$

para todo $\mathbf{b}$. En consecuencia, $P = I_3$.

13.6. La aproximación óptima de un vector a un subespacio de dimensión finita: las aproximaciones de Fourier

La aproximación óptima a un subespacio de dimensión finita

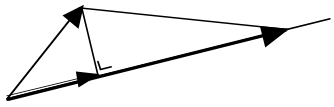
En la siguiente propiedad enunciamos y demostramos que la proyección ortogonal de un vector sobre una recta resuelve el problema de determinar cuál es la *mínima distancia* de ese vector a la recta.

Propiedad de la mejor aproximación: En el espacio euclídeo E consideramos dos vectores $\mathbf{u}$ y $\mathbf{v}$, este último no nulo, entonces el vector $\mathbf{p}$, proyección ortogonal de $\mathbf{u}$ sobre $\mathbf{v}$, verifica:

$$\|\mathbf{u} - \mathbf{p}\| \leq \|\mathbf{u} - t\mathbf{v}\|, \text{ para todo escalar } t;$$

es decir, $\mathbf{p}$ es el vector de la dirección de $\mathbf{v}$ que está más próximo a $\mathbf{u}$.

***Demostración*:** Calculemos la longitud de $\mathbf{u} - t\mathbf{v}$ mediante el teorema de Pitágoras, es decir, descomponiéndolo en suma de vectores perpendiculares:



$$\mathbf{u} - t\mathbf{v} = (\mathbf{u} - \mathbf{p}) + (\mathbf{p} - t\mathbf{v}) = \mathbf{n} + \underbrace{(\mathbf{p} - t\mathbf{v})}_{\text{Dirección } \mathbf{v}},$$

donde $\mathbf{n} \perp (\mathbf{p} - t\mathbf{v})$.

En consecuencia, $\|\mathbf{u}-t\mathbf{v}\|^2=\|\mathbf{n}\|^2+\|\mathbf{p}-t\mathbf{v}\|^2=\|\mathbf{u}-\mathbf{p}\|^2+\|\mathbf{p}-t\mathbf{v}\|^2\geq\|\mathbf{u}-\mathbf{p}\|^2$ de donde se deduce lo que queríamos demostrar: $\mathbf{p}$ es el vector de la dirección de $\mathbf{v}$ más cercano a $\mathbf{u}$. □

Del mismo modo que la proyección ortogonal sobre un vector conduce a resolver el problema de la mínima distancia de un vector a la recta generada por otro, la proyección ortogonal sobre un subespacio resuelve un problema de minimizar distancias que se enuncia con precisión en la siguiente propiedad.

Propiedad de mejor aproximación: Sea E un espacio vectorial euclídeo y W un subespacio de dimensión finita.
La mínima distancia de un vector $\mathbf{u}\in E$ al subespacio W es la longitud del vector $\mathbf{u}-\mathrm{Proy}_W\mathbf{u}$, es decir, de todos los vectores de W el más cercano a $\mathbf{u}$ es $\mathrm{Proy}_W\mathbf{u}$.
Simbólicamente se expresa así, el vector $\mathrm{Proy}_W\mathbf{u}$ es el único de W que verifica

$$\|\mathbf{u}-\mathrm{Proy}_W\mathbf{u}\|\leq\|\mathbf{u}-\mathbf{w}\|, \text{ para todo } \mathbf{w}\in W.$$

Demostración: Para simplificar la notación escribimos $\mathbf{p}=\mathrm{Proy}_W\mathbf{u}$ y $\mathbf{n}=\mathbf{u}-\mathrm{Proy}_W\mathbf{u}$ y recordamos que: $\mathbf{u}=\mathbf{p}+\mathbf{n}$, donde $\mathbf{n}$ es perpendicular a W.

Sea $\mathbf{w}\in W$, al igual que en el caso en que W era subespacio unidimensional, expresamos el vector diferencia $\mathbf{u}-\mathbf{w}$ como suma de dos vectores perpendiculares $\mathbf{u}-\mathbf{w}=\underbrace{(\mathbf{u}-\mathbf{p})}_{\mathbf{n}}+(\mathbf{p}-\mathbf{w})$ y, por el teorema de Pitágoras, tenemos

$$\|\mathbf{u}-\mathbf{w}\|^2=\|\mathbf{u}-\mathbf{p}\|^2+\|\mathbf{p}-\mathbf{w}\|^2\geq\|\mathbf{u}-\mathbf{p}\|^2,$$

es decir, $\|\mathbf{u}-\mathrm{Proy}_W\mathbf{u}\|^2\leq\|\mathbf{u}-\mathbf{w}\|^2$.

