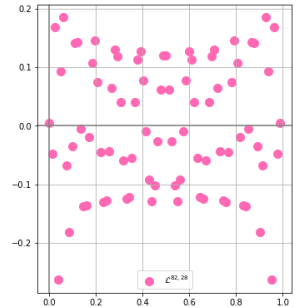
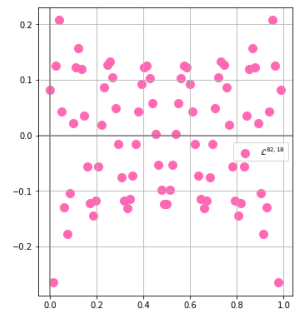


BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

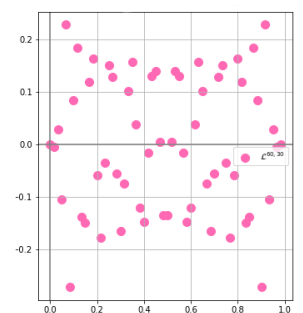


Polinomios de Legendre discretos para el estudio morfológico de señales finitas



TESIS
que para obtener el grado de
LICENCIADA EN MATEMÁTICAS
presenta
DENISSE AMÉLIE SOPHIE BERNÈS CARMONA

Asesor: DR. MOISÉS SOTO BAJO
Co-asesor: DR. JAVIER HERRERA VEGA



Puebla, Pue. Mayo, 2023. Quita número en esta página.

Agradecimientos

;))

23 Noviembre- Pendientes

- **Cambia las figuras.** Fija un color (rosa) para TODAS las gráficas de señales $x..$
- **revisar literatura**
- **implementar en python los polinomios discretos de Legendre:** Listo! Ahora debería de intentar usar algunas librerías que encontré sobre polinomios discretos de Legendre.
- hacer un mapa sobre cómo se relacionan unos resultados con otros.
- **Dimensión, grado, posición.**
- Cambia todos los entornos de demostración
- revisa que siempre digas que la dimensión del polinomio discreto es n .
- ¿Deberían adjuntar en la tesis una copia de Survey? Porque cito muchas veces fórmulas de ahí.
- **Sí es importante poner la dependencia a n de los espacios W_k . Es muy confuso si no lo haces así.**
- cambia el formato de las captions para figuras.
- revisa que usas t_i y no x_i para puntos de la malla.
- nemu
- Deberías ver este video: https://www.youtube.com/watch?v=q5Nsc4LyhdE&ab_channel=SeminarioGAMMA
- poner macadores para el final de un ejemplo.
- Cambiar comas por puntos en el entorno align. Creo que esto se corrige si pones la cita dentro de un entorno de texto.
- marcador al final de ejemplos. ◇
- Cambia los símbolos ¡,¿ por langle y rangle.
- Preguntar al Dr. Gabriel por información.
- ve cuándo es la primera vez que citas en TFA y da la formulación dada en rotman.
- Di que algunos autores dicen que el grado del polinomio cero no está definido, pero para nosotros es cero.
- checa que las discretizaciones de los de legendre no son los que tengo yo.

Resumen

Motivados por la búsqueda de un sistema de representación de \mathbb{R}^n adecuado para el estudio morfológico de señales finitas (c.f. lista de deseos ??), para cada $n \in \mathbb{N}$ mayor a uno, definimos (c.f. capítulo ??) n vectores de \mathbb{R}^n , que llamamos **polinomios discretos de Legendre** (aka PDL), cada uno asociado a un grado entero k de 0 a $n-1$, que juntos conforman una base ortonormal de \mathbb{R}^n , la **base de Legendre discreta n -dimensional** y que, como mostramos (c.f. capítulo ??), cumple satisfactoriamente nuestros requisitos.

Encontramos en la literatura que estos objetos ya se habían definido antes (c.f. [NS74]), pero no encontramos una construcción rigurosa de estos como la dada por nosotros ni aplicaciones de los PDL para estudios de morfología.

Hacemos también un estudio de simetrías de los PDL (c.f. capítulo ??) que, junto con las fórmulas dadas en [NS74], será usado para definir algoritmos (c.f. capítulo ??) para programar a los PDL de cualquier dimensión n . Las implementaciones en Python pueden encontrarse en el repositorio [ref](#).

Se realizó además un análisis espectral de los PDL (c.f. capítulo ??) que nos llevó a plantear una interesante conjetura en la que establecemos una relación (que depende sólo de los parámetros de dimensión n y grado k) entre frecuencias y polinomios discretos de Legendre [explica mejor esto](#).

Para que los conceptos y herramientas que usamos a lo largo del desarrollo de este trabajo queden claras, incluimos al final un apéndice de teoría en el que plasmamos algunos resultados o definiciones que son estrictamente necesarios para el desarrollo de este trabajo de tesis. Preferimos no limitarnos a citar referencias que abordan estos temas, pues definiciones presentes en unas son diferentes a las usadas por otras; además, para algunos resultados clásicos (como el teorema de Gram-Schmidt) necesitábamos dar formulaciones distintas a las canónicas pero útiles para nosotros.

Mostrada con teoría y ejemplos cómo las bases de Legendre discretas son un sistema de representación particularmente útil para hacer un estudio cuantitativo de morfología (y, potencialmente, un análisis espectral también) de señales finitas, esbozamos cómo esto da lugar a trabajos a futuro en esta línea.

Este trabajo fue escrito en L^AT_EX. El lenguaje de programación de nuestra elección para el desarrollo de la parte computacional fue Python. La mayoría de las imágenes aquí presentadas fueron, a su vez, programadas en Python y, en algunas ocasiones, retocadas (o dibujadas completamente) en Krita. El formato adoptado se basa en el diseño de los libros de Edward Tufte.

Índice general

A. Apéndice: teoría	2
A.1. Polinomios y teorema fundamental del álgebra	2
A.2. Definiciones básicas de espacios de Hilbert	4
A.3. Hiperplanos de espacios con producto punto de dimensión finita	6
A.4. El teorema de la proyección ortogonal	7
A.5. Bases ortonormales de espacios de Hilbert	10
A.5.1. El concepto de redes como una generalización del de sucesión .	11
A.5.2. Caracterización y propiedades de las bases ortonormales de un espacio de Hilbert	13
A.6. El teorema de Gram-Schmidt	18
A.7. Ángulo entre elementos de un espacio con producto punto	20
A.8. Definición de ángulo entre un punto y un subespacio cerrado de un espacio euclídeo	21
A.8.1. Caso particular en el que el subespacio en cuestión es un plano	23
A.9. Similitud coseno	24
A.10. La transformada discreta de Fourier y estudios espectrales de señales finitas	25
A.10.1. La exponencial compleja y raíces n -ésimas de la unidad . . .	25
A.10.2. La transformada discreta de Fourier	27
A.10.3. Versión real de la DFT	30
B. Notaciones y abreviaciones	34

Apéndice A

Apéndice: teoría

En este apéndice, establecemos las definiciones y resultados del álgebra y de la teoría de espacios de Hilbert estrictamente necesarios para el desarrollo de este trabajo de tesis. A pesar de que los teoremas aquí expuestos son clásicos, preferimos no limitarnos a simplemente citarlos ya que queremos puntos de vista de estos tal vez un poco distintos a los usuales, pero útiles para nosotros.

A menos que se diga explícitamente lo contrario, las elecciones del orden en la exposición del material, al igual que las demostraciones aquí contenidas, son personales.

A.1. Polinomios y teorema fundamental del álgebra

En general, si R es un anillo, se define a partir de él un nuevo anillo denominado en **anillo de polinomios con coeficientes en R** ; a pesar de que la construcción algebraica de este anillo no se hace, por lo general, definiendo funciones de R en R , ... [referencia libro de álgebra a construcción](#).

Definición A.1.1

Sean K un anillo, $f(t) = \sum_{k=0}^n a_k t^k \in K[t]$ con a_n un elemento no cero del anillo K .

- Al elemento $a_n \in K$ se le llama el **coeficiente principal** del polinomio f , y
- a $n \in \mathbb{N}$ se le llama el **grado de f** y se le denota por $\partial(f)$.

A todo elemento $r \in K$ tal que $f(r) = 0$ se le llamará una **raíz de f** .

Nota A.1.2

En A.1.1 se definió el coeficiente principal y grado de todo polinomio no cero; por lo general, los algebraistas prefieren definir el grado del polinomio cero como $-\infty$ (c.f. [Jac85], p.128) o dejarlo indefinido (c.f. [Rot10], p.126). Nosotros preferimos definir el grado del polinomio cero como cero (que es el grado de todo polinomio constante). Esto lo hacemos simplemente para no tener que distinguir al polinomio cero como un caso especial, pues eso simplifica las formulaciones

de los resultados que necesitamos.

Proposición A.1.3

(c.f. [Rot10], p.127) Si K es un dominio entero, entonces, para cualesquiera polinomios $f(t), g(t) \in K[t]$ no cero se tiene que el producto $f(t) \cdot g(t)$ de estos dos polinomios no es el polinomio cero y que

$$\partial(f \cdot g) = \partial(f) \cdot \partial(g).$$

A nosotros nos interesará tomar siempre al anillo de coeficientes K como el campo \mathbb{R} o el campo \mathbb{C} .

Es una consecuencia inmediata del algoritmo de la división (c.f. [Rot10] p. 131) el que, si K es un campo y $f(x) \in K[x]$ es un polinomio de grado $n > 0$, entonces f tiene a lo más n raíces (el argumento es relacionar raíces de un polinomio con divisores lineales de este y argumentar que, por A.1.3, si el grado de f es n entonces f no puede factorizarse como el producto de más de n factores lineales, luego, no puede tener más de n raíces).

Se tiene pues una cota superior para la cantidad de raíces de un polinomio basada en su grado, sin embargo, esta cota bien puede no alcanzarse. De hecho, como muestra el siguiente ejemplo, puede ser que un polinomio no constante con coeficientes en un campo no tenga ninguna raíz.

Ejemplo 1. Considere al anillo \mathbb{Z}_5 de enteros módulo 5. Como 5 es un número primo, \mathbb{Z}_5 es un campo (c.f. proposición 3.12 de [Rot10]). Sea $f(t) = t^2 - 2 \in \mathbb{Z}_5[t]$; evaluar a este polinomio en cada uno de los cinco elementos de \mathbb{Z}_5 nunca da lugar al cero del campo, luego, f es un polinomio de grado dos con coeficientes en \mathbb{Z}_5 sin raíces. \diamond

Ejemplo 2. El ejemplo canónico de polinomio con coeficientes reales sin raíces (reales) es $p(t) = t^2 + 1 \in \mathbb{R}[t]$. No puede tener raíces por ser el cuadrado de todo número real no negativo. \diamond

Teorema A.1.4

(**Teorema fundamental del álgebra, versión dada por [Kur68], p.149**): Todo polinomio $f(t) \in \mathbb{C}[t]$ de coeficientes complejos y grado al menos uno tiene al menos una raíz compleja.

Como se hace notar en la referencia, es difícil exagerar la importancia que tiene este teorema no sólo en el álgebra, sino en la matemática en general; se resalta también el hecho de que, hasta el momento, todas las pruebas de este teorema hacen uso no sólo de la estructura algebraica del dominio entero $\mathbb{C}[t]$, sino de propiedades de otra índole, por ejemplo topológicas, de \mathbb{C} visto como un \mathbb{C} -espacio vectorial (obtenidas, por ejemplo, al definir en \mathbb{C} una norma, c.f. (A.6)). Hay, sin embargo, construcciones puramente algebraicas del campo \mathbb{C} basadas en la idea de “completar” al campo de los números reales (c.f. [God66], *L’anneau $K[\sqrt{d}]$* , p.152).

De un simple argumento inductivo, usando el teorema A.1.4 se demuestra que el que todo polinomio no constante con coeficientes en \mathbb{C} tenga al menos una raíz compleja **equivale** a que tenga tantas raíces como su grado.

Teorema A.1.5

Todo polinomio $f(t) \in \mathbb{C}[t]$ de grado $n > 0$ tiene exactamente n raíces comple-

jas (contando multiplicidades).

Demostración. Procedemos por inducción sobre n , el grado del polinomio. Sea f un polinomio con coeficientes complejos y de grado $n = 1$. Según el teorema A.1.4, f tiene al menos una raíz $r_1 \in \mathbb{C}$, luego, el polinomio lineal $x - r_1$ divide a f . Ahora bien, si f tuviese una segunda raíz r_2 , el polinomio de grado dos $(x - r_1)(x - r_2)$ dividiría a f , pero esto contradice el que todos los divisores de un polinomio f con coeficientes en un campo tengan grado menor o igual al de f . Con esto comprobamos la validez del teorema para $n = 1$.

Supongamos ahora el teorema cierto para todo polinomio de grado $n \geq 1$. Sea $f \in \mathbb{C}[t]$ un polinomio de grado $n + 1$; según A.1.4, f tiene al menos una raíz r_1 , luego,

$$f(x) = (x - r_1)g(x), \quad (\text{A.1})$$

con $g(x)$ algún polinomio de coeficientes complejos y grado n (c.f. proposición A.1.3). Por hipótesis de inducción, g tiene exactamente n raíces complejas (contando multiplicidades); puesto que, según la ecuación (A.1), las raíces de f son r_1 y las raíces de g (pues, en un campo, el producto de dos elementos del campo es cero sí y sólo si alguno de estos es cero), concluimos, como queríamos, que, salvo multiplicidades, f tiene $n + 1$ raíces complejas. \square

La siguiente consecuencia inmediata del teorema fundamental del álgebra es una de las piezas angulares del trabajo desarrollado en esta tesis.

Proposición A.1.6

Sea $f(t) \in \mathbb{R}[t]$ un polinomio con coeficientes reales de grado a lo más n . Si f tiene más de n raíces reales, entonces f es el polinomio cero.

