

Espacios vectoriales lineales

La ruta de este capítulo

Este es el capítulo central de esta obra, en el cual discutiremos los conceptos fundamentales sobre espacios vectoriales abstractos, sus propiedades y sus múltiples expresiones: vectores de \mathbb{R}^n , matrices, polinomios, funciones continuas (entre otras). Además, introduciremos la notación de *bra* y *ket* que facilita mucho la organización de los conceptos y, que nos acompañará por el resto de este libro. Haremos una constante referencia a los conceptos que fueron discutidos en el capítulo anterior en el marco de los vectores en \mathbb{R}^3 . Como en el capítulo 1, cada una de las secciones presenta ejemplos con la utilización de la herramienta de cómputo algebraico **Maxima**, que consideramos parte fundamental de estas notas y nos permite mostrar el alcance de los conceptos abstractos.

Para empezar, en la próxima sección, iniciamos con una discusión somera sobre grupos y sus propiedades. Seguimos con la sección 2.1.3, en la cual presentamos el concepto de espacio vectorial lineal para describir varios de sus escenarios que usualmente se presentan desconectados. Hacemos un esfuerzo por ilustrar, bajo un mismo enfoque, su aplicación desde \mathbb{R}^n hasta los espacios vectoriales de funciones continuas $\mathcal{C}_{[a,b]}$. Luego, en la sección 2.2.3.1, abordamos los conceptos de distancia (métrica), norma y, finalmente el de producto interno. Esta última definición nos permite equipar los espacios vectoriales, no solo con norma y distancia, sino también dotarlos de geometría, definiendo ángulos entre vectores abstractos. Nos detenemos un momento para discutir el significado de ángulo entre vectores pertenecientes a espacios vectoriales reales y complejos. En la sección 2.3 mostramos el concepto de variedades lineales, discutimos el criterio de independencia lineal y, a partir de éste definimos las bases para los espacios vectoriales. Apoyándonos en el criterio de ortogonalidad construimos bases ortogonales para varios espacios vectoriales, discutimos los subespacios vectoriales ortogonales y las proyecciones ortogonales. El capítulo lo finalizamos mostrando la utilidad de expresar funciones como combinación lineal de vectores ortogonales y, cómo este tipo de expansiones constituye la mejor aproximación a la función.

2.1 Grupos, campos y espacios vectoriales

En los cursos básicos de matemáticas nos enseñaron el concepto de conjunto: una colección de elementos de una misma naturaleza, y aprendimos una variedad de operaciones aplicadas a los elementos que conforman los conjuntos y a los conjuntos mismos. Es probable que en esos tiempos nos quedara la sensación de la poca

utilidad de todos esos conceptos de la teoría de conjuntos, pero como veremos en esta sección, la noción de conjuntos es fundamental para desarrollar todas las ideas que conforman lo que se conoce como el *álgebra abstracta*. En el estudio de las estructuras algebraicas se incluyen las teorías sobre: grupos, anillos, campos, espacios vectoriales, redes, álgebras; que merecen ahora nuestra atención. Comenzaremos con la estructura de grupo, y nos daremos cuenta de que una buena cantidad de objetos matemáticos, en apariencia muy diferentes, tienen incorporadas la estructura de grupo. La noción de grupo no llevará entonces al importante concepto de espacios vectoriales abstractos que discutiremos en la sección 2.1.3 y a la definición de los espacios métricos, sección 2.2.1.

2.1.1 Grupos

Considere el siguiente conjunto no vacío $G = \{g_1, g_2, g_3, \dots, g_n, \dots\}$ y la operación interna \square (la ley del grupo). Entonces los elementos del conjunto forman un grupo respecto a la operación \square si $\forall g_i \in G$ se cumplen las siguientes condiciones:

1. Cerrada respecto a la operación \square : $\{g_i \in G, g_j \in G\} \Rightarrow \exists g_k = g_i \square g_j \in G$.
2. Asociativa respecto a la operación \square : $g_k \square (g_i \square g_j) = (g_k \square g_i) \square g_j$.
3. Existencia de un elemento neutro: $\exists \hat{g} \in G \Rightarrow g_i \square \hat{g} = g_i = \hat{g} \square g_i$.
4. Existencia de un elemento inverso: $g_i \in G \Rightarrow \exists g_i^{-1} \in G \Rightarrow g_i \square g_i^{-1} = g_i^{-1} \square g_i = \hat{g}$.
Si adicionalmente se cumple que:
5. Conmutativa respecto a la operación \square : $g_i \square g_j \equiv g_j \square g_i$; el grupo se denomina *grupo abeliano*¹
Suele denotarse al grupo G como (G, \square) para indicar el tipo de operación.

2.1.1.1 Algunos grupos:

- Los enteros $\mathbb{Z} = \{\dots - 3 - 2, -1, 0, 1, 2, 3, \dots\}$ respecto a la suma.
- Los racionales $\mathbb{Q} = \{a/b\}$ respecto a la suma y a la multiplicación.
- Los números complejos $\mathbb{C} = \{e^{i\theta}\}$ respecto a la multiplicación.
- Las rotaciones en 2 dimensiones (2D), y las rotaciones en 3D (grupo no-abeliano).
- Las matrices de dimensión $n \times m$ respecto a la suma (grupo abeliano).

Consideremos el siguiente grupo de tres elementos: $G = \{1, a, b\}$, con $\hat{g} \equiv 1$ y la operación \square . Por construcción, si queremos que la operación de dos de los elementos provea un tercero distinto, entonces la ÚNICA “tabla de multiplicación” posible será:

\square	1	a	b
1	1	a	b
a	a	b	1
b	b	1	a

Notemos que el grupo más simple será aquel conformado únicamente por el elemento neutro: $G = \{1\}$ y, como veremos más adelante, se puede definir subgrupos si un subconjunto de los elementos de un grupo $g_i \in G$ también forman un grupo.

Podemos resumir, sin demostración, las principales propiedades que presenta un grupo (G, \square) .

1. El elemento identidad o neutro es único.

¹Niels Henrik Abe (1802-1829) Pionero en el desarrollo de diferentes ramas de la matemática moderna, la teoría de las integrales de funciones algebraicas. https://es.wikipedia.org/wiki/Niels_Henrik_Abel

2. Para todo $g \in G$ existe un inverso único $g^{-1} \in G$.
3. Para todo elemento $g \in G$ se cumple que: $(g^{-1})^{-1} = g$.
4. Para cualesquiera g_1 y $g_2 \in G$, se cumple que: $(g_1 \square g_2)^{-1} = g_1^{-1} \square g_2^{-1}$.

El número de los elementos de un grupo puede ser finito o infinito. En el primer caso se denominan *grupos finitos* y el número de elementos que contenga, el cardinal del conjunto, se conoce como el *orden del grupo*. Un grupo finito que se construye a partir de una operación con un único miembro se denomina *grupo cíclico*, y el caso más elemental es $G = \{1, g, g^2, g^3, \dots, g^{n-1}\}$. Obviamente hemos definido aquí: $g^2 = g \square g$ y $g^3 = g^2 \square g = g \square g \square g$ y así consecutivamente hasta ejecutarse $n - 1$ veces, entonces se retoma el elemento identidad, esto es: $g^{n-1} \square g = g^n = 1$.

2.1.1.2 Subgrupos

Dado un grupo (G, \square) y H un subconjunto de G , $H \subseteq G$. Si H es un grupo bajo la operación \square , definida para G , entonces diremos que H es un subgrupo. De manera equivalente: si $H \subseteq G$ entonces H es un subgrupo si para cualesquiera h_1 y $h_2 \in H$ se cumple:

1. $h_1 \square h_2 \in H$.
2. $h^{-1} \in H \forall h \in H$.

Es decir, $1 \in H$ y H es cerrado para la misma operación que define G y para sus inversos.

Existe una condición necesaria para que un subconjunto de un grupo finito G sea un subgrupo, esta condición se conoce como el teorema de Lagrange, que dice lo siguiente:

Teorema 2.1 (Teorema de Lagrange)

Si (G, \square) es un grupo finito y H un subgrupo de G , entonces el orden de H divide al orden de G .

Por otro lado, sea (G, \square) un grupo y $g \in G$, se puede demostrar que el conjunto $H = \{g^n\}$, con n perteneciente al conjunto de los números enteros, es un subgrupo de G . Con la propiedad adicional de que H es el subgrupo, que contiene a g , más pequeño de G .

Como mencionamos con anterioridad, un grupo finito que se construye a partir de operar con un único miembro se le denomina *grupo cíclico*, en este caso se dice que H es un subgrupo cíclico generado por g . A este elemento se le llama el generador de H y se acostumbra a denotarlo con $H = \langle g \rangle$.

Para finalizar, definiremos el *orden de un elemento* de un grupo. Sea un elemento $g \in G$, el orden de g es el menor entero positivo n tal que: $g^n = 1$. Cuando este entero positivo no existe se dice que g tiene un orden infinito.

Ejemplo 2.1 El conjunto de las rotaciones del espacio ordinario cuando rotamos un ángulo ϕ alrededor del eje z , ecuación (1.13), forman un grupo.

Denominaremos a este grupo por $R_{z\phi} = \{G_{\phi_i}\}$, con $0 \leq \phi \leq 2\pi$. Donde con G_{ϕ} estaremos indicando un giro alrededor del eje z y la operación que nos define las rotaciones la definiremos de la siguiente forma:

$$G_{\phi_1} \square G_{\phi_2} = G_{\phi_1 + \phi_2}.$$

Entonces podemos ver que:

1. \square es una operación cerrada ya que dos rotaciones resulta en otra rotación:

$$G_{\phi_1} \square G_{\phi_2} = G_{\phi_1 + \phi_2} = G_{\phi_3} \in R_{z\phi}.$$

2. \square es asociativa:

$$(G_{\phi_1} \square G_{\phi_2}) \square G_{\phi_3} = G_{\phi_1} \square (G_{\phi_2} \square G_{\phi_3}).$$

3. Existe el elemento neutro, \mathcal{G}_{ϕ_0} , que no hace ninguna rotación:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_0} = \mathcal{G}_{\phi_0} \square \mathcal{G}_{\phi_1} = \mathcal{G}_{\phi_1}.$$

4. Existe el elemento inverso, \mathcal{G}_{ϕ_-} , ya que podemos rotar en un sentido y en sentido inverso:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_-1} = \mathcal{G}_{\phi_-1} \square \mathcal{G}_{\phi_1} = \mathcal{G}_{\phi_0}.$$

5. Es conmutativa ya que las rotaciones no se afectan por el orden en que son producidas:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_2} = \mathcal{G}_{\phi_2} \square \mathcal{G}_{\phi_1}.$$

Ejemplo 2.2 Consideremos el grupo $\mathbf{G} = \mathbb{Z} \times \mathbb{Z}$, es decir, el conjunto de duplas (x_i, y_i) de números enteros, y la operación:

$$(x_1, y_1) \square (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

Aquí el símbolo $+$ es la suma convencional de números enteros.

Probemos que tenemos un grupo.

1. La operación \square es cerrada para la suma:

$$(x_1, y_1) \square (x_2, y_2) = (x_1 + x_2, y_1 + y_2) = (x_3, y_3) \in \mathbf{G}.$$

2. La operación \square es asociativa:

$$\begin{aligned} ((x_1, y_1) \square (x_2, y_2)) \square (x_3, y_3) &= (x_1 + x_2, y_1 + y_2) \square (x_3, y_3) = ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3) \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)) = (x_1, y_1) \square ((x_2, y_2) \square (x_3, y_3)). \end{aligned}$$

3. Existe el elemento neutro, $(0, 0)$:

$$(x_1, y_1) \square (0, 0) = (0, 0) \square (x_1, y_1) = (x_1, y_1).$$

4. Existe el elemento inverso, $(-x_1, -y_1)$:

$$(x_1, y_1) \square (-x_1, -y_1) = (-x_1, -y_1) \square (x_1, y_1) = (0, 0).$$

5. La operación \square es conmutativa:

$$(x_1, y_1) \square (x_2, y_2) = (x_2, y_2) \square (x_1, y_1).$$

Ejemplo 2.3 Consideremos el conjunto de las permutaciones de 3 objetos, cuyos elementos pueden ser representados como se muestra a continuación:

$$\begin{aligned} \mathbb{P}_0 &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}; \mathbb{P}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}; \mathbb{P}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}; \\ \mathbb{P}_3 &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}; \mathbb{P}_4 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}; \mathbb{P}_5 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}; \end{aligned}$$

y la operación de composición de permutaciones:

$$\mathbb{P}_1 \odot \mathbb{P}_3 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \odot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \mathbb{P}_4,$$

es decir, luego de intercambiar las primera y segunda posición, intercambio la segunda y la tercera. La tabla de multiplicación del grupo quedará:

\odot	\mathbb{P}_0	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_4	\mathbb{P}_5
\mathbb{P}_0	\mathbb{P}_0	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_4	\mathbb{P}_5
\mathbb{P}_1	\mathbb{P}_1	\mathbb{P}_0	\mathbb{P}_5	\mathbb{P}_4	\mathbb{P}_3	\mathbb{P}_2
\mathbb{P}_2	\mathbb{P}_2	\mathbb{P}_4	\mathbb{P}_0	\mathbb{P}_5	\mathbb{P}_1	\mathbb{P}_3
\mathbb{P}_3	\mathbb{P}_3	\mathbb{P}_5	\mathbb{P}_4	\mathbb{P}_0	\mathbb{P}_2	\mathbb{P}_1
\mathbb{P}_4	\mathbb{P}_4	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_1	\mathbb{P}_5	\mathbb{P}_0
\mathbb{P}_5	\mathbb{P}_5	\mathbb{P}_3	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_0	\mathbb{P}_4

Ver el ejercicio 3 de 2.1.6.

2.1.1.3 Grupos Isomorfos

Muchas veces podemos comparar la estructura entre grupos buscando relaciones entre ellos, esto lo podemos hacer estudiando, por ejemplo, cómo “se multiplican”. La búsqueda de estas relaciones no llevan al concepto de *isomorfismo* de grupos. Podemos decir que dos grupos son isomorfos porque tienen aspectos en común dentro de sus estructuras algebraicas, es decir, tienen las mismas propiedades. De esta manera, los isomorfismo permiten clasificar los grupos por sus relaciones de equivalencia.

A continuación, veremos algunos ejemplos de grupos y sus tablas de “multiplicación”.

Ejemplo 2.4 Consideremos el siguiente conjunto y una operación que define la “multiplicación”² módulo \times , es decir, $(a \cdot b)_{\text{mod } \times}$

1. $G_{\text{mod } 8} = \{1, 3, 5, 7\}$. Tenemos la operación multiplicación módulo 8 y la tabla de multiplicación del grupo es

$\times \text{mod } 8$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

Para entender como funciona la tabla anterior, recordemos la regla que estamos utilizando. Por ejemplo, para los elementos 3 y 7, tenemos lo siguiente: $3 \cdot 7 = 21$ y el residuo de dividir 21 entre 8 es 5, por lo tanto

$$(3 \cdot 7)_{\text{mod } 8} = (21)_{\text{mod } 8} = 5 \Leftarrow 21 \div 8 = 2 \times 8 + 5.$$

y se denota: $21 \text{ mod } 8 = 5$.

2. $G_{\text{mod } 24} = \{1, 5, 7, 11\}$, con la operación multiplicación módulo 24:

²Vamos a considerar lo que se conoce como *Aritmética Modular*. Sabemos que para la división de dos enteros a y b ($b \neq 0$) existe un único entero c tal que $a = cb + r$ con $0 \leq r < |b|$. Y como es conocido: a , b , c y r son: el dividendo, el divisor, el cociente y el resto, respectivamente. Cuando $r = 0$ se dice que a es divisible por b . En esta “aritmética” tan particular se establece una relación de congruencia sobre los enteros: para un entero positivo n , dos enteros a y b se llaman congruentes módulo n ($\text{mod } n$) si a y b tienen el mismo resto r cuando se dividen entre n , y se denota $a \equiv b \text{ mod } n$.

$\times \text{mod} 24$	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

Diremos que los grupos $G_{\text{mod}8}$ y $G_{\text{mod}24}$ son *isomorfos* porque tienen tablas equivalentes de multiplicación. Esto es, dado un grupo genérico $G = \{1, A, B, C\}$ o de manera equivalente $\{1, \circ, \square, \triangle\}$ sus tablas de multiplicación serán las que se muestran a continuación:

\times	1	A	B	C
1	1	A	B	C
A	A	1	C	B
B	B	C	1	A
C	C	B	A	1

 \Leftrightarrow

\times	1	\circ	\square	\triangle
1	1	\circ	\square	\triangle
\circ	\circ	1	\triangle	\square
\square	\square	\triangle	1	\circ
\triangle	\triangle	\square	\circ	1

Note que $A^{-1} = A$, de la misma manera $\triangle = \triangle^{-1}$, es decir, que siempre la operación de dos elementos da uno distinto a los operados. Esto último también se puede representar de la manera siguiente: $A \times A^{-1} = 1$, es decir, el resultado es el elemento neutro del grupo. Tenemos entonces una operación sobre los elementos del grupo que es biyectiva.

Consideremos ahora el siguiente ejemplo con otros dos grupos isomorfos

Ejemplo 2.5

1. $G_{\text{mod}5} = \{1, 2, 3, 4\}$, y la operación multiplicación módulo 5:

$\times \text{mod} 5$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

 \Leftrightarrow

$\times \text{mod} 5$	1	2	4	3
1	1	2	4	3
2	2	4	3	1
4	4	3	1	2
3	3	1	2	4

Note que las tablas anteriores se diferencian por el hecho de que se cambió el orden de los elementos, por lo tanto, son dos tablas de multiplicación diferentes porque los elementos de los grupos están dispuestos de manera diferente.

2. $G_{\times} = \{1, i, -1, -i\}$, y la siguiente operación multiplicación:

\times	1.	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Por lo tanto, los grupos G_{\times} y $G_{\text{mod}5}$ son isomorfos con una tabla de multiplicación como la que se muestra a continuación:

\times	1	A	B	C
1	1	A	B	C
A	A	B	C	1
B	B	C	1	A
C	C	1	A	B

2.1.2 Campo

Definiremos como un campo (o cuerpo) el conjunto $K = \{\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n, \dots\}$ sobre el cual están definidas dos operaciones: suma (+) y multiplicación (\cdot) y que satisfacen las siguientes propiedades:

1. Forman un grupo abeliano respecto a la suma (+) con el elemento neutro representado por el cero: 0.
2. Forman un grupo abeliano respecto a la multiplicación (\cdot). Se excluye el cero y se denota el elemento neutro de la multiplicación como 1.
3. Es distributiva respecto a la suma (+). Dados α_i, α_j y α_k , se tiene que: $\alpha_i \cdot (\alpha_j + \alpha_k) = \alpha_i \cdot \alpha_j + \alpha_i \cdot \alpha_k$.

Ejemplos típicos de campos lo constituyen el conjunto de los números racionales \mathbb{Q} , los números reales \mathbb{R} y los números complejos \mathbb{C} . Normalmente se refiere estos campos como *campos escalares*.

La noción de campo nos permite introducir el importante concepto de espacio vectorial.

2.1.3 Espacios vectoriales lineales

Sea el conjunto de objetos $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_i\rangle, \dots\}$. Se denominará V un espacio vectorial lineal y sus elementos $|v_i\rangle$ vectores, si existe una operación suma, \boxplus , respecto a la cual los elementos $|v_i\rangle \in V$ forman un grupo abeliano y una operación multiplicación por un elemento de un campo, $K = \{\alpha, \beta, \gamma, \dots\}$, tal que³:

1. V es cerrado bajo la operación \boxplus : $|v_k\rangle = |v_i\rangle \boxplus |v_j\rangle \in V \quad \forall |v_i\rangle, |v_j\rangle \in V$.
2. La operación \boxplus es conmutativa: $|v_i\rangle \boxplus |v_j\rangle = |v_j\rangle \boxplus |v_i\rangle \quad \forall |v_i\rangle, |v_j\rangle \in V$.
3. La operación \boxplus es asociativa: $(|v_i\rangle \boxplus |v_j\rangle) \boxplus |v_k\rangle = |v_i\rangle \boxplus (|v_j\rangle \boxplus |v_k\rangle) \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in V$.
4. Existe un único elemento neutro $|0\rangle$: $|0\rangle \boxplus |v_i\rangle = |v_i\rangle \boxplus |0\rangle = |v_i\rangle \quad \forall |v_i\rangle \in V$.
5. Existe un elemento simétrico para cada elemento de V : $\exists |-v_i\rangle / |v_i\rangle \boxplus |-v_i\rangle = |0\rangle \quad \forall |v_i\rangle \in V$.
6. V es cerrado bajo el producto por un número: $\alpha |v_i\rangle \in V \quad \forall \alpha \in K$ y cualquier $|v_i\rangle \in V$,
7. $\alpha (\beta |v_i\rangle) = (\alpha\beta) |v_i\rangle \quad \forall |v_i\rangle \in V$ y $\alpha, \beta \in K$.
8. $(\alpha + \beta) |v_i\rangle = \alpha |v_i\rangle \boxplus \beta |v_i\rangle \quad \forall |v_i\rangle \in V$ y $\alpha, \beta \in K$.
9. $\alpha (|v_i\rangle \boxplus |v_j\rangle) = \alpha |v_i\rangle \boxplus \alpha |v_j\rangle \quad \forall |v_i\rangle, |v_j\rangle \in V$ y $\alpha \in K$.
10. $1 |v_i\rangle = |v_i\rangle \quad \forall |v_i\rangle \in V$ y $1 \in K$.

Es importante resaltar lo siguiente: existen dos objetos neutros, el vector $|0\rangle \in V$ y el elemento $0 \in K$ y también dos operaciones producto diferentes, el producto de dos números dentro de K y el producto de un $\alpha \in K$ por un vector $|v\rangle \in V$.

Notemos también que podemos definir subespacios S vectoriales dentro de los espacios vectoriales. Ellos serán aquellos conjuntos de vectores que cumplan con los requisitos anteriores pero además cerrados dentro de los mismos conjuntos de vectores. Se puede ver entonces que la condición necesaria y suficiente para que

³También suele decirse: un espacio vectorial V sobre K .

$S \subseteq V$ sea un subespacio vectorial de V es que para cualesquier $|u_i\rangle$ y $|v_i\rangle$ de S y para cualesquier α y β de K se tiene que: $\alpha |u_i\rangle + \beta |v_i\rangle \in S$.