Para comprobar que $\mathrm{Proy}_W\mathbf{u}$ es el único vector de W en el que se alcanza la mínima distancia del vector $\mathbf{u}\in E$ al subespacio W, suponemos que hay otro vector de W, digamos $\mathbf{v}$, que también verifica $\|\mathbf{u}-\mathbf{v}\|\leq\|\mathbf{u}-\mathbf{w}\|$, para todo $\mathbf{w}\in W$; el objetivo es deducir que necesariamente $\mathbf{v}=\mathrm{Proy}_W\mathbf{u}$. Hemos supuesto que la mínima distancia de $\mathbf{u}\in E$ al subespacio W admite dos expresiones equivalentes: $\|\mathbf{u}-\mathbf{v}\|=\|\mathbf{u}-\mathrm{Proy}_W\mathbf{u}\|$, le proponemos al lector que utilice esa igualdad junto con la desigualdad triangular para deducir que $\|\mathbf{v}-\mathrm{Proy}_W\mathbf{u}\|=0$, es decir, $\mathbf{v}=\mathrm{Proy}_W\mathbf{u}$. □

El haber conseguido dar respuesta, en un espacio euclídeo, al problema de averiguar la mejor aproximación de un vector a un subespacio de dimensión finita (problema de minimos cuadrados) tiene diferentes aplicaciones en contextos muy diversos. Mencionamos dos.

En las ciencias experimentales surgen con cierta frecuencia sistemas lineales $A\mathbf{x}=\mathbf{b}$, con más ecuaciones que incógnitas, que, en general, son incompatibles debido a los errores en las mediciones; ahora bien, el problema lineal ha de tener solución, lo cual se consigue proponiendo como mejor elección aquella en que los valores del vector de incógnitas $\mathbf{x}$ sean los que hacen que $A\mathbf{x}$ esté lo más cerca posible de $\mathbf{b}$.

Por otro lado, aproximar una función continua a trozos y periódica mediante *polinomios trigonométricos* es utilizado con frecuencia en los campos de la Técnica en los que se necesita usar oscilaciones diferentes de las sinusoidales.

Dedicamos más esfuerzo al problema de las aproximaciones de Fourier en el siguiente apartado de esta sección y damos tan solo una ligera idea del problema de mínimos cuadrados aplicado a un sistema incompatible en el siguiente ejercicio.

Ejercicio 16: El sistema lineal $\begin{cases} \frac{2}{3}x-\frac{1}{3}y=0 \\ \frac{2}{3}x+\frac{2}{3}y=0 \\ -\frac{1}{3}x+\frac{2}{3}y=1 \end{cases}$, de más ecuaciones que incógnitas, es incompatible, es decir, el vector de términos independientes $\mathbf{b}=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ no está en el espacio columna de la matriz de coeficientes, es decir, no está en el plano generado por $\begin{pmatrix} 2/3 \\ 2/3 \\ -1/3 \end{pmatrix}$ y $\begin{pmatrix} -1/3 \\ 2/3 \\ 2/3 \end{pmatrix}$. La cuestión es averiguar cuál es el vector del espacio columna más próximo al vector de términos independientes. La respuesta es el vector proyección de $\mathbf{b}$ sobre $\mathcal{C}ol(A)$.

Si denotamos mediante A la matriz de coeficientes, matriz de tamaño 3×2, se pide

a. ¿Cuál es la proyección del vector de términos independientes $\mathbf{b}$ sobre el espacio columna de A? Como las dos columnas de A son ortonormales, la proyección requerida es

$$\text{Proy}_{\mathcal{col}(A)}\mathbf{b} = (\mathbf{b}\bullet\mathbf{c}_1)\mathbf{c}_1 + (\mathbf{b}\bullet\mathbf{c}_2)\mathbf{c}_2 = -\frac{1}{3}\mathbf{c}_1 + \frac{2}{3}\mathbf{c}_2 = \left(-\frac{4}{9},\ \frac{2}{9},\ \frac{5}{9}\right).$$

b. ¿Cuáles son los mejores valores que podríamos asignar a x e y? (mejores en el sentido de que minimicen la distancia del término independiente al espacio columna de A).
La mejor solución es precisamente el vector proyección acabado de obtener. Es decir, los mejores valores que podemos dar a las incógnitas son $x = -\frac{1}{3}$ e $y = \frac{2}{3}$, pues con ellos se obtiene el vector $\left(-\frac{4}{9},\ \frac{2}{9},\ \frac{5}{9}\right)$ del espacio columna de A más cercano a $\mathbf{b} = (0,0,1)$.