Demostración. En efecto, si f tuviese grado mayor a cero, entonces, según el teorema A.1.5, f tendría exactamente n raíces reales, pero, por hipótesis, f tiene más de n ; así, f debe tener grado cero, o sea, debe ser un polinomio constante. Puesto que el único polinomio constante con al menos una raíz es el polinomio cero, concluimos que $f = 0$. \square

A.2. Definiciones básicas de espacios de Hilbert

Sea $F \in \{\mathbb{R}, \mathbb{C}\}$. La definición central [aaa](#) de producto punto; el tener definida una tal función en V dota al espacio de nociones ; es posible ya no sólo sumar y escalar elementos de V , sino también hablar de **ángulos** (en particular, nos provee de una noción de **ortogonalidad**) y de **distancias**.

Definición A.2.1

Sea V un F -espacio vectorial. A una función $\langle \cdot, \cdot \rangle : V \times V \longrightarrow F$ tal que, para cualesquiera $x, y, z \in V$ y $a, b \in F$ se cumple que

- $\langle a \cdot x + b \cdot y, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- $\langle x, x \rangle \geq 0$
- $\langle x, x \rangle = 0$ si y sólo si $x = 0$

Si $z = a + ib \in \mathbb{C}$, al **conjugado complejo** de z , o sea, al complejo $a - ib$, se le denota como \bar{z} .

se le llamara un **producto punto** (o **producto interior**) definido en V .

En la sección A.7 mostramos cómo se puede definir la noción de ángulo cuando se cuenta con un producto punto definido en el espacio.

Observe que, según el segundo punto de la definición A.2.1, para todo $x \in V$ se tiene que el conjugado complejo de $\langle x, x \rangle$ coincide con $\langle x, x \rangle$, luego ¹, $\langle x, x \rangle \in \mathbb{R}$; se usa este hecho y el cuarto punto de la definición para definir, a partir de un producto punto, una norma ² en el espacio, luego, una noción de longitud (y, por lo tanto, de distancia también).

Proposición A.2.2

(La norma inducida por un producto interior) Dado un espacio vectorial con producto punto $(V, \langle \cdot, \cdot \rangle)$, la función $\| \cdot \| : V \rightarrow [0, \infty[$ definida como

$$\forall v \in V : \|v\| = \sqrt{\langle v, v \rangle} \quad (\text{A.2})$$

satisface la definición de norma. A tal función se le llama la **norma inducida por el producto interior**.

Provistos de una noción de norma, siempre se puede definir la **distancia** entre dos elementos del espacio $x, y \in V$ como

$$d(x, y) := \|x - y\|; \quad (\text{A.3})$$

dotamos así a V de estructura de **espacio métrico**. ³

Teniendo una estructura de espacio métrico en el espacio se tiene también una de **topología** definida de forma natural como sigue:

$$\tau = \{B_\epsilon(x) : x \in V, \epsilon > 0\}, \quad (\text{A.4})$$

donde

$$B_\epsilon(x) = \{y \in V : d(x, y) < \epsilon\}.$$

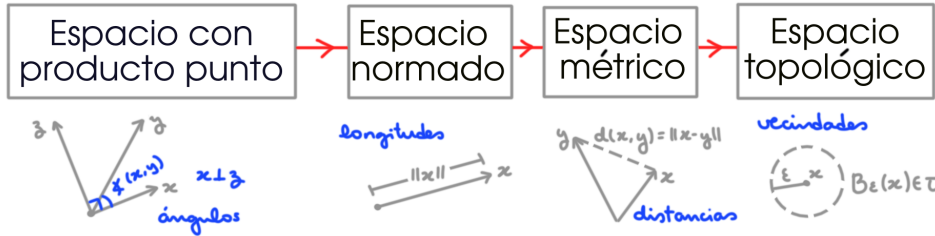


Figura A.1: El definir un producto punto en un espacio vectorial permite considerar en este estructuras de espacio normado, métrico y topológico. Se tienen pues definiciones matemáticas de conceptos como “ángulo”, “longitud”, “distancia” y “vecindad”.

En general, si X es un conjunto no vacío y $d : X \times X \rightarrow [0, \infty[$ es una función de distancia definida en él, entonces toda **sucesión de Cauchy** $(x_m)_{m \in \mathbb{N}}$ del espacio (es decir, toda sucesión tal que, para todo $\epsilon > 0$ sea siempre posible encontrar un natural N tal que, para cualesquiera $m, n > N$ se cumpla que $|x_m - x_n| < \epsilon$) trivialmente converge a un punto de X ; si, recíprocamente, toda sucesión de Cauchy es convergente, se dice que el espacio métrico (X, d) es **completo**.

¹ Por eso tiene sentido comparar a $\langle x, x \rangle$ con cero en el tercer punto de la definición (no se puede definir una relación de orden en \mathbb{C} , c.f. [JEM99] p.6).

² Puede consultar la definición de norma en [Car], p.40

³ Consulte la definición de espacio métrico en [Car], p.37.

Definición A.2.3

Sea $(V, \langle \cdot, \cdot \rangle)$ un F -espacio vectorial con producto punto. Sea $\| \cdot \|$ la norma definida en A.2.2 y d la distancia definida a partir de esta como en (A.3). Si el espacio métrico (V, d) es completo, entonces decimos que $(V, \langle \cdot, \cdot \rangle)$ es un **espacio de Hilbert**.

Ejemplo 3. Sea $n \geq 1$.

En el \mathbb{R} -espacio vectorial \mathbb{R}^n se tiene definido un producto punto canónico como sigue:

$$\forall x = (x_m)_{0 \leq m \leq n-1}, y = (y_m)_{0 \leq m \leq n-1} \in \mathbb{R}^n : \quad \langle x, y \rangle = \sum_{m=0}^{n-1} x_m y_m. \quad (\text{A.5})$$

En el \mathbb{C} -espacio vectorial \mathbb{C}^n el producto canónico se define como sigue:

$$\forall x = (x_m)_{0 \leq m \leq n-1}, y = (y_m)_{0 \leq m \leq n-1} \in \mathbb{C}^n : \quad \langle x, y \rangle = \sum_{m=0}^{n-1} x_m \overline{y_m}. \quad (\text{A.6})$$

Puesto que ambos espacios son finito dimensionales (de hecho, de dimensión n), según [Mseb] son espacios de Hilbert. \diamond

A.3. Hiperplanos de espacios con producto punto de dimensión finita

En general, si V es cualquier F -espacio vectorial (con F un campo cualquiera) y W es un subespacio de V cuya dimensión difiere de la del espacio ambiente V por uno, W se denomina un “hiperplano” de V .

Como a nosotros sólo nos interesa el escenario en el que V es finito dimensional y en él se ha definido un producto punto, damos la definición para este caso particular y algunas de sus consecuencias a continuación.

Definición A.3.1

Sea V un \mathbb{R} -espacio vectorial, con $\dim(V) = n < \infty$ y con un producto punto $\langle \cdot, \cdot \rangle$ definido en él¹. A todo subespacio W de V con $\dim(W) = n - 1$ le llamaremos un **hiperplano de V** .

Si W es un hiperplano de V , puesto que $V = W + W^\perp$ (c.f. [Fri02], p.355, ejercicio 13), W^\perp tiene dimensión uno; fijando pues un vector u de W^\perp , tenemos que $W^\perp = \text{span}\{u\}$. En base a este podemos definir (gracias a la linealidad del producto punto) al siguiente funcional⁴:

$$\begin{aligned} h_u : V &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle x, u \rangle. \end{aligned}$$

Observe que

- $h_u(x) > 0$ si y sólo si $x = au$ para algún $a > 0$,
- $h_u(x) = 0$ si y sólo si $x \in W$,

⁴ Por “funcional” entendemos toda transformación lineal de un F -espacio vectorial V a F (c.f. [Kre]).

- $h_u(x) < 0$ si y sólo si $x = au$ para algún $a < 0$.

Así, en base al hiperplano W , via el funcional h_u podemos dividir al espacio ambiente V en tres subconjuntos ajenos dos a dos, a saber,

$$R_I = \{x \in V : h_u(x) > 0\}, \quad R_{II} = W, \quad R_{III} = \{x \in V : h_u(x) < 0\}, \quad (\text{A.7})$$

siendo el funcional h_u el criterio usado para determinar la pertenencia de un $x \in V$ a alguna de estas regiones.

Es fácil comprobar que, si se escoge otro vector $\tilde{x} \in W^\perp$, las regiones análogas a las (A.7) obtenidas ahora con el funcional $h_{\tilde{x}}$ coinciden con las dadas en (A.7) (aunque, obviamente, pueden diferir en el orden en que estas se escriban), pues todos los elementos de W^\perp son múltiplos escalares uno del otro. En este sentido decimos que **todo hiperplano divide en tres regiones ajenas al espacio ambiente**.

A.4. El teorema de la proyección ortogonal

En general, si V es un espacio métrico y W es un subconjunto de este, puede definirse la distancia de un punto $x \in V$ a W como el ínfimo del conjunto de distancias entre x y puntos de W :

$$d(x, W) = \inf\{\|x - y\| : y \in W\}.$$

Tal ínfimo existe por estar el conjunto descrito acotado inferiormente, por ejemplo, por el cero. El siguiente teorema nos dice que, en el contexto de espacios de Hilbert, si W es un subespacio de V que además es cerrado ⁵, tal ínfimo se alcanza, es decir, existe un punto (además, único) \hat{x} de W que minimiza la distancia a x .

Teorema A.4.1

(*de la proyección ortogonal* [Nim]): Sean V un espacio de Hilbert, W un subespacio cerrado de V . Para todo $x \in V$ existe un único $\hat{x} \in W$ tal que

$$\|x - \hat{x}\| = \inf_{y \in W} \|x - y\|;$$

además, \hat{x} es el único elemento de W tal que $x - \hat{x} \in W^\perp$.

En este trabajo, el espacio con producto punto particular que nos concierne es \mathbb{R}^n con el producto punto usual; por ser este un espacio finito-dimensional, cualquier subespacio de este es cerrado (c.f. teorema 2.4-3 de [Kre]), luego, el teorema de la proyección siempre aplica.

En general, si W es un subespacio cerrado de un espacio de Hilbert V , gracias a la unicidad establecida en el teorema A.4.1, podemos definir la función **proyección a W** como sigue:

$$\begin{aligned} \Pi_W : V &\longrightarrow W \\ x &\longmapsto \hat{x} \end{aligned}, \quad (\text{A.8})$$

donde \hat{x} es el único vector del que se habla en el teorema A.4.1.

Puede pensarse a esta como la función que a cada elemento x del espacio le asocia “su representante” del espacio W más cercano. Otra consecuencia de la unicidad establecida en el teorema A.4.1 es la linealidad de la función Π_W .

⁵ Cuidado aquí: hay dos nociones distintas de cerradura involucradas. Como el espacio métrico V es un espacio vectorial normado, con “subespacio” nos referimos a un subespacio vectorial de V , no a un mero subconjunto de este; es decir, requerimos que W sea cerrado bajo las operaciones de espacio vectorial. Se pide además que W (pensado como subconjunto del espacio métrico V) sea cerrado en el sentido topológico (c.f. [Mun00] p.93).

Corolario A.4.2

Sean V es un espacio de Hilbert, W un subespacio cerrado de V cuyo complemento ortogonal también es cerrado en V . Entonces,

- Para todo $x \in V$, $x = \Pi_W(x) + \Pi_{W^\perp}(x)$.
- $\Pi_W : V \longrightarrow W$ es un operador autoadjunto.

Demostración. Sea $x \in V$ cualquiera; según el teorema de la proyección A.4.1,

$$x - \Pi_W(x) \in W^\perp; \quad (\text{A.9})$$

además,

$$x - (x - \Pi_W(x)) = \Pi_W(x) \in W \subseteq (W^\perp)^\perp; \quad (\text{A.10})$$

puesto que, según el teorema de la proyección, $\Pi_{W^\perp}(x)$ se caracteriza por ser el único elemento de W^\perp tal que $x - \Pi_{W^\perp}(x) \in (W^\perp)^\perp$, concluimos por (A.9) y (A.10) la igualdad

$$x - \Pi_W(x) = \Pi_{W^\perp}(x).$$

Esto demuestra el primer punto del corolario. Para demostrar el segundo, o sea, que para cualesquiera $x, y \in V$ se tiene que

$$\langle x, \Pi_W(y) \rangle = \langle \Pi_W(x), y \rangle,$$

sólo observe que, usando el primer punto ya demostrado, tenemos que

$$\begin{aligned} \langle x, \Pi_W(y) \rangle &= \langle \Pi_W(x) + \Pi_{W^\perp}(x), \Pi_W(y) \rangle \\ &= \langle \Pi_W(x), \Pi_W(y) \rangle + \langle \Pi_{W^\perp}(x), \Pi_W(y) \rangle \\ &= \langle \Pi_W(x), \Pi_W(y) \rangle, \end{aligned}$$

donde la última igualdad se da porque $\Pi_W(y) \in W$ y $\Pi_{W^\perp}(x) \in W^\perp$ (luego, son vectores mutuamente perpendiculares), y, análogamente,

$$\langle \Pi_W(x), y \rangle = \langle \Pi_W(x), \Pi_W(y) \rangle.$$

□

Nota A.4.3

De hecho, una vez establecida la continuidad del producto punto más adelante en la proposición A.4.5, podremos quitar en la formulación del corolario A.4.2 la hipótesis de que W^\perp sea cerrado, pues esto será consecuencia de que W lo sea; si $(a_n)_{n \in \mathbb{N}}$ es una sucesión en W^\perp convergente a algún $a \in V$, entonces, para todo $w \in W$, $\langle a, w \rangle = \lim_{n \rightarrow \infty} \langle a_n, w \rangle = \lim_{n \rightarrow \infty} 0 = 0$, luego, $a \in W^\perp$.