2.1.4 Algunos espacios vectoriales

1. Los conjunto de los números reales $V = \mathbb{R}$ y el conjunto de los números complejos $V = \mathbb{C}$ con el campo K de reales o complejos y definidas las operaciones ordinarias de suma y multiplicación.

Cuando el campo K es el conjunto de los números reales se dirá que es un *espacio vectorial real de números reales* si $V \equiv \mathbb{R}$, pero si $V \equiv \mathbb{C}$ se dirá un *espacio vectorial real de números complejos*. Por su parte, si $K \equiv \mathbb{C}$ diremos que es un espacio vectorial complejo. Siempre se asociará el campo de escalares al espacio vectorial: se dirá que es un espacio vectorial sobre el campo de los escalares. Es decir, si el campo es real (complejo) se dirá que el espacio vectorial es real (complejo).

2. El conjunto $V \equiv \mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$, vale decir el producto cartesiano de \mathbb{R} , cuyos elementos son n -uplas de números, con la operación suma ordinaria de vectores en n -dimensionales y la multiplicación por números.

$$\begin{aligned} |x\rangle &= (x_1, x_2, x_3, \cdots, x_n) \quad \wedge \quad |y\rangle = (y_1, y_2, y_3, \cdots, y_n), \\ |x\rangle \boxplus |y\rangle &\equiv (x_1 + y_1, x_2 + y_2, x_3 + y_3, \cdots, x_n + y_n), \\ \alpha |x\rangle &= (\alpha x_1, \alpha x_2, \alpha x_3, \cdots, \alpha x_n). \end{aligned}$$

Este espacio vectorial es de dimensión finita.

Igualmente, será un espacio vectorial $V \equiv \mathbb{C}^n = \mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}$, en donde los elementos $x_i \in \mathbb{C}$. Si para este caso el campo, sobre el cual se define el espacio vectorial \mathbb{C}^n es real, tendremos un espacio vectorial real de números complejos.

Es obvio que en el caso $V \equiv \mathbb{R}$, para el cual $|x\rangle_1 = (x_1, 0, 0, \cdots, 0)$ y $|y\rangle_1 = (y_1, 0, 0, \cdots, 0)$ o cualquier espacio de vectores formados por las componentes, i.e. $|x\rangle_i = (0, 0, 0, \cdots, x_i, \cdots, 0)$ y $|y\rangle_i = (0, 0, 0, \cdots, y_i, \cdots, 0)$ formarán subespacios vectoriales dentro de \mathbb{R}^n .

En el caso específico de \mathbb{R}^3 , y en donde hemos desarrollado todo un conjunto de conceptos matemáticos, es bueno comentar sobre la equivalencia que hemos tomado como obvia entre dos definiciones diferentes:

- Los vectores $\mathbf{a} = (a^1, a^2, a^3)$ y $\mathbf{b} = (b^1, b^2, b^3)$ con sus respectivas operaciones para la suma $\mathbf{a} + \mathbf{b} = (a^1 + b^1, a^2 + b^2, a^3 + b^3)$ y multiplicación por un escalar $\lambda \mathbf{a} = (\lambda a^1, \lambda a^2, \lambda a^3)$, con $\{a^i\}$, $\{b^i\}$ y $\lambda \in \mathbb{R}$.
- Los vectores geométricos, es decir, segmentos orientados en el espacio con un origen común y donde la suma se definió mediante la regla del paralelogramo y el producto por un escalar como el alargamiento o acortamiento de los segmentos o flechas con el posible cambio de dirección.

Aunque ambos son vectores de \mathbb{R}^3 es bueno tener claro que se trata de objetos diferentes que viven en espacios vectoriales diferentes. Se puede decir también que son dos representaciones diferentes para los vectores. Es buena la ocasión para señalar que existe una manera de pasar de una representación a otra, esto se hace por medio de una función que asigne a una tríada (x^1, x^2, x^3) una y sólo una flecha (y viceversa), conservando por supuesto las operaciones de suma y multiplicación por números. A este tipo de funciones se les denomina un *isomorfismo*.

Para finalizar, también es posible ver \mathbb{R}^3 simplemente como un conjunto de puntos donde se pueden definir superficies inmersas, como por ejemplo una esfera, S^2 , centrada en el origen y de radio R . En un punto q sobre la esfera es posible generar un plano tangente a la esfera y en ese punto construir un espacio vectorial con todos los vectores cuyo origen está en q , y por lo tanto, son tangentes a la esfera.

En el lenguaje de la geometría diferencial y las variedades a este conjunto de vectores se le denota con: $T_q\mathbb{S}^2$. Lo anteriormente dicho se puede generalizar para \mathbb{R}^n .

3. El espacio E^∞ constituido por vectores $|x\rangle = (x_1, x_2, x_3, \dots, x_n, \dots)$ contables pero con infinitas componentes.

$$\begin{aligned} |x\rangle &= (x_1, x_2, x_3, \dots, x_n, \dots) \quad \wedge \quad |y\rangle = (y_1, y_2, y_3, \dots, y_n, \dots), \\ |x\rangle \boxplus |y\rangle &\equiv (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots, x_n + y_n, \dots) \\ \alpha |x\rangle &= (\alpha x_1, \alpha x_2, \alpha x_3, \dots, \alpha x_n, \dots), \end{aligned}$$

con la siguiente restricción:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n x_i = L, \quad \text{con } L \text{ finito.}$$

4. El conjunto de las matrices $n \times n$, reales o complejas, con el campo K real o complejo.

$$\begin{aligned} |u\rangle &= M_{ab} \quad \wedge \quad |v\rangle = N_{ab}, \\ |u\rangle \boxplus |v\rangle &\equiv M_{ab} + N_{ab} = (M + N)_{ab}, \\ \alpha |u\rangle &= \alpha M_{ab} = (\alpha M)_{ab}. \end{aligned}$$

Es también obvio que se podrán formar subespacios vectoriales cuyos elementos sean matrices de dimensión menor a $n \times n$.

5. El conjunto de los vectores geométricos, o vectores cartesianos, en 2 y 3 dimensiones, con las propiedades habituales de suma vectorial y multiplicación por un escalar.
6. El conjunto de todos los polinomios con coeficientes reales: $\mathcal{P} = \{a_0, a_1x, a_2x^2, \dots, a_nx^n, \dots\}$, con \boxplus la suma ordinaria entre polinomios y la multiplicación ordinaria de polinomios con números.
7. Espacios Funcionales (de los cuales los polinomios son un caso particular). En estos espacios los vectores serán funciones, la suma será la suma ordinaria entre funciones y la multiplicación por un escalar también será la multiplicación ordinaria de una función por un elemento de un campo:

$$\begin{aligned} |f\rangle &= f(x) \quad \wedge \quad |g\rangle = g(x), \\ |f\rangle \boxplus |g\rangle &\equiv f(x) + g(x) \equiv (f + g)(x), \\ \alpha |f\rangle &= (\alpha f)(x) \equiv \alpha f(x). \end{aligned}$$

8. El conjunto de todas las funciones continuas e infinitamente diferenciables, definidas en el intervalo $[a, b] : C_{[a,b]}^\infty$.
9. El conjunto de todas las funciones complejas de variable real, $\psi(x)$, definidas en $[a, b]$, de cuadrado integrable (es decir para las cuales $\int_a^b dx |\psi(x)|^2$ sea finita). Este espacio se denomina comúnmente \mathcal{L}^2 y puede ser definido dentro de un rango $[a, b]$, finito o infinito, y para más de una variable.
10. El conjunto de todas las soluciones de un sistema de ecuaciones lineales y homogéneas, por ejemplo:

$$\begin{aligned} a + 3b + c &= 0 \\ 4a + 2b + 2c &= 0 \end{aligned}$$

Este sistema tiene como solución al conjunto: $\{a, b = a/2, c = -5a/2\}$. La suma de estos elementos y la multiplicación por un número conforman un espacio vectorial en el sentido de que son soluciones del sistema de ecuaciones.

2.1.4.1 Subespacios

Supongamos que tenemos un conjunto S de elementos de un espacio vectorial lineal V que satisface las siguientes propiedades:

1. El vector neutro de V está en S .
2. Si $|s_1\rangle, |s_2\rangle \in S$, entonces $|s_1\rangle \boxplus |s_2\rangle \in S$.
3. Si $|s\rangle \in S$ y α es un elemento del campo K , entonces $\alpha |s\rangle \in S$.

De esta manera, las operaciones que hacen de V un espacio vectorial también están definidas para S . Se puede demostrar que S es también un espacio vectorial lineal, es decir, satisface el conjunto de axiomas 2.1.3. Al conjunto $S \subset V$ que satisface las propiedades 1 y 2 más los axiomas 2.1.3 se le denomina un subespacio vectorial lineal (o simplemente subespacio) del espacio vectorial lineal V .

Notemos que el conjunto conformado con el vector neutro como único elemento: $S = \{|0\rangle\} \in V$ es el subespacio vectorial más pequeño de V . Por otro lado, el espacio completo V es el subespacio más grande posible de V . Se acostumbra llamar a estos dos subespacios los subespacios triviales de V .

Notemos que el espacio vectorial \mathbb{R}^2 no es un subespacio de \mathbb{R}^3 porque \mathbb{R}^2 ni siquiera es un subconjunto de \mathbb{R}^3 . De la misma manera, un plano en \mathbb{R}^3 que no contenga el origen no es un subespacio de \mathbb{R}^3 , porque el plano no contiene el vector cero de \mathbb{R}^3 .

2.1.5 La importancia de la notación

En los ejemplos antes mencionados hemos utilizado para representar un vector abstracto la notación de $|v\rangle$ y con éstos construimos un espacio vectorial abstracto $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$. Un espacio vectorial abstracto será un conjunto de elementos genéricos que satisfacen ciertos axiomas. Dependiendo del conjunto de axiomas tendremos distintos tipos de espacios abstractos, la teoría desarrolla las consecuencias lógicas que resultan de esos axiomas. En matemática el concepto de espacios abstractos es reciente (1928) y, aparentemente, se le debe a Maurice Fréchet⁴.

Los elementos de esos espacios se dejan sin especificar a propósito. Ese vector abstracto puede representar, vectores en \mathbb{R}^n , matrices $n \times n$ o funciones continuas. La notación $|v\rangle$, que se denomina un *ket* y al cual le corresponde un *bra* $\langle u|$ proviene del vocablo inglés *braket* que significa corchete y será evidente más adelante cuando construyamos escalares *braket* $\langle u|v\rangle$. Esta útil notación la ideó Paul Dirac⁵, uno de los físicos más influyentes en el desarrollo de la Física del siglo XX. En Mecánica Cuántica un estado cuántico particular suele representarse por una función de onda $\psi(x)$, que depende de la variable posición x o de una función alternativa que puede depender de la variable momentum p . En la notación de Dirac, el símbolo $|\psi\rangle$ denota el estado cuántico sin referirse a la función en particular y también sirve para distinguir a los vectores de los escalares (números complejos) que vienen a ser los elementos fundamentales en el espacio de Hilbert de la Mecánica Cuántica.

⁴Maurice Fréchet (1878-1973). Versátil matemático francés, con importantes contribuciones en espacios métricos, topología y creador del concepto de espacios abstractos. https://es.wikipedia.org/wiki/Maurice_Fréchet

⁵Paul Adrien Maurice Dirac (1902-1984). Además de contribuir de manera determinante en la comprensión de la Mecánica Cuántica, es uno de los creadores de la Mecánica Cuántica Relativista la cual ayudó a comprender el papel que juega el espín en las partículas subatómicas. https://es.wikipedia.org/wiki/Paul_Dirac

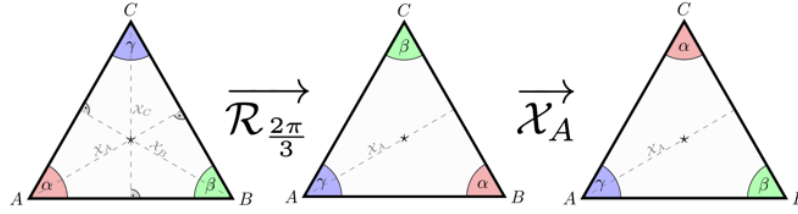


Figura 2.1: Transformaciones que dejan invariante un triángulo equilátero. Concatenación de una rotación, $\mathcal{R}_{\frac{2\pi}{3}}$ con una reflexión, \mathcal{X}_A , respecto a un plano que pasa por A

2.1.6 Ejercicios

1. Diga cuales de los siguientes grupos son abelianos: $\{\mathbb{Z}, +\}$; $\{\mathbb{Z}_n, + \forall, n \in \mathbb{N}\}$; $\{\mathbb{N}, +\}$ y el grupo $\{\mathbb{Z}, \cdot\}$.
2. Sea S el conjunto de todos los números reales excluyendo -1 y defina la operación \square tal que:

$$a \square b = a + b + ab.$$

Donde $+$ es la suma estándar entre números reales. Entonces:

- (a). Muestre si $[S, \square]$ forman grupo.
 - (b). Encuentre la solución en S para la ecuación $2 \square x \square 3 = 7$.
3. Considere un triángulo equilátero que se muestra en la figura 2.1. Se pueden identificar operaciones de rotación alrededor de un eje perpendicular a la figura que pasa por su baricentro \star y, reflexiones respecto a planos, \mathcal{X}_A , \mathcal{X}_B y \mathcal{X}_C , que dejan invariante la figura del triángulo. El lector puede consultar los ejemplos 2.1, 2.2 y 2.3 para fijar ideas.

Ahora bien, se puede definir la operación concatenación de rotaciones y reflexiones que dejan igualmente invariante al triángulo, tal y como mostramos en la mencionada figura 2.1. Note que lo ilustrado en la figura, puede esquematizarse como:

$$(A \alpha, B \beta, C \gamma) \xrightarrow{\mathcal{R}_{\frac{2\pi}{3}}} (A \gamma, B \alpha, C \beta) \xrightarrow{\mathcal{X}_A} (A \gamma, B \beta, C \alpha).$$

- (a). Construya la tabla de multiplicación para G_Δ , vale decir $G_\Delta = \{\mathcal{I}, \{\mathcal{R}_i\}, \{\bar{\mathcal{R}}_j\}, \{\mathcal{X}_k\}\}$ y la operación es concatenación tal y como mostramos en la figura 2.1. Donde \mathcal{I} es la operación identidad, $\{\mathcal{R}_i\}$ es un conjunto de rotaciones en sentido horario, mientras que $\{\bar{\mathcal{R}}_j\}$ es un conjunto de rotaciones en el sentido antihorario, y $\{\mathcal{X}_k\}$ el conjunto de las reflexiones que dejan invariante el triángulo.
- (b). Muestre que el conjunto de estas operaciones forman el grupo: G_Δ .
- (c). Identifique cada una de las \mathcal{R}_i y $\bar{\mathcal{R}}_j$, y muestre además, que forman un subgrupo cíclico de orden 3. De igual modo identifique las reflexiones y muestre que, cada una de las reflexiones y la identidad, $\{\mathcal{I}, \mathcal{X}_i\}$, forman también un subgrupo cíclico, pero de orden 2.
- (d). Considere las siguientes matrices:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$\mathbb{C} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathbb{E} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

Muestre que forman grupo bajo la multiplicación de matrices y que ese grupo es isomorfo a G_{Δ} .

- (e). Considere el conjunto de permutaciones de 3 objetos y la operación composición de permutaciones que discutimos como ejemplo en la sección 2.3 ¿Es ese grupo isomorfo a G_{Δ} ? Justifique su respuesta.
- (f). ¿Qué puede decir de las operaciones simetrías que dejan invariante un triángulo isósceles? ¿formarán grupo? ¿y si el triángulo es escaleno, cuales son las operaciones de simetría que lo dejan invariante?
4. Considere las siguientes funciones:

$$f_1(x) = x, \quad f_2(x) = \frac{1}{x}, \quad f_3(x) = \frac{1}{1-x}, \quad f_4(x) = \frac{x-1}{x}, \quad f_5(x) = 1-x, \quad f_6(x) = \frac{x}{x-1}.$$

Muestre que forman grupo bajo la operación: $f_i(x) \odot f_j(x) = f_i(f_j(x))$, y que ese grupo es isomorfo a G_{Δ} , del ejercicio anterior.

5. Definamos una operación binaria \blacksquare como:

$$x \blacksquare y = x + y + \alpha xy,$$

con $x, y, \alpha \in \mathbb{R}$ y además $\alpha \neq 0$.

- (a). Demuestre que \blacksquare es asociativa.
- (b). Muestre que \blacksquare genera un grupo en $\{\mathbb{R} - (\frac{-1}{\alpha})\}$. Es decir, $\forall x, y \in \mathbb{R} \wedge x \neq \frac{-1}{\alpha}, y \neq \frac{-1}{\alpha}$, entonces: $x \blacksquare y$ forma un grupo.
6. Muestre que el siguiente conjunto de transformaciones en el plano xy forman un grupo y construya su tabla de multiplicación.
- (a). $I = \{x \rightarrow x \wedge y \rightarrow y\}$.
- (b). $I = \{x \rightarrow -x \wedge y \rightarrow -y\}$.
- (c). $I_x = \{x \rightarrow -x \wedge y \rightarrow y\}$.
- (d). $I_y = \{x \rightarrow x \wedge y \rightarrow -y\}$.
7. Considere un conjunto S conformado únicamente por números reales positivos. Consideremos las siguientes reglas sobre S : Por “suma” de dos números entenderemos su producto en el sentido usual, y el “producto” de un elemento $r \in S$ y un número real λ entenderemos r elevado a la potencia de λ , en el sentido usual ¿ S es un espacio vectorial?
8. Considere el conjunto de vectores en el plano conformado por vectores localizados en el origen y cuyos puntos finales permanecen siempre en el primer cuadrante ¿Este conjunto es un espacio vectorial?
9. Muestre que también serán espacios vectoriales:
- (a). El conjunto de todas las funciones $f = f(x)$ definidas en $x = 1$ con $f(1) = 0$. Si $f(1) = c$ ¿Tendremos igual un espacio vectorial? ¿Por qué?
- (b). Los vectores $(x, y, z) \in V^3$ tal que sus componentes satisfacen el siguiente sistema de ecuaciones lineales:

$$a_{11}x + a_{12}y + a_{13}z = 0$$

$$a_{21}x + a_{22}y + a_{23}z = 0$$

$$a_{31}x + a_{32}y + a_{33}z = 0.$$

10. Sea \mathcal{P}_n el conjunto de todos los polinomios de grado n , en x , con coeficientes reales:

$$|p_n\rangle \Rightarrow p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} = \sum_{i=0}^{n-1} a_i x^i.$$

- (a). Demostrar que \mathcal{P}_n es un espacio vectorial respecto a la suma de polinomios y a la multiplicación de polinomios por un número (número real).
- (b). Si los coeficientes a_i son enteros ¿ \mathcal{P}_n será un espacio vectorial? ¿Por qué?
- (c). ¿Cuál de los siguientes subconjuntos de \mathcal{P}_n es un subespacio vectorial?
 - I. El polinomio cero y todos los polinomios de grado $n - 1$.
 - II. El polinomio cero y todos los polinomios de grado par.
 - III. Todos los polinomios que tienen a x como un factor (grado $n > 1$).
 - IV. Todos los polinomios que tienen a $x - 1$ como un factor.
- 11. Un subespacio \mathcal{P} es generado por: $|x_1\rangle = x^3 + 2x + 1$, $|x_2\rangle = x^2 - 2$, $|x_3\rangle = x^3 + x$ ¿Cuál(es) de los siguientes polinomios pertenece al subespacio \mathcal{P} ?
 - (a). $x^2 - 2x + 1$.
 - (b). $x^4 + 1$.
 - (c). $-\frac{1}{2}x^3 + \frac{5}{2}x^2 - x - 1$.
 - (d). $x - 5$.
- 12. Resuelva los ejercicios anteriores utilizando **Maxima**.

Practicando con Maxima

En este módulo utilizaremos algunas de las herramientas disponibles para incorporar el álgebra discreta. Comenzaremos con la introducción a los conjuntos. Podemos operar con conjuntos pero primero debemos definirlos. Existen varias posibilidades, como mostramos a continuación:

```
(%i1) A:{1,2,3,4,5,6,7,8,9};
(%o1) {1, 2, 3, 4, 5, 6, 7, 8, 9}

(%i2) B:set(1,3,5,7,9);
(%o2) {1, 3, 5, 7, 9}

(%i3) C:set(2,4,6,8);
(%o3) {2, 4, 6, 8}

(%i4) D:makeset(i/j, [i,j], [[1,3], [2,3], [3,3], [4,3]]);
(%o4) {1/3, 2/3, 1, 4/3}
```

Notemos que es igual definir los conjuntos con llaves, con la función **set** o **makeset**. Podemos preguntar si determinado elemento perteneces, o no, a un conjunto.

```
(%i5) elementp(7,A);
(%o5) true

(%i6) elementp(7,C);
(%o6) false
```

Operaciones elementales con conjuntos:

```
(%i7) UBC:union(B,C);
```

```
(%o7) {1, 2, 3, 4, 5, 6, 7, 8, 9}
```

```
(%i8) is(A=UBC);
(%o8) true
```

```
(%i9) Cv:intersection(B,C);
(%o9) {}
```

Para **Maxima** el conjunto vacío es {}.

```
(%i10) setdifference(A,C);
(%o10) {1, 3, 5, 7, 9}
```

Esto es, el conjunto con los elementos del conjunto A que no pertenecen al conjunto C .