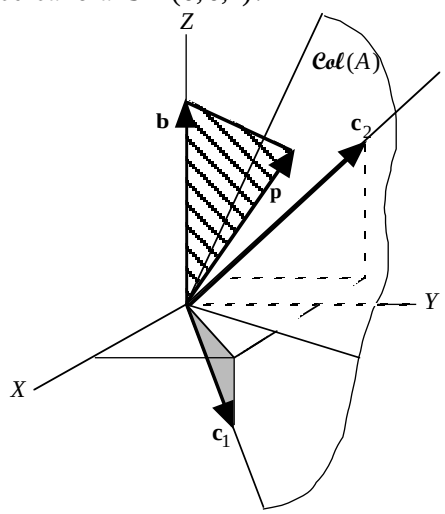


FIGURA 15: Representación del espacio columna de A, del vector $\mathbf{b}$ y de su proyección sobre ese subespacio.

Como hemos dicho, este último ejercicio no es más que la descripción, en términos muy sencillos, del problema de construir una «solución razonable» a los sistemas lineales $A\mathbf{x} = \mathbf{b}$ incompatibles. Aunque no vamos a dedicarnos a estudiar el método de los mínimos cuadrados, método que calcula el vector $\mathbf{x}$ tal que la longitud de $A\mathbf{x} - \mathbf{b}$ sea mínima, sí queremos darle al lector una idea de las etapas del camino que lleva a conseguir ese vector $\mathbf{x}$.

Primero, un vector $\mathbf{x}$ tal que la longitud de $A\mathbf{x}-\mathbf{b}$ sea mínima es solución de $A^tA\mathbf{x}=A^t\mathbf{b}$, llamado *sistema de ecuaciones normales* que, por cierto, siempre es compatible.

Segundo, puede haber más de un vector $\mathbf{x}$ en las condiciones anteriores; de hecho, la unicidad de la solución es equivalente a que las columnas de A sean linealmente independientes. Por tanto, aparecen dos posibilidades: si las columnas de A forman un sistema libre, entonces la «mejor solución» del sistema incompatible $A\mathbf{x}=\mathbf{b}$ es la única solución del sistema normal $A^tA\mathbf{x}=A^t\mathbf{b}$; y, en caso contrario, hay que elegir una de las múltiples soluciones de $A^tA\mathbf{x}=A^t\mathbf{b}$, el criterio de elección suele variar según el contexto, pero podemos apuntar que uno de los más comúnmente utilizados es elegir aquella solución que tenga la longitud más pequeña.

Las aproximaciones de Fourier de funciones periódicas

Supongamos que f está en el espacio euclídeo $\mathcal{PC}(2\pi)$ de todas las funciones periódicas de período 2π, continuas en $[-\pi,\pi]$ salvo a lo sumo en un número finito de puntos de discontinuidad de salto finito.

Recuérdese que el producto escalar definido en el espacio euclídeo $\mathcal{PC}(2\pi)$ es $f\bullet g=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$ y que, con ese producto escalar, el sistema trigonométrico $\mathcal{S}=\{1,\cos t,\sin t,\cos 2t,\sin 2t,\ldots,\cos nt,\sin nt\}$ es un sistema ortogonal formado por funciones de norma uno salvo la primera cuya norma es $\sqrt{2}$.

El sistema trigonométrico $\mathcal{S}$ genera el subespacio de $\mathcal{PC}(2\pi)$ formado por los llamados *polinomios trigonométricos de orden* n, es decir, por las funciones de la forma:

$$p_n(t)=\frac{a_0}{2}+(a_1\cos t+b_1\sin t)+(a_2\cos 2t+b_2\sin 2t)+\ldots\ldots+(a_n\cos nt+b_n\sin nt).$$