Ya que estamos hablando de proyecciones ortogonales, usemos estas para demostrar una de las desigualdades más importantes con las que se cuenta en un espacio con producto punto.

Teorema A.4.4

(Desigualdad de Cauchy-Schwarz) Sea V un \mathbb{R} -espacio vectorial con pro-

$$\left| \begin{array}{l} \text{ducto punto } \langle \cdot, \cdot \rangle. \\ \forall x, y \in V : \quad |\langle x, y \rangle| \leq \|x\| \cdot \|y\|, \\ \text{siendo } \|\cdot\| \text{ la norma inducida por el producto punto.} \end{array} \right. \quad (\text{A.11})$$

Demostración. (basada en la demostración ofrecida en [Lan97], p. 292)

- Si alguno de los vectores x o y es cero, la igualdad se da trivialmente (el vector cero tiene norma cero y es ortogonal a cualquier vector).
- Si $x = ay$ para algun real a , también se da la igualdad:

$$|\langle x, y \rangle| = |\langle ay, y \rangle| = |a| \cdot |\langle y, y \rangle| = |a| \cdot \|y\|^2 = \|ay\| \cdot \|y\| = \|x\| \cdot \|y\|.$$

- Supongamos por último que no ocurre ninguno de los casos de los puntos anteriores, o sea, que x y y son linealmente independientes.

Si $c := \frac{\langle x, y \rangle}{\langle y, y \rangle}$, según el teorema de Gram-Schmidt A.6.1, $x - cy$ y cy son vectores ortogonales cuya suma es x ; así, por el teorema de pitágoras,

$$\begin{aligned} \|x\|^2 &= \|x - cy\|^2 + \|cy\|^2 \\ &\geq \|cy\|^2 = |c|^2 \cdot \|y\|^2 \\ &= \frac{\langle x, y \rangle^2}{\|y\|^4} \|y\|^2 = \frac{\langle x, y \rangle^2}{\|y\|^2}, \end{aligned}$$

revisa eso, hay un pequeño error. luego,

$$\langle x, y \rangle^2 \leq \|x\|^2 \cdot \|y\|^2.$$

Tomando raíces cuadradas en ambos lados de la desigualdad llegamos a la desigualdad deseada. \square

Esta desigualdad, que relaciona la norma de los vectores con el valor absoluto de su producto punto, permite establecer la continuidad del producto punto.

Proposición A.4.5

La función $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ es continua.

Demostración. Demostrar la continuidad de $\langle \cdot, \cdot \rangle$ significa probar la veracidad de la implicación

$$(x_n, y_n) \rightarrow (x, y) \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle.$$

Hagamos esto. Puesto que

$$x_n \rightarrow x \quad \text{y} \quad y_n \rightarrow y,$$

los módulos de los vectores x_n y y_n están acotados, digamos por las constantes K y M .

Por la bilinealidad del producto punto, para toda n

$$\langle x_n - x, y_n - y \rangle = \langle x_n, y_n \rangle - \langle x_n, y \rangle - \langle x, y_n \rangle + \langle x, y \rangle,$$

luego, por la desigualdad triangular y la de Cauchy-Schwarz (A.4.4),

$$\begin{aligned} 0 \leq |\langle x_n, y_n \rangle - \langle x, y \rangle| &\leq |\langle x_n - x, y_n - y \rangle| + |\langle x_n, y \rangle| + |\langle x, y_n \rangle| \\ &\leq \|x_n - x\| \cdot \|y_n - y\| + \|x_n\| \cdot \|y\| + \|x\| \cdot \|y_n\| \\ &\leq \|x_n - x\| \cdot \|y_n - y\| + K \cdot \|y_n\| + M \cdot \|x\|; \end{aligned}$$

Como $x_n \rightarrow x$ y $y_n \rightarrow y$, el lado derecho de la desigualdad tiende a cero conforme $n \rightarrow \infty$. Por el teorema de la compresión, concluimos que

$$\lim_{n \rightarrow \infty} |\langle x_n, y_n \rangle - \langle x, y \rangle| = 0,$$

o sea, que $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle x, y \rangle$. □

A.5. Bases ortonormales de espacios de Hilbert

ver si, junto con G-S, puedo subir esto.

En general, dado un F -espacio vectorial V , si uno acepta el axioma de elección, siempre cuenta con bases para este espacio, es decir, de subconjuntos de V que son tanto linealmente independientes (término abreviado como “l.i.”) como generadores de todo el espacio; la importancia de estas es obvia pues, por definición, una base permite representar de forma única a un elemento $v \in V$ por medio de una colección de escalares.

Puede demostrarse que, dado $A \subseteq V$, el que A sea base de V equivale a que A sea un maximal linealmente independiente (c.f. sección 1.7 de [Fri02]), es decir, que todo subconjunto linealmente independiente de V que contenga a A de hecho coincide con A .

El tener además en V definido un producto punto dota de estructura extra al espacio, en particular, como dijimos ya, nos provee de una noción de longitud (c.f. proposición A.2.2) y también de ortogonalidad (ref), luego, en este contexto podemos también hablar de subconjuntos B ortonormales maximales. Puesto que la ortogonalidad implica trivialmente la independencia lineal, es natural plantearse la siguiente pregunta: en un F -espacio vectorial V con producto punto cualquiera, ¿es cierto que todo ortonormal maximal es también maximal l.i., o sea, una base del espacio?

Antes de proceder, conviene dejar por escrito estos dos conceptos.

Definición A.5.1

Sea V un espacio vectorial con producto punto. Si A es un subconjunto de V que es maximal con respecto a la propiedad de

- ser linealmente independiente, entonces lo llamaremos **base de Hamel** de V
- ser ortonormal, entonces lo llamaremos **base ortonormal** de V (y abreviamos este nombre como **BON**).

Como se establece fácilmente en la siguiente proposición, en el caso finito dimensional, la respuesta a la pregunta planteada es positiva.

Proposición A.5.2

(en espacios de Hilbert de dimensión finita, toda BON es base de

Hamel) Sea V un espacio finito dimensional con producto punto. Si A es un subconjunto ortonormal maximal de V , entonces también es maximal l.i. (o sea, base de V)

Demostración. Supongamos que $A \subseteq V$ es maximal respecto a la propiedad de ser ortonormal pero no a la de ser l.i., es decir, que existe un subconjunto A' l.i. que contiene a A propiamente. Digamos que $A' = A \cup B$. Podemos ortonormalizar a este con el proceso de Gram-Schmidt para obtener un subconjunto ortonormal A de V que genera el mismo espacio que genera A' ; por las fórmulas del teorema A.6.1 y el hecho de que los elementos de A sean ortogonales entre sí es claro que los primeros $|A|$ elementos son los elementos de A (“intactos”); así A'' es un subconjunto ortonormal de V que contiene propiamente a A (contradicción). \square

Como se comenta en [Msea], en el caso en el que el espacio de Hilbert V no sea infinito-dimensional, **una base Hamel de este espacio no puede ser una base ortonormal del espacio.**

Así, en general los sistemas de representación descritos en la definición A.5.1 no son equivalentes. A pesar de que

- todo espacio de Hilbert tiene bases de Hamel, aunque no necesariamente bases ortonormales (c.f. [Msec]), y que
- si \mathcal{B} es una base Hamel uno *siempre* puede representar cualquier vector x del espacio como combinación lineal *finita* de elementos de \mathcal{B} , mientras que, si \mathcal{B} es una BON, se tienen no igualdades entre x y combinaciones lineales de elementos de \mathcal{B} , sino aproximaciones del primero a partir de los segundos (c.f. inciso d) del teorema A.5.9),

como mostraremos más adelante, en el contexto de espacios de Hilbert, conviene mucho más trabajar con BONs (cuando existen para el espacio particular) que con bases de Hamel.

Es por eso que en la mayoría de los libros de análisis funcional, el nombre “base” suele ser sinónimo de “base ortonormal” (pudiendo llegar a confundir a un lector primerizo, y llevarlo a pensar en bases de Hamel). Puesto que nosotros no queremos dar lugar a confusiones de este estilo, nos abstenemos de usar el nombre “base” y en su lugar ocupamos los nombres introducidos en la definición A.5.1.

A.5.1. El concepto de redes como una generalización del de sucesión

Para dar unas equivalencias útiles a la propiedad de ser base ortonormal de un espacio de Hilbert, requeriremos del concepto de “red”, usado en topología como una extensión de la noción de sucesión.

Puede recordar la definición de “orden parcial” en [Mun00] p. 187; el único tipo de conjunto parcialmente ordenado (abreviado “POSET”) que aquí nos interesa es aquel ordenado por la relación “ \subseteq ” (o sea, por la contención).

Definición A.5.3

Sea (\mathcal{E}, \leq) un orden parcial. Si para cualesquiera $\alpha, \beta \in \mathcal{E}$ existe un $\gamma \in \mathcal{E}$ tal que $\alpha \leq \gamma$ y $\beta \leq \gamma$, diremos que (\mathcal{E}, \leq) es un **conjunto dirigido**.

Definición A.5.4

Si (V, τ) es un espacio topológico y (\mathcal{E}, \leq) es un conjunto dirigido, entonces a

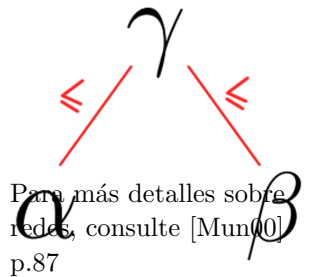


Figura A.2: Para que un POSET (\mathcal{E}, \leq) sea llamado “conjunto dirigido” debe cumplirse que, dados dos elementos arbitrarios de \mathcal{E} siempre sea posible encontrar un tercer elemento “arriba” de ambos.

toda función $f : \mathcal{E} \rightarrow V$ se le llamará una **red**.

Definición A.5.5

Si $f : \mathcal{E} \rightarrow V$ es una red, decimos que f **converge** a un elemento del espacio $v \in V$ si para cualquier vecindad U de v (i.e. cualquier elemento de τ que contenga a v) existe un $\alpha \in \mathcal{E}$ tal que

$$\forall \beta \in \mathcal{E} : \alpha \leq \beta \Rightarrow f(\beta) \in U.$$

Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert; dado un subconjunto $W = \{v_i : i \in I\}$ de V , queremos “sumar todos los elementos de W ”; puesto que trabajamos en un espacio vectorial, la acción de sumar una cantidad *finita* de elementos de V tiene sentido, pero como *a priori* no se ha supuesto nada sobre la cardinalidad del conjunto de índices I , en general no siempre tiene sentido intentar sumar los elementos de W ; definimos rigurosamente esta acción a partir de una de una red que definimos como sigue; consideramos al conjunto

$$\mathcal{P} := \{\alpha \subseteq I : \alpha \text{ es finito}\}. \quad (\text{A.12})$$

Observe que \mathcal{P} consta exactamente de los subconjuntos α de I para los que la expresión

$$\sum_{i \in \alpha} v_i$$

tiene sentido. Teniendo ya un espacio topológico (V, τ) (con la topología τ definida como en (A.4)) y el conjunto dirigido (\mathcal{P}, \subseteq) , definimos la siguiente red;

$$\begin{aligned} f_{\mathcal{P}} : \mathcal{P} &\rightarrow V \\ \alpha &\mapsto \sum_{i \in \alpha} v_i. \end{aligned}$$

El que esta red converja (c.f. definición A.5.5) a un $v \in V$ significa que para todo $\epsilon > 0$ exista $\alpha \subseteq I$ finito tal que, si β es otro subconjunto finito de I que contenga a α , entonces

$$\left| \sum_{i \in \beta} v_i - v \right| < \epsilon;$$

o sea, que $\sum_{i \in \alpha} v_i$ se encuentra a distancia ϵ de v y que, no importa cuántos sumandos se agregen, la suma resultante sigue siendo un vector que se encuentra a distancia ϵ de v .

Notación A.5.6

Si la red $f_{\mathcal{P}}$ converge, denotaremos a su límite como $\sum \{v_i : i \in I\}$.

Nota A.5.7

Si el conjunto de índices I es infinito numerable, ya teníamos una forma natural de definir la suma de los elementos de $W = \{v_i : i \in I\}$; haciendo $I = \mathbb{N}$ (mejor dicho, identificando a I con \mathbb{N} via una biyección), podíamos preguntarnos por la existencia del límite de la sucesión de sumas parciales $\sum_{i=1}^n v_i$; **debe tener**

cuidado, pues la convergencia de la **red** $f_{\mathcal{P}}$ definida como $f_{\mathcal{P}}(\alpha) = \sum_{i \in \alpha} v_i$ (con $\alpha \subseteq \mathbb{N}$ cualquier subconjunto finito de \mathbb{N}) **no** equivale a la convergencia de la **sucesión** $\left(\sum_{i=1}^n v_i \right)_{n \in \mathbb{N}}$ (c.f. p. 16 de [Con90]).

De existir, al límite de la red lo denotamos por $\sum \{v_i : i \in \mathbb{N}\}$, mientras que al límite de la sucesión lo denotamos por $\sum_{i=1}^{\infty} v_i$.

A.5.2. Caracterización y propiedades de las bases ortonormales de un espacio de Hilbert

Enunciamos y demostramos el teorema A.5.9, en el que damos algunas equivalencias a ser base ortonormal.

Para la demostración necesitamos antes del siguiente lema, cuya demostración puede consultarse en [Con90], p. 16.