Si queremos el conjunto de todos los subconjuntos del conjunto $\{a, b, c\}$ le podemos pedir al programa que nos lo muestre:

```
(%i11) powerset(a,b,c);
(%o11) {{}, {a}, {a, b}, {a, b, c}, {a, c}, {b}, {b, c}, {c}}
```

El producto cartesiano de los conjuntos A y B es el conjunto conformado por los pares (a, b) :

$$A \times B = \{(a, b) / a \in A, b \in B\}.$$

```
(%i12) cartesian_product(B,C);
(%o12) {[1, 2], [1, 4], [1, 6], [1, 8], [3, 2], [3, 4], [3, 6], [3, 8], [5, 2], [5, 4], [5, 6], [5, 8], [7, 2],
      [7, 4], [7, 6], [7, 8], [9, 2], [9, 4], [9, 6], [9, 8]}
```

Le podemos pedir al programa la suma de los pares del producto cartesiano anterior:

```
(%i13) makeset(a+b, [a,b], cartesian_product(B,C));
(%o13) {3, 5, 7, 9, 11, 13, 15, 17}
```

Maxima trata a los conjuntos y a las listas como objetos de distinta naturaleza, lo que permite trabajar con conjuntos cuyos elementos puedan ser también conjuntos o listas, es decir, subconjuntos.

```
(%i14) lista:makelist(i,i,1,25);
(%o14) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]
```

La lista anterior la convertiremos en un conjunto, para este fin debemos utilizar el comando **setify**.

```
(%i15) E:setify(lista);
(%o15) {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25}
```

De este último conjunto podemos construir el subconjunto conformado por los números primos:

```
(%i16) Primos:subset(E,primep);
(%o16) {2, 3, 5, 7, 11, 13, 17, 19, 23}
```

Con **cardinality** podemos saber cuantos elementos contiene un conjunto:

```
(%i17) cardinality(%);
```

```
(%o17) 9
```

La función de **Maxima** que nos permite calcular las tablas que utilizamos en la sección 2.1 es la función **mod**. Veamos el siguiente ejemplo:

```
(%i18) Gm8:1,3,5,7;
(%o18) {1, 3, 5, 7}
```

Todos los productos posibles entre los elementos de este conjunto son:

```
(%i19) makeset(a,b,[a,b], cartesian_product(Gm8,Gm8));
(%o19) {1, 3, 5, 7, 9, 15, 21, 25, 35, 49}
```

Los modulo 8 para algunos de los números anteriores son:

```
(%i20) mod(21,8); mod(35,8); mod(49,8);
(%o20) 5
(%o21) 3
(%o22) 1
```

Para generar el grupo $G_{\text{mod}5}$ escribimos:

```
(%i23) setify(makelist(mod(i,5),i,1,4));
(%o23) {1, 2, 3, 4}
```

En la sección 2.1.1.2, definimos el orden de un elemento $g \in G$. Consideremos el siguiente conjunto de números enteros:

$$\mathbb{Z}_{15} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}.$$

Se podría demostrar que el orden de $g \in G$ es igual al número de elementos de $\langle g \rangle$ y lo dejaremos como ejercicio.

Vamos a calcular el orden de todos los elementos de \mathbb{Z}_{15} , sabiendo que el orden de cada uno de esos elementos divide a 15, que es el cardinal de \mathbb{Z}_{15} .

Probemos primero con el número $6 \in \mathbb{Z}_{15}$.

```
(%i24) setify(makelist(mod(6*i,15),i,0,14));
(%o24) {0, 3, 6, 9, 12}
```

El orden de $6 \in \mathbb{Z}_{15}$ es:

```
(%i25) length(%);
(%o25) 5
```

En la siguiente instrucción lo haremos para todos los elementos de \mathbb{Z}_{15} .

```
(%i26) makelist([j,length(setify(makelist(mod(j*i,15),i,0,14)))],j,0,14);
(%o26) [[0, 1], [1, 15], [2, 15], [3, 5], [4, 15], [5, 3], [6, 5], [7, 15], [8, 15], [9, 5], [10, 3], [11, 15], [12, 5],
[13, 15], [14, 15]]
```

La salida no es más que una lista $[x, y]$ con cada elemento $x \in \mathbb{Z}_{15}$ y su orden y . Por lo tanto, el conjunto de órdenes es: $\{1, 3, 5, 15\}$, todos divisores de 15, como estipula el teorema de Lagrange.

2.2 Espacios métricos, normados y con producto interno

En esta sección vamos a introducir una función de distancia, de manera que si tenemos un par de puntos o elementos de un espacio vectorial podemos hablar de que existe una cierta distancia entre ellos. Se dice que la función distancia induce una topología sobre el espacio vectorial. Comenzaremos definiendo el concepto de métrica sobre espacios vectoriales y con esta estructura llegar a la noción de norma.

2.2.1 Métricas y espacios métricos

El concepto de métrica surge de la generalización de la idea de distancia entre dos puntos de la recta real. Un espacio vectorial será métrico si podemos definir una función:

$$d : V \times V \rightarrow \mathbb{R} / \forall |x\rangle, |y\rangle, |z\rangle \in V,$$

tal que se cumple lo siguiente:

1. $d(|x\rangle, |y\rangle) \geq 0$, si $d(|x\rangle, |y\rangle) = 0 \Rightarrow |x\rangle \equiv |y\rangle$.
2. $d(|x\rangle, |y\rangle) \equiv d(|y\rangle, |x\rangle)$.
3. $d(|x\rangle, |y\rangle) \leq d(|x\rangle, |z\rangle) + d(|y\rangle, |z\rangle)$ (La desigualdad triangular).

Así, diremos que (V, K, \oplus, d) es un espacio vectorial lineal, métrico.

Podemos enumerar algunos ejemplos de espacios métricos:

1. Espacios reales \mathbb{R}^n .

(a). Para \mathbb{R} , es decir la recta real, la definición de métrica es: $d(|x\rangle, |y\rangle) \equiv |x - y|$.

(b). Para \mathbb{R}^2 , es decir el plano, una definición de métrica es: $d(|x\rangle, |y\rangle) \equiv \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. También podemos construir otra definición de métrica como: $d(|x\rangle, |y\rangle) \equiv |x_1 - y_1| + |x_2 - y_2|$. La primera de estas métricas se conoce como métrica euclídea y la segunda como métrica Manhattan o métrica de taxistas. Es claro como el mismo espacio vectorial genera varios espacios métricos, dependiendo de la definición de métrica. Para estos casos particulares, las métricas “miden” el desplazamiento entre dos puntos de forma distinta: en drones (métrica euclídea) o vehículos terrestre en ciudades.

(c). En general para espacios reales \mathbb{R}^n una posible definición de métrica será:

$$d(|x\rangle, |y\rangle) \equiv \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots + (x_n - y_n)^2}.$$

Esta definición de métrica no es más que una generalización del teorema de Pitágoras y se denomina “distancia euclidiana”.

2. Espacios unitarios n -dimensionales, o espacios complejos, \mathbb{C}^n . La definición de distancia puede construirse como:

$$d(|x\rangle, |y\rangle) \equiv \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \cdots + |x_n - y_n|^2},$$

y es claro que se recupera la idea de distancia en el plano complejo: $d(|x\rangle, |y\rangle) \equiv |x - y|$.

3. Para los espacios de funciones $C_{[a,b]}^\infty$ una posible definición de distancia sería:

$$d(|f\rangle, |g\rangle) \equiv \max_{t \in [a,b]} |f(t) - g(t)|.$$

4. La métrica trivial o discreta

$$d(|x\rangle, |y\rangle) = \begin{cases} 1 & \text{si } x \neq y \\ 0 & \text{si } x = y \end{cases}$$

Ejemplo 2.6 Como vimos en la sección anterior, en el campo de los números complejos el valor absoluto de

$z = x + iy$ es $|z| = \sqrt{x^2 + y^2}$. La métrica que podemos asociar a este espacio vectorial viene dada por:

$$d(|z_1\rangle, |z_2\rangle) \equiv ||z_1\rangle - |z_2\rangle|| = \sqrt{\langle z_1 - z_2 | z_1 - z_2 \rangle} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

con: $|z_1\rangle = x_1 + iy_1$ y $|z_2\rangle = x_2 + iy_2$.

◀

2.2.2 Normas y espacios normados

El concepto de distancia (o métrica) es la propiedad más elemental que se le puede exigir a un espacio vectorial. Mucho más interesante son aquellos espacios vectoriales que están equipados con la idea de norma y, a partir de allí, se obtiene la distancia. La norma tiene que ver con el “tamaño” del vector y la métrica con la distancia entre vectores.

La norma, $\mathcal{N}(|v_i\rangle) \equiv ||v_i\rangle||$, de un espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ será una función:

$$\mathcal{N} : V \rightarrow \mathbb{R} / \forall |v_i\rangle \in V,$$

que cumple con:

1. $||v_i\rangle|| \geq 0$, si $||v_i\rangle|| = 0 \Rightarrow |v_i\rangle \equiv |0\rangle$.
2. $||\alpha |v_i\rangle|| = |\alpha| ||v_i\rangle||$.
3. $||v_i\rangle + |v_j\rangle|| \leq ||v_i\rangle|| + ||v_j\rangle||$ (Desigualdad triangular).

La definición de norma induce una métrica de la forma:

$$d(|v_i\rangle, |v_j\rangle) \equiv ||v_i\rangle - |v_j\rangle||. \quad (2.1)$$

Denotaremos un espacio vectorial normado⁶ como $(V, K, \boxplus; ||\cdot||)$, espacio que también es conocido como un espacio de Banach⁷. Esto significa que todo espacio normado es a su vez un espacio métrico, pero no todo espacio métrico es normado.

De la definición de distancia (2.1) resulta que la métrica así definida es invariante bajo traslaciones de vectores. Esto es, si: $|\tilde{x}\rangle = |x\rangle + |a\rangle \wedge |\tilde{y}\rangle = |y\rangle + |a\rangle$, entonces, $d(|x\rangle, |y\rangle) \equiv d(|\tilde{x}\rangle, |\tilde{y}\rangle)$. Y además es homogénea: $d(\lambda |x\rangle, \lambda |y\rangle) = |\lambda| d(|x\rangle, |y\rangle)$.

Como ejemplos de espacios normados podemos mostrar los siguientes:

1. Los espacios reales, \mathbb{R}^n y los espacios complejos \mathbb{C}^n . Para estos espacios de Banach, la norma se define como:

$$||x\rangle|| = \sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2 + \dots + |x_n|^2} = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

Para un espacio en \mathbb{R}^3 se cumple que $||x\rangle|| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, por lo tanto, la idea de norma generaliza la noción de “tamaño” del vector $|x\rangle$. Es claro que la definición de distancia se construye a partir de la norma de la forma:

$$d(|x\rangle, |y\rangle) \equiv ||x\rangle - |y\rangle|| = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \dots + |x_n - y_n|^2}.$$

Este espacio se llama espacio normado euclidiano de dimensión n .

2. Para el espacio lineal de matrices $n \times n$, reales o complejas, con el campo K real o complejo, una definición

⁶El concepto de espacio normado fue formulado en 1922 de manera independiente por: S. Banach, H. Hahn y N. Wiener.

⁷Stefan Banach (1892-1945) Matemático polaco, uno de los fundadores del análisis funcional moderno, con sus mayores contribuciones a la teoría de espacios topológicos. https://es.wikipedia.org/wiki/Stefan_Banach

de norma es:

$$\|M\| = \sum_{a=1}^m \sum_{b=1}^n |M_{ab}| ,$$

y la correspondiente definición de distancia:

$$d(|x\rangle, |y\rangle) \equiv \|M - N\| = \sum_{a=1}^m \sum_{b=1}^n |M_{ab} - N_{ab}| .$$

3. Para los espacios funcionales $\mathcal{C}_{[a,b]}^\infty$ una posible definición de norma sería:

$$\|f\| = \max_{t \in [a,b]} |f(t)| ,$$

otra posible definición puede ser:

$$\|f\| = \sqrt{\int_a^b |f(t)|^2 dt} .$$

2.2.3 Espacios euclidianos

El siguiente paso en la construcción de espacios vectoriales más ricos es equiparlo con la definición de producto interno, y a partir de esta definición, construir el concepto de norma y con éste el de distancia. La idea de producto interno generaliza el concepto de producto escalar de vectores en \mathbb{R}^3 e incorpora a los espacios vectoriales abstractos el concepto de ortogonalidad y descomposición ortogonal. Históricamente, la teoría de espacios vectoriales con producto interno es anterior a la teoría de espacios métricos y espacios de Banach y se le debe a D. Hilbert⁸. Adicionalmente, la semejanza entre la geometría euclidiana y la geométrica de \mathbb{R}^n ha hecho que espacios en los cuales se puedan definir, distancia, ángulos, a partir de una definición de producto interno, se denominen también espacios euclidianos.

2.2.3.1 Producto interno

En un espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$, la definición del producto interno de dos vectores la denotaremos como $\langle v_i | v_j \rangle$ y es una aplicación:

$$\mathcal{I}(|v_i\rangle, |v_j\rangle) : V \times V \rightarrow K, \forall |v_i\rangle, |v_j\rangle \in V .$$

Es decir, asocia a ese par de vectores con un elemento del campo K .

Las propiedades que definen el producto interno son:

1. $\langle v_i | v_i \rangle \equiv \| |v_i\rangle \|^2 \in K \wedge \langle v_i | v_i \rangle \geq 0 \quad \forall |v_i\rangle \in V$, si $\langle v_i | v_i \rangle = 0 \Rightarrow |v_i\rangle \equiv |0\rangle$.
2. $\langle v_i | v_j \rangle = \langle v_j | v_i \rangle^* \quad \forall |v_i\rangle, |v_j\rangle \in V$.
3. $\langle v_i | \alpha v_j + \beta v_k \rangle = \alpha \langle v_i | v_j \rangle + \beta \langle v_i | v_k \rangle \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in V \wedge \alpha, \beta \in K$.
4. $\langle \alpha v_i + \beta v_j | v_k \rangle = \alpha^* \langle v_i | v_k \rangle + \beta^* \langle v_j | v_k \rangle \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in V \wedge \alpha, \beta \in K$.
5. $\langle v_i | 0 \rangle = \langle 0 | v_i \rangle = 0$.

La segunda y cuarta propiedad resultan del hecho de que si el campo es complejo $K = \mathbb{C}$, entonces:

$$\langle \alpha v_i | \alpha v_i \rangle = \alpha^2 \langle v_i | v_i \rangle = i^2 \langle v_i | v_i \rangle = - \langle v_i | v_i \rangle ,$$

lo cual contradice el hecho de que la norma tiene que ser positiva. Por eso la necesidad de tomar el complejo conjugado. Se dice entonces, que el producto escalar es *antilineal* respecto al primer factor y *lineal* respecto al

⁸David Hilbert(1862-1943) Matemático alemán defensor de la axiomática como enfoque primordial de los problemas científicos. Hizo importantes contribuciones en distintas áreas de la matemática, como: Invariantes, Campos de Números Algebraicos, Análisis Funcional, Ecuaciones Integrales, Física-Matemática y Cálculo en Variaciones. https://es.wikipedia.org/wiki/David_Hilbert

segundo.

A partir de la definición de producto interno se construyen los conceptos de norma y distancia:

$$\| |v_i\rangle \| = \sqrt{\langle v_i | v_i \rangle} \quad \text{y} \quad d(|v_i\rangle, |v_j\rangle) \equiv \| |v_i\rangle - |v_j\rangle \| = \sqrt{\langle v_i - v_j | v_i - v_j \rangle}.$$

2.2.3.2 La desigualdad de Cauchy-Schwarz: los ángulos entre vectores reales y complejos

Todo producto interno $\langle v_i | v_j \rangle$ definido en un espacio vectorial normado $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ cumple con la desigualdad de Cauchy-Schwarz:

$$|\langle v_i | v_j \rangle|^2 \leq \langle v_i | v_i \rangle \langle v_j | v_j \rangle \iff |\langle v_i | v_j \rangle| \leq \| |v_i\rangle \| \| |v_j\rangle \| . \quad (2.2)$$

Es claro que si $|v_i\rangle = |0\rangle \wedge |v_j\rangle = |0\rangle$ se cumple la igualdad y es trivial la afirmación.

Para demostrar la desigualdad, tomemos dos vectores $|v_i\rangle \wedge |v_j\rangle$ cualesquiera, entonces podemos construir un tercero: $|v_k\rangle = \alpha |v_i\rangle + \beta |v_j\rangle$ (α y β tendrán valores particulares), por lo tanto:

$$\langle v_k | v_k \rangle \equiv \langle \alpha v_i + \beta v_j | \alpha v_i + \beta v_j \rangle \geq 0,$$

esto significa que:

$$\begin{aligned} \langle \alpha v_i + \beta v_j | \alpha v_i + \beta v_j \rangle &= \langle \alpha v_i | \alpha v_i \rangle + \langle \alpha v_i | \beta v_j \rangle + \langle \beta v_j | \alpha v_i \rangle + \langle \beta v_j | \beta v_j \rangle \geq 0 \\ &= |\alpha|^2 \langle v_i | v_i \rangle + \alpha^* \beta \langle v_i | v_j \rangle + \beta^* \alpha \langle v_j | v_i \rangle + |\beta|^2 \langle v_j | v_j \rangle \geq 0. \end{aligned}$$

Si $\alpha = \langle v_j | v_j \rangle$, se tiene que:

$$\begin{aligned} \langle v_j | v_j \rangle \langle v_i | v_i \rangle + \beta \langle v_i | v_j \rangle + \beta^* \langle v_j | v_i \rangle + |\beta|^2 &\geq 0 \\ \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq -\beta \langle v_i | v_j \rangle - \beta^* \langle v_j | v_i \rangle - |\beta|^2, \end{aligned}$$

seguidamente seleccionamos: $\beta = -\langle v_j | v_i \rangle$, y por lo tanto: $\beta^* = -\langle v_i | v_j \rangle$, consecuentemente:

$$\begin{aligned} \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq \langle v_j | v_i \rangle \langle v_i | v_j \rangle + \langle v_i | v_j \rangle \langle v_j | v_i \rangle - \langle v_j | v_i \rangle \langle v_i | v_j \rangle \\ \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq \langle v_i | v_j \rangle \langle v_j | v_i \rangle = |\langle v_i | v_j \rangle|^2. \quad \blacksquare \end{aligned}$$

Ahora consideremos lo siguiente, de la desigualdad de Cauchy-Schwarz y la definición de norma se desprende que:

$$\frac{|\langle v_i | v_j \rangle|^2}{\| |v_i\rangle \|^2 \| |v_j\rangle \|^2} \leq 1 \Rightarrow -1 \leq \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|} \leq 1,$$

por lo tanto podemos definir el “ángulo” entre los vectores abstractos $|v_i\rangle \wedge |v_j\rangle$ como:

$$\cos(\Theta_G) = \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|}, \quad (2.3)$$

donde hemos denotado como Θ_G el ángulo genérico que forman los vectores reales o complejos.

Si estamos considerando espacios vectoriales reales -en los cuales el campo corresponde a los números reales- entonces el ángulo definido entre vectores abstractos reales corresponde al que intuitivamente siempre hemos considerado para los vectores cartesianos y que discutimos en la sección 1.1.4.1,

$$\cos(\Theta_R) = \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|}, \quad \text{con} \quad 0 \leq \Theta_R \leq \pi,$$

donde se toma $\Theta_R = 0$ para vectores colineales y $\Theta_R = \pi$ para vectores opuestos (antilineales). Si bien es cierto que esta definición coincide con la de los vectores cartesianos, hay que resaltar que la estamos extendiendo para cualquier vector abstracto, vale decir: funciones reales, matrices, y todos los objetos matemáticos que cumplan con las reglas para los espacios vectoriales expuestas en 2.1.3.

Para el caso de espacios vectoriales complejos la situación es más sutil y significa definir un ángulo entre dos vectores abstractos y complejos, sin embargo podemos abordar el problema suponiendo:

1. un espacio complejo n -dimensional de n -uplas de números complejos $|z\rangle \leftrightarrow (z_1, z_2, \dots, z_n)$ asociado (isomorfo) a un espacio vectorial real de $2n$ dimensiones, con vectores representados por $2n$ -uplas de números reales $|w\rangle \leftrightarrow (\text{Re}(z_1), \text{Re}(z_2), \dots, \text{Re}(z_n), \text{Im}(z_1), \text{Im}(z_2), \dots, \text{Im}(z_n))$, donde hemos representado $\text{Re}(z_j)$ y $\text{Im}(z_j)$ como las partes reales e imaginarias de z_j , respectivamente o,
2. directamente a partir de una definición de producto interno entre vectores complejos implementado por:

$$\langle w_i | v_j \rangle = \sum_{j=1}^n w_j^* v_j.$$

Ambas aproximaciones no son del todo independientes pero igualmente justificadas⁹.

Consideremos el segundo caso, esto es: directamente a partir de una definición de producto interno entre vectores complejos. Para este caso consideramos un ángulo complejo, y $\cos(\Theta_C)$ una función de variable compleja, que puede ser expresada en su forma polar como:

$$\cos(\Theta_C) = \frac{|\langle v_i | v_j \rangle|}{\|v_i\| \|v_j\|} \Rightarrow \cos(\Theta_C) = \rho e^{i\phi}, \quad \text{con } \rho = |\cos(\Theta_C)| < 1.$$

Entonces podemos asociar $\rho = \cos(\Theta_H)$ y definir Θ_H , en el rango $0 \leq \Theta_H \leq \pi/2$, como el *ángulo hermítico* entre los vectores complejos $|v_i\rangle$ y $|v_j\rangle$. Mientras que ϕ , definido en $-\pi \leq \phi \leq \pi$, corresponde al pseudo ángulo de Kasner, que representa la orientación o rotación del ángulo hermítico y no tiene mayor significado al cuantificar el ángulo entre esos dos vectores. Esta diferencia de significados puede intuirse cuando multiplicamos $|v_i\rangle$ y $|v_j\rangle$, por una constante compleja: $|\tilde{v}_i\rangle \rightarrow \alpha_i |v_i\rangle$ y comprobamos que el ángulo Θ_H permanece inalterado y no así el ángulo de Kasner¹⁰.

2.2.3.3 Teoremas del coseno y de Pitágoras

A partir de la definición de norma se obtiene:

$$\begin{aligned} \| |v_i\rangle - |v_j\rangle \|^2 &= \langle v_i - v_j | v_i - v_j \rangle = \langle v_i | v_i \rangle - \langle v_i | v_j \rangle - \langle v_i | v_j \rangle^* + \langle v_j | v_j \rangle \\ &= \langle v_i | v_i \rangle + \langle v_j | v_j \rangle - 2 \text{Re}(\langle v_i | v_j \rangle), \end{aligned}$$

con lo cual hemos generalizado el teorema del coseno para un espacio vectorial abstracto:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta_G). \quad (2.4)$$

De tal forma que para espacios vectoriales reales tendremos:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta), \quad \text{con } 0 \leq \Theta \leq \pi,$$

y para espacios vectoriales complejos:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta_H) \cos(\phi), \quad \text{con } 0 \leq \Theta_H \leq \pi/2.$$

Para el caso que los vectores $|v_i\rangle \wedge |v_j\rangle$ sean ortogonales, esto es $\langle v_i | v_j \rangle = 0$, tendremos el teorema de Pitágoras generalizado:

$$\| |v_i\rangle - |v_j\rangle \|^2 \equiv \| |v_i\rangle + |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2. \quad (2.5)$$

Veamos algunos ejemplos de espacios vectoriales con producto interno.