El teorema de la aproximación óptima aplicado al espacio euclídeo $\mathcal{PC}(2\pi)$ y al sistema ortogonal descrito en las líneas anteriores dice: *«El polinomio trigonométrico de orden* n *que más se aproxima a* f *es la suma de las proyecciones de* f *sobre cada una de las funciones del sistema*

ortogonal $S=\{1,\cos t,\sin t,\cos 2t,\sin 2t,\ldots,\cos nt,\sin nt\}$». Dicho polinomio se llama *aproximación de Fourier de* f *de orden* n y es

$$F_n(t)=\frac{a_0}{2}+\underbrace{a_1\cos t+b_1\sin t}_{\text{armónico de frecuencia 1}}+\underbrace{a_2\cos 2t+b_2\sin 2t}_{\text{armónico de frecuencia 2}}+\ldots\ldots+\underbrace{a_n\cos nt+b_n\sin nt}_{\text{armónico de frecuencia } n}=$$

$$=\frac{a_0}{2}+\sum_{k=1}^{n}\left(\underbrace{a_k\cos kt+b_k\sin kt}_{\text{armónico de frecuencia } k}\right)$$

que, como se observa, es un polinomio trigonométrico de orden n, suma de armónicos de frecuencias $1,\ 2,\ldots,\ n$.

Sus coeficientes, que se llaman *coeficientes de Fourier de* f, se calculan aplicando la fórmula $\text{Proy}_g f=\dfrac{f\bullet g}{\|g\|^2}$ de los coeficientes de la proyección ortogonal. Resultan las siguientes expresiones:

$$a_0=f\bullet 1=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\,dt;$$

$$a_1=f\bullet c_1=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\cos t\,dt,\qquad b_1=f\bullet s_1=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\sin t\,dt,$$

$$a_2=f\bullet c_2=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\cos 2t\,dt,\qquad b_2=f\bullet s_2=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\sin 2t\,dt,$$

……

y, en general, $a_n=f\bullet c_n=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(t)\cos nt\,dt,\qquad b_n=f\bullet s_n=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(t)\sin nt\,dt.$

El polinomio trigonométrico $F_n(t)$ aproxima a f mejor que cualquier otro polinomio trigonométrico de orden n, en el sentido de que $\|f-F_n\|$ es la más pequeña posible: $\|f-F_n\|\le\|f-p_n\|$, para todo p_n polinomio trigonométrico de orden n.

La *serie de Fourier* de f es la suma infinita de todas las proyecciones:

$$\frac{a_0}{2}+(a_1\cos t+b_1\sin t)+(a_2\cos 2t+b_n\sin 2t)+\ldots\ldots=\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kt+b_k\sin kt).$$

Una observación antes de abordar un ejemplo. Las aproximaciones de Fourier admiten otra expresión trigonométrica equivalente, que consiste en expresar cada armónico en la forma alternativa

$$a_k \cos kt + b_k \sin kt = A_k \sin(kt + \varphi_k),$$

lo cual nos lleva a escribir la proyección o aproximación de Fourier así:

$$F_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(\underbrace{A_k \sin(kt + \varphi_k)}_{\text{armónico de frecuencia } k} \right).$$

Ejercicio 17: Consideremos el espacio euclídeo $\mathcal{PC}(2\pi)$ y obtengamos la proyección, y, por tanto, los coeficientes de Fourier, de la extensión periódica de $E(t) = \begin{cases} 0, & \text{si } -\pi < t < 0 \\ V, & \text{si } 0 \le t < \pi \end{cases}$ sobre cada una de las funciones del sistema ortogonal $\mathcal{S} = \{c_0(t) = 1, c_1(t) = \cos t, s_1(t) = \sin t, c_2(t) = \cos 2t, s_2(t) = \sin 2t\}$.

a. La proyección de $E(t)$ sobre $c_0(t)$ es la función

$$p_0(t) = \frac{E \bullet c_0}{\|c_0\|^2} c_0(t) = \frac{\frac{1}{\pi}\int_0^{\pi} V\,dt}{\frac{1}{\pi}\int_{-\pi}^{\pi} dt} c_0(t) = \frac{V}{2}$$

Así pues, $\frac{V}{2}$ es la función constante más próxima a $E(t)$.

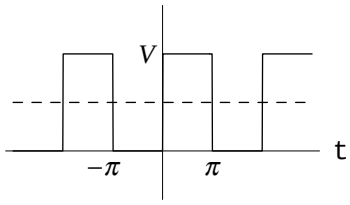


FIGURA 16: La línea discontinua representa el polinomio trigonométrico de orden cero de la onda rectangular.

b. La proyección de $E(t)$ sobre $c_1(t)$ es la función

$$p_1(t) = \frac{E \bullet c_1}{\|c_1\|^2} \cos t = \frac{\frac{1}{\pi}\int_0^{\pi} V \cos t\,dt}{\frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2 t\,dt} \cos t = \frac{0}{1} \cos t = 0,$$

es decir, ambas funciones son perpendiculares y, por eso, la proyección es la función nula.