Lema A.5.8

(Lema 4.12 de [Con90]) Sean $(V, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert, $\mathcal{E} = \{e_i : i \in I\}$ un subconjunto ortonormal de V . Si $x \in V$ es un elemento arbitrario del espacio, $W := \{\langle x, e_i \rangle : i \in I\}$ y $\mathcal{P} := \{\alpha \subseteq I : \alpha \text{ finito}\}$, entonces la red $f_{\mathcal{P}} : \mathcal{P} \rightarrow V$ definida como $f(\alpha) = \sum_{i \in \alpha} \langle x, e_i \rangle e_i$ converge a un elemento de V que denotamos por $\sum \{\langle x, e \rangle e : e \in \mathcal{E}\}$.

Teorema A.5.9

(Teorema 4.13 de [Con90]) Sea \mathcal{E} un subconjunto ortonormal del espacio de Hilbert $(V, \langle \cdot, \cdot \rangle)$. Las siguientes proposiciones son equivalentes entre sí:

- a) \mathcal{E} es una BON de V (en el sentido de la definición A.5.1).
- b) Si $x \in V$ es ortogonal a todo elemento de \mathcal{E} entonces $x = 0$.
- c) $\overline{\text{span}}(\mathcal{E}) = V$.
- d) Para todo $x \in V$ se cumple que

$$x = \sum \{\langle x, e \rangle e \mid e \in \mathcal{E}\}$$

- e) Para cualesquiera $x, y \in V$,

$$\langle x, y \rangle = \sum \{\langle x, e \rangle \langle e, y \rangle \mid e \in \mathcal{E}\}$$

- f) **(Identidad de Parseval)** Para todo $x \in V$ se cumple que

$$\|x\|^2 = \sum \{\langle x, e \rangle^2 \mid e \in \mathcal{E}\}.$$

Demostración.

- a) \Rightarrow b) Suponer la existencia de un x no cero ortogonal a todo elemento de \mathcal{E} nos permite considerar al subconjunto ortonormal

$$\mathcal{E} \cup \left\{ \frac{x}{\|x\|} \right\}$$

de V que contiene propiamente a \mathcal{E} , contradiciendo así la maximalidad de \mathcal{E} respecto a la propiedad de ser ortonormal.

- b) \Rightarrow c) Sea el subespacio $W := \overline{\text{span}}(\mathcal{E})$ de V . Por ser W un subespacio cerrado de V , podemos proyectar sobre él (c.f. teorema A.4.1). Para mostrar que tenemos la igualdad entre V y W , tomemos un $x \in V$ cualquiera. Según el teorema de la proyección ortogonal A.4.1, $x - \Pi_W(x)$ es ortogonal a todo elemento de W , en particular, será ortogonal a todo elemento de \mathcal{E} (pues $\mathcal{E} \subseteq W$); por hipótesis, esto implica que $x - \Pi_W(x)$ sea el vector cero; así,

$$x = \Pi_W(x) \in W.$$

- c) \Rightarrow b) Sea $v \in V$ ortogonal a todo elemento de \mathcal{E} ; mostremos que v es cero. Como $v \in V = \overline{\text{span}}(\mathcal{E})$, existe $(b_n)_{n \in \mathbb{N}}$ una sucesión de $\text{span}(\mathcal{E})$ tal que $\lim_{n \rightarrow \infty} b_n = v$ (c.f. lema 21.2 de [Mun00]). Usando el que para toda n el producto punto $\langle b_n, h \rangle$ sea cero y la continuidad establecida en la proposición A.4.5, llegamos a que

$$\langle x, x \rangle = \langle x, \lim_{n \rightarrow \infty} b_n \rangle = \lim_{n \rightarrow \infty} \langle x, b_n \rangle = \lim_{n \rightarrow \infty} 0 = 0,$$

de donde concluimos (c.f. (A.2) y definición A.2.1) que x es cero.

- b) \Rightarrow d) Según el lema A.5.8, la red $f_{\mathcal{E}}$ converge a un vector, que denotamos por $\sum \{\langle x, e \rangle e | e \in \mathcal{E}\}$; resta ver que tal vector es x o, equivalentemente, que $\sum \{\langle x, e \rangle e | e \in \mathcal{E}\} - x$ es el vector cero; según nuestra hipótesis, podremos concluir esto si demostramos que para todo $e \in \mathcal{E}$ ocurre

$$\langle \sum \{\langle x, e \rangle e | e \in \mathcal{E}\} - x, e \rangle = 0. \quad (\text{A.13})$$

Fijemos pues un $e \in \mathcal{E}$ y sea $\epsilon > 0$. Por definición de convergencia de redes, sabemos que existe $F \subseteq \mathcal{E}$ finito con $e \in F$ tal que

$$|L_F| < \epsilon, \quad (\text{A.14})$$

donde $L_F := \sum_{f \in F} \langle x, f \rangle f - \sum \{\langle x, e \rangle e | e \in \mathcal{E}\} \in V$, o sea,

$$\sum \{\langle x, e \rangle e | e \in \mathcal{E}\} = \sum_{f \in F} \langle x, f \rangle f - L_F \quad (\text{A.15})$$

Por la bilinealidad del producto punto y por ser e ortogonal a todo elemento de $\mathcal{E} - \{e\}$ tenemos que

$$\langle \sum_{f \in F} \langle x, f \rangle f, e \rangle = \sum_{f \in F} \langle x, f \rangle \langle f, e \rangle = \langle x, e \rangle \langle e, e \rangle = \langle x, e \rangle. \quad (\text{A.16})$$

El inciso c) del teorema A.5.9 corresponde a la definición de BON dada en [AK]; así, una BON es también un subconjunto ortonormal tal que el menor subespacio cerrado de V que contiene a \mathcal{E} es todo V . Compare este inciso con el hecho de que \mathcal{E} es una base de Hamel de V si y sólo si E es linealmente independiente y $\text{span}(\mathcal{E}) = V$.

Así,

$$\begin{aligned}
0 &\leq |\langle \sum \{ \langle x, e \rangle e \mid e \in \mathcal{E} \} - x, e \rangle| \\
&= |\langle \sum \{ \langle x, e \rangle e \mid e \in \mathcal{E} \}, e \rangle - \langle x, e \rangle| \\
&\text{(por (A.15))} = |\langle \sum_{f \in F} \langle x, f \rangle f - L_F, e \rangle - \langle x, e \rangle| \\
&= |\langle \sum_{f \in F} \langle x, f \rangle f, e \rangle - \langle L_F, e \rangle - \langle x, e \rangle| \\
&\text{(por (A.16))} = |\langle x, e \rangle - \langle L_F, e \rangle - \langle x, e \rangle| \\
&= |\langle L_F, e \rangle| \\
&\text{(Cauchy-Schwarz)} \leq \|L_F\| \cdot \|e\| = \|L_F\| \\
&\text{(por (A.14))} \leq \epsilon;
\end{aligned}$$

demostramos así que

$$\forall \epsilon > 0 : 0 \leq |\langle \sum \{ \langle x, e \rangle e \mid e \in \mathcal{E} \} - x, e \rangle| \leq \epsilon;$$

de esto concluimos que $|\langle \sum \{ \langle x, e \rangle e \mid e \in \mathcal{E} \} - x, e \rangle|$ es cero, de donde se deduce, como queríamos, (A.13).

d) \Rightarrow e) Según nuestra hipótesis,

$$x = \sum \{ \langle x, e \rangle e \mid e \in \mathcal{E} \}, \quad y = \sum \{ \langle y, e \rangle e \mid e \in \mathcal{E} \}.$$

Para $F \subseteq \mathcal{E}$ finito, sean

$$X_F := \sum_{e \in F} \langle x, e \rangle e - x, \quad Y_F := \sum_{e \in F} \langle y, e \rangle e - y.$$

Por la ortonormalidad supuesta en el conjunto \mathcal{E} ,

$$\left\langle \sum_{e \in F} \langle x, e \rangle e, \sum_{e \in F} \langle y, e \rangle e \right\rangle = \sum_{e \in F} \langle x, e \rangle \langle y, e \rangle; \quad (\text{A.17})$$

así,

$$\begin{aligned}
\left| \sum_{e \in F} \langle x, e \rangle \langle y, e \rangle - \langle x, y \rangle \right| &\stackrel{\text{por (A.17)}}{=} \left| \left\langle \sum_{e \in F} \langle x, e \rangle e, \sum_{e \in F} \langle y, e \rangle e \right\rangle - \langle x, y \rangle \right| \\
&= |\langle X_F + x, Y_F + y \rangle - \langle x, y \rangle| \\
&= |\langle X_F, Y_F \rangle + \langle X_F, y \rangle + \langle x, Y_F \rangle| \\
&\text{(desigualdad triangular)} \leq |\langle X_F, Y_F \rangle| + |\langle X_F, y \rangle| + |\langle x, Y_F \rangle| \\
&\text{(Cauchy-Schwarz)} \leq \|X_F\| \cdot \|Y_F\| + \|X_F\| \cdot \|y\| + \|x\| \cdot \|Y_F\|;
\end{aligned}$$

puesto que $\|x\|$ y $\|y\|$ son constantes y $\|X_F\|$ y $\|Y_F\|$ pueden hacerse tan pequeñas como se quiera (escogiendo apropiadamente al conjunto finito F), concluimos que $\sum \{ \langle x, e \rangle \langle y, e \rangle : e \in \mathcal{E} \} = \langle x, y \rangle$.

$e) \Rightarrow f)$ Para $x \in V$,

$$\|x\|^2 = \langle x, x \rangle = \sum \{\langle x, e \rangle^2 | e \in \mathcal{E}\}.$$

$f) \Rightarrow a)$ Supongamos que \mathcal{E} puede extenderse aún más para ser ortonormal, es decir, que existe $e_0 \in V$ unitario tal que para todo $e \in \mathcal{E}$ se tenga que $\langle e_0, e \rangle = 0$. Llegamos entonces a la siguiente contradicción:

$$1 = \|e_0\|^2 = \sum \{\langle e_0, e \rangle^2 | e \in \mathcal{E}\} = \sum \{0\} = 0.$$

□

Nota A.5.10

(sobre la identidad de Parseval en el caso finito dimensional) Puesto que, para el desarrollo teórico de este trabajo de tesis, se tratará siempre con espacios de Hilbert finito dimensionales, luego, con bases ortonormales finitas, conviene reescribir algunos puntos del teorema A.5.9 en este caso particular. Digamos que $\dim(V) = n$; si $\mathcal{E} := \{e_k : 0 \leq k \leq n-1\}$ es una BON del espacio de Hilbert V , entonces, para todo $x \in V$, se tiene que

$$x = \sum_{k=0}^{n-1} \langle x, e_k \rangle^2 e_k \quad (\text{A.18})$$

y además

$$\|x\|^2 = \sum_{k=0}^{n-1} \langle x, e_k \rangle^2; \quad (\text{Identidad de Parseval}) \quad (\text{A.19})$$

así, si se usa a una BON \mathcal{E} del espacio como sistema de representación (o sea, si se representa a todo x del espacio por medio de sus n coeficientes respecto a \mathcal{E}), se tiene que

- tales coeficientes son de hecho los productos punto entre x y los elementos de la base \mathcal{E} (c.f. (A.18)), y
- se tiene una relación lineal muy sencilla entre el cuadrado de la norma de x y el cuadrado de los coeficientes de la representación (c.f. (A.18)).

Estos dos hechos permiten **usar la magnitud de un coeficiente en (A.18) para valorar la contribución del respectivo elemento de la base \mathcal{B} para sintetizar a x** ; en efecto, si, por ejemplo, el coeficiente $\langle x, e_{k_1} \rangle$ es “muy cercano a cero”, digamos $\langle x, e_{k_1} \rangle = \epsilon$, entonces el vector $x_{k_1} := \sum_{\substack{k=0, \\ k \neq k_1}}^{n-1} \langle x, e_k \rangle^2 e_k$ es, en la misma medida, cercano al vector inicial x , pues

$$\|x - x_{k_1}\| = \|\langle x, e_{k_1} \rangle^2 e_{k_1}\| = \langle x, e_{k_1} \rangle^2 = \epsilon.$$

Esta última nota da fuertes razones sobre la utilidad de las bases ortonormales como sistema de representación en espacios de Hilbert. Veamos ahora cómo las BON también son útiles para expresar proyecciones de un vector sobre subespacios cerrados en términos del vector y la base ortonormal.

Para la tarea necesitamos el siguiente resultado.

Proposición A.5.11

([AK], thm. 6 p. 149) Si $\{e_k | k \in \mathbb{N}\}$ es un sistema ortonormal de un espacio de Hilbert V y $x \in V$, entonces la expresión

$$\left\| x - \sum_{k=1}^{\infty} a_k e_k \right\|$$

alcanza su mínimo para la elección de coeficientes $a_k = \langle x, e_k \rangle$.

Nota: En esta proposición consideramos la situación de un subespacio con BON numerable, pero en la demostración del resultado ya se incluye el caso de una BON finita.

Demostración. Si $(a_k)_{k \in \mathbb{N}}$ es una sucesión de coeficientes, como es costumbre, en caso de existir el límite de sumas parciales $S_n := \sum_{k=1}^n a_k e_k$, este es un vector que denotamos por $\sum_{k=1}^{\infty} a_k e_k$. Además, por la continuidad de la norma,

$$\left\| x - \sum_{k=1}^{\infty} a_k e_k \right\| = \lim_{n \rightarrow \infty} \|x - S_n\|$$

por lo que, si encontramos una elección de coeficientes $(a_k)_{k \in \mathbb{N}}$ que haga que la sucesión $(S_n)_{n \in \mathbb{N}}$ converja y tal que las normas $\|x - S_n\|$ (o, equivalentemente, los números $\|x - S_n\|^2$) sean mínimas, esa será la elección buscada.