1. Espacios euclidianos reales, \mathbb{R}^n y espacios euclidianos complejos \mathbb{C}^n .

Los vectores en estos espacios euclidianos pueden ser representados por

$$|x\rangle = (x_1, x_2, \dots, x_n) \wedge |y\rangle = (y_1, y_2, \dots, y_n),$$

⁹Scharnhorst, K. (2001). Angles in complex vector spaces. *Acta Applicandae Mathematica*, **69** (1), 95-103.

¹⁰Puede consultarse Reju, V. G., Koh, S. N., y Soon, Y. (2009). An algorithm for mixing matrix estimation in instantaneous blind source separation. *Signal Processing*, **89** (9), 1762-1773.

y el **producto interno** queda definido por:

$$\langle x | y \rangle = x_1^* y_1 + x_2^* y_2 + x_3^* y_3 + \cdots + x_n^* y_n = \sum_{i=1}^n x_i^* y_i, \quad (2.6)$$

es claro que esta definición de producto interno coincide, para \mathbb{R}^2 (y \mathbb{R}^3) con la idea de producto escalar convencional que consideramos en las secciones 1.1.4.1 y 1.2.6.1, vale decir:

$$\left. \begin{aligned} \mathbf{a} &= a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} \\ \mathbf{b} &= b_x \hat{\mathbf{i}} + b_y \hat{\mathbf{j}} \end{aligned} \right\} \Rightarrow \mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y.$$

Ahora bien, el lector puede comprobar que para vectores en \mathbb{R}^2 también se puede proveer una definición de producto interno diferente:

$$\mathbf{a} \otimes \mathbf{b} = 2a_x b_x + a_x b_y + a_y b_x + a_y b_y,$$

igualmente válida, con lo cual es claro que en un mismo espacio vectorial pueden coexistir diferentes productos internos.

Por su parte, la **norma** es:

$$\| |x\rangle \| = \sqrt{\langle x | x \rangle} = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}.$$

La **distancia** también recupera la idea intuitiva de distancia euclidiana:

$$\begin{aligned} d(|x\rangle, |y\rangle) &\equiv \| |x\rangle - |y\rangle \| = \sqrt{\langle x - y | x - y \rangle} \\ d(|x\rangle, |y\rangle) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots + (x_n - y_n)^2}. \end{aligned}$$

El teorema del coseno queda como:

$$\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 + 2 \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2} \cos(\Theta),$$

mientras que el teorema de Pitágoras es:

$$\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2,$$

es obvio que para \mathbb{R}^2 tanto el teorema del coseno como el teorema de Pitágoras retoman su forma tradicional.

Finalmente la desigualdad de Cauchy-Schwarz se expresa de la siguiente manera:

$$|\langle x | y \rangle| \leq \| |x\rangle \| \| |y\rangle \| \Rightarrow \left| \sum_{i=1}^n x_i y_i \right|^2 \leq \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2.$$

2. Para los espacios de funciones continuas $C_{[a,b]}^\infty$ una posible definición de **producto interno** sería:

$$\langle f | g \rangle = \int_a^b dx f^*(x) g(x),$$

de la cual se deriva la siguiente expresión para la **norma**:

$$\| |f\rangle \|^2 = \langle f | f \rangle = \int_a^b dx |f(x)|^2.$$

La **distancia** entre funciones quedará definida como:

$$\begin{aligned} d(|f\rangle, |g\rangle) &\equiv \| |f\rangle - |g\rangle \| \equiv \sqrt{\langle f - g | f - g \rangle} = \sqrt{\langle f | f \rangle - \langle f | g \rangle - \langle f | g \rangle^* + \langle g | g \rangle} \\ d(|f\rangle, |g\rangle) &= \sqrt{\int_a^b dx |f(x) - g(x)|^2} \\ &= \sqrt{\int_a^b dx |f(x)|^2 - 2 \operatorname{Re} \left(\int_a^b dx f^*(x) g(x) \right) + \int_a^b dx |g(x)|^2}. \end{aligned}$$

El teorema del coseno puede ser escritos como:

$$\begin{aligned} \int_a^b dx |f(x) + g(x)|^2 &= \int_a^b dx |f(x)|^2 + \int_a^b dx |g(x)|^2 \\ &\quad + 2 \left(\int_a^b dx |f(x)|^2 \right)^{\frac{1}{2}} \left(\int_a^b dx |g(x)|^2 \right)^{\frac{1}{2}} \cos(\Theta), \end{aligned}$$

donde:

$$\cos(\Theta) = \frac{\int_a^b dx f^*(x) g(x)}{\left(\int_a^b dx |f(x)|^2 \right)^{\frac{1}{2}} \left(\int_a^b dx |g(x)|^2 \right)^{\frac{1}{2}}}.$$

Y como era de esperarse el teorema de Pitágoras queda:

$$\int_a^b dx |f(x) + g(x)|^2 = \int_a^b dx |f(x)|^2 + \int_a^b dx |g(x)|^2,$$

para funciones $f(x)$ y $g(x)$ ortogonales, mientras que para este caso, la desigualdad de Cauchy-Schwarz se expresa:

$$\left| \int_a^b dx f^*(x) g(x) \right|^2 \leq \int_a^b dx |f(x)|^2 \int_a^b dx |g(x)|^2.$$

Ejemplo 2.7 Consideramos el espacio vectorial de polinomios de grado $g \leq n$ definidos en el intervalo $[0, 1]$ o en el intervalo $[-1, 1]$ según el caso. Suponiendo las siguientes definiciones de producto interno en \mathcal{P}^n :

$$\langle q_n | p_n \rangle = \int_{-1}^1 p(x) q(x) dx \quad \text{y} \quad \langle q_n | p_n \rangle = \int_0^1 p(x) q(x) dx.$$

Vamos a encontrar la distancia y el ángulo entre los vectores: $|x_1\rangle = x(x-1)$ y $|x_2\rangle = x$.

En general, la definición de distancia es:

$$d(|x_1\rangle, |x_2\rangle) = \sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle},$$

por lo tanto para $\langle q_n | p_n \rangle = \int_{-1}^1 p(x) q(x) dx$ la distancia será:

$$\sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle} = \sqrt{\int_{-1}^1 [x(x-1) - x]^2 dx} = \frac{1}{15} \sqrt{690},$$

y para $\langle q_n | p_n \rangle = \int_0^1 p(x) q(x) dx$, será:

$$\sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle} = \sqrt{\int_0^1 (x(x-1) - x)^2 dx} = \frac{2}{15} \sqrt{30}.$$

Con respecto a los ángulos:

$$\theta = \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right).$$

Para $\langle q_n | p_n \rangle = \int_{-1}^1 p(x) q(x) dx$ tenemos:

$$\begin{aligned}\theta &= \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right) = \arccos \left(\frac{\int_{-1}^1 (x(x-1)) x \, dx}{\sqrt{\int_{-1}^1 (x(x-1))^2 \, dx} \sqrt{\int_{-1}^1 x^2 \, dx}} \right) \\ &= \arccos \left(-\frac{1}{12} \sqrt{15} \sqrt{6} \right) = 2,4825 \text{ rad}.\end{aligned}$$

Para $\langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx$

$$\begin{aligned}\theta &= \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right) = \arccos \left(\frac{\int_0^1 (x(x-1)) (x) \, dx}{\sqrt{\int_0^1 (x(x-1))^2 \, dx} \sqrt{\int_0^1 x^2 \, dx}} \right) \\ &= \arccos \left(-\frac{1}{12} \sqrt{15} \sqrt{2} \right) = 2,4825 \text{ rad} \quad \text{¡El mismo ángulo!}\end{aligned}$$

◀

2.2.4 Ejercicios

- Consideremos el espacio vectorial conformado por los vectores geométricos en \mathbb{R}^3 . ¿Serán espacios euclidianos para las siguientes definiciones de producto interno?
 - El producto de las longitudes de los vectores.
 - El producto de las longitudes por el cubo del coseno del ángulo entre ellos.
 - El producto como dos veces el producto escalar usual entre vectores.
- Considerando estas definiciones de producto interno en \mathcal{P}_n :

$$a) \quad \langle q_n | p_n \rangle = \int_{-1}^1 p(x)q(x)dx, \quad b) \quad \langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx.$$

- Encuentre los ángulos en el “triángulo” formado por los vectores: $|x_1\rangle = 1$, $|x_2\rangle = t$, $|x_3\rangle = 1 - t$.
- Encuentre la distancia y el ángulo entre los siguientes pares de vectores en \mathcal{P}_3 :
 - $|x_1\rangle = 1$; $|x_2\rangle = x$.
 - $|x_1\rangle = 2x$; $|x_2\rangle = x^2$.

- Sea E' un subespacio euclidiano de dimensión k , $E' \subset E$, y sea $|v\rangle$ un vector que no necesariamente es un elemento E' . Podemos plantearnos el problema de representar $|v\rangle$ de la forma: $|v\rangle = |g\rangle + |h\rangle$; donde $|g\rangle \in E'$ y $|h\rangle$ es ortogonal a $|g\rangle$. La existencia de la expansión anterior nos muestra que el espacio total E , de dimensión n , es la suma directa de los subespacios E' y su complemento ortogonal E^\perp de dimensión $n - k$.

Nota: El conjunto de los vectores de E que son ortogonales a todos los de E' :

$$E^\perp = \{|v\rangle \in E \mid \langle v | g \rangle = 0 \, \forall \, |g\rangle \in E'\}$$

será un subespacio de E y se denomina complemento ortogonal de E' .

Encuentre el vector $|v\rangle$, como la suma del vector $|g\rangle$, expandido por los vectores $|g_i\rangle$, y el vector perpendicular $|h\rangle$ cuando:

- $|h\rangle = (5, 2, -2, 2)$, $|g_1\rangle = (2, 1, 1, -\alpha)$, $|g_2\rangle = (1, \beta, 3, 0)$.
- $|h\rangle = (-3, 5, 9, 3)$, $|g_1\rangle = (1, 1, 1, \gamma)$, $|g_2\rangle = (2\eta, -1, 1, 1)$, $|g_3\rangle = (2, -7\delta, -1, -1)$.

4. Considere el espacio vectorial de matrices complejas 2×2 con la siguiente definición

$$\langle a | b \rangle \equiv \text{Tr}(\mathbb{A}^\dagger \mathbb{B}) \equiv (A^\dagger)_j^i B_i^j \equiv (A^*)^j_i B_i^j.$$

Compruebe si ésta es una buena definición de producto interno.

Nota: Las matrices complejas 2×2 están conformadas por cuatro números complejos, $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, $z_3 = x_3 + iy_3$, y $z_4 = x_4 + iy_4$. Entonces, hemos denotado por $(A^\dagger)_j^i$ la matriz adjunta de A_j^i . Vale decir, una matriz adjunta de otra será su traspuesta conjugada. Por traspuesta entendemos cambiar filas por columnas manteniendo intacta la diagonal, $A_j^i \rightarrow A_i^j$ (ver sección 4.4.3). Por otro lado, hemos definido como $\text{Tr}(\mathbb{A})$ la traza de una matriz como la suma de los elementos de la diagonal, $\text{Tr}(\mathbb{A}) = A_i^i$. A continuación ejemplificamos estas dos definiciones:

$$\mathbb{A} \leftrightarrow \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \Rightarrow \mathbb{A}^\dagger \leftrightarrow \begin{pmatrix} z_1^* & z_3^* \\ z_2^* & z_4^* \end{pmatrix} \quad \text{y} \quad \text{Tr}(\mathbb{A}) = z_1 + z_4.$$

Ambas definiciones serán discutidas con rigor en las secciones 4.2.3 y 4.3.4, respectivamente. Sin embargo, este par de definiciones funcionales –la adjunta de una matriz es su traspuesta conjugada y su traza será la suma de los elementos de la diagonal– serán suficiente para desarrollar este ejercicio.

5. Sean $|p_n\rangle = p(x) = \sum_{i=0}^{n-1} a_i x^i$; $|q_n\rangle = q(x) = \sum_{i=0}^{n-1} b_i x^i \in \mathcal{P}_n$ y considere la siguiente definición:

$$\langle q_n | p_n \rangle \equiv a_0 b_0 + a_1 b_1 + a_2 b_2 + \dots + a_{n-1} b_{n-1} = \sum_{i=0}^{n-1} a_i b_i$$

- (a). Muestre que ésta es una buena definición de producto interno.
 (b). Con esta definición de producto interior ¿Se puede considerar \mathcal{P}_n un subespacio de $\mathcal{C}_{[a,b]}$? ¿Por qué?
 6. Los vectores en \mathbb{R}^3 en coordenadas cartesianas los definimos como $\mathbf{a} = a^i |e_i\rangle = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}}$ y definimos una “tabla de multiplicación” entre ellos de la forma $\langle e^i | e_j \rangle = \delta_j^i$ con $i, j = 1, 2, 3$, esto es:

$\langle e^i e_j \rangle$	$\hat{\mathbf{i}}$	$\hat{\mathbf{j}}$	$\hat{\mathbf{k}}$
$\hat{\mathbf{i}}$	1	0	0
$\hat{\mathbf{j}}$	0	1	0
$\hat{\mathbf{k}}$	0	0	1

Un cuaternión cartesiano puede escribirse de manera análoga a los vectores cartesianos, vale decir:

$$|a\rangle = a^\alpha |q_\alpha\rangle = a^0 + a^i |q_i\rangle = a_0 + a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}},$$

con $\alpha = 0, 1, 2, 3$ y donde las a^i (con $i = 1, 2, 3$) son números reales que representan las componentes vectoriales en coordenadas cartesianas de los cuaterniones, mientras que la a^0 , también un número real se le llama componente escalar¹¹.

Los cuaterniones fueron inventados por el matemático irlandés William Rowan Hamilton a mediados del siglo XIX, y por decirlo de alguna manera, son híbridos o generalizaciones a un plano hipercomplejo.

Un vector cartesiano es un cuaternión con la componente escalar nula. Hoy encontramos aplicaciones del álgebra de cuaterniones en Física¹² y más recientemente ha tenido alguna utilización en computación

¹¹Recuerde que estamos utilizando la convención de Einstein: $c^\alpha |q_\alpha\rangle \equiv c^0 + \sum_{j=1}^3 c^j |q_j\rangle$. Es decir, hemos supuesto que: $|q_0\rangle \equiv 1$, la unidad en los números reales. Adicionalmente, nótese que los índices griegos α, β, \dots toman los valores 0, 1, 2, 3, mientras que los latinos que acompañan a los vectores cartesianos toman los siguiente valores $j, k, l = 1, 2, 3$.

¹²Hace algunas décadas se dio una discusión sobre la importancia de utilizar esta representación en Física Cuántica. Pueden consultar:

- Berezin, A. V., Kurochkin, Y. A., & Tolkachev, E. A. (1989). Quaternions in relativistic physics. Nauka i Tekhnika, Minsk.
- Girard, P. R. (1984). The quaternion group and modern physics. European Journal of Physics, 5(1), 25.
- Horwitz, L. P., & Biedenharn, L. C. (1984). Quaternion quantum mechanics: second quantization and gauge fields. Annals of Physics, 157(2), 432-488.

gráfica que discutiremos en el próximo problema dentro del contexto del álgebra geométrica y las álgebras de Grassman.

Basándonos en este esquema podemos definir la “tabla de multiplicación” para los cuaterniones cartesianos como:

$ q_i\rangle \odot q_j\rangle$	1	$ q_1\rangle$	$ q_2\rangle$	$ q_3\rangle$
1	1	$ q_1\rangle$	$ q_2\rangle$	$ q_3\rangle$
$ q_1\rangle$	$ q_1\rangle$	-1	$ q_3\rangle$	$- q_2\rangle$
$ q_2\rangle$	$ q_2\rangle$	$- q_3\rangle$	-1	$ q_1\rangle$
$ q_3\rangle$	$ q_3\rangle$	$ q_2\rangle$	$- q_1\rangle$	-1

Nótese que por el hecho de que:

$$|q_j\rangle \odot |q_j\rangle = -1 \Rightarrow |q_1\rangle \odot |q_1\rangle = |q_2\rangle \odot |q_2\rangle = |q_3\rangle \odot |q_3\rangle = -1,$$

se puede pensar que un cuaternión es la generalización de los números complejos a más de una dimensión (un número hipercomplejo), donde la parte imaginaria tendría tres dimensiones y no una como es costumbre.

Esto es,

$$|a\rangle = a^\alpha |q_\alpha\rangle = a^0 \underbrace{|q_0\rangle}_1 + a^j |q_j\rangle = a^0 + \underbrace{a^1 |q_1\rangle + a^2 |q_2\rangle + a^3 |q_3\rangle}_{\text{“parte compleja”}}.$$

Siendo consistente con esa visión de generalización de un número complejo, definiremos el conjugado de un cuaternión como:

$$|b\rangle^\star = b^0 |q_0\rangle - b^j |q_j\rangle,$$

con $j = 1, 2, 3$.

Es decir, en analogía con los números complejos el conjugado de un cuaternión cambia el signo de su “parte compleja vectorial”.

Igualmente, definiremos la suma entre cuaterniones de la siguiente manera:

$$\left. \begin{array}{l} |a\rangle = a^\alpha |q_\alpha\rangle \\ |b\rangle = b^\alpha |q_\alpha\rangle \end{array} \right\} \Rightarrow |c\rangle = c^\alpha |q_\alpha\rangle = |a\rangle + |b\rangle = (a^\alpha + b^\alpha) |q_\alpha\rangle \Rightarrow c^\alpha = (a^\alpha + b^\alpha).$$

Esto quiere decir que los vectores se suman componente a componente. Mientras que la multiplicación por un escalar queda definida por $\alpha |c\rangle = \alpha c^\alpha |q_\alpha\rangle$, es decir se multiplica el escalar por cada componente. Con la información anterior, responda las siguientes preguntas:

- Compruebe si los cuaterniones, $|a\rangle$, forman un espacio vectorial respecto a esa operación suma y esa multiplicación por escalares, análoga a la de los vectores en \mathbb{R}^3 en coordenada cartesianas.
- Dados dos cuaterniones cualesquiera $|b\rangle \equiv (b^0, \mathbf{b})$ y $|r\rangle \equiv (r^0, \mathbf{r})$, y su tabla de multiplicación, muestre que el producto entre esos cuaterniones $|d\rangle = |b\rangle \odot |r\rangle$ podrá representarse como:

$$|d\rangle = |b\rangle \odot |r\rangle \longleftrightarrow (d^0, \mathbf{d}) = (b^0 r^0 - \mathbf{b} \cdot \mathbf{r}, r^0 \mathbf{b} + b^0 \mathbf{r} + \mathbf{b} \times \mathbf{r}),$$

donde \cdot y \times corresponden con los productos escalares y vectoriales tridimensionales de siempre.

- Ahora con índices: dados $|b\rangle = b^\alpha |q_\alpha\rangle$ y $|r\rangle = r^\alpha |q_\alpha\rangle$, compruebe si el producto $|d\rangle = |b\rangle \odot |r\rangle$ puede ser siempre escrito de la forma:

$$|d\rangle = |b\rangle \odot |r\rangle = a |q_0\rangle + S^{(\alpha j)} \delta_\alpha^0 |q_j\rangle + A^{[jk]i} b_j r_k |q_i\rangle.$$

donde a representa un número, $S^{(\alpha j)} \delta_\alpha^0$ (recuerde que los índices latinos toman los valores $j, k, l =$

1, 2, 3, mientras $\alpha = 0, 1, 2, 3$, donde $S^{(ij)}$ indica $S^{ji} = S^{ij}$, que la cantidad S^{ij} es simétrica, y por lo tanto $(S^{\alpha j} \delta_\alpha^0 + S^{j\alpha} \delta_\alpha^0) |q_j\rangle$.

Mientras $A^{[jk]i}$ representa un conjunto de objetos antisimétricos en j y k :¹³

$$A^{[jk]i} \rightarrow A^{jki} = -A^{kji} \rightarrow (A^{jki} b_j r_k - A^{kji} b_j r_k) |q_i\rangle.$$

- (d). Identifique las cantidades: a , $S^{(ij)}$ y $A^{[jk]i}$, en términos de las componentes de los cuaterniones.
¿El producto de cuaterniones $|d\rangle = |a\rangle \odot |r\rangle$ será un vector, pseudovector o ninguna de las anteriores?
Explique por qué.

- (e). Compruebe si las matrices de Pauli y la identidad

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_0 \equiv \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

pueden representar la base de los cuaterniones $\{|q_1\rangle, |q_2\rangle, |q_3\rangle, |q_0\rangle\}$. Seguidamente muestre que matrices complejas 2×2 del tipo:

$$|b\rangle \longleftrightarrow \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix},$$

pueden ser consideradas como cuaterniones, donde $z = x + iy$ y $w = a + ib$ son números complejos. Las Matrices de Pauli aparecen en mecánica cuántica cuando se tiene en cuenta la interacción del espín de una partícula con un campo electromagnético externo y en estas notas las consideraremos en varios momentos (ver, por ejemplo los ejercicios de las secciones 4.3.8 y 4.7.5)

- (f). Muestre que una representación posible para la base de cuaterniones son: la matriz identidad y las matrices reales 4×4 de la forma:

$$|q_1\rangle = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad |q_2\rangle = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad |q_3\rangle = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

- (g). Compruebe si la siguiente es una buena definición de producto interno:

$$\langle a | b \rangle = |a\rangle^{\star} \odot |b\rangle.$$

- (h). Modifique un poco la definición anterior de tal forma que:

$$\langle a | b \rangle = \frac{1}{2} \left[\langle a | b \rangle - |q_1\rangle \odot \langle a | b \rangle \odot |q_1\rangle \right],$$

y compruebe si esta definición compleja del producto interno cumple con todas las propiedades.

Nótese que un cuaternión de la forma $|f\rangle = f^0 + f^1 |q_1\rangle$ es un número complejo convencional.

- (i). Compruebe si la siguiente es una buena definición de norma para los cuaterniones:

$$n(|b\rangle) = ||a\rangle|| = \sqrt{\langle a | a \rangle} = \sqrt{|a\rangle^{\star} \odot |a\rangle}.$$

- (j). Compruebe si un cuaternión definido por:

$$\overline{|a\rangle} = \frac{|a\rangle^{\star}}{||a\rangle||^2},$$

puede ser considerado como el inverso o elemento simétrico de $|a\rangle$, respecto a la multiplicación \odot .

- (k). Compruebe si los cuaterniones $|a\rangle$ forman un grupo respecto a una operación multiplicación \odot .