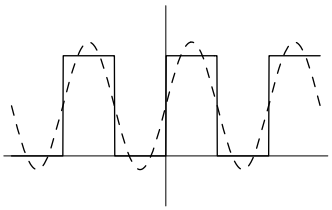
c. La proyección de $E(t)$ sobre $s_1(t)$ es la función

$$q_1(t)=\frac{E\bullet s_1}{\|s_1\|^2}\sin t=\frac{\frac{1}{\pi}\int_0^{\pi}V\sin t\,dt}{\frac{1}{\pi}\int_{-\pi}^{\pi}\sin^2 t\,dt}\sin t=\frac{2/\pi}{1}V\sin t=\frac{2V}{\pi}\sin t.$$

En este caso, el múltiplo de $s_1(t)=\sin t$ más próximo a $E(t)$ es la función $\frac{2V}{\pi}\sin t$.

d. Del mismo modo, se obtienen las proyecciones de $E(t)$ sobre $c_2(t)=\cos 2t$ y sobre $s_2(t)=\sin 2t$ y resultan ser cero.

La suma de todas las proyecciones calculadas es el polinomio o aproximación de Fourier de $E(t)$ de orden dos: $F_2(t)=\frac{V}{2}+\frac{2V}{\pi}\sin t$.



Ejercicio 18: Calcula la norma o longitud de la función $E(t)$ del ejercicio anterior. ¿Cuál es la longitud de cada una de sus proyecciones? ¿y la de la suma?

— Sobre la norma o longitud de la función $E(t)$:

$$\|E(t)\|^2=E(t)\bullet E(t)=\frac{1}{\pi}\int_0^{\pi}V^2\,dt=V^2 \quad\Rightarrow\quad \|E(t)\|=V.$$

— La longitud de cada una de sus proyecciones:

$$\|p_0\|^2=\frac{V}{2}\bullet\frac{V}{2}=\frac{1}{\pi}\int_{-\pi}^{\pi}\frac{V^2}{4}dt=\frac{V^2}{2} \quad\Rightarrow\quad \|p_0\|=\frac{V}{\sqrt{2}};$$

$$\|p_1\|^2=0;$$

$$\|q_1\|^2=\frac{1}{\pi}\int_{-\pi}^{\pi}\left(\frac{2V}{\pi}\sin t\right)^2dt=\left(\frac{2V}{\pi}\right)^2\frac{1}{\pi}\int_{-\pi}^{\pi}\sin^2 t\,dt=\left(\frac{2V}{\pi}\right)^2\|\sin t\|^2=\left(\frac{2V}{\pi}\right)^2$$

$$\Rightarrow\quad \|q_1\|=\frac{2V}{\pi}.$$

Nótese que la longitud de q_1 es precisamente su amplitud.

— La longitud de la suma de las tres proyecciones se obtiene fácilmente usando el hecho de que los tres sumandos son ortogonales entre sí y que, por el teorema de Pitágoras:

$$\| p_0 + p_1 + q_1 \|^2 = \| p_0 \|^2 + \| p_1 \|^2 + \| q_1 \|^2 = \frac{V^2}{2} + 0 + \left(\frac{2V}{\pi}\right)^2;$$

en definitiva, $\| p_0 + p_1 + q_1 \| = \sqrt{\| p_0 \|^2 + \| p_1 \|^2 + \| q_1 \|^2} = V\sqrt{\frac{1}{2} + \frac{4}{\pi^2}}$.

En caso de que el orden de la aproximación de Fourier fuera n, entonces tendríamos $\| F_n \|^2 = \| p_0 \|^2 + \sum_{\substack{k=1 \\ k\ impar}}^{k=n} \| q_k \|^2 = \frac{V^2}{2} + \sum_{\substack{k=1 \\ k\ impar}}^{k=n} \left(\frac{2V}{k\pi}\right)^2$.