Observe que, para toda n ,

$$\begin{aligned} \|x - S_n\|^2 &= \langle x - S_n, x - S_n \rangle \\ (\text{bilinealidad}) &= \langle x, x \rangle - 2 \langle x, S_n \rangle + \langle S_n, S_n \rangle \\ (\text{ortonormalidad}) &= \|x\|^2 - 2 \langle x, \sum_{k=1}^n a_k e_k \rangle + \sum_{k=1}^n a_k^2 \\ &= \|x\|^2 - 2 \sum_{k=1}^n a_k c_k + \sum_{k=1}^n a_k^2, \end{aligned}$$

donde $c_k := \langle x, e_k \rangle$. Así,

$$\|x - S_n\|^2 = \|x\|^2 - \sum_{k=1}^n c_k^2 + \sum_{k=1}^n (a_k - c_k)^2;$$

claro que la sucesión de coeficientes que minimiza esta expresión es $(c_k)_{k \in \mathbb{N}}$; además, no es difícil ver que esta candidata hace que la sucesión $(S_n)_{n \in \mathbb{N}}$ de sumas parciales converja, pues tenemos la siguiente acotación válida para toda n :

$$\sum_{k=1}^n c_k^2 \leq \|x\|^2 - \|x - S_n\|^2 \leq \underbrace{\|x\|^2}_{cte}. \quad (\text{A.20})$$

□

Corolario A.5.12

(Dando explícitamente a la proyección de un vector a un subespacio

cerrado respecto a una BON de este último) Con la notación e hipótesis de la proposición A.5.11, para todo $x \in V$ tenemos que

$$\Pi_W(x) = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k,$$

donde $W := \overline{\{e_k | k \in \mathbb{N}\}}$.

Corolario A.5.13

(Caso particular de A.5.12 para espacios de dimensión finita) Sean V un espacio de Hilbert, W un subespacio cerrado de V y de dimensión finita. Si $\mathcal{B} = \{e_k : 0 \leq k \leq n\}$ es una BON de W , entonces, para todo $x \in V$,

$$\Pi_W(x) = \sum_{k=0}^{n-1} \langle x, e_k \rangle e_k.$$

Corolario A.5.14

Si V es un espacio con producto punto y $\mathcal{B} = (e_k)_{k \in \Delta}$ es una BON de este a lo más numerable, entonces, para todo $x \in V$, $x = \sum_{k \in \Delta} \langle x, e_k \rangle e_k$

Demostración. La proyección de un $x \in V$ sobre V es trivialmente x (x es el elemento de V más cercano a sí mismo); según el corolario A.5.12, esta proyección es la serie propuesta. \square

Note que, en el caso en el que Δ sea infinito numerable, el orden en que se tome la serie de arriba no importa. Esto es porque, independientemente del orden que se elija, siempre se tratará con una BON del espacio.

Corolario A.5.15

(Desigualdad de Bessel) Si $\{e_k | k \in \mathbb{N}\}$ es un sistema ortonormal en el espacio con producto punto V , entonces, para todo $x \in V$,

$$\sum_{k=1}^{\infty} \langle x, e_k \rangle^2 \leq \|x\|^2.$$

Demostración. Esto es una consecuencia inmediata de la acotación (A.20) establecida en la demostración del teorema A.5.11. \square

A.6. El teorema de Gram-Schmidt

Dado un espacio de Hilbert V y W un subespacio de dimensión finita de V , en ocasiones nos interesará contar no sólo con una base de Hamel W , sino con una base ortonormal de este (recuerde que, en el caso finito dimensional, toda BON es una base Hamel).

El siguiente resultado nos permite lograr justamente eso; la esencia del teorema de Gram-Schmidt es el reemplazar una base $S = \{v_0, \dots, v_{n-1}\}$ de un subespacio $W := \text{span}(S)$ de dimensión finita de V por una base ortogonal S' para este, que

fácilmente puede normalizarse multiplicando a cada elemento de S' por el recíproco de su norma. Al proceso descrito a continuación lo llamaremos el **proceso de ortonormalización de Gram-Schmidt**, y lo abreviaremos como “G-S”.

Teorema A.6.1

(de **Gram-Schmidt**, [Fri02] p.344): Sean V un espacio vectorial con producto punto, $S = \{v_k : 0 \leq k \leq n-1\}$ un subconjunto linealmente independiente de V . Sean los vectores

$$\begin{aligned}\xi_0 &:= v_0, \\ \xi_k &:= v_k - \sum_{j=0}^{k-1} \frac{\langle v_k, \xi_j \rangle}{\langle \xi_j, \xi_j \rangle} \xi_j, \quad k = 1, \dots, n-1.\end{aligned}$$

El subconjunto $S' := \{\xi_k : 0 \leq k \leq n-1\}$ de V es ortogonal y genera el mismo espacio que S .

La ventaja de la formulación del teorema de Gram-Schmidt dada en A.6.1 es que esta da explícitamente la forma de calcular a los elementos de la base ortogonal para el subespacio finito-dimensional, pero, para nuestros fines, una formulación que involucre proyecciones ortogonales sobre espacios será preferible. De este modo la geometría que hay detrás del proceso puede vislumbrarse mejor.

Proposición A.6.2

(**versión del teorema de Gram-Schmidt en términos de proyecciones**)
Sean V un espacio vectorial con producto punto,

$$S = \{v_k : 0 \leq k \leq n-1\}$$

un subconjunto linealmente independiente de V y

$$S' = \{\xi_k : 0 \leq k \leq n-1\}$$

el subconjunto ortogonal que resulta de aplicar el proceso de Gram-Schmidt A.6.1 a S .

Si para cada $0 \leq k \leq n-1$ se define al subespacio

$$W_k = \text{span}(v_0, \dots, v_k)$$

de V , entonces, para toda $1 \leq k \leq n-1$ se tiene que

$$\xi_k = v_k - \Pi_{W_{k-1}}(v_k).$$

Demostración.

Sea $1 \leq k \leq n-1$. Según el teorema de Gram-Schmidt (A.6.1),

$$W_{k-1} = \text{span}(\xi_0, \dots, \xi_{k-1}). \quad (\text{A.21})$$

Si mostramos que

- el vector $v_k - \xi_k$ es elemento del espacio W_{k-1} , y que

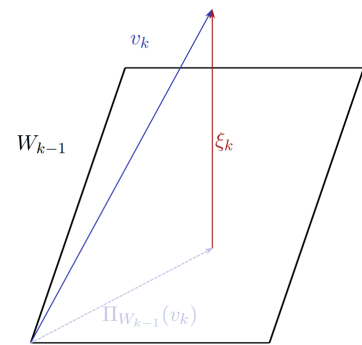


Figura A.3: Ilustrando el proceso de Gram-Schmidt formulado en términos de proyecciones.

- $\xi_k = v_k - (v_k - \xi_k)$ es elemento de W_{k-1}^\perp ,

por la unicidad establecida en el teorema de la proyección ortogonal A.4.1 podremos concluir la igualdad deseada.

Lo primero es claro, pues, según las fórmulas dadas en el teorema A.6.1,

$$v_k - \xi_k = \sum_{j=0}^{k-1} \left(\frac{v_k \cdot \xi_j}{\xi_j \cdot \xi_j} \right) \xi_j \in W_{k-1}.$$

Lo segundo se sigue de observar que ξ_k es ortogonal a los vectores ξ_0, \dots, ξ_{k-1} ; como estos conforman una base para W_{k-1} (c.f. (A.21)), concluimos que $\xi_k \in W_{k-1}^\perp$.

□

Comentemos una consecuencia sencilla del teorema de Gram-Schmidt.

Corolario A.6.3

Todo espacio de Hilbert finito dimensional tiene bases ortonormales.

Demostración. Sólo observe que, si el espacio de Hilbert V es finito dimensional, uno puede aplicar el proceso de ortonormalización de Gram-Schmidt para obtener, a partir de una base de Hamel de V , una base ortonormal. □

Así, a pesar de que, como comentamos en A.5, en general uno no puede asegurar que un espacio de Hilbert tenga bases ortonormales, en el caso finito dimensional siempre se tienen estos sistemas de representación; **deberías decir por qué te gusta llamar “sistemas de representación” a las bases** recuerde las ventajas de estos sobre simples bases de Hamel discutidas en la nota A.5.10.

A.7. Ángulo entre elementos de un espacio con producto punto

Es gracias a la desigualdad de Cauchy-Schwarz que es posible introducir una de las nociones más valiosas con las que se cuenta en un espacio vectorial con producto punto, a saber, la de **ángulo** entre dos vectores.

Si v y w son dos elementos no cero de V , entonces, según el teorema A.4.4,

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} \leq 1,$$

luego, puesto que la función coseno es una biyección del intervalo $[0, \pi]$ al intervalo $[-1, 1]$, tenemos que existe un único elemento $\theta \in [0, \pi]$ tal que

$$\cos(\theta) = \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|};$$

a tal número θ se le denomina el **ángulo formado por los vectores v y w** , y lo denotamos por $\angle(v, w)$

Si alguno de los vectores v o w fuese cero, definimos el ángulo entre estos como cero.

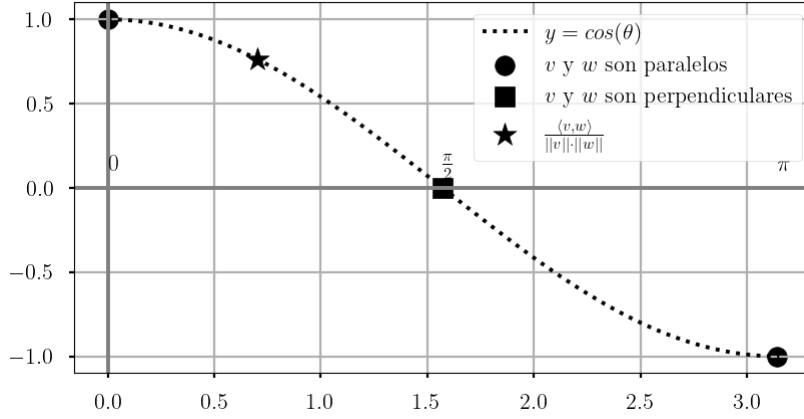


Figura A.4: Si θ es el ángulo entre v y w , los casos extremos en los que $\theta \in \{0, \pi\}$ son aquellos en los que v y w son múltiplos escalares uno del otro, mientras que, si $\cos(\theta) = 0$, o sea, si $\theta = \frac{\pi}{2}$, entonces v y w son perpendiculares.

Terminamos esta sección con una sencilla observación, que es consecuencia de la linealidad del producto punto y de la norma.

Observación A.7.1

(El ángulo entre dos vectores no se ve afectado por multiplicación por un mismo escalar) Sean V un espacio de Hilbert, $u, v \in V$ no cero. Si $a \in F$ es un escalar no cero, entonces $\angle(v, w) = \angle(a \cdot v, a \cdot w)$

A.8. Definición de ángulo entre un punto y un subespacio cerrado de un espacio euclideo

Ya vimos en A.7 cómo definir el ángulo entre dos elementos de un espacio de Hilbert V .

Si además W es un subespacio cerrado de V (lo que, recuerde, siempre ocurre en caso de que V sea finito dimensional, c.f. teorema 2.4-3 de [Kre]), entonces también es posible definir el ángulo entre un punto $x \in V$ del espacio y W .

Definición A.8.1

Sea $(V, \langle \cdot, \cdot \rangle)$ un espacio con producto punto. Sean $W \leq V$ un subespacio cerrado de V y $x \in V$. Definimos el **ángulo entre x y W** como el ángulo que forma x con su proyección a W , es decir,

$$\angle(x, W) := \angle(x, \Pi_W(x)).$$

Una definición alternativa del ángulo entre un punto y un subespacio, así como una caracterización sencilla del coseno de este, se dan a continuación.

Proposición A.8.2

Si V , W y x son como en la definición A.8.1, entonces

- $\angle(x, W) = \min\{\angle(x, w) : w \in W\}$, y
- $\cos(\angle(x, W)) = \frac{\|\Pi_W(x)\|}{\|x\|}$.

Demostración.

- Sea w un elemento cualquiera de W . Puesto que el ángulo entre dos vectores es preservado bajo multiplicación por escalares (c.f. observación A.7.1), sin pérdida de generalidad podemos suponer que

$$\|w\| = \|\Pi_W(x)\|. \quad (\text{A.22})$$

De la definición del vector $\Pi_W(x)$ se sigue que

$$\|x - \Pi_W(x)\|^2 \leq \|x - w\|^2;$$

expresando ambos lados de la desigualdad como un producto punto (c.f. (A.2)) y aplicando la bilinealidad del producto punto, llegamos a que

$$\langle x, x \rangle - 2\langle x, \Pi_W(x) \rangle + \langle \Pi_W(x), \Pi_W(x) \rangle \leq \langle x, x \rangle - 2\langle x, w \rangle + \langle w, w \rangle;$$

usando (A.22), podemos simplificar esta última desigualdad para llegar a

$$\langle x, w \rangle \leq \langle x, \Pi_W(x) \rangle,$$

de donde se sigue, usando nuevamente (A.22), que

$$\cos(\angle(x, w)) = \frac{\langle x, w \rangle}{\|x\| \cdot \|w\|} \leq \frac{\langle x, \Pi_W(x) \rangle}{\|x\| \cdot \|\Pi_W(x)\|} = \cos(\angle(x, \Pi_W(x)));$$

del que la función coseno sea decreciente en el intervalo $[0, \pi]$ se concluye de esta última desigualdad que $\angle(x, \Pi_W(x)) \leq \angle(x, w)$.