- (l). Los vectores en \mathbb{R}^3 en coordenadas cartesianas, $|v\rangle$, pueden ser representados como cuaterniones, donde la parte escalar es nula $v^0 = 0 \rightarrow |v\rangle = v^j |q_j\rangle$. Compruebe si el siguiente producto conserva

¹³Para familiarizarse con las expresiones vectoriales con la notación de índices puede consultar la sección 1.4.

la norma:

$$|v'\rangle = \overline{|a\rangle} \odot |v\rangle \odot |a\rangle.$$

$$\text{Estos es: } \| |v'\rangle \|^2 = (v^1)^2 + (v^2)^2 + (v^3)^2 \equiv (v^1)^2 + (v^2)^2 + (v^3)^2 = \| |v\rangle \|^2.$$

7. En el mismo espíritu de los cuaterniones considerados previamente estudiemos el siguiente problema.

Consideremos otra vez el espacio \mathbb{R}^3 expandido por la base ortonormal estándar $\{\hat{i}, \hat{j}, \hat{k}\}$. Supongamos en este espacio el producto de dos vectores \mathbf{a} y \mathbf{b} (o $|a\rangle$ y $|b\rangle$ en la notación de vectores abstractos de Dirac) definido a la manera del *álgebra geométrica*. Esto es:

$$|a\rangle \odot |b\rangle \equiv \mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b},$$

con $\mathbf{a} \cdot \mathbf{b}$ el producto escalar estándar de \mathbb{R}^3 , representando la parte conmutativa de \mathbf{ab} y $\mathbf{a} \wedge \mathbf{b}$ su parte anticonmutativa. Esta última parte se relaciona con el producto vectorial estándar de la representación de Gibbs como: $\mathbf{a} \wedge \mathbf{b} = i\mathbf{a} \times \mathbf{b}$, con $i = \hat{i} \wedge \hat{j} \cdot \hat{k}$ un pseudoescalar.

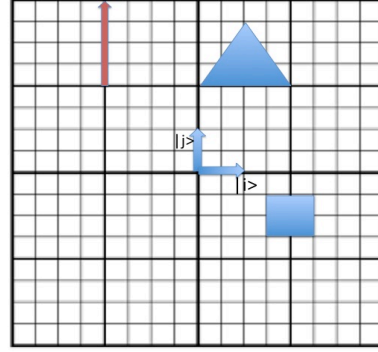


Figura 2.2: Se ilustran los vectores ortonormales estándares $\{\hat{i}, \hat{j}\}$ y como siempre $\hat{k} = \hat{i} \times \hat{j}$.

Este formalismo ha tenido cierto impacto en computación gráfica¹⁴ y lo vamos a utilizar para modelar transformaciones de objetos en el espacio. Considere el caso 2D representado en la figura 2.2, en el formalismo de *álgebra geométrica* el objeto triángulo lo asociamos con vector abstracto $|\triangle\rangle$, mientras que el objeto cuadrado lo representaremos como $|\square\rangle$, y el tercer objeto por $|\uparrow\rangle$.

Entonces:

- Expresé los vectores: $|\triangle\rangle, |\square\rangle, |\uparrow\rangle$ en términos de la base ortonormal estándar $\{\hat{i}, \hat{j}, \hat{k}\}$ (o $|i\rangle, |j\rangle, |k\rangle$) en el formalismo de *álgebra geométrica*.
- Expresé $|\triangle\rangle \odot |\uparrow\rangle$ en término de la base geométrica. Vale decir, si $|O_g\rangle$ es un objeto geométrico este podrá representarse en términos de una base: $|\epsilon\rangle, |\sigma\rangle, |B\rangle, |i\rangle$, donde $|\epsilon\rangle$ es un escalar; $|\sigma\rangle$ es un vector, $|B\rangle$ un bivector y, finalmente $|i\rangle$ un pseudoescalar.
- Encuentre la norma de cada uno de ellos y la distancia entre $|\triangle\rangle$ y $|\uparrow\rangle$.
- Considere ahora la operación: $\mathbb{A}|\triangle\rangle = |\tilde{\triangle}\rangle$.

I. Si $\mathbb{A}_{|j\rangle}$ es el operador reflexión respecto $|j\rangle$, exprese en términos geométricos $|\tilde{\triangle}\rangle = \mathbb{A}_{|j\rangle} |\triangle\rangle$ y $\mathbb{A}_{|j\rangle} (|\triangle\rangle \odot |\uparrow\rangle)$.

¹⁴Esta representación para objetos físicos ha tenido cierto éxito en Física clásica y cuántica; se puede consultar:

- Hestenes, D., & Sobczyk, G. (2012). *Clifford algebra to geometric calculus: a unified language for mathematics and physics* (Vol. 5). Springer Science & Business Media.
- Hestenes, D. (1971). Vectors, spinors, and complex numbers in classical and quantum physics. *Am. J. Phys.*, 39(9), 1013-1027.
- Hestenes, D. (2003). Spacetime physics with geometric algebra. *American Journal of Physics*, 71(7), 691-714.
- Dressel, J., Bliokh, K. Y., & Nori, F. (2015). Spacetime algebra as a powerful tool for electromagnetism. *Physics Reports*, 589, 1-71.
- Goldman, R. (2002). On the algebraic and geometric foundations of computer graphics. *ACM Transactions on Graphics (TOG)*, 21(1), 52-86.
- Hildenbrand (2011). *From Grassmann's vision to geometric algebra computing*. Springer Basel.
- Hildenbrand, D., Fontijne, D., Perwass, C., & Dorst, L. (2004). Geometric algebra and its application to computer graphics. In *Tutorial notes of the EUROGRAPHICS conference*.
- Vince, J. (2008). *Geometric algebra for computer graphics*. Springer Science & Business Media

y las referencias allí citadas.

- II. Si $\mathbb{A}_{\theta, |B\rangle}$ es el operador de rotación alrededor de un bivector $|B\rangle$ un ángulo θ . Encuentre $\mathbb{A}_{\pi/4, |B\rangle}(|\Delta\rangle \odot |\Uparrow\rangle)$ con $|B\rangle = |j\rangle + |k\rangle$.
- III. ¿Cómo interpreta Ud. la ecuación de autovalores $\mathbb{A}|\Delta\rangle = 2|\Delta\rangle$?
- (e). Considere el caso 3D en el cual $|\hat{\Delta}\rangle$ representa un tetraedro regular con la base representada por la figura 2.2 y $|\hat{\square}\rangle$, un cubo, también con su base representada en el plano de la misma figura.
- I. Expresé los vectores: $|\hat{\Delta}\rangle, |\hat{\square}\rangle$ en términos de la base ortonormal estándar $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ y el formalismo de álgebra geométrica.
- II. Expresé $(|\hat{\Delta}\rangle \odot |\Uparrow\rangle) \odot |\Delta\rangle$ en término de la base geométrica.
- III. Encuentre la norma de $|\hat{\Delta}\rangle$ y la distancia entre $|\hat{\Delta}\rangle$ y $\mathbb{A}_{\pi/2, |-j\rangle}|\hat{\square}\rangle$.
- IV. Expresé $[\mathbb{A}_{\pi/2, |-j\rangle}, \mathbb{B}_{\pi/4, |i\rangle}]|\hat{\square}\rangle$, con $[\mathbb{A}_{\pi/2, |-j\rangle}, \mathbb{B}_{\pi/4, |i\rangle}] = \mathbb{A}_{\pi/2, |-j\rangle}\mathbb{B}_{\pi/4, |i\rangle} - \mathbb{B}_{\pi/4, |i\rangle}\mathbb{A}_{\pi/2, |-j\rangle}$.
8. En Geometría Diferencial podemos considerar a \mathbb{R}^3 como un conjunto de puntos (una variedad o espacio topológico), es decir, \mathbb{R}^3 no es un espacio vectorial como se definió con anterioridad. En un punto q cualquiera podemos generar un plano \mathbb{R}^2 , entendido también como un conjunto de puntos, y definir el siguiente conjunto $T_q\mathbb{R}^2 = \{\text{el conjunto de todos los vectores geométricos (flechas) con origen en } q \text{ que son tangentes al plano } \mathbb{R}^2\}$. Notemos que la idea la podemos extender para cualquier superficie, por ejemplo, una esfera \mathbb{S}^2 (embebida en \mathbb{R}^3) y sobre la esfera seleccionar un punto arbitrario q . A partir de este punto q generar un plano y construir el conjunto $T_q\mathbb{S}^2 = \{\text{el conjunto de todos los vectores geométricos con origen en } q \text{ que son tangentes a la esfera } \mathbb{S}^2\}$.
- (a). ¿Es $T_q\mathbb{R}^2$ un espacio vectorial?
- (b). Claramente podemos seleccionar otro punto arbitrario, pero diferente de q , digamos p y construir el conjunto $T_p\mathbb{R}^2$. ¿Son estos dos espacios isomorfos?
9. Consideremos ahora el conjunto de todas las funciones diferenciables $f: \mathbb{R}^3 \rightarrow \mathbb{R}$. Esto es, que existen todas las derivadas parciales en $q \in \mathbb{R}^3$:

$$\partial_x f(x^i), \partial_y f(x^i), \partial_z f(x^i) \in \mathbb{R} \quad \forall (x^i) \in \mathbb{R}^3.$$

Note que hemos supuesto un sistema de coordenadas cartesiano $x^i = (x, y, z)$ en \mathbb{R}^3 y que las derivadas son evaluadas en q .

Consideremos el espacio tangente $T_q\mathbb{R}^3$ en un punto q arbitrario y un vector $|u_q\rangle \in T_q\mathbb{R}^3$ con componentes (u^1, u^2, u^3) , podemos definir el siguiente operador (operador derivada direccional) en q siguiendo a $|u_q\rangle$:

$$|U_q\rangle = (u^1\partial_x + u^2\partial_y + u^3\partial_z)_q.$$

Por lo tanto:

$$|U_q\rangle f \equiv (u^1\partial_x + u^2\partial_y + u^3\partial_z)_q f = u^1(\partial_x f)_q + u^2(\partial_y f)_q + u^3(\partial_z f)_q.$$

Es posible entonces construir el conjunto de los operadores derivadas direccionales que actúan sobre las funciones f :

$$\mathcal{D}_q(\mathbb{R}^3) = \{|U_q\rangle = (u^1\partial_x + u^2\partial_y + u^3\partial_z)_q \quad \forall |u_q\rangle \in T_q\mathbb{R}^3\}.$$

c) ¿Es \mathcal{D}_q un espacio vectorial?

d) ¿Los espacios \mathcal{D}_q y $T_q\mathbb{R}^3$ son isomorfos?

Practicando con Maxima

1. **Espacios y subespacios vectoriales** Sea el espacio vectorial $V = K^n$, definido en $K = \mathbb{R}$ y donde n es un entero positivo. Consideremos el caso $n = 4$. El producto de un elemento de K^4 , digamos $|x\rangle = (x_1, x_2, x_3, x_4)$ por un escalar $\alpha \in K$ resulta en otro elemento de K^4 .

Primero introducimos los elementos como listas:

```
(%i1) X:[x1,x2,x3,x4]; Y:[y1,y2,y3,y4]; Z:[z1,z2,z3,z4];
(%o1) [x1, x2, x3, x4]
(%o2) [y1, y2, y3, y4]
(%o3) [z1, z2, z3, z4]

(%i4) alpha*X=Y;
(%o4) [\alpha x1, \alpha x2, \alpha x3, \alpha x4] = [y1, y2, y3, y4]
```

El resultado es un elemento del espacio vectorial K^4 .

La suma de $|x\rangle = (x_1, x_2, x_3, x_4)$ y $|y\rangle = (y_1, y_2, y_3, y_4)$ será:

```
(%i5) X+Y=Z;
(%o5) [y1 + x1, y2 + x2, y3 + x3, y4 + x4] = [z1, z2, z3, z4]
```

con $(z_1, z_2, z_3, z_4) \in K^4$.

Podemos ver rápidamente que el conjunto de vectores que tienen la forma $(x_1, x_2, x_3, 0)$ conforman un subespacio de K^4 , ya que:

```
(%i6) map(":",[x4,y4,z4],[0,0,0]);
(%o6) [0, 0, 0]

(%i7) X:[x1,x2,x3,x4]; Y:[y1,y2,y3,y4]; Z:[z1,z2,z3,z4];
(%o7) [x1, x2, x3, 0]
(%o8) [y1, y2, y3, 0]
(%o9) [z1, z2, z3, 0]

(%i10) alpha*X+beta*Y=Z;
(%o10) [\beta y1 + \alpha x1, \beta y2 + \alpha x2, \beta y3 + \alpha x3, 0] = [z1, z2, z3, 0]
```

Para recuperar las variables x_4, y_4, z_4 escribimos:

```
(%i11) kill(x4,y4,z4)$
(%i12) X:[x1,x2,x3,x4]; Y:[y1,y2,y3,y4]; Z:[z1,z2,z3,z4];
(%o12) [x1, x2, x3, x4]
(%o13) [y1, y2, y3, y4]
(%o14) [z1, z2, z3, z4]
```

Para calcular el producto interno entre vectores es necesario utilizar la biblioteca **eigen**.

```
(%i15) load("eigen")$
(%i16) innerproduct(X,Y);
(%o16) x4 y4 + x3 y3 + x2 y2 + x1 y1
```

Consideremos ahora $V = K^n$, definido en $K = \mathbb{C}$, con $n = 3$. Por lo tanto, los vectores serán ahora de la siguiente forma: $z = (x_1 + iy_1, x_2 + iy_2, x_3 + iy_3)$.

```
(%i17) Z1:[x1+%i*y1,x2+%i*y2,x3+%i*y3]; Z2:[u1+%i*v1,u2+%i*v2,u3+%i*v3];
(%o17) [i y1 + x1, i y2 + x2, i y3 + x3]
(%o18) [i v1 + u1, i v2 + u2, i v3 + u3]
```

Y los escalares de la forma $\alpha = a + ib$.

```
(%i19) alpha:a+%i*b;
(%o19) i b + a
```

El producto por el escalar α es:

```
(%i20) Z3:alpha*Z1;
(%o20) [(i b + a) (i y1 + x1), (i b + a) (i y2 + x2), (i b + a) (i y3 + x3)]

(%i21) map(rectform,Z3);
(%o21) [i (a y1 + b x1) - b y1 + a x1, i (a y2 + b x2) - b y2 + a x2, i (a y3 + b x3) - b y3 + a x3]

(%i22) map(realpart,Z3),factor; map(imagpart,Z3),factor;
(%o22) [- (b y1 - a x1), - (b y2 - a x2), - (b y3 - a x3)]
(%o23) [a y1 + b x1, a y2 + b x2, a y3 + b x3]
```

Calculemos ahora el producto interno:

$$Z_1 \cdot Z_2 = (x_1 + iy_1)^*(u_1 + iv_1) + (x_2 + iy_2)^*(u_2 + iv_2) + (x_3 + iy_3)^*(u_3 + iv_3).$$

```
(%i24) Z4:innerproduct(Z1,Z2);
(%o24) (v3 - iu3) y3 + (v2 - iu2) y2 + (v1 - iu1) y1 + (iv3 + u3) x3 + (iv2 + u2) x2 + (iv1 + u1) x1

(%i25) Re:map(realpart,Z4)$ Im:map(imagpart,Z4)$
(%i26) Re+%i*Im;
(%o26) i (-u3y3 - u2y2 - u1y1 + v3x3 + v2x2 + v1x1) + v3y3 + v2y2 + v1y1 + u3x3 + u2x2 + u1x1

(%i27) kill(all)$
```

2. **Producto de polinomios** Consideremos el siguiente producto escalar entre elementos de un espacio vectorial de polinomios:

$$\langle p_i | p_j \rangle = \int_a^b p_i(x) p_j(x) dx.$$

Vamos a encontrar la distancia y el ángulo entre los vectores $|x_1\rangle = x(x-1)$ y $|x_2\rangle = x$ en dos intervalos diferentes: $[0, 1]$ y $[-1, 1]$

Debemos introducir los objetos a multiplicar:

```
(%i1) P1:x*(x-1); P2:x;
(%o1) (x - 1) x
(%o2) x
```

Ahora calculamos las distancias entre los vectores para ambos intervalos. Haremos gala de algunas posibilidades que ofrece el programa para escribir las expresiones.

(%i3) `sqrt('integrate(((P1-P2)^2),x,-1,1))=sqrt(integrate(((P1-P2)^2),x,-1,1));`

(%o3) $\sqrt{\int_{-1}^1 ((x-1)x-x)^2 dx} = \frac{\sqrt{46}}{\sqrt{15}}$

(%i4) `sqrt('integrate(((P1-P2)^2),x,0,1))=sqrt(integrate(((P1-P2)^2),x,0,1));`

(%o4) $\sqrt{\int_0^1 ((x-1)x-x)^2 dx} = \frac{2^{\frac{3}{2}}}{\sqrt{15}}$

El ángulo entre los polinomios definidos en el intervalo $[-1, 1]$ es:

(%i5) `'integrate((P1*P2),x,-1,1)/(sqrt('integrate((P1*P1),x,-1,1))*sqrt('integrate((P2*P2),x,-1,1)));`

(%o5)
$$\frac{\int_{-1}^1 (x-1)x^2 dx}{\sqrt{\int_{-1}^1 x^2 dx} \sqrt{\int_{-1}^1 (x-1)^2 x^2 dx}}$$

(%i6) `ev(%i5,integrate);`

(%o6)
$$-\frac{\sqrt{15}}{2^{\frac{3}{2}}\sqrt{3}}$$

(%i7) `acos(%i6),numer;`

(%o7) 2,482534617763384

Y ahora, el ángulo entre los polinomios definidos en el intervalo $[0, 1]$:

(%i8) `'integrate((P1*P2),x,0,1)/(sqrt('integrate((P1*P1),x,0,1))*sqrt('integrate((P2*P2),x,0,1)));`

(%o8)
$$\frac{\int_0^1 (x-1)x^2 dx}{\sqrt{\int_0^1 x^2 dx} \sqrt{\int_0^1 (x-1)^2 x^2 dx}}$$

(%i9) `ev(%i8,integrate);`

(%o9)
$$-\frac{\sqrt{30}}{4\sqrt{3}}$$

(%i10) `acos(%i9),numer;`

(%o10) 2,482534617763384

2.3 Variedades lineales

Si tenemos una cantidad de vectores: $\{|w_1\rangle, |w_2\rangle, |w_3\rangle, \dots\} \in V$, entonces una variedad lineal generada por este conjunto se entenderá como el conjunto de todas las combinaciones lineales finitas:

$$\alpha |w_1\rangle + \beta |w_2\rangle + \gamma |w_3\rangle + \dots,$$

donde los coeficientes: $\alpha, \beta, \gamma, \dots$ pertenecen al campo K .

Se puede comprobar que esta variedad lineal es un subespacio de V . Claramente, todo subespacio que contenga los vectores $\{|w_1\rangle, |w_2\rangle, |w_3\rangle, \dots\}$ también contiene todas sus combinaciones lineales, por lo tanto, la variedad lineal generada de esta manera es el subespacio más pequeño que contiene al conjunto de estos vectores.

Una variedad lineal sencilla de construir en $V = \mathbb{R}^3$ la constituye un par de vectores no colineales. La variedad lineal generada por estos dos vectores será el conjunto de todos los vectores paralelos al plano determinado por este par de vectores no colineales. Mientras que la variedad lineal generada por los vectores: $\{\hat{i}, \hat{j}, \hat{k}\} \in \mathbb{R}^3$ es el mismo espacio entero \mathbb{R}^3 .

Pasemos ahora a considerar el problema de construir una base para una variedad lineal y a partir de allí determinar la dimensión de la variedad.

2.3.1 Dependencia/Independencia lineal

Siguiendo la misma línea de razonamiento que en las secciones 1.1.3 y 1.2.5, generalizamos el concepto de dependencia e independencia lineal de \mathbb{R}^2 y \mathbb{R}^3 .

Dada la ecuación

$$|0\rangle = C_1 |v_1\rangle + C_2 |v_2\rangle + C_3 |v_3\rangle \cdots + C_n |v_n\rangle = \sum_{i=1}^n C_i |v_i\rangle, \quad (2.7)$$

las cantidades C_i son llamados los coeficientes de la combinación lineal.

Podemos afirmar que:

- Si esta ecuación se cumple para algún conjunto de $\{C_i\}$ no nulos, se dirá que el conjunto de vectores correspondiente $\{|v_i\rangle\}$ son **linealmente dependientes**.
- Por el contrario, si esta ecuación **sólo** puede ser satisfecha para todos los $C_i = 0$, entonces se dirá que el conjunto de vectores correspondiente $\{|v_i\rangle\}$ son **linealmente independientes**.

Notemos que la suma que aparece aquí es necesariamente una suma finita, y cuando un determinado conjunto de vectores es linealmente dependiente, entonces uno de ellos se puede escribir como combinación lineal de los demás.

2.3.2 Bases de un espacio vectorial

Ahora bien, dado un espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle \cdots, |v_n\rangle\}$, si encontramos que el conjunto de $\{|v_n\rangle\}$ es linealmente dependiente, entonces siempre es posible despejar uno de los vectores en términos de los demás, vale decir:

$$|v_n\rangle = \bar{C}_1 |v_1\rangle + \bar{C}_2 |v_2\rangle + \bar{C}_3 |v_3\rangle \cdots + \bar{C}_{n-1} |v_{n-1}\rangle = \sum_{i=1}^{n-1} \bar{C}_i |v_i\rangle.$$

Seguidamente podemos proceder a comprobar si $\{|v_1\rangle, |v_2\rangle, |v_3\rangle \cdots, |v_{n-1}\rangle\}$ es un conjunto de vectores linealmente independientes, es decir: $\bar{C}_1 = \bar{C}_2 = \bar{C}_3 = \cdots = \bar{C}_{n-1} = 0$.

En caso de no serlo se procede otra vez a despejar uno de los vectores en términos de los anteriores y aplicar el criterio de independencia lineal:

$$|v_{n-1}\rangle = \tilde{C}_1 |v_1\rangle + \tilde{C}_2 |v_2\rangle + \tilde{C}_3 |v_3\rangle \cdots + \tilde{C}_{n-2} |v_{n-2}\rangle = \sum_{i=1}^{n-2} \tilde{C}_i |v_i\rangle ,$$

nuevamente se comprueba si se cumple: $\tilde{C}_1 = \tilde{C}_2 = \tilde{C}_3 = \cdots = \tilde{C}_{n-2} = 0$.

En caso contrario, se repite este procedimiento hasta encontrar un conjunto: $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_{n-j}\rangle\}$ de vectores linealmente independientes. Esto es: $\check{C}_1 = \check{C}_2 = \check{C}_3 = \cdots = \check{C}_{n-j} = 0$.