— Por último, averigüemos la relación entre las amplitudes o normas de cada armónico y la de la función total:

$$\frac{\| p_0 \|^2}{\| E \|^2} = \frac{V^2/2}{V^2} = 0.5 = 50\%,$$

$$\frac{\| q_1 \|^2}{\| E \|^2} = \frac{\left(\frac{2V}{\pi}\right)^2}{V^2} = \frac{4}{\pi} \approx 0.4053 = 40.53\%,$$

$$\frac{\| q_3 \|^2}{\| E \|^2} = \frac{\left(\frac{2V}{3\pi}\right)^2}{V^2} = \frac{4}{9\pi} \approx 0.045 = 4.5\%, \text{ etc.}$$

Ejercicio 19: En el espacio de la funciones periódicas, de período 2π, y continuas a trozos con el producto escalar $f \bullet g = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx$, se considera el sistema trigonométrico $\mathcal{S} = \{1, \cos t, \sin t, \cos 2t, \sin 2t\}$. Para la función $I(t) = t$, $t \in (-\pi, \pi]$, se pide:

a) ¿Cuál es la norma de $I(t)$?

$$\|I(t)\|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2\,dt = \frac{2}{3}\pi^2 \quad \Rightarrow \quad \|I(t)\| = \pi\sqrt{\frac{2}{3}} \approx 2'565.$$

b) Calcula el trabajo realizado por la corriente de intensidad $I(t)$ al pasar por una resistencia de R ohmios, $W = R\int_{-\pi}^{\pi} [I(t)]^2\,dt$, y compáralo con la norma

obtenida en el apartado a):

$$W = R\int_{-\pi}^{\pi} [I(t)]^2\,dt = R\int_{-\pi}^{\pi} t^2\,dt = R\pi\|I(t)\|^2 = \frac{2}{3}R\pi^3.$$

c) Calcula el polinomio trigonométrico de orden dos

$$F_2(t) = \frac{a_0}{2} + a_1\cos t + b_1\sin t + a_2\cos 2t + b_2\sin 2t$$

que más se aproxima a $I(t)$.

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} t\,dt = \text{(función impar)} = 0,$$

$$a_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos t\,dt = \text{(función impar)} = 0,$$

$$b_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin t\,dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} t\sin t\,dt = 2,$$

$$a_2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos 2t\,dt = \text{(función impar)} = 0,$$

$$b_2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin 2t\,dt = \text{(función par)} = \frac{2}{\pi}\int_{0}^{\pi} t\sin 2t\,dt = -1.$$

Por tanto, $F_2(t) = 2\sin t - \sin 2t$.

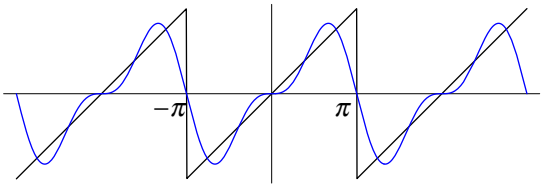


FIGURA 18: Gráfica conjunta de la onda en diente de sierra y de su aproximación de Fourier de orden dos.

d) Averigua el trabajo realizado por la parte constante, por el primer armónico y por el segundo, ¿cuál es el trabajo realizado por $F_2(t)$?

— El primer armónico realiza el trabajo:

$$W_1 = R\int_{-\pi}^{\pi} [2\sin t]^2\,dt = R\pi\|2\sin t\|^2 = 4\pi R,$$

que, comparado con el trabajo total de la onda $\frac{W_1}{W} = \frac{4\pi R}{\frac{2}{3}\pi^3 R} = \frac{6}{\pi^2} \approx 0.6079$, significa que el primer armónico realiza casi el 61% del trabajo total.

— En cuanto al trabajo del segundo armónico:

$$W_2 = R\int_{-\pi}^{\pi}[-\sin 2t]^2\,dt = R\pi\|-\sin 2t\|^2 = \pi R,$$

se tiene que $\dfrac{W_2}{W} = \dfrac{\pi R}{\frac{2}{3}\pi^3 R} = \dfrac{1.5}{\pi^2} \approx 0.1520$, es decir, que el segundo armónico realiza aproximadamente el 15'2% del trabajo total.

— El trabajo realizado por $F_2(t) = 2\sin t - \sin 2t$ es

$$T = R\int_{-\pi}^{\pi}[2\sin t - \sin 2t]^2\,dt \overset{\text{(Teorema de Pitágoras)}}{=} R\left[\pi\|2\sin t\|^2 + \pi\|-\sin 2t\|^2\right] =$$

$$= W_1 + W_2 = 5\pi R.$$

En definitiva, el trabajo realizado por los dos primeros armónicos es la suma del realizado por cada uno de ellos y eso supone que la aproximación $F_2(t) = 2\sin t - \sin 2t$ realiza casi el 76'2% del trabajo total de la onda en diente de sierra.