- Según el corolario A.4.2, x puede expresarse como la suma entre su proyección a W y su proyección a W^\perp , luego, según la definición A.8.1 y la linealidad del producto punto, concluimos que

$$\begin{aligned} \cos(\angle(x, W)) &= \cos(\angle(x, \Pi_W(x))) \\ &= \frac{\langle x, \Pi_W(x) \rangle}{\|x\| \cdot \|\Pi_W(x)\|} \\ &= \frac{\langle \Pi_W(x) + \Pi_{W^\perp}(x), \Pi_W(x) \rangle}{\|x\| \cdot \|\Pi_W(x)\|} \\ &= \frac{\langle \Pi_W(x), \Pi_W(x) \rangle}{\|x\| \cdot \|\Pi_W(x)\|} \\ &= \frac{\|\Pi_W(x)\|^2}{\|x\| \cdot \|\Pi_W(x)\|} \\ &= \frac{\|\Pi_W(x)\|}{\|x\|}. \end{aligned}$$

□

Nota A.8.3

Observe que, según el segundo punto de la proposición A.8.2, el coseno entre un vector x y un subespacio cerrado W es siempre no negativo; esto nos permite acotar el rango en el que se encuentra el ángulo (que, por definición, es $[0, \pi]$)

y afirmar que $\angle(x, W) \in [0, \frac{\pi}{2}]$.

A.8.1. Caso particular en el que el subespacio en cuestión es un plano

Necesitaremos concentrarnos en el caso particular en el que el subespacio cerrado W es un plano,⁶ por lo que elaboremos un poco más la teoría de la sección A.7 para este caso particular.

La situación es la siguiente: V es un \mathbb{R} -espacio de Hilbert, u y v son elementos de V , unitarios y linealmente independientes entre sí. El espacio que ellos generan es pues un plano, digamos,

$$P := \text{span}\{u, v\}.$$

Dado $x \in V$ unitario, el coseno del ángulo entre x y W es, según la proposición A.8.2,

$$\cos(\angle(x, P)) = \|\Pi_P(x)\|; \quad (\text{A.23})$$

para lograr expresar el lado derecho de la igualdad en términos sólo de u , v y x (que son los elementos básicos de nuestra discusión), conviene primero considerar una base ortonormal del espacio P .

Observación A.8.4

Si $u, v \in V$ son unitarios y linealmente independientes, y P es el plano que generan, entonces $\{u, z\}$, donde

$$z := \frac{v - \langle u, v \rangle u}{\|v - \langle u, v \rangle u\|} \quad (\text{A.24})$$

es una BON de P

Demostración. Basta aplicar el teorema de Gram-Schmidt A.6.1. \square

Teniendo una BON de P , según el corolario A.5.13, se tiene la siguiente expresión para la proyección de x en P ;

$$\Pi_P(x) = \langle x, u \rangle u + \langle x, z \rangle z; \quad (\text{A.25})$$

puesto que, según la definición (A.24) de z este vector es función de u y v , fácilmente se puede derivar, a partir de (A.25), una expresión de $\Pi_P(x)$ en función sólo de x , u y v . Se plasman las fórmulas concretas (sin demostración, pues esta involucra simple aritmética) en la siguiente proposición.

Proposición A.8.5

Sean V un espacio de Hilbert, $x \in V$ unitario, $u, v \in V$ linealmente independientes. Si P es el plano que generan, entonces,

$$\Pi_P(x) = \frac{\langle x, u \rangle - \langle u, v \rangle \langle x, v \rangle}{1 - \langle u, v \rangle^2} u + \frac{\langle x, v \rangle - \langle u, v \rangle \langle x, u \rangle}{1 - \langle u, v \rangle^2} v \quad (\text{A.26})$$

y

$$\|\Pi_P(x)\|^2 = \frac{1}{1 - \langle u, v \rangle^2} (\langle x, u \rangle^2 + \langle x, v \rangle^2 - 2\langle x, u \rangle^2 \langle x, v \rangle^2 \langle u, v \rangle^2). \quad (\text{A.27})$$

⁶ O sea, un subespacio de dimensión 2.

A.9. Similitud coseno

Sea V un espacio de Hilbert. Sean $W \subseteq V$ un subespacio cerrado de V y x un elemento cualquiera de V . Hay dos formas “naturales” de intentar medir la cercanía de x a W :

- Usando la distancia euclídea de x a W** , es decir, tomando la norma del vector $x - \Pi_W(x)$ como una medida de qué tanto pertenece x a W .
- Usando el ángulo que x forma con W** : como ya vimos en la sección A.8, el ángulo $\angle(x, W) \in [0, \frac{\pi}{2}]$ entre x y W está definido como el ángulo entre x y su proyección $\Pi_W(x)$ a W , luego,
 - entre más cercano a cero sea $\angle(x, W)$, x se aleja más del plano W , pues es casi perpendicular al elemento de W más cercano a x , y
 - entre más cercano a uno sea $\angle(x, W)$, más se acerca x a ser paralelo a su representante en W , luego, más se acerca x a pertenecer a W .

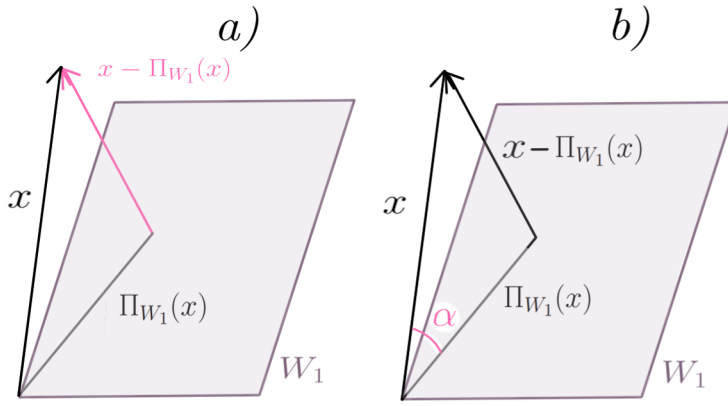


Figura A.5: Esquema de dos formas en las que uno puede intentar medir qué tan afín es una señal x .

Este segundo enfoque es conocido como “**cosine similarity**” en inglés (c.f. [Cos]). Una ventaja, que será clave para nosotros, del segundo método sobre el primero se expone en el siguiente resultado.

Proposición A.9.1

Sean V un espacio euclídeo, $W \subseteq V$ cerrado, $x \in V$, $a \in \mathbb{R} - \{0\}$. Si $|a| \neq 1$, la distancia euclídea de $a \cdot x$ a W es distinta a la de x a W , mientras que los ángulos que forman $a \cdot x$ y x con W coinciden.

Demostración. En efecto,

$$\|a \cdot x - \Pi_W(a \cdot x)\| = |a| \cdot \|x - \Pi_W(x)\|,$$

y

$$\begin{aligned} \angle(a \cdot x, W_{n,1}) &:= \angle(a \cdot x, \Pi_{W_{n,1}}(a \cdot x)) \\ &= \frac{\langle a \cdot x, \Pi_{W_{n,1}}(a \cdot x) \rangle}{\|a \cdot x\| \cdot \|\Pi_{W_{n,1}}(a \cdot x)\|} \\ &= \frac{a^2 \langle x, \Pi_{W_{n,1}}(x) \rangle}{a^2 \|x\| \cdot \|\Pi_{W_{n,1}}(x)\|} \end{aligned}$$

$$=\angle(x, W_{n,1}).$$

□

Así el criterio *b)* no se ve afectado por multiplicación escalar, mientras que el *a)* sí.

A.10. La transformada discreta de Fourier y estudios espectrales de señales finitas

En esta sección vamos a introducir una base ortonormal de \mathbb{C}^n y otra para \mathbb{R}^n (que llamaremos “bases de Fourier complejas y reales”), que están definidas en términos de discretizaciones de sinusoides de frecuencias enteras, y que son herramientas clásicas para hacer lo que comúnmente se denomina un **análisis espectral** de señales finitas.

Puesto que la definición de estas herramientas requiere de algunas nociones del análisis complejo (en particular, de la definición de la exponencial compleja y de las raíces n –ésimas de la unidad), damos brevemente algunas definiciones y resultados necesarios para definir las bases de Fourier.

El producto punto que estamos considerando en el \mathbb{C} –espacio vectorial \mathbb{C}^n es el que definimos en (A.6).

A.10.1. La exponencial compleja y raíces n –ésimas de la unidad

La definición de la función exponencial compleja tiene diversas motivaciones, puede consultar algunas de estas en [JEM99] nosotros sólo damos la definición de esta así como algunas propiedades de ella que se usarán en lo que sigue.

Definición A.10.1

Si $y \in \mathbb{R}$, entonces por $\exp(iy)$ denotamos al número complejo de módulo uno y argumento y , es decir,

$$\exp(iy) := \cos(y) + i\sin(y). \quad (\text{A.28})$$

Para todo número complejo $z = a + bi$, definimos $\exp(z)$ como sigue:

$$\exp(z) := e^x(\cos(y) + i\sin(y)). \quad (\text{A.29})$$

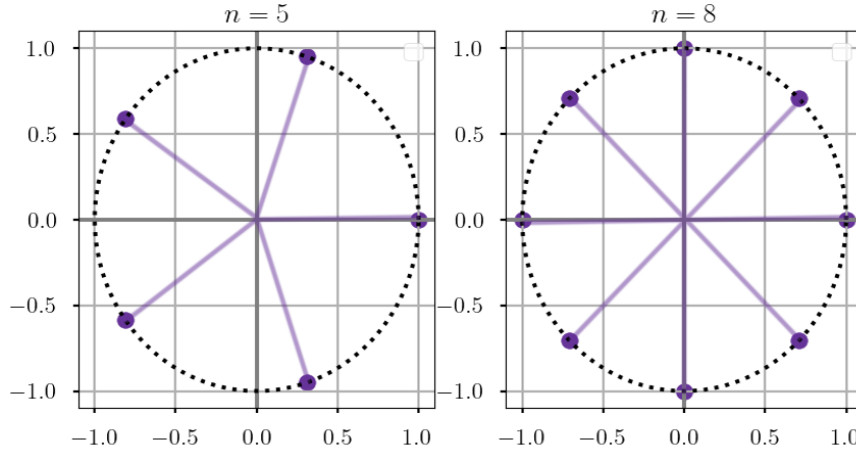


Figura A.6: Para construir gráficamente a las raíces n -ésimas de la unidad, se debe dividir, a partir del punto $(0, 1)$, a la circunferencia unitaria en n partes iguales. Según esta construcción y la interpretación geométrica de la multiplicación compleja, es claro que multiplicando a $z_{n,1}$ consigo misma se obtienen a todas las demás raíces n -ésimas.

Proposición A.10.2

(Algunas propiedades de la exponencial compleja) Sean $z_1, z_2 \in \mathbb{C}$, $\omega \in \mathbb{Z}$.

- $\frac{\exp(z_1)}{\exp(z_2)} = \exp(z_1 - z_2)$
- $\exp(z_1 + z_2) = \exp(z_1) \cdot \exp(z_2)$
- $(\exp(z_1))^\omega = \exp(\omega z_1)$
- $\exp(z) = 1$ si y sólo si $z = 2K\pi i$ para algún $K \in \mathbb{Z}$
- para todo $y \in \mathbb{R}$,

$$\cos(y) = \frac{\exp(iy) + \exp(-iy)}{2} \quad (\text{A.30})$$

y

$$\sin(y) = \frac{\exp(iy) - \exp(-iy)}{2i}. \quad (\text{A.31})$$

Cambiaste el orden de los items. Checa cuando los citas más abajo.

Definición A.10.3

Sea $n \in \mathbb{N}$. A las n raíces del polinomio $p_n(t) = t^n - 1$ se les denominará las **raíces n -ésimas de la unidad**.

Las raíces n -ésimas de la unidad son pues los números complejos tales que, elevados a la potencia n , son iguales a 1; según el teorema fundamental del álgebra A.1.4, sí hay números complejos que satisfacen la definición A.10.3, y además son a lo más n . Es fácil establecer, como hacemos a continuación, fórmulas explícitas para estos números, que de hecho son exactamente n .

Proposición A.10.4

Sea $n \in \mathbb{N}$, $n \geq 2$. Hay exactamente n raíces n -ésimas de la unidad, y estas son los números complejos

$$z_{n,\omega} := \exp\left(\frac{2\pi i}{n}\omega\right), \quad \text{con } \omega \in \{0, 1, \dots, n-1\}. \quad (\text{A.32})$$

Demostración. Por las propiedades expresadas en la proposición A.10.2, es fácil ver que $z_{n,1} := \exp\left(\frac{2\pi i}{n}\right)$ es raíz n -ésima de la unidad, pues

$$(z_{n,1})^n = \exp(2\pi i) = 1.$$

Además, para todo $\omega \in \{0, \dots, n-1\}$, el número

$$z_{n,\omega} := (z_{n,1})^\omega = \exp\left(\frac{2\pi i}{n}\omega\right)$$

también es raíz n -ésima de la unidad, ya que

$$(z_{n,\omega})^n = ((z_{n,1})^\omega)^n = ((z_{n,1})^n)^\omega = 1^\omega = 1.$$

Note ahora que los n números complejos $z_{n,\omega}$ son todos distintos entre sí, pues si ω_1 y ω_2 son enteros entre 0 y $n-1$ tales que $z_{n,\omega_1} = z_{n,\omega_2}$, o sea, tales que $\exp\left(\frac{2\pi i}{n}\omega_1\right) = \exp\left(\frac{2\pi i}{n}\omega_2\right)$, entonces, según el tercer punto de la proposición A.10.2, $1 = \exp\left(\frac{2\pi i}{n}(\omega_1 - \omega_2)\right)$, luego, según el primer punto de esta misma proposición, $\frac{\omega_1 - \omega_2}{n}$ es entero, o sea, n divide a $\omega_1 - \omega_2$; por el rango de ω_1 y ω_2 , esto sólo ocurre si $\omega_1 - \omega_2$ es cero, o sea, si ω_1 y ω_2 son iguales. \square

A.10.2. La transformada discreta de Fourier

Ya tenemos todo lo necesario para definir a la base de Fourier compleja de la que hablamos al inicio.