Por lo tanto:

$$|v_{n-j+1}\rangle = \check{C}_1 |v_1\rangle + \check{C}_2 |v_2\rangle + \check{C}_3 |v_3\rangle \cdots + \check{C}_{n-j} |v_{n-j}\rangle = \sum_{i=1}^{n-j} \check{C}_i |v_i\rangle .$$

Diremos entonces que $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_{n-j}\rangle\}$ forman una base para el espacio vectorial \mathbf{V} .

Es importante señalar, que la dimensión de \mathbf{V} será el número de vectores linealmente independientes, que para este caso será: $\dim \mathbf{V} = n - j$.

Entonces, se puede comprobar que, dado un vector arbitrario $|x\rangle \in \mathbf{V}$, se tiene que:

$$|x\rangle = \sum_{i=1}^{n-j} C_i |v_i\rangle \quad \forall \quad |x\rangle \in \mathbf{V},$$

y el conjunto $\{C_1, C_2, C_3, \dots, C_{n-j}\}$ será único.

Diremos que el número mínimo de vectores: $|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_{n-j}\rangle$ que expanden \mathbf{V} conforman una base de ese espacio vectorial, y que el número finito de cantidades $C_1, C_2, C_3, \dots, C_{n-j}$, constituyen las *componentes* de $|x\rangle$ relativas a la base $\{|v_1\rangle, |v_2\rangle, \dots, |v_{n-j}\rangle\}$.

Queda claro que el vector cero, $|0\rangle$, es linealmente dependiente, y cualquier conjunto de vectores que lo contenga es un conjunto linealmente dependiente de vectores.

De lo anteriormente expuesto se puede concretar la siguiente definición para una base de un espacio vectorial \mathbf{V} :

Definición 2.1 (Bases de un espacio vectorial)

Un conjunto finito de vectores:

$$\mathcal{B} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\} \in \mathbf{V},$$

se les denominará base del espacio vectorial \mathbf{V} si los vectores: $|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle$, son linealmente independientes y expanden \mathbf{V} . Diremos entonces que \mathcal{B} es un sistema generador de \mathbf{V} .

Es fácil darse cuenta que si \mathbf{V} lo expanden n vectores linealmente independientes, cualquier otro vector $|x\rangle \in \mathbf{V}$ será linealmente dependiente. Igualmente, todas las bases de un espacio vectorial \mathbf{V} , de dimensión finita, tendrán el mismo número de elementos y ese número de elemento será la dimensión del espacio, es decir, la $\dim \mathbf{V} = \text{número de vectores que forman una base del espacio}$.

La base más familiar en el espacio tridimensional real es el conjunto de vectores ortogonales y unitarios: $\{\hat{i}, \hat{j}, \hat{k}\}$. Por lo tanto, como ya sabemos, la dimensión del espacio vectorial $\mathbf{V} = \mathbb{R}^3$ es 3. Al conjunto de vectores $\{\hat{i}, \hat{j}, \hat{k}\}$ le podemos asociar tres ejes coordenados: $\{x, y, z\}$ (decimos que le anclamos un sistema de coordenadas), de manera que las componentes C_1, C_2, C_3 de un vector $|v\rangle$ respecto a esta base son las proyecciones de $|v\rangle$ a lo largo de los ejes coordenados.

El concepto de base de un espacio vectorial es de fundamental importancia, ya que una vez especificada la base las operaciones sobre los elementos del espacio vectorial abstracto se pueden realizar ahora sobre los

números que representan las componentes del vector con respecto a la base. Esto significa que cuando sumamos dos vectores de un espacio vectorial abstracto V , sus componentes (respecto a una base) son sumadas. Cuando multiplicamos un vector de V por un elemento α del campo K , todas sus componentes son multiplicadas por α .

Adicionalmente, dentro de un espacio vectorial V se pueden encontrar subespacios y dentro de esos subespacios un conjunto de vectores base.

Vale decir, si $\forall |x\rangle \in V$:

$$|x\rangle = \underbrace{C_1 |v_1\rangle \cdots + C_{n-j} |v_{n-j}\rangle}_{S_1} + \underbrace{C_{n-j+1} |v_{n-j+1}\rangle \cdots + C_{n-k} |v_{n-k}\rangle}_{S_2} + \underbrace{C_{n-k+1} |v_{n-k+1}\rangle \cdots + C_n |v_n\rangle}_{S_3},$$

con: $|x\rangle = |x_1\rangle + |x_2\rangle + |x_3\rangle$ y $|x_1\rangle \in S_1$; $|x_2\rangle \in S_2$; $|x_3\rangle \in S_3$. Entonces diremos que V es la suma directa de S_1, S_2 y S_3 y lo denotaremos como: $V = S_1 \oplus S_2 \oplus S_3$.

También es bueno señalar que, una vez fijada una base, las componentes de un vector según esa base, son únicas y que dada una base de un espacio vectorial, se pueden construir otras bases diferentes de ésta, como veremos más adelante.

Ejemplo 2.8 Independencia lineal: Presentamos otros ejemplos sobre dependencia/independencia lineal.

1. Si consideramos el espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ serán ejemplos de **independencia lineal**:

- $|v_k\rangle \equiv f(t) = t^k$, para $k = 1, 2, 3, \dots$. Es claro que un polinomio de grado $n+1$, no podrá ser expresado en términos un polinomio de grado n , en otras palabras: $t^{n+1} \neq \sum_{i=0}^n C_i t^i$.
- $|v_k\rangle \equiv f(t) = e^{a_k t}$, con a_1, a_2, a_3, \dots coeficientes constantes. También salta a la vista que no podremos expresar una de esas funciones exponenciales como una combinación lineal.

2. Si consideramos: $|v_1\rangle = \cos^2(t)$, $|v_2\rangle = \sin^2(t)$ y $|v_3\rangle = 1$, es claro que $|v_1\rangle, |v_2\rangle$ y $|v_3\rangle$ son **linealmente dependientes** por cuanto: $|v_1\rangle + |v_2\rangle = |v_3\rangle$. Nótese que si:

$$|v_1\rangle = \cos(t), |v_2\rangle = \sin(t) \text{ y } |v_3\rangle = 1,$$

entonces $|v_1\rangle, |v_2\rangle$ y $|v_3\rangle$ serán vectores **linealmente independientes**.

3. Consideremos ahora otro ejemplo en \mathcal{P}^3 ,

$$|x_1\rangle = 1, \quad |x_2\rangle = x - 1, \quad |x_3\rangle = x^2, \quad |x_4\rangle = x^2 + 2x + 1.$$

Podemos ver que este conjunto es linealmente dependiente ya que siempre podremos expresar:

$$|x_4\rangle = 3|x_1\rangle + 2|x_2\rangle + |x_3\rangle.$$

◀

2.3.3 La matriz y el determinante de Gram

Para un conjunto de vectores: $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\} \in V$, existe una forma directa de comprobar la independencia lineal del conjunto. Dado un vector $|x\rangle \in V$ tendremos que $|x\rangle = C^i |v_i\rangle$, entonces, al multiplicar ambos lados por $\langle v^i |$, resulta:

$$\begin{aligned} \langle v^1 | x \rangle &= C^1 \langle v^1 | v_1 \rangle + C^2 \langle v^1 | v_2 \rangle + C^3 \langle v^1 | v_3 \rangle + \cdots + C^n \langle v^1 | v_n \rangle \\ \langle v^2 | x \rangle &= C^1 \langle v^2 | v_1 \rangle + C^2 \langle v^2 | v_2 \rangle + C^3 \langle v^2 | v_3 \rangle + \cdots + C^n \langle v^2 | v_n \rangle \\ &\vdots \\ \langle v^n | x \rangle &= C^1 \langle v^n | v_1 \rangle + C^2 \langle v^n | v_2 \rangle + C^3 \langle v^n | v_3 \rangle + \cdots + C^n \langle v^n | v_n \rangle \end{aligned}$$

donde las $C^1, C^2, C^3, \dots, C^n$ son las incógnitas. Por lo cual, para que este sistema tenga solución se impone que el determinante:

$$\begin{vmatrix} \langle v^1 | v_1 \rangle & \langle v^1 | v_2 \rangle & \langle v^1 | v_3 \rangle & \cdots & \langle v^1 | v_n \rangle \\ \langle v^2 | v_1 \rangle & \langle v^2 | v_2 \rangle & \langle v^2 | v_3 \rangle & \cdots & \langle v^2 | v_n \rangle \\ \vdots & & \ddots & & \vdots \\ \langle v^n | v_1 \rangle & \langle v^n | v_2 \rangle & \langle v^n | v_3 \rangle & \cdots & \langle v^n | v_n \rangle \end{vmatrix} \neq 0.$$

Esto es, que el determinante de Gram¹⁵ sea distinto de cero implica que el conjunto: $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$ es linealmente independiente. La inversa también es cierta.

La matriz de Gram aparece de manera natural cuando estudiamos el producto interno en un espacio vectorial. Sea V^n un espacio vectorial de dimensión n con una base $\mathcal{B} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$. Consideremos también que en V^n existe un producto escalar $\langle x | y \rangle$. Entonces, para dos vectores cualesquiera $|x\rangle$ y $|y\rangle$ expandidos en esa base \mathcal{B} resulta lo siguiente:

$$\begin{aligned} \langle x | y \rangle &= \left\langle \sum_{i=1}^n x_i v_i \left| \sum_{j=1}^n y_j v_j \right. \right\rangle = \sum_{i=1}^n x_i^* \left[\sum_{j=1}^n y_j \langle v^i | v_j \rangle \right] \\ &= (x_1^* \ x_2^* \ x_3^* \cdots x_n^*) \begin{pmatrix} \sum_{j=1}^n y_j \langle v^1 | v_j \rangle \\ \sum_{j=1}^n y_j \langle v^2 | v_j \rangle \\ \vdots \\ \sum_{j=1}^n y_j \langle v^n | v_j \rangle \end{pmatrix} \\ &= (x_1^* \ x_2^* \ x_3^* \cdots x_n^*) \begin{pmatrix} \langle v^1 | v_1 \rangle & \langle v^1 | v_2 \rangle & \cdots & \langle v^1 | v_n \rangle \\ \langle v^2 | v_1 \rangle & \langle v^2 | v_2 \rangle & \cdots & \langle v^2 | v_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle v^n | v_1 \rangle & \langle v^n | v_2 \rangle & \cdots & \langle v^n | v_n \rangle \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \end{aligned}$$

Por lo tanto:

$$\langle x | y \rangle = (|x\rangle^t)^* G |y\rangle \quad (2.8)$$

La matriz G se llama la matriz de Gram para el producto escalar respecto a la base \mathcal{B} . En el caso de espacios vectoriales reales la matriz de Gram es simétrica: $G_{ij} = G_{ji}$, y en el caso complejo es hermítica o autoadjunta: $(G_{ij}^t)^* = G_{ij}$, es decir, es igual a su transpuesta compleja conjugada.

Ejemplo 2.9 El espacio vectorial V^n tendrá dimensión n y una de las posibles bases $\{|v_i\rangle\}$ será:

$$|v_1\rangle = (1, 0, 0, \dots, 0), \quad |v_2\rangle = (0, 1, 0, \dots, 0), \quad |v_3\rangle = (0, 0, 1, \dots, 0), \dots, \quad |v_n\rangle = (0, 0, 0, \dots, 1).$$

Esta base se conoce con el nombre de base canónica, la base con la que aprendimos el álgebra vectorial en \mathbb{R}^3 : $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$.

Ejemplo 2.10 Es fácil ver que la matriz de Gram para el producto escalar respecto a la base canónica $\{\hat{\mathbf{i}}_i\}$ en

¹⁵Jørgen Pedersen Gram (1850-1916) Es conocido mayormente por el método de ortogonalización, pero aparentemente fue ideado por Laplace y utilizado también por Cauchy en 1836. https://en.wikipedia.org/wiki/J\T1\orgen_Pedersen_Gram

\mathbb{R}^n es la matriz identidad. Para dos vectores $|x\rangle$ y $|y\rangle$ tenemos

$$\begin{aligned}\langle x | y \rangle &= (x_1 \ x_2 \ x_3 \ \cdots \ x_n) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.\end{aligned}$$

Ejemplo 2.11 El espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ tendrá como una de las posibles bases al conjunto: $\{1, t, t^2, t^3, \dots, t^n\}$, porque cualquier polinomio de grado $\leq n$ podrá ser expresado como combinación lineal de estos $n + 1$ vectores. Más aún, el espacio de **todos** los polinomios, \mathcal{P}^∞ , tendrá como una posible base al conjunto de funciones: $\{1, t, t^2, t^3, \dots, t^n, \dots\}$. En este caso \mathcal{P}^∞ será infinito dimensional.

Ejemplo 2.12 Consideremos el espacio de polinomios \mathcal{P}^3 y la base $\{1, t, t^2\}$ con el producto interno definido como:

$$\langle p, q \rangle = \frac{1}{2} \int_{-1}^1 p(t)q(t)dt.$$

Por lo tanto:

$$\langle 1, 1 \rangle = \frac{1}{2} \int_{-1}^1 dt = 1, \quad \langle 1, t \rangle = \frac{1}{2} \int_{-1}^1 t dt = 0, \quad \langle 1, t^2 \rangle = \frac{1}{2} \int_{-1}^1 t^2 dt = \frac{1}{3}, \dots$$

De esta manera, la matriz de Gram respecto a la base dada es:

$$G_{\{1, t, t^2\}} = \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} \end{pmatrix}.$$

2.3.4 Ortogonalidad y bases ortogonales

En un espacio vectorial con producto interno, dos vectores $|v_1\rangle$ y $|v_2\rangle$ serán ortogonales si su producto interno se anula

$$|v_1\rangle \perp |v_2\rangle \Leftrightarrow \langle v_2 | v_1 \rangle = 0.$$

Se denomina un conjunto **ortogonal** de vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ si:

$$\langle e^i | e_j \rangle = \delta_j^i \|e_j\|^2, \quad i, j = 1, 2, 3, \dots, n \quad \text{y con} \quad \delta_j^i = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

En este caso, es fácil ver que la matriz de Gram es una matriz en diagonal.

Además, se denominará un conjunto **ortonormal** de vectores $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle\}$ si:

$$\langle \hat{e}^i | \hat{e}_j \rangle = \delta_j^i, \quad i, j = 1, 2, 3, \dots, n,$$

es decir, cuando $\|e_j\|^2 = 1$ y la matriz de Gram será la matriz identidad.

Un conjunto ortogonal de vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\} \in \mathbf{V}$ es linealmente independiente. Más aún, para el caso particular de un espacio euclidiano este conjunto conforma una base ortogonal para \mathbf{V} . La demostración es sencilla, para un determinado espacio vectorial \mathbf{V} una combinación lineal de los vectores:

$\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ se anula. Veamos:

$$\sum_{i=1}^n C^i |e_i\rangle = |0\rangle \Rightarrow \begin{cases} \langle e^1 | [\sum_{i=1}^n C^i |e_i\rangle] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^1 = 0 & \Rightarrow C^1 = 0 \\ \langle e^2 | [\sum_{i=1}^n C^i |e_i\rangle] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^2 = 0 & \Rightarrow C^2 = 0 \\ \langle e^3 | [\sum_{i=1}^n C^i |e_i\rangle] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^3 = 0 & \Rightarrow C^3 = 0 \\ \vdots & \ddots & \vdots \\ \langle e^n | [\sum_{i=1}^n C^i |e_i\rangle] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^n = 0 & \Rightarrow C^n = 0 \end{cases}$$

entonces, queda claro que: $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ son un conjunto de vectores linealmente independientes.

Si la dimensión de V es n ($\dim V = n$) y tenemos n vectores linealmente independientes, entonces esos n vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ forman una base ortogonal para V , y por lo tanto, las componentes de un vector en esa base se pueden expresar de manera simple:

$$\forall |x\rangle \in V \Rightarrow |x\rangle = \sum_{i=1}^n C^i |e_i\rangle \Rightarrow \langle e^j | x \rangle = \langle e^j | \left[\sum_{i=1}^n C^i |e_i\rangle \right] \Rightarrow C^j = \frac{\langle e^j | x \rangle}{\langle e^j | e_j \rangle} = \frac{\langle e^j | x \rangle}{\| |e_j\rangle \|^2}.$$

En el caso de un conjunto ortonormal de vectores $\{|\hat{e}_i\rangle\} \in V^n$, las componentes de cualquier vector quedan determinadas de una forma todavía más simple y con consecuencias mucho más impactantes:

$$\text{Si } \| |\hat{e}_i\rangle \|^2 = 1 \Rightarrow C^i = \langle \hat{e}^i | x \rangle \Rightarrow |x\rangle = \sum_{i=1}^n C^i |\hat{e}_i\rangle = \sum_{i=1}^n \langle \hat{e}^i | x \rangle |\hat{e}_i\rangle \equiv \underbrace{\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i | x \rangle}_1.$$

Es bueno recalcar la relación de cierre¹⁶

$$\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i| = 1, \quad (2.9)$$

con lo cual es trivial demostrar la fórmula de Parseval:

$$\forall |x\rangle, |y\rangle \in V \Rightarrow \langle y | x \rangle \equiv \langle y | \left(\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i| \right) | x \rangle = \sum_{i=1}^n \langle y | \hat{e}_i \rangle \langle \hat{e}^i | x \rangle = \sum_{i=1}^n \langle y | \hat{e}_i \rangle \langle x | \hat{e}_i \rangle^*,$$

para el caso de $|x\rangle \equiv |y\rangle$ se llega a la generalización del teorema de Pitágoras:

$$\langle x | x \rangle \equiv \| |x\rangle \|^2 = \sum_{i=1}^n |\langle x | \hat{e}_i \rangle|^2.$$

Consideremos un par de ejemplos de bases ortogonales en espacio funcionales:

Ejemplo 2.13 Funciones trigonométricas: Uno de los ejemplos más emblemáticos es el caso de las funciones continuas, reales de variable real y definidas en $[0, 2\pi]$, $\mathcal{C}_{[0, 2\pi]}^\infty$, con el producto interno definido por:

$$\langle f | g \rangle = \int_0^{2\pi} dx f(x) g(x), \quad (2.10)$$

esto es, el conjunto de funciones $\{|e_i\rangle\}$ representadas por:

$$|e_0\rangle = 1, \quad |e_{2n-1}\rangle = \cos(nx) \quad \text{y} \quad |e_{2n}\rangle = \sin(nx), \quad \text{con } n = 1, 2, 3, \dots \quad (2.11)$$

¹⁶La relación de cierre expresa una propiedad importante de los vectores base: si los $|\hat{e}_i\rangle$ se multiplican por la derecha por los $\langle \hat{e}^i|$, el resultado, luego de sumar para todos los vectores, es el operador lineal unitario. Se dice también que el conjunto de vectores $\{|\hat{e}_i\rangle\}$ forman un conjunto completo.

Es claro que $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle, \dots\}$ es un conjunto de funciones ortogonales por cuanto:

$$\langle e^n | e_m \rangle = \delta_m^n \| |e_n\rangle \|^2 \Rightarrow \begin{cases} 0 & \text{si } n \neq m \Rightarrow \begin{cases} \int_0^{2\pi} dx \sin(nx) \sin(mx) = 0 \\ \int_0^{2\pi} dx \cos(nx) \sin(mx) = 0 \\ \int_0^{2\pi} dx \cos(nx) \cos(mx) = 0 \end{cases} \\ \| |e_n\rangle \|^2 & \text{si } n = m \Rightarrow \begin{cases} \int_0^{2\pi} dx = 2\pi & \text{si } n = m = 0 \\ \int_0^{2\pi} dx \cos^2(nx) = \pi & \text{si } n = m = 2k-1 \\ \int_0^{2\pi} dx \sin^2(nx) = \pi & \text{si } n = m = 2k \end{cases} \end{cases}$$

con $k = 1, 2, 3, \dots$.

Claramente, podremos construir una base ortonormal de funciones: $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, \dots, |\hat{e}_n\rangle, \dots\}$ de la forma:

$$|\hat{e}_0\rangle = \frac{1}{\sqrt{2\pi}}, \quad |\hat{e}_{2n-1}\rangle = \frac{1}{\sqrt{\pi}} \cos(nx) \quad \text{y} \quad |\hat{e}_{2n}\rangle = \frac{1}{\sqrt{\pi}} \sin(nx).$$

Por lo tanto, dada una función definida en el intervalo $[0, 2\pi]$, podremos expresarla en términos de la base ortogonal como:

$$|f\rangle = \sum_{i=1}^{\infty} C^i |e_i\rangle \Rightarrow C^i = \langle e^i | f \rangle = \begin{cases} \int_0^{2\pi} dx f(x) = C^0 & \text{si } i = 0 \\ \int_0^{2\pi} dx f(x) \cos(nx) = C^{2n-1} & \text{si } i = 2n-1 \\ \int_0^{2\pi} dx f(x) \sin(nx) = C^{2n} & \text{si } i = 2n \end{cases}$$

donde los C^i son los coeficientes de Fourier.

◀

Ejemplo 2.14 Polinomios de Legendre: Otro de los ejemplos típicos son los llamados polinomios de Legendre, $P_n(x)$, definidos en el intervalo $[-1, 1]$. Estos polinomios pueden ser generados a partir de la Fórmula de Rodrigues¹⁷

$$P_n(x) = \frac{1}{n! 2^n} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n = 0, 1, 2, \dots \quad (2.12)$$

con $P_0(x) = 1$.