13.7. Proceso de ortogonalización de Gram-Schmidt

La última etapa del camino consiste en aprender a construir bases ortonormales, clave en la construcción de las proyecciones ortogonales, pues hasta ahora nos hemos limitado a mostrarlas como datos en cada problema. El método que nos ocupa en esta sección «convierte» una base cualquiera en una base ortonormal; por eso lleva el nombre de *proceso de ortogonalización,* además se añaden los nombres de los dos matemáticos que lo inventaron: *Gram* y *Schmidt.*

La verdad es que el proceso no contiene ninguna novedad y es posible que el lector enseguida vea en él una repetición de las ideas desarrolladas en el estudio de las proyecciones ortogonales. Describámoslo sin más demora.

Sea E un espacio euclídeo y W un subespacio de dimensión finita. Consideremos una base $\{\mathbf{w}_1,\ldots,\mathbf{w}_k\}$ cualquiera de W. El procedimiento de ortogonalización de Gram-Schmidt consiste en ir reemplazando gradualmente

los vectores $\mathbf{w}_1,\ldots,\ \mathbf{w}_k$ por los vectores ortogonales $\mathbf{u}_1,\ldots,\ \mathbf{u}_k$ de modo que sean una base.

Paso1: Mantenemos el mismo primer vector $\mathbf{u}_1 = \mathbf{w}_1$. Se cumple:

- $\mathbf{u}_1$ genera el mismo subespacio que $\mathbf{w}_1$.

Paso 2: Reemplazamos el segundo vector $\mathbf{w}_2$ por su componente ortogonal al subespacio generado por $\mathbf{u}_1$, es decir,

$$\mathbf{u}_2 = \mathbf{w}_2 - \frac{\mathbf{w}_2 \bullet \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1.$$

Se cumple:

- $\mathbf{u}_1$ y $\mathbf{u}_2$ generan el mismo subespacio que $\mathbf{w}_1$ y $\mathbf{w}_2$;
- $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales.

Paso 3: Reemplazamos el tercer vector $\mathbf{w}_3$ por su componente ortogonal al subespacio generado por $\mathbf{u}_1$ y $\mathbf{u}_2$:

$$\mathbf{u}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \bullet \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\mathbf{w}_3 \bullet \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\mathbf{u}_2.$$

Se cumple:

- $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$ generan el mismo subespacio que $\mathbf{w}_1$, $\mathbf{w}_2$ y $\mathbf{w}_3$;
- $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$ son ortogonales.

Y así sucesivamente, hasta el k-ésimo vector.

Paso k: Reemplazamos el k-ésimo vector $\mathbf{w}_k$ por su componente ortogonal respecto al subespacio generado por $\mathbf{u}_1,\ldots,\ \mathbf{u}_{k-1}$:

$$\mathbf{u}_k = \mathbf{w}_k - \frac{\mathbf{w}_k \bullet \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\mathbf{w}_k \bullet \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\mathbf{u}_2 - \ldots - \frac{\mathbf{w}_k \bullet \mathbf{u}_{k-1}}{\|\mathbf{u}_{k-1}\|^2}\mathbf{u}_{k-1}.$$

Se cumple:

- $\mathbf{u}_1, \ldots, \mathbf{u}_k$ generan el mismo subespacio que $\mathbf{w}_1, \ldots, \mathbf{w}_k$;
- $\mathbf{u}_1, \ldots, \mathbf{u}_k$ son ortogonales entre sí.

Finalmente, si deseamos obtener una base ortonormal, tan solo es necesario que normalicemos los vectores $\mathbf{u}_1,\ldots,\ \mathbf{u}_k$ construidos.

Ejercicio 18: ¿Qué múltiplo de $\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ debemos restar de $\mathbf{a}_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$ para obtener un vector ortogonal a $\mathbf{a}_1$?

Se pregunta por el segundo paso del proceso de ortogonalización y la respuesta es

$$\mathbf{n} = \mathbf{a}_2 - \frac{\mathbf{a}_2 \bullet \mathbf{a}_1}{\|\mathbf{a}_1\|^2}\mathbf{a}_1 = \begin{pmatrix} 4 \\ 0 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Observar que $\{\mathbf{a}_1, \mathbf{n}\}$ es una base ortogonal del plano.

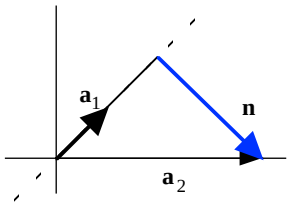


FIGURA 19: Primer paso del proceso de Gram-Schmidt.