Proposición A.10.5

Sea $n \in \mathbb{N}$. El conjunto

$$\mathcal{B}_n := \left\{ e_{n,\omega} = \left(\frac{1}{\sqrt{n}} \exp\left(2\pi i \omega \frac{m}{n}\right) \right)_{0 \leq m \leq n-1} : 0 \leq \omega \leq n-1 \right\} \quad (\text{A.33})$$

es una base ortonormal del \mathbb{C} -espacio vectorial \mathbb{C}^n .

Demostración. Calculemos el producto punto de dos elementos e_{n,ω_1} y e_{n,ω_2} del conjunto (A.33); si $\omega := \omega_1 - \omega_2$,

$$\begin{aligned} \langle e_{n,\omega_1}, e_{n,\omega_2} \rangle &= \frac{1}{n} \sum_{m=0}^{n-1} \exp\left(2\pi i \frac{m}{n} \omega_1\right) \cdot \overline{\exp\left(2\pi i \frac{m}{n} \omega_2\right)} \\ &= \frac{1}{n} \sum_{m=0}^{n-1} \exp\left(2\pi i \frac{m}{n} (\omega_1 - \omega_2)\right) \\ &= \frac{1}{n} \sum_{m=0}^{n-1} \exp\left(2\pi i \frac{\omega}{n} m\right) \\ &= \frac{1}{n} \sum_{m=0}^{n-1} \exp\left(2\pi i \frac{\omega}{n} m\right)^m \\ &= \frac{1}{n} \sum_{m=0}^{n-1} (z_{n,\omega})^m; \end{aligned}$$

esta última es una suma geométrica.

- Si $\omega_1 \neq \omega_2$, entonces n no puede dividir a $\omega = \omega_1 - \omega_2$ (pues, por el rango en el que se encuentran ω_1 y ω_2 , $w \in [-(n-1), n-1]$, y el único múltiplo de n en este intervalo es cero), luego, $z_{n,\omega} \neq 1$. En este caso se tiene entonces que

$$\langle e_{n,w_1}, e_{n,w_2} \rangle = \frac{1}{n} \sum_{m=0}^{n-1} (z_{n,\omega})^m = \frac{1}{n} \cdot \frac{(z_{n,\omega})^n - 1}{z_{n,\omega} - 1} = \frac{1}{n} \cdot \frac{1 - 1}{z_{n,\omega} - 1} = 0.$$

- Si $\omega_1 = \omega_2$, entonces $\omega = 0$, y

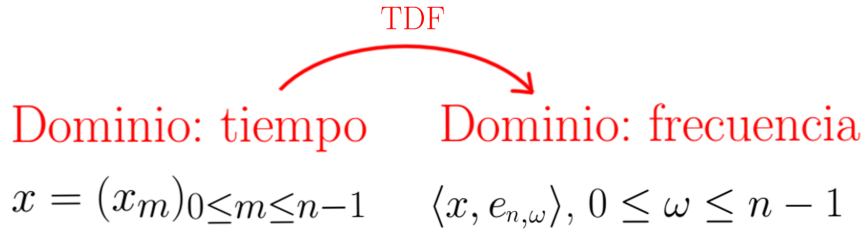
$$\langle e_{n,w_1}, e_{n,w_2} \rangle = \frac{1}{n} \sum_{m=0}^{n-1} (z_{n,0})^m = \frac{1}{n} \sum_{m=0}^{n-1} 1 = \frac{1}{n} \cdot n = 1.$$

Demostremos así que los elementos de \mathcal{B}_n tienen norma uno y que además son ortogonales dos a dos, luego, \mathcal{B}_n es un subconjunto l.i. de \mathbb{C}^n ; como \mathbb{C}^n es un \mathbb{C} -espacio vectorial de dimensión n , concluimos lo deseado. \square

Por ser (A.33) una BON de \mathbb{C}^n , siempre es posible expresar a un vector $x = (x_m)_{0 \leq m \leq n-1} \in \mathbb{C}^n$ como combinación lineal de los elementos de (A.33) y además los coeficientes están dados por los productos puntos de x y los elementos de (A.33).

Definición A.10.6

Al proceso de calcular los coeficientes de x respecto a \mathcal{B}_n se le conoce como el **cálculo de la transformada discreta de Fourier de x** .



Por sus siglas en inglés, a la transformada discreta de Fourier también se le denomina “DFT”.

Figura A.7: Usualmente uno representa a una señal discreta x de dimensión n con n mediciones complejas; en este caso, el dominio de la representación es el tiempo. Pero también se puede representar unívocamente a x con sus coeficientes respecto a la base de frecuencias \mathcal{B}_n ; en este caso, puesto que cada coeficiente da el peso que tiene la respectiva frecuencia para construir la señal original x , decimos que el dominio de la representación es el de frecuencia.

Calcular entonces la transformada discreta de Fourier de x consiste en calcular a los productos punto $\langle x, e_{n,\omega} \rangle$, que son

$$\begin{aligned} \langle x, e_{n,\omega} \rangle &= \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} x_m \exp\left(2\pi i \omega \frac{m}{n}\right) \\ &= \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} x_m \left(\exp\left(\frac{2\pi i}{n} \omega\right) \right)^m \\ &= A_x(z_{n,\omega}), \end{aligned}$$

donde $z_{n,\omega}$ es como en (A.32) y $A_x = A_x(t) \in \mathbb{C}[t]$ es el polinomio de coeficientes complejos definido a partir de x como sigue:

$$A_x(t) := \sum_{m=0}^{n-1} \frac{x_m}{\sqrt{n}} t \in \mathbb{C}[t]; \quad (\text{A.34})$$

así, calcular la transformada discreta de Fourier de x es lo mismo que evaluar al polinomio A_x de grado $n-1$ definido en (A.34) en todas las raíces n -ésimas de la unidad.

Nota A.10.7

Según este último párrafo, calcular transformadas discretas de Fourier requiere de algoritmos eficientes para evaluar polinomios en raíces n -ésimas de la unidad. Usando propiedades de las raíces n -ésimas de la unidad (por ejemplo, véase el lema 30.5 de [Alg]) es posible usar recursión para disminuir el tiempo de cómputo. Al algoritmo estándar usado para esto se le conoce como la **transformada rápida de Fourier** (abreviado como “FFT” por sus siglas en inglés); puede consultar los detalles técnicos en el capítulo 30 de [Alg]

Al usar a \mathcal{B}_n como sistema de representación en \mathbb{C}^n , lo que estamos haciendo es representar a un $x = (x_m)_{0 \leq m \leq n-1}$ como combinación lineal de los vectores

$$e_{n,\omega} = \frac{1}{\sqrt{n}} \left(\cos\left(2\pi\omega\frac{m}{n}\right) + i \sin\left(2\pi\omega\frac{m}{n}\right) \right)_{0 \leq m \leq n-1}, \quad 0 \leq \omega \leq n-1;$$

observe que las partes reales de las entradas de $e_{n,\omega} \in \mathbb{C}^n$ se obtienen de tomar n muestras uniformes de la función $c_\omega(t) := \frac{1}{\sqrt{n}} \cos(2\pi\omega t)$ (o sea, de la función coseno de amplitud $\frac{1}{\sqrt{n}}$, frecuencia ω y desfase 0) y, similarmente, las partes imaginarias de las entradas se obtienen muestreando uniformemente a la función $s_\omega(t) := \frac{1}{\sqrt{n}} \sin(2\pi\omega t)$.

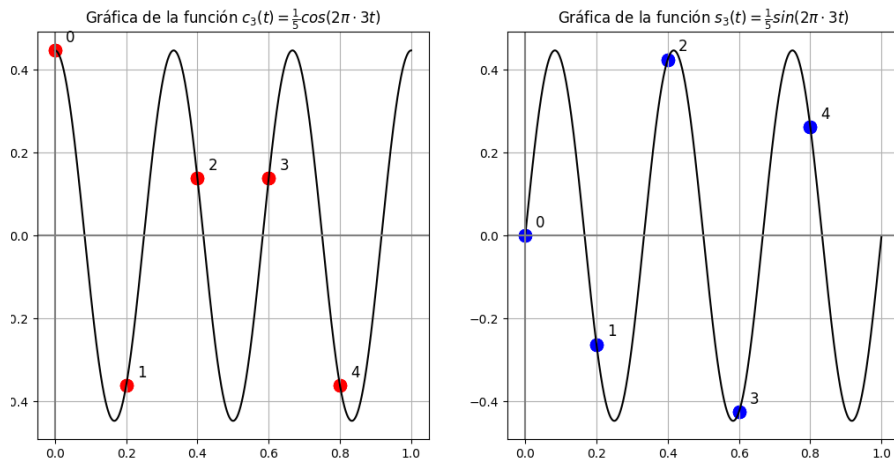


Figura A.8: Por ejemplo, si $n = 5$, para construir al vector e_3 de la base \mathcal{B}_n se muestrean uniformemente cosenos y senos de frecuencia 3 como se muestra en la figura; los puntos rojos representan las partes reales de las entradas y los azules las imaginarias.

Según esto, el vector $e_{n,\omega}$ se construye a partir de funciones de frecuencia ω ; considerando esto y el que \mathcal{B}_n sea una BON de \mathbb{C}^n (luego, el que se valga la identidad de Parseval, c.f. nota A.5.10), tenemos que la síntesis

$$x = \sum_{\omega=0}^{n-1} \langle x, e_{n,\omega} \rangle e_{n,\omega}$$

es una expresión de x en términos de vectores de frecuencia ω y que los respectivos coeficientes $\langle x, e_{n,\omega} \rangle$ indican qué tanto contribuye la frecuencia ω para construir a x .

Es por eso que al proceso de considerar representaciones de señales complejas finitas respecto a las bases de Fourier se le conoce como **realizar un análisis espectral**.

A.10.3. Versión real de la DFT

En el caso en el que todas las entradas de un vector $x = (x_m)_{0 \leq m \leq n-1}$ sean reales, se puede definir una base ortonormal de \mathbb{R}^n , análoga a la BON \mathcal{B}_n de \mathbb{C}^n construida en A.10.5, a partir de muestreos uniformes de sinusoides de frecuencias enteras.

Proposición A.10.8

Sean $n \in \mathbb{N}$ mayor a uno, $M = \lceil \frac{n}{2} \rceil$. Para cualquier $\omega > 0$, sean los vectores

$$c_{n,\omega} := \left(\sqrt{\frac{2}{n}} \cos \left(2\pi\omega \frac{m}{n} \right) \right)_{0 \leq m \leq n-1} \quad \text{y} \quad s_{n,\omega} := \left(\sqrt{\frac{2}{n}} \sin \left(2\pi\omega \frac{m}{n} \right) \right)_{0 \leq m \leq n-1}. \quad (\text{A.35})$$

El subconjunto \mathcal{F}_n de \mathbb{R}^n definido como

- $\mathcal{F}_n := \{c_{n,0}, c_{n,1}, s_{n,1}, \dots, c_{n,M-1}, s_{n,M-1}, c_{n,M}\}$ si n es par (o sea, si $n = 2M$), y como
- $\mathcal{F}_n := \{c_{n,0}, c_{n,1}, s_{n,1}, \dots, c_{n,M-1}, s_{n,M-1}\}$ si n es impar (o sea, si $n = 2M - 1$)

es una base ortonormal del \mathbb{R} -espacio vectorial \mathbb{R}^n .

Demostración. Supongamos n par. Si $0 \leq \omega_1, \omega_2 \leq M$ son enteros, entonces $\omega_1 + \omega_2$ sólo es divisible por n si ambos números son iguales a M . Si suponemos a ω_1 y ω_2 distintos, entonces

$$\begin{aligned} \langle c_{\omega_1}, c_{\omega_2} \rangle &= \frac{1}{n} \sum_{m=0}^{n-1} \cos \left(2\pi\omega_1 \frac{m}{n} \right) \cdot \cos \left(2\pi\omega_2 \frac{m}{n} \right) \\ &= \frac{1}{2n} \left(\cos \left(2\pi(\omega_1 + \omega_2) \frac{m}{n} \right) + \cos \left(2\pi(\omega_1 - \omega_2) \frac{m}{n} \right) \right) \\ &= \frac{1}{4n} \left(\sum_{m=0}^{n-1} (\exp(2\pi m(\omega_1 + \omega_2)i/n) + \exp(-2\pi m(\omega_1 + \omega_2)i/n)) \right. \\ &\quad \left. + \exp(2\pi m(\omega_1 - \omega_2)i/n) + \exp(-2\pi m(\omega_1 - \omega_2)i/n) \right) \\ (\text{suma geométrica}) &= \frac{\exp(2\pi i(\omega_1 + \omega_2)) - 1}{4n(\exp(2\pi i(\omega_1 + \omega_2)/n) - 1)} + \frac{\exp(-2\pi i(\omega_1 + \omega_2)) - 1}{4n(\exp(-2\pi i(\omega_1 + \omega_2)/n) - 1)} \\ &\quad + \frac{\exp(2\pi i(\omega_1 - \omega_2)) - 1}{4n(\exp(2\pi i(\omega_1 - \omega_2)/n) - 1)} + \frac{\exp(-2\pi i(\omega_1 - \omega_2)) - 1}{4n(\exp(-2\pi i(\omega_1 - \omega_2)/n) - 1)}; \end{aligned}$$

puesto que $\omega_1 + \omega_2$ y $\omega_1 - \omega_2$ son ambos enteros, según la proposición A.10.2 las

exponenciales de los numeradores de esta última expresión son todas iguales a uno, luego, $\langle c_{n,\omega_1}, c_{n,\omega_2} \rangle = 0$.