Algunos de estos polinomios son los siguientes:

$$P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{x}{2}(5x^2 - 3), \quad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3), \dots$$

los polinomios de Legendre son soluciones de la ecuación diferencial:

$$(1 - x^2) y'' - 2x y' + \lambda(\lambda + 1)y = 0.$$

Es fácil comprobar que los polinomios de Legendre $|P_\alpha\rangle = P_\alpha(x)$ son mutuamente ortogonales con un producto interno definido como,

$$\langle P_n | P_m \rangle = \int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2n+1} \delta_{nm}, \quad (2.13)$$

y con una norma definida por,

$$\|P_n\|^2 = \langle P_n | P_n \rangle = \int_{-1}^1 P_n^2(x) dx = \frac{2}{2n+1}. \quad (2.14)$$

¹⁷Benjamin Olinde Rodrigues (1794-1851). https://es.wikipedia.org/wiki/Olinde_Rodrigues

Por lo tanto, cualquier función en el intervalo $[-1, 1]$ puede ser expresada en esa base.

$$f(x) = |f\rangle = \sum_{k=0}^{\infty} a_k |P_k\rangle = \sum_{k=0}^{\infty} \frac{\langle P_k | f \rangle}{\langle P_k | P_k \rangle} |P_k\rangle. \quad (2.15)$$

Si $f(x)$ es un polinomio, entonces

$$f(x) = \sum_{n=0}^m b_n x^n = \sum_{k=0}^{\infty} a^k |P_k\rangle = \sum_{n=0}^{\infty} a^n P_n(x),$$

y los coeficientes a^n se determinan fácilmente a través de un sistema de ecuaciones algebraicas. Por ejemplo, para el caso de $f(x) = x^2$ tendremos

$$\begin{aligned} f(x) = x^2 &= a^0 P_0(x) + a^1 P_1(x) + a^2 P_2(x) \\ &= a^0 + a^1 x + \frac{1}{2} a^2 (3x^2 - 1) = \frac{1}{3} P_0(x) + \frac{2}{3} P_2(x). \end{aligned}$$

Quedará como ejercicio demostrar que para el caso

$$g(x) = \sqrt{\frac{1-x}{2}} = \sum_{k=0}^{\infty} \frac{\langle P_k | g \rangle}{\langle P_k | P_k \rangle} |P_k\rangle = \frac{2}{3} P_0(x) - 2 \sum_{n=1}^{\infty} \frac{P_n(x)}{(2n-1)(2n+3)},$$

con

$$\langle P_k | g \rangle = \int_{-1}^1 g(x) P_k(x) dx = \int_{-1}^1 \sqrt{\frac{1-x}{2}} P_k(x) dx.$$

◀

2.3.5 Ortogonalización

Hemos visto que un conjunto de vectores ortogonales forman una base para un espacio vectorial. Ahora bien, siempre es posible construir un conjunto de vectores ortogonales a partir de un conjunto de vectores linealmente independientes. El método de “ortogonalización” se conoce como el método de Gram-Schmidt¹⁸, en honor de estos dos matemáticos alemanes que no lo inventaron pero hicieron famoso al método que, al parecer, se le debe al matemático francés P.S. Laplace.

Consideremos, por ejemplo, un conjunto de vectores linealmente independientes, $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ que expanden un espacio euclidiano real de dimensión finita, E^n . Entonces, siempre se podrá construir un conjunto ortogonal de vectores, $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$, que también expandan E^n , de la siguiente forma:

1. Hacemos coincidir uno de los vectores de la base dada con nuestro primer vector ortogonal:

$$|e_1\rangle \equiv |v_1\rangle.$$

2. A partir de $|e_1\rangle$ y otro de los vectores dados construimos un segundo vector:

$$|e_2\rangle \equiv |v_2\rangle - \frac{\langle e_1 | v_2 \rangle}{\langle e_1 | e_1 \rangle} |e_1\rangle \quad \backslash \quad \langle e_2 | e_1 \rangle = 0.$$

3. De esta manera podemos seguir calculando los otros vectores de la base ortogonal hasta completar los n vectores dados:

¹⁸Erhard Schmidt (1876-1959). Matemático alemán fundador del primer instituto de matemáticas aplicadas de Berlín. Alumno de Hilbert, Schmidt hizo sus mayores contribuciones en el campo de ecuaciones integrales y teoría de funciones en el espacio de Hilbert.
https://es.wikipedia.org/wiki/Erhard_Schmidt

$$\begin{aligned}
|e_3\rangle &\equiv |v_3\rangle - \frac{\langle e^2|v_3\rangle}{\langle e^2|e_2\rangle} |e_2\rangle - \frac{\langle e^1|v_3\rangle}{\langle e^1|e_1\rangle} |e_1\rangle & \setminus & \begin{cases} \langle e^3|e_1\rangle = 0 \\ \langle e^3|e_2\rangle = 0 \end{cases} \\
|e_4\rangle &\equiv |v_4\rangle - \frac{\langle e^3|v_4\rangle}{\langle e^3|e_3\rangle} |e_3\rangle - \frac{\langle e^2|v_4\rangle}{\langle e^2|e_2\rangle} |e_2\rangle - \frac{\langle e^1|v_4\rangle}{\langle e^1|e_1\rangle} |e_1\rangle & \setminus & \begin{cases} \langle e^4|e_1\rangle = 0 \\ \langle e^4|e_2\rangle = 0 \\ \langle e^4|e_3\rangle = 0 \end{cases} \\
\vdots & & & \vdots \\
|e_n\rangle &\equiv |v_n\rangle - \sum_{i=1}^{n-1} \frac{\langle e^i|v_n\rangle}{\langle e^i|e_i\rangle} |e_i\rangle & \setminus & \begin{cases} \langle e^n|e_1\rangle = 0 \\ \langle e^n|e_2\rangle = 0 \\ \vdots \\ \langle e^n|e_{n-1}\rangle = 0 \end{cases}
\end{aligned}$$

Entonces, siempre es posible construir una base ortogonal a partir de un conjunto de vectores linealmente independientes y una definición de producto interno. Esta base será única en E^n —para cada definición de producto interno— y si existe otra, sus vectores serán proporcionales. Más aún, cada espacio vectorial V^n de dimensión finita tendrá una base ortogonal asociada¹⁹.

Ejemplo 2.15 Consideremos un par de ejemplos del método de ortogonalización sencillos, y uno más elaborado:

1. Para el caso de \mathbb{R}^2 es muy claro. Si tenemos dos vectores $|v_1\rangle$ y $|v_2\rangle$ linealmente independientes:

$$|v_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

y si elegimos $|e_1\rangle \equiv |v_1\rangle$, entonces, $|e_2\rangle$ vendrá dado por:

$$|e_2\rangle \equiv |v_1\rangle - \frac{\langle e^1|v_1\rangle}{\langle e^1|e_1\rangle} |e_1\rangle \Rightarrow |e_2\rangle \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

tal y como se esperaba, el otro vector ortogonal es el canónico.

2. Un subespacio de V^4 , expandido por los siguientes vectores

$$|v_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \quad |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

¹⁹Hemos construido la base ortogonal para un espacio de dimensión finita, pero el procedimiento es válido para espacios de dimensión infinita

tendrá una base ortogonal asociada dada por:

$$|e_1\rangle \equiv |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}; |e_2\rangle \equiv |v_2\rangle - \frac{\langle e^1 | v_2 \rangle}{\langle e^1 | e_1 \rangle} |e_1\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix} - (-1) \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}$$

$$|e_3\rangle \equiv |v_1\rangle - \frac{\langle e^2 | v_1 \rangle}{\langle e^2 | e_2 \rangle} |e_2\rangle - \frac{\langle e^1 | v_1 \rangle}{\langle e^1 | e_1 \rangle} |e_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix} - \left(\frac{9}{12}\right) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix} - (1) \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} \\ \frac{5}{4} \\ -\frac{7}{4} \\ -\frac{1}{4} \end{pmatrix}.$$

Y la base ortonormal asociada será:

$$|\hat{e}_1\rangle = \frac{|e_1\rangle}{\sqrt{\langle e^1 | e_1 \rangle}} = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}; |\hat{e}_2\rangle = \frac{|e_2\rangle}{\sqrt{\langle e^2 | e_2 \rangle}} = \frac{\sqrt{3}}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}; |\hat{e}_3\rangle = \frac{|e_3\rangle}{\sqrt{\langle e^3 | e_3 \rangle}} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -\frac{7}{5} \\ -\frac{1}{5} \end{pmatrix}.$$

En este ejemplo hemos mostrado que $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ son linealmente independientes y, por lo tanto, base de un subespacio de V^4 . Cabría preguntarse ¿Cómo construimos un cuarto vector, linealmente independiente a los anteriores, que expanda todo V^4 ?

3. Suponga el espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ definidos en el intervalo $[-1, 1]$. Este espacio vectorial tendrá como una de las posibles bases al conjunto $\{|\pi_i\rangle\} = \{1, t, t^2, t^3, \dots, t^n\}$ con el producto interno definido por:²⁰

$$\langle \pi^i | \pi_j \rangle = \int_{-1}^1 dt \pi_i(t) \pi_j(t).$$

Procederemos a construir una base ortogonal $\{|P_i\rangle\}$, y tomaremos como vector de inicio a $|\pi_0\rangle$:

$$|P_0\rangle \equiv |\pi_0\rangle = 1.$$

El siguiente vector será:

$$|P_1\rangle \equiv |\pi_1\rangle - \frac{\langle \pi^1 | P_0 \rangle}{\langle P^0 | P_0 \rangle} |P_0\rangle = t \Leftarrow \begin{cases} \langle \pi^1 | P_0 \rangle = \int_{-1}^1 dt t = 0 \\ \langle P^0 | P_0 \rangle = \int_{-1}^1 dt = 2 \end{cases}$$

²⁰En este punto es importante señalar la importancia de la definición de producto interno. Si esta definición hubiera sido $\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x) \sqrt{1-x^2}$ o $\langle f | g \rangle = \int_{-1}^1 dx \frac{f(x) g(x)}{\sqrt{1-x^2}}$ las bases correspondientes a estas definiciones de producto interno serían distintas.

El siguiente:

$$|P_2\rangle \equiv |\pi_2\rangle - \frac{\langle \pi^2 | P_1 \rangle}{\langle P^1 | P_1 \rangle} |P_1\rangle - \frac{\langle \pi^2 | P_0 \rangle}{\langle P^0 | P_0 \rangle} |P_0\rangle = t^2 - \frac{1}{3} \Leftrightarrow \begin{cases} \langle \pi^2 | P_0 \rangle = \int_{-1}^1 dt t^2 = \frac{2}{3} \\ \langle \pi^2 | P_1 \rangle = \int_{-1}^1 dt t^3 = 0 \\ \langle P^1 | P_1 \rangle = \int_{-1}^1 dt t^2 = \frac{2}{3} \end{cases}$$

Para el cuarto:

$$|P_3\rangle \equiv |\pi_3\rangle - \frac{\langle \pi^3 | P_2 \rangle}{\langle P^2 | P_2 \rangle} |P_2\rangle - \frac{\langle \pi^3 | P_1 \rangle}{\langle P^1 | P_1 \rangle} |P_1\rangle - \frac{\langle \pi^3 | P_0 \rangle}{\langle P^0 | P_0 \rangle} |P_0\rangle = t^3 - \frac{3}{5}t$$

$$\begin{cases} \langle \pi^3 | P_0 \rangle = \int_{-1}^1 dt t^3 = 0, & \langle \pi^3 | P_1 \rangle = \int_{-1}^1 dt t^4 = \frac{2}{5} \\ \langle \pi^3 | P_2 \rangle = \int_{-1}^1 dt t^3 [t^2 - \frac{1}{3}] = 0, & \langle P^2 | P_2 \rangle = \int_{-1}^1 dt [t^2 - \frac{1}{3}]^2 = \frac{8}{45} \end{cases}$$

Queda como ejercicio comprobar la ortogonalidad de los vectores recién calculados:

$$\langle P^0 | P_1 \rangle = \langle P^1 | P_2 \rangle = \langle P^2 | P_3 \rangle = \dots = 0.$$

Esta base ortogonal está formada por los polinomios de Legendre, discutidos en la sección 2.3.4. Es decir, si ortogonalizamos una base de monomios $\{|\pi_i\rangle\}$ mediante la definición de producto interno: $\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x)$, obtendremos la base de polinomios ortogonales de Legendre.

Podemos resumir los cálculos anteriores, construyendo también la base ortonormal a partir de los monomios $\{|\pi_i\rangle\}$ como se muestra a continuación:

$ \pi_n\rangle$	$ P_n\rangle$	\hat{P}_n
1	1	$\frac{1}{\sqrt{2}}$
t	t	$\sqrt{\frac{3}{2}} t$
t^2	$t^2 - \frac{1}{3}$	$\frac{1}{2} \sqrt{\frac{5}{2}} (3t^2 - 1)$
t^3	$t^3 - \frac{3}{5}t$	$\frac{1}{2} \sqrt{\frac{7}{2}} (5t^3 - 3t)$
t^4	$t^4 - \frac{6}{7}t^2 + \frac{3}{35}$	$\frac{3}{8} \sqrt{\frac{1}{2}} (35t^4 - 30t^2 + 3)$
\vdots	\vdots	\vdots

◀

2.3.6 Ejercicios

- Diga si los siguientes conjuntos de vectores en \mathcal{P}^3 son o no linealmente independientes.
 - $|x_1\rangle = 2x$, $|x_2\rangle = x^2 + 1$, $|x_3\rangle = x + 1$, $|x_4\rangle = x^2 - 1$.
 - $|x_1\rangle = x(x - 1)$, $|x_2\rangle = x$, $|x_3\rangle = x^3$, $|x_4\rangle = 2x^3 - x^2$.
- Probar que los polinomios: $|x_1\rangle = 1$, $|x_2\rangle = x$, $|x_3\rangle = \frac{3}{2}x^2 - \frac{1}{2}$ y $|x_4\rangle = \frac{5}{2}x^3 - \frac{3}{2}x$, forman una base en \mathcal{P}_4 . Expresar $|p\rangle = x^2$, $|q\rangle = x^3$ en función de esa base.
- Encontrar la proyección perpendicular de los siguientes vectores en $\mathcal{C}_{[-1,1]}$ (espacio de funciones continuas en el intervalo $[-1,1]$ y con el producto interno definido por: $\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x)$) al subespacio generado por los polinomios: $\{1, x, x^2 - 1\}$. Calcular la distancia de cada una de estas funciones al subespacio mencionado.
 - $f(x) = x^n$, n entero.
 - $f(x) = \sin(x)$.

- (c). $f(x) = 3x^2$.
4. Utilizando el método de Gram-Schmidt encuentre una base ortonormal para los siguientes conjuntos de vectores:
- (a). $|v_1\rangle = (1, 0, 1)$, $|v_2\rangle = (0, 1, 1)$ y $|v_3\rangle = (1, 0, 0)$. En \mathbb{R}^3 .
- (b). $|v_1\rangle = (2, -4, 5, 2)$, $|v_2\rangle = (-6, 5, -1, -6)$ y $|v_3\rangle = (-10, 13, -4, -3)$. En \mathbb{R}^4 .
5. Considere el espacio vectorial de las matrices complejas 2×2 hermiticas. Tal y como demostraremos con rigor en la sección 4.3.2.2 y lo detallamos en la sección 4.4.9, una matriz hermitica (o autoadjunta) será igual a su adjunta. Esto es, una matriz será igual a su traspuesta conjugada $(A^\dagger)_j^i \rightarrow (A^*)_i^j \equiv A_i^j$:

$$\mathbb{A} \leftrightarrow \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} = \mathbb{A}^\dagger = \begin{pmatrix} z_1^* & z_3^* \\ z_2^* & z_4^* \end{pmatrix} \quad \text{es decir} \quad \begin{cases} z_1^* = z_1 & \text{real} \\ z_4^* = z_4 & \text{real} \\ z_2^* = z_3 & \text{complejos} \end{cases}.$$

Entonces

- (a). Muestre que las matrices de Pauli $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ presentadas en los ejercicios de la sección 2.2.4 forman una base para ese espacio vectorial.
- (b). Compruebe que esa base es ortogonal bajo la definición de producto interno $\langle a | b \rangle \equiv \text{Tr}(\mathbb{A}^\dagger \mathbb{B})$ que introdujimos en los ejercicios de esa misma sección.
- (c). Explore si se pueden construir subespacios vectoriales de matrices reales e imaginarias puras.
6. Utilizando **Maxima** reproduzca el ejemplo 2. que expusimos en la página 120. Es decir, suponga el espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ definidos en el intervalo $[-1, 1]$. Este espacio vectorial tendrá como una de las posibles bases al conjunto $\{|\pi_i\rangle\} = \{1, t, t^2, t^3, \dots, t^n\}$, pero en este caso con el producto interno definido por:

$$\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x) \sqrt{1-x^2}.$$

Encuentre la base ortogonal correspondiente. A esta nueva base se le conoce como polinomios de Chebyshev de segunda especie²¹.

7. Otra vez, utilizando **Maxima**, reproduzca el ejercicio anterior, pero con la definición de producto interno:

$$\langle f | g \rangle = \int_{-1}^1 dx \frac{f(x) g(x)}{\sqrt{1-x^2}}.$$

A esta nueva base se le conoce como polinomios de Chebyshev de primera especie.

Practicando con Maxima

1. **Independencia Lineal:** En 2.3.5 vimos que si en la ecuación

$$|0\rangle = C_1 |v_1\rangle + C_2 |v_2\rangle + C_3 |v_3\rangle \cdots + C_n |v_n\rangle$$

con todos los $C_i = 0$, entonces se dirá que el conjunto de vectores son linealmente independientes. Para el segundo ejemplo de esa sección (ejemplo 2.) se obtuvo el siguiente sistema de ecuaciones:

$$\begin{array}{rrcr} C_1 & +2C_2 & -C_3 & = 0 \\ 3C_1 & & +C_3 & = 0 \\ -C_1 & +C_2 & & = 0 \\ 2C_1 & +3C_2 & & = 0 \end{array}$$

que procedemos a resolver usando el comando **linsolve**:

²¹https://en.wikipedia.org/wiki/Chebyshev_polynomials

```
(%i1) linsolve([C1+2*C2-C3=0, 3*C1+C3=0, -C1+C2=0, 2*C1+3*C2=0], [C1,C2,C3]);
solve: dependent equations eliminated: (4)
```

```
(%o1) [C1 = 0, C2 = 0, C3 = 0]
```

2. **Bases para espacios vectoriales:** En este ejercicio aprenderemos a calcular una base a partir de un conjunto de vectores perteneciente a un determinado espacio vectorial. Por ejemplo, si en \mathbb{R}^5 tenemos el siguiente conjunto de vectores:

$$|v_1\rangle = (1, 2, 3, 4, 5), |v_2\rangle = (0, -1, 1, 2, 3), |v_3\rangle = (3, 2, 1, 0, -1), |v_4\rangle = (-4, -3, -2, -1, 0).$$

```
(%i1) v1:[1,2,3,4,5]; v2:[0,-1,1,2,3]; v3:[3,2,1,0,-1]; v4:[-4,-3,-2,-1,0];
```

```
(%o1) [1, 2, 3, 4, 5]
```

```
(%o2) [0, -1, 1, 2, 3]
```

```
(%o3) [3, 2, 1, 0, -1]
```

```
(%o4) [-4, -3, -2, -1, 0]
```

Con los vectores dados construimos la siguiente matriz

```
(%i5) M:matrix(v1,v2,v3,v4);
```

```
(%o5) (1 2 3 4 5)
      (0 -1 1 2 3)
      (3 2 1 0 -1)
      (-4 -3 -2 -1 0)
```

Como veremos más adelante, el Rango de una matriz indica el número máximo de vectores fila o columna linealmente independientes.

```
(%i6) rank(M);
```

```
(%o6) 3
```

Podemos aplicar el método de eliminación gaussiana a la matriz M para obtener una nueva matriz escalonada. El cálculo además se hace normalizando el primer elemento no nulo de cada fila. Esto se logra con el comando **echelon(M)**.

```
(%i7) echelon(M);
```

```
(%o7) (1 2/3 1/3 0 -1/3)
      (0 1 -1 -2 -3)
      (0 0 1 5/3 7/3)
      (0 0 0 0 0)
```

Por lo tanto, cada fila de la matriz anterior conformará un conjunto de vectores linealmente independiente.

Los podemos aislar de la matriz con el comando **row(M,i)**, que devolverá la i -ésima fila de la matriz M .

```
(%i8) e1:row(%o7,1);
```

```
(%o8) (1 2/3 1/3 0 -1/3)
```

```
(%i9) e2:row(%o7,2);
```

```
(%o9) (0 1 -1 -2 -3)
```

```
(%i10) e3:row(%o7,3);
```

$$(\%o10) \begin{pmatrix} 0 & 0 & 1 & \frac{5}{3} & \frac{7}{3} \end{pmatrix}$$

Es trivial verificar que el conjunto: $\{e_1, e_2, e_3\}$ es linealmente independientes.

(%i11) solve(alpha*[1,2/3,1/3,0,-1/3]+beta*[0,1,-1,-2,-3]+gamma*[0,0,5/3,7/3],[alpha,beta,gamma]);
solve: dependent equations eliminated: (3)

(%o11) $[[\alpha = 0, \beta = 0, \gamma = 0]]$

Consideremos otros ejemplos, dados los vectores $\mathbf{a} = (1, 3)$ y $\mathbf{b} = (-1, 1)$ ¿Serán linealmente independientes?

(%i12) solve(alpha*[1,3]+beta*[-1,1],[alpha,beta]);

(%o12) $[[\alpha = 0, \beta = 0]]$

Los vectores $\mathbf{a} = (1, 2, 3)$ y $\mathbf{b} = (4, 8, 12)$ ¿Serán linealmente independientes?

(%i13) solve(alpha*[1,2,3]+beta*[4,8,12],[alpha,beta]);

solve: dependent equations eliminated: (2 3)

(%o13) $[[\alpha = -4 \%r_4, \beta = \%r_4]]$

Aquí $\alpha = -4\beta$.

Sea ahora $\{\mathbf{e}_i\} = \{(1, 1, 1), (1, 2, 1), (0, 0, 2)\}$ una base para \mathbb{R}^3 . Vamos a calcular las componentes del vector $\mathbf{a} = (3, 2, 1)$ respecto de esa base.

Primero verifiquemos si efectivamente es una base:

(%i14) solve(alpha*[1,1,1]+beta*[1,2,1]+gamma*[0,0,2],[alpha,beta,gamma]);

(%o14) $[[\alpha = 0, \beta = 0, \gamma = 0]]$

(%i15) solve([3,2,1]-ax*[1,1,1]-ay*[1,2,1]-az*[0,0,2],[ax,ay,az]);

(%o15) $[[ax = 4, ay = -1, az = -1]]$

Por lo tanto, $\mathbf{a} = 4\mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3$.

En la base canónica las componentes del vector \mathbf{a} serán:

(%i16) solve([3,2,1]-ax*[1,0,0]-ay*[0,1,0]-az*[0,0,1],[ax,ay,az]);

(%o16) $[[ax = 3, ay = 2, az = 1]]$

Es decir, $\mathbf{a} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$.