Ejercicio 19: Aplicar el proceso de ortogonalización de Gram-Schmidt a los vectores $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ y $\mathbf{c} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

El primer vector: $\mathbf{u}_1 = \mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$;

el segundo vector:

$$\mathbf{u}_2 = \mathbf{b} - \mathrm{Proy}_{\mathbf{u}_1}\mathbf{b} = \mathbf{b} - \frac{\mathbf{b} \bullet \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix};$$

y el tercero:

$$\mathbf{u}_3 = \mathbf{c} - \mathrm{Proy}_{\mathbf{u}_1}\mathbf{c} - \mathrm{Proy}_{\mathbf{u}_2}\mathbf{c} =$$

$$= \mathbf{c} - \frac{\mathbf{c} \bullet \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\mathbf{c} \bullet \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

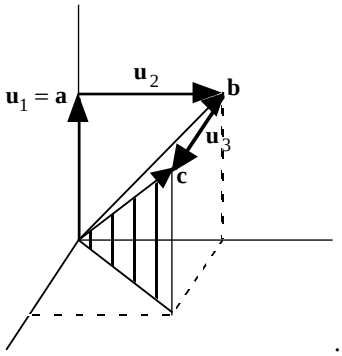


FIGURA 20: Proceso de Gram-Schmidt en el espacio.

Ejercicio 20: Construye la proyección ortogonal del vector $\mathbf{b} = (1,0,1)$ sobre el subespacio generado por los vectores del sistema $\mathcal{S} = \{(1,1,0), (0,1,1)\}$.

Necesitamos una base ortogonal del subespacio generado por $\mathcal{S}$:

1.º Mantenemos el mismo primer vector $\mathbf{u}_1 = (1,1,0)$.

2.º Reemplazamos el segundo vector $(0,1,1)$ por su componente ortogonal respecto al subespacio generado $\mathbf{u}_1$:

$$\mathbf{u}_2 = (0,1,1) - \frac{(0,1,1)\cdot \mathbf{u}_1}{\mathbf{u}_1 \bullet \mathbf{u}_1}\mathbf{u}_1 = (0,1,1) - \frac{1}{2}(1,1,0) = \left(-\tfrac{1}{2}, \tfrac{1}{2}, 1\right).$$

La proyección ortogonal de $\mathbf{b} = (1,0,1)$ sobre el subespacio generado por S es:

$$\text{Proy}_S \mathbf{b} = \frac{\mathbf{b}\cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 + \frac{\mathbf{b}\cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2}\mathbf{u}_2 = \frac{1}{2}(1,1,0) + \frac{\cancel{\frac{1}{2}}}{\cancel{\frac{3}{2}}}\left(-\tfrac{1}{2}, \tfrac{1}{2}, 1\right) = \left(\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{3}\right).$$

Ejercicio 21: Calcula el cuarto polinomio de Legendre, es decir, un polinomio de grado tres y ortogonal al sistema $\left\{1,\ x,\ x^2 - \frac{1}{3}\right\}$ según el producto escalar $p \cdot q = \int_{-1}^{1} p(x)q(x)dx$.

Podemos resolver el problema de dos formas distintas. O bien buscamos un polinomio cúbico $p(x) = a + bx + cx^2 + dx^3$ resolviendo las tres ecuaciones de la perpendicularidad $\begin{cases} p(x)\bullet 1 = 0 \\ p(x)\bullet x = 0 \\ p(x)\bullet\left(x^2 - \frac{1}{3}\right) = 0 \end{cases}$, o bien observamos que $\left\{1,\ x,\ x^2 - \frac{1}{3}\right\}$ es un sistema ortogonal y calculamos la componente ortogonal del polinomio $p(x) = x^3$ respecto al subespacio de los polinomios de grado no superior a dos.

Por este segundo método, una vez probado que $\left\{1,\ x,\ x^2 - \frac{1}{3}\right\}$ es un conjunto ortogonal (tarea que dejamos al lector), deducimos que el tercer polinomio de Legendre es

$$q(x) = x^3 - \frac{x^3 \bullet 1}{\|1\|^2}1 - \frac{x^3 \bullet x}{\|x\|^2}x - \frac{x^3 \bullet \left(x^2 - \frac{1}{3}\right)}{\left\|x^2 - \frac{1}{3}\right\|^2}\left(x^2 - \tfrac{1}{3}\right) = x^3 - \tfrac{3}{5}x.$$