Con argumentos similares se prueba que todos los elementos de \mathcal{F}_n tienen norma uno, así como la ortogonalidad entre dos elementos distintos del conjunto \mathcal{F}_n , por lo tanto, la independencia lineal de este conjunto, luego, el que \mathcal{F}_n sea base (ortonormal) de \mathbb{R}^n .

□

Definición A.10.9

Sea $n \in \mathbb{N}$, $n \geq 2$. Llamaremos a la BON \mathcal{F}_n de \mathbb{R}^n definida en A.10.8 la **base de Fourier real de dimensión n** .

Observe que \mathcal{F}_n , a diferencia de $\mathcal{B}_n \subseteq \mathbb{C}^n$, considera frecuencias enteras no mayores a $M := \lceil \frac{n}{2} \rceil$ (cuando n es par) o a $M - 1$ (cuando n es impar), mientras que en \mathcal{B}_n se consideran las frecuencias enteras entre 0 y $n - 1$ (inclusivo); así, si decidimos representar a una señal $x \in \mathbb{R}^n$ en base a $\mathcal{F}_n \subseteq \mathbb{R}^n$ y no en base a $\mathcal{B}_n \subseteq \mathbb{C}^n$, sintetizaremos a x respecto a frecuencias enteras acotadas por M o por $M - 1$ (dependiendo de la paridad de n), y no respecto a frecuencias menores a n .

Ejemplo 4. Consideremos a la señal

$$x = (-0.5, -8, -5.3, 15, -0.3, 6, 4) \in \mathbb{R}^7. \quad (\text{A.36})$$

Según la construcción de \mathcal{F}_7 (c.f. proposición A.10.8), una expresión de x respecto a \mathcal{F}_7 es una síntesis de x a partir de señales de frecuencias $\omega = 0, 1, 2, 3$. En la imagen de abajo se muestran los coeficientes de x respecto a \mathcal{F}_7 .

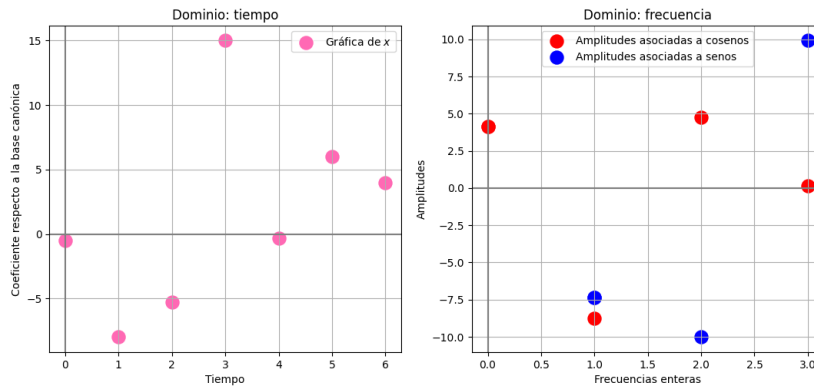


Figura A.9: Se muestran la gráfica de x junto con la gráfica de los coeficientes de x respecto a la BON \mathcal{F}_7 . Observe que, por definición, sólo un vector de \mathcal{F}_7 tiene frecuencia cero (i.e. es constante), mientras que para las otras frecuencias tenemos dos vectores de la misma frecuencia, uno construido a partir de un coseno y otro a partir de un seno.

Redondeando los coeficientes, se tiene la siguiente descomposición de x ;

$$x = 4.12c_0 - 8.76c_1 - 7.35s_1 + 4.77c_2 - 10s_2 + 0.14c_3 + 9.91s_3.$$

A continuación mostramos las gráficas de los sinusoides que fueron discretizados para obtener los vectores de frecuencia 0, 1, 2 y 3 en los que descompusimos a x .

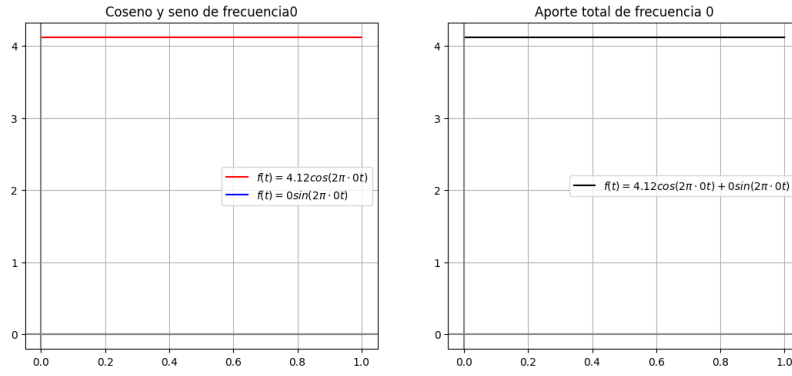


Figura A.10: Aporte de frecuencia 0.

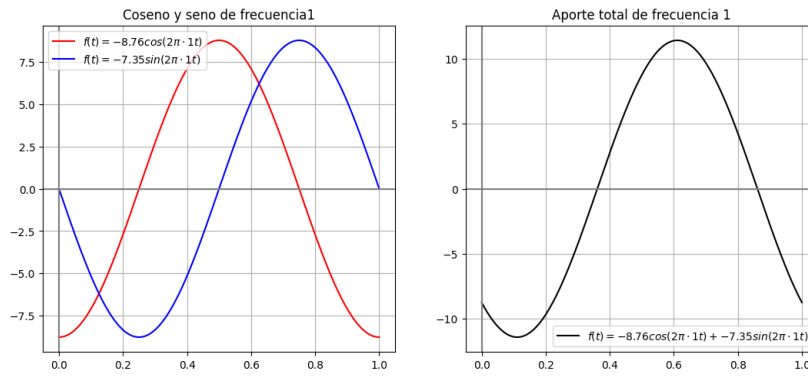


Figura A.11: Aporte de frecuencia 1.

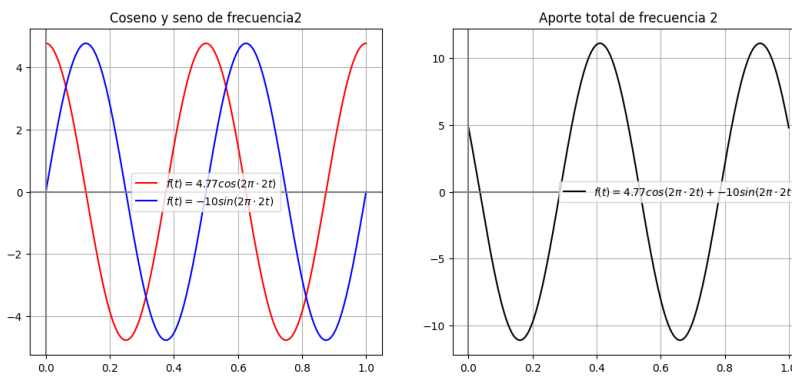


Figura A.12: Aporte de frecuencia 2.

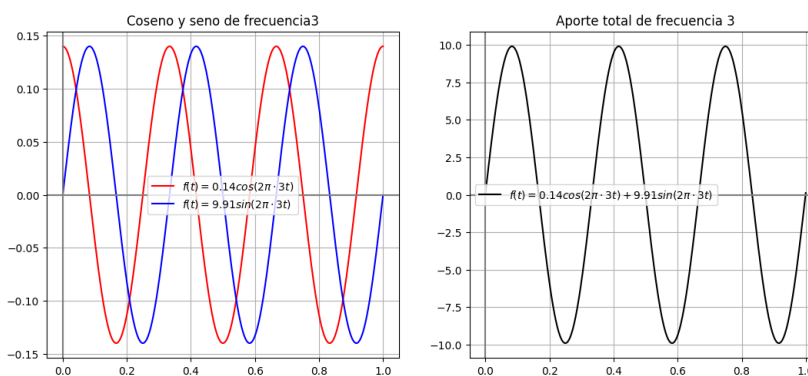


Figura A.13: Aporte de frecuencia 3.

Sumando todas las gráficas de la derecha, obviamente obtenemos una función

de cosenos y senos tal que, al muestrearla uniformemente en $[0,1]$, obtenemos al vector x (A.36).

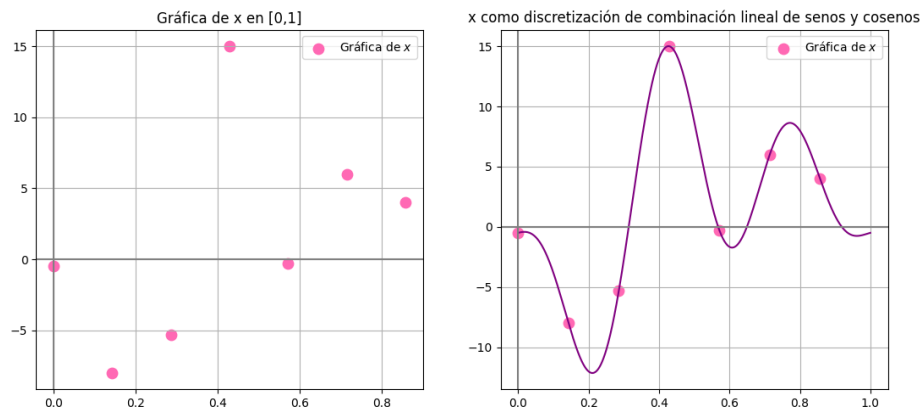


Figura A.14: En morado se muestra la gráfica de la función suma de las gráficas derechas en las figuras anteriores.

◇

Deberías poner imágenes de espectros!

Apéndice B

Notaciones y abreviaciones

Notaciones

No es necesario ponerlo en forma de lista, también puedes formar párrafos.
Poner una referencia para cuando el término aparezca por primera vez en el texto.

Fading factorial:

$$x = (x_j)_{j=0}^{n-1}$$

$:=$

$$\equiv \pmod{n}$$

$//$

\mathbb{N}

$\overline{\mathbb{N}}$

\mathbb{Z}

\mathbb{R}

\mathbb{R}^+

\mathbb{R}_0^+

$$\text{span}(W)$$

(con W subconjunto de un espacio vectorial) el subespacio de V que consta de todas las combinaciones lineales de los vectores de W

$$\mathbb{R}[x]$$

El espacio vectorial real de los polinomios con coeficientes reales

$$\partial(f)$$

Grado de un polinomio f

$$\langle, \rangle$$

Usado para denotar un producto punto

$$\| \quad \|$$

Usado para denotar una norma

$$\Pi_W$$

Proyección sobre el subespacio cerrado W (inserta referencia)

$$A^B$$

El conjunto de funciones de B a A .

$$B^\perp$$

Complemento ortogonal de un subconjunto B de un espacio con producto interno

$$\oplus$$

Suma directa

$$\leq$$

Uso esto para subespacios?

$$\angle$$

ángulos.

$$\boxplus$$

Suma ortogonal

$$\boxminus$$

Diferencia directa de dos subconjuntos A y B , $A \boxminus B := A \cap B$

Abreviaciones

G-S	Gram-Schmidt
C-S	Cauchy-Schwarz
BON	Base ortonormal
DLOP	Abreviación (adoptada de [NS74]) de “polinomio ortogonal discreto de Legendre”
sii	si y sólo si
l.i.	linealmente independiente

Bibliografía

- [AK] S.V. Fomin A.N. Kolmogorov. *Introductory real analysis*. Dover Publications.
- [Alg] *Introduction to algorithms*. The MIT Press, 2009.
- [Car] *Real analysis*. Cambridge University Press, 2000.
- [Con90] John B Conway. *A course in functional analysis*. 2nd edition. Graduate Texts in Mathematics. Springer, 1990.
- [Cos] *Cosine similarity*. https://en.wikipedia.org/wiki/Cosine_similarity. Accessed: 2023-04-11.
- [Fri02] Stephen H. Friedberg. *Linear Algebra*. 4th edition. Pearson, 2002.
- [God66] Roger Godement. *Cours d'algèbre*. Hermann éditeurs, 1966.
- [Jac85] Nathan Jacobson. *Basic Algebra I*. 2nd edition. W H Freeman y Company, 1985.
- [JEM99] Michael J. Hoffman Jerrold E. Marsden. *Basic complex analysis*. 3rd edition. W H Freeman y Company, 1999.
- [Kre] *Introductory functional analysis with applications*. John Wiley y Sons, 1991.
- [Kur68] A. G. Kurosch. *Curso de álgebra superior*. edition. Editorial MIR, 1968.
- [Lan97] Serge Lang. *Undergraduate Analysis*. 2nd edition. Undergraduate Texts in Mathematics. Springer New York, NY, 1997.
- [Msea] *An orthonormal set cannot be a basis in an infinite dimension vector space?* <https://math.stackexchange.com/questions/13641/an-orthonormal-set-cannot-be-a-basis-in-an-infinite-dimension-vector-space?>. Accessed: 2023-04-11.
- [Mseb] *Proof that every finite dimensional normed vector space is complete*. <https://math.stackexchange.com/questions/168275/proof-that-every-finite-dimensional-normed-vector-space-is-complete>. Accessed: 2023-04-11.
- [Msec] *Why isn't every Hamel basis a Schauder basis?* <https://math.stackexchange.com/questions/1653691/why-isnt-every-hamel-basis-a-schauder-basis?>. Accessed: 2023-04-11.
- [Mun00] James R. Munkres. *Topology*. 2dn edition. Pearson College Div, 2000.
- [Nim] Kristoffer Nimark. "The Projection Theorem". <http://www.kris-nimark.net/pdf/ProjectionTheorem.pdf>.

- [NS74] Charles P Neuman y Dave I Schonbach. “Discrete (Legendre) orthogonal polynomials—a survey”. En: *International Journal for Numerical Methods in Engineering* 8.4 (1974), págs. 743-770.
- [Rot10] Joseph J. Rotman. *Advanced Modern Algebra*. 2nd edition. Graduate Studies in Mathematics. American Mathematical Society, 2010.