(%i17) kill(all)\$

3. Ortogonalización con Maxima Anteriormente hicimos los cálculos para hallar una base ortogonal a partir del siguiente conjunto de vectores linealmente independientes:

$$|v_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}; \quad |v_2\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}; \quad |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

La función **gramschmidt(x)** es la que ejecuta el algoritmo de orthogonalización de Gram-Schmidt sobre

el objeto x . Si x es una matriz el algoritmo actúa sobre sus filas. Antes de usar la función es necesario cargar la biblioteca **eigen**.

```
(%i1) load("eigen")$
(%i2) v:matrix([-1, 1, 0, 0], [2, 0, 1, 3], [1, 3, -1, 2]);
(%o2) 
$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 3 \\ 1 & 3 & -1 & 2 \end{pmatrix}$$

```

Notemos que el vector $|v_3\rangle$ es el que hemos puesto en la primera fila de la matriz. Ahora procedemos al cálculo:

```
(%i3) e:gramschmidt(v);
(%o3) 
$$\left[ [-1, 1, 0, 0], [1, 1, 1, 3], \left[ \frac{5}{2^2}, \frac{5}{2^2}, -\frac{7}{2^2}, -\frac{1}{2^2} \right] \right]$$

```

El resultado viene en forma de una lista: $[e[1], e[2], e[3]]$. Simplificando resulta:

```
(%i4) e:expand(%);
(%o4) 
$$\left[ [-1, 1, 0, 0], [1, 1, 1, 3], \left[ \frac{5}{4}, \frac{5}{4}, -\frac{7}{4}, -\frac{1}{4} \right] \right]$$

```

Podemos verificar que son ortogonales:

```
(%i5) map(innerproduct, [e[1], e[2], e[3]], [e[2], e[3], e[1]]);
(%o5) [0, 0, 0]
```

Normalizamos ahora cada uno de los vectores:

```
(%i6) e[1]/sqrt(innerproduct(e[1],e[1]));
(%o6) 
$$\left[ -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right]$$

(%i7) e[2]/sqrt(innerproduct(e[2],e[2]));
(%o7) 
$$\left[ \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{\sqrt{3}}{2} \right]$$

(%i8) e[3]/sqrt(innerproduct(e[3],e[3]));
(%o8) 
$$\left[ \frac{1}{2}, \frac{1}{2}, -\frac{7}{10}, -\frac{1}{10} \right]$$

```

La función **gramschmidt**(x,F) nos permite definir un producto interno diferente a **innerproduct**. Veamos como se hace para el otro ejemplo donde el conjunto de vectores linealmente independientes estaba dado por: $\{1, t, t^2, t^3, \dots, t^n\}$.

```
(%i9) producto(f,g):=integrate(f*g, t, a, b);
(%o9) 
$$\text{producto}(f,g) := \int_a^b f g \, dt$$

(%i10) e:gramschmidt([1, t, t^2, t^3, t^4], producto), a=-1, b=1;
(%o10) 
$$\left[ 1, t, \frac{3t^2-1}{3}, \frac{t(5t^2-3)}{5}, \frac{35t^4-30t^2+3}{35} \right]$$

(%i11) e:expand(e);
```

$$(\%o11) \left[1, t, t^2 - \frac{1}{3}, t^3 - \frac{3t}{5}, t^4 - \frac{6t^2}{7} + \frac{3}{35} \right]$$

Verifiquemos si son ortogonales:

(%i12) map(producto, [e[1], e[2], e[3]], [e[2], e[3], e[1]]), a=-1, b=1;

(%o12) [0, 0, 0]

Para normalizar cada uno de los vectores procedemos como en el caso anterior pero primero hacemos $a = -1$ y $b = 1$.

(%i13) a: -1 b: 1

(%i14) e[1]/sqrt(producto(e[1], e[1]));

(%o14) $\frac{1}{\sqrt{2}}$

(%i15) e[2]/sqrt(producto(e[2], e[2]));

(%o15) $\frac{\sqrt{3}t}{\sqrt{2}}$

(%i16) e[3]/sqrt(producto(e[3], e[3]));

(%o16) $\frac{3\sqrt{5}\left(t^2 - \frac{1}{3}\right)}{2^{\frac{3}{2}}}$

(%i17) e[4]/sqrt(producto(e[4], e[4]));

(%o17) $\frac{5\sqrt{7}\left(t^3 - \frac{3t}{5}\right)}{2^{\frac{3}{2}}}$

2.4 Aproximación de funciones

Armados de los conocimientos previos podemos pasar a aplicarlos en un intento de aproximar funciones. La aproximación de una función tiene varias facetas y seguramente en cursos anteriores hemos hecho este tipo de aproximaciones una buena cantidad de veces cuando necesitábamos convertir una expresión, que nos resultaba muy complicada, en otras más sencillas y casi equivalentes. Por ejemplo, cuando aplicamos la aproximación lineal: $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$, con x muy cercano a x_0 ¿Cuántas veces no hemos utilizado la aproximación: $\sin(x) \approx x$? También aproximamos funciones cuando en el laboratorio nos veíamos en la necesidad de “ajustar con la mejor curva” una serie de puntos experimentales.

Consideremos primero algunos conceptos importantes.

2.4.1 Complementos ortogonales y descomposición ortogonal

Volvamos a retomar la idea de los subespacios y su complemento ortogonal, ejercicio 3, de la sección 2.2. Sea un subespacio $S \subset V$. Un elemento $|\bar{v}_i\rangle \in V$ se dice ortogonal a S si $\langle s^k | \bar{v}_i \rangle = 0 \forall |s_k\rangle \in S$, es decir, es ortogonal a todos los elementos de S . El conjunto $\{|\bar{v}_1\rangle, |\bar{v}_2\rangle, |\bar{v}_3\rangle, \dots, |\bar{v}_m\rangle\}$ de todos los elementos ortogonales a S , se denomina S -perpendicular y se denota como S^\perp . Es fácil demostrar que S^\perp es un subespacio, aún si S no lo es.

Dado un espacio euclidiano de dimensión infinita $V : \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle, \dots\}$ y un subespacio $S \subset V$ con dimensión finita: $\dim S = m$. Entonces, $\forall |v_k\rangle \in V$ puede expresarse como la suma de dos vectores $|s_k\rangle \in S \wedge |s_k\rangle^\perp \in S^\perp$. Esto es:

$$|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp, \quad |s_k\rangle \in S \quad \wedge \quad |s_k\rangle^\perp \in S^\perp.$$

La norma de $|v_k\rangle$ se calcula a través del teorema de Pitágoras generalizado:

$$\| |v_k\rangle \|^2 = \| |s_k\rangle \|^2 + \| |s_k\rangle^\perp \|^2.$$

La demostración es sencilla. Primero se prueba que la descomposición ortogonal $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es siempre posible. Para ello recordamos que $S \subset V$ es de dimensión finita, por lo tanto existe una base ortonormal $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_m\rangle\}$ para S . Es decir, dado un $|v_k\rangle$ definimos los elementos $|s_k\rangle$ y $|s_k\rangle^\perp$ como sigue:

$$|s_k\rangle = \sum_{i=1}^m \langle v_k | \hat{e}_i \rangle |\hat{e}_i\rangle \quad \wedge \quad |s_k\rangle^\perp = |v_k\rangle - |s_k\rangle. \quad (2.16)$$

Nótese que $\langle v_k | \hat{e}_i \rangle |\hat{e}_i\rangle$ es la proyección de $|v_k\rangle$ a lo largo de $|\hat{e}_i\rangle$ y $|s_k\rangle$ se expresa como la combinación lineal de la base de S , por lo tanto, está en S . Por otro lado:

$${}^\perp \langle s^k | \hat{e}_i \rangle = \langle v^k - s^k | \hat{e}_i \rangle = \langle v^k | \hat{e}_i \rangle - \langle s^k | \hat{e}_i \rangle = \langle v^k | \hat{e}_i \rangle - \left[\sum_{j=1}^m \langle v^k | \hat{e}_j \rangle \langle \hat{e}^j | \right] |\hat{e}_i\rangle = 0 \Rightarrow |s_k\rangle^\perp \perp |\hat{e}_j\rangle,$$

lo cual indica que $|s_k\rangle^\perp \in S^\perp$.

Podemos ir un poco más allá y ver que la descomposición $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es única en V . Para ello suponemos que existen dos posibles descomposiciones, vale decir:

$$|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp \quad \wedge \quad |v_k\rangle = |t_k\rangle + |t_k\rangle^\perp, \quad \text{con } |s_k\rangle \wedge |t_k\rangle \in S \quad \wedge \quad |s_k\rangle^\perp \wedge |t_k\rangle^\perp \in S^\perp.$$

Por lo tanto:

$$|v_k\rangle - |v_k\rangle = (|s_k\rangle + |s_k\rangle^\perp) - (|t_k\rangle + |t_k\rangle^\perp) = 0 \Rightarrow |s_k\rangle - |t_k\rangle = |t_k\rangle^\perp - |s_k\rangle^\perp.$$

Pero $|s_k\rangle - |t_k\rangle \in S$, es decir, ortogonal a todos los elementos de S^\perp y $|s_k\rangle - |t_k\rangle = |t_k\rangle^\perp - |s_k\rangle^\perp$.

Con lo cual $|s_k\rangle - |t_k\rangle \equiv |0\rangle$, que es el único elemento que es ortogonal a sí mismo y en consecuencia la descomposición $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es única.

Finalmente, con la definición de norma:

$$\| |v_k\rangle \|^2 = \| |s_k\rangle + |s_k\rangle^\perp \|^2 = (\langle s^k | + \langle s^k |^\perp) (|s_k\rangle + |s_k\rangle^\perp) = \langle s^k | s_k \rangle + \langle s^k |^\perp s_k \rangle^\perp = \| |s_k\rangle \|^2 + \| |s_k\rangle^\perp \|^2.$$

Así, dado S^m un subespacio de V de dimensión finita y dado un $|v_k\rangle \in V$ el elemento:

$$|s_k\rangle \in S \Rightarrow |s_k\rangle = \sum_{i=1}^m \langle v_k | e_i \rangle |e_i\rangle,$$

será la proyección de $|v_k\rangle$ en S .

En general, dado un vector $|x\rangle \in V$ y un subespacio de V con dimensión finita, $S^m \subset V$ y $\dim S = m$, entonces la distancia de $|x\rangle$ a S^m es la norma de la componente de $|x\rangle$, perpendicular a S^m . Más aún, esa distancia será mínima y $|x\rangle_{S^m}$ la proyección de $|x\rangle$, en S^m será el elemento de S^m más próximo a $|x\rangle$ y, por la mejor aproximación.

Ejemplo 2.16 Sea el espacio vectorial $V : \mathbb{R}^3$ con los vectores base: $|x_1\rangle = \hat{i}$, $|x_2\rangle = \hat{k}$ y $|x_3\rangle = \hat{j} + \hat{k}$ y consideramos el subespacio $S : \{x + y = 0\} \subset \mathbb{R}^3$. Podemos calcular una base en S parametrizándola de la manera siguiente:

$$x + y = 0 \Rightarrow \begin{cases} x = -\mu \\ y = \mu \\ z = \nu \end{cases}$$

Por lo que una base en S puede ser: $|e_1\rangle = -\hat{i} + \hat{j}$ y $|e_2\rangle = \hat{k}$. Si tomamos un vector $|x\rangle \in \mathbb{R}^3$ con coordenadas (x, y, z) , para que éste vector $|x\rangle \in S^\perp$ se tiene que cumplir que: $\langle x | e_k \rangle = 0$. Podemos utilizar la matriz de Gram, que conocemos para la base dada, para calcular el producto escalar a través de las ecuaciones:

$$\begin{aligned} \langle x | e_1 \rangle &= (x, y, z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0 \Rightarrow -x + y + z = 0 \\ \langle x | e_2 \rangle &= (x, y, z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \Rightarrow y + 2z = 0 \end{aligned}$$

El par de ecuaciones: $-x + y + z = 0$ y $y + 2z = 0$ vienen a ser representación cartesiana de S^\perp . La matriz de los coeficientes de ese sistema de ecuaciones es:

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow \begin{cases} x + z = 0 \\ y + 2z = 0 \end{cases}$$

por lo que una base para esa representación cartesiana puede ser:

$$\begin{cases} x + z = 0 \\ y + 2z = 0 \end{cases} \Rightarrow \begin{cases} x = -\mu \\ y = -2\mu \\ z = \mu \end{cases},$$

es decir, $S^\perp : \{(-1, -2, 1)\}$.

Notemos que partimos de un espacio vectorial V de dimensión 3, la dimensión de S es 2 y por lo tanto, la de S^\perp es 1.

El lector debe comprobar que efectivamente $\{(-1, -2, 1)\}$ es ortogonal a cada uno de los vectores base de $S : \{(-1, 1, 0), (0, 0, 1)\}$.



2.4.2 Condiciones para la aproximación de funciones

Sea $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle, \dots\}$ un espacio euclidiano de dimensión infinita, y un subespacio $S^m \subset V$, con dimensión finita, $\dim S = m$, y sea un elemento $|v_i\rangle \in V$. La proyección de $|v_i\rangle$ en S^m , $|s_i\rangle$, será el elemento de S^m más próximo a $|v_i\rangle$. En otras palabras:

$$\| |v_i\rangle - |s_i\rangle \| \leq \| |v_i\rangle - |t_i\rangle \| \quad \forall |t_i\rangle \in S.$$

La demostración se sigue así:

$$|v_i\rangle - |t_i\rangle = (|v_i\rangle - |s_i\rangle) + (|s_i\rangle - |t_i\rangle) \Rightarrow \| |v_i\rangle - |t_i\rangle \|^2 = \| |v_i\rangle - |s_i\rangle \|^2 + \| |s_i\rangle - |t_i\rangle \|^2,$$

ya que $|v_i\rangle - |s_i\rangle = |s_k\rangle^\perp \in S^\perp \wedge |s_i\rangle - |t_i\rangle \in S$, y vale el teorema de Pitágoras generalizado.

Ahora bien, como:

$$\| |s_i\rangle - |t_i\rangle \|^2 \geq 0 \Rightarrow \| |v_i\rangle - |t_i\rangle \|^2 \geq \| |v_i\rangle - |s_i\rangle \|^2 \Rightarrow \| |v_i\rangle - |t_i\rangle \| \geq \| |v_i\rangle - |s_i\rangle \|.$$

En la sección 2.3.4 consideramos la expansión de funciones continuas, reales de variable real, definidas en $[0, 2\pi]$, $C_{[0, 2\pi]}^\infty$, mediante funciones trigonométricas y con el producto interno definido por:

$$\langle f | g \rangle = \int_0^{2\pi} dx f(x) g(x).$$

En ese entonces consideramos, para ese espacio vectorial, una base ortogonal definida por:

$$|e_0\rangle = 1, \quad |e_{2n-1}\rangle = \cos(nx) \quad y \quad |e_{2n}\rangle = \sin(nx), \quad \text{con } n = 1, 2, 3, \dots$$

Por lo tanto, cualquier función definida en el intervalo $[0, 2\pi]$ puede expresarse en términos de esta base como mostramos a continuación:

$$|f\rangle = \sum_{i=1}^{\infty} C^i |e_i\rangle.$$

Los C^i son los coeficientes de Fourier. Es decir, cualquier función puede ser expresada como una serie de Fourier de la forma:

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)],$$

donde:

$$a_k = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \cos(kx) \quad \wedge \quad b_k = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \sin(kx).$$

Es claro que para la aproximación de funciones por funciones trigonométricas cuyos coeficientes son los coeficientes de Fourier constituyen la mejor aproximación. Por lo tanto, de todas las funciones $\mathcal{F}(x) \in C_{[0, 2\pi]}^\infty$ las funciones trigonométricas, $\mathcal{T}(x)$ minimizan la desviación cuadrática media:

$$\int_0^{2\pi} dx (f(x) - \mathcal{P}(x))^2 \geq \int_0^{2\pi} dx (f(x) - \mathcal{T}(x))^2.$$

2.4.3 El método de mínimos cuadrados

Una de las aplicaciones más importantes en la aproximación de funciones es el método de mínimos cuadrados. La idea es determinar el valor más aproximado de una cantidad física, c , a partir de un conjunto de medidas experimentales: $\{x_1, x_2, x_3, \dots, x_n\}$. Para ello asociamos ese conjunto de medidas con las componentes de un vector $|x\rangle \equiv (x_1, x_2, x_3, \dots, x_n)$ en \mathbb{R}^n y supondremos que su mejor aproximación, que llamaremos $c|1\rangle \equiv (c, c, c, \dots, c)$, será la proyección perpendicular de $|x\rangle$ (las medidas) sobre el subespacio generado por

$|1\rangle$. Esto es:

$$|x\rangle = c|1\rangle \Rightarrow c = \frac{\langle x|1\rangle}{\langle 1|1\rangle} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n},$$

que no es otra cosa que construir -de una manera sofisticada- el promedio aritmético de las medidas. Es claro que la proyección perpendicular de $|x\rangle$ sobre $|1\rangle$ hace mínima la distancia entre el subespacio perpendicular generado por $|1\rangle$ y el vector $|x\rangle$, con ello también hará mínimo su cuadrado:

$$[d(|x\rangle, c|1\rangle)]^2 = \|x - c\|^2 = \langle x - c | x - c \rangle = \sum_{i=1}^n (x_i - c)^2.$$

Esta manera sofisticada, que se deriva del formalismo utilizado, muestra el significado del promedio aritmético como medida más cercana al valor “real” de una cantidad obtenida a partir de una serie de medidas experimentales.

Obviamente, este problema se puede generalizar para el caso de dos (o n) cantidades si extendemos la dimensión del espacio y los resultados experimentales se expresarán como un vector de $2n$ dimensiones

$$|x\rangle = (x_{11}, x_{12}, x_{13}, \cdots, x_{1n}, x_{21}, x_{22}, x_{23}, \cdots, x_{2n}),$$

mientras que los vectores que representan las cantidades más aproximadas serán:

$$c_1|1_1\rangle = \left(\underbrace{c_1, c_1, c_1, \cdots, c_1}_n, \underbrace{0, 0, 0, \cdots, 0}_n \right) \quad \wedge \quad c_2|1_2\rangle = (0, 0, 0, \cdots, 0, c_2, c_2, c_2, \cdots, c_2).$$

Ahora $\{|1_1\rangle, |1_2\rangle\}$ expanden un subespacio vectorial sobre el cual $|x\rangle$ tiene como proyección ortogonal a: $c_1|1_1\rangle + c_2|1_2\rangle$ y consecuentemente $|x - c_1|1_1\rangle - c_2|1_2\rangle$ será perpendicular a $\{|1_1\rangle, |1_2\rangle\}$. Por lo tanto:

$$c_1 = \frac{\langle x|1_1\rangle}{\langle 1^1|1_1\rangle} = \frac{x_{11} + x_{12} + x_{13} + \cdots + x_{1n}}{n} \quad \wedge \quad c_2 = \frac{\langle x|1_2\rangle}{\langle 1^2|1_2\rangle} = \frac{x_{21} + x_{22} + x_{23} + \cdots + x_{2n}}{n}.$$

Quizá la consecuencia más conocida de esta forma de aproximar funciones es el “ajuste” de un conjunto de datos experimentales $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \cdots, (x_n, y_n)\}$ a la ecuación de una recta $y = cx$. En este caso, el planteamiento del problema se reduce a encontrar el vector $c|x\rangle$, en el subespacio $S(|x\rangle)$, que esté lo más cercano posible al vector $|y\rangle$.

Como en el caso anterior, queremos que la distancia entre $|y\rangle$ y su valor más aproximado $c|x\rangle$ sea la menor posible. Por lo tanto, $\|cx - y\|^2$ será la menor posible y $|cx - y\rangle$ será perpendicular a $S(|x\rangle)$, por lo que:

$$\langle x | cx - y \rangle = 0 \Rightarrow c = \frac{\langle x | y \rangle}{\langle x | x \rangle} = \frac{x_1y_1 + x_2y_2 + x_3y_3 + \cdots + x_ny_n}{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}.$$

Si la recta a “ajustar” es $y = cx + b$ el procedimiento será uno equivalente: proyectar sobre los vectores y obtener ecuaciones. Si representamos $|b\rangle = b|1\rangle$, tendremos:

$$|y\rangle = c|x\rangle + |b\rangle \Rightarrow \begin{cases} \langle x | y \rangle = c \langle x | x \rangle + \langle x | b \rangle \Rightarrow \sum_{i=1}^n x_i y_i = c \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \\ \langle b | y \rangle = c \langle b | x \rangle + \langle b | b \rangle \Rightarrow \sum_{i=1}^n y_i = c \sum_{i=1}^n x_i + bn \end{cases}$$

Que no es otra cosa que un sistema de dos ecuaciones con dos incógnitas: c y b . Por simplicidad y conveniencia continuaremos con el análisis del caso $b = 0$, vale decir, aproximando las medidas experimentales a una recta que pase por el origen de coordenadas $y = cx$. El lector podrá extender el procedimiento para caso $b \neq 0$.

Para tratar de entender (y extender) lo antes expuesto, consideremos tres ejemplos que muestran la versatilidad del método y la ventaja de disponer de una clara notación. Primeramente, mostraremos el caso más utilizado para construir el mejor ajuste lineal a un conjunto de datos experimentales. Buscaremos la mejor recta que describe ese conjunto de puntos. Luego mostraremos la aplicación del método para buscar la mejor

función multilineal, es decir, que ajustaremos la mejor función de n variables con una contribución lineal de sus argumentos: $y = y(x_1, x_2, \dots, x_n) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$. Este caso tendrá una complicación adicional, por cuando realizaremos varias mediciones de las variables. En los ejemplos de la sección 2.18 analizaremos varios casos en los cuales extendemos el método.

2.4.4 Interpolación polinomial de puntos experimentales

Muchas veces nos encontramos con la situación en la cual tenemos un conjunto de (digamos n) medidas o puntos experimentales $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ y para modelar ese experimento quisiéramos una función que ajuste estos puntos, de manera que: $\{(x_1, y_1 = f(x_1)), (x_2, y_2 = f(x_2)), \dots, (x_n, y_n = f(x_n))\}$.

El tener una función nos provee la gran ventaja de poder intuir o aproximar los puntos que no hemos medido. La función candidata más inmediata es un polinomio y debemos definir el grado del polinomio y la estrategia que aproxime esos puntos. Puede ser que el polinomio no sea lineal y entonces queramos ajustar esos puntos a un polinomio tal que éste pase por los puntos experimentales.

Queda entonces por decidir la estrategia, esto es, construimos la función como “trozos” de polinomios que ajusten a subconjuntos: $\{(x_1, y_1 = f(x_1)), (x_2, y_2 = f(x_2)), \dots, (x_m, y_m = f(x_m))\}$ con $m < n$ de los puntos experimentales. En este caso tendremos una función de interpolación para cada conjunto de puntos. También podemos ajustar la función a todo el conjunto de puntos experimentales de manera que el máximo grado del polinomio que los interpole será de grado $n - 1$.

Para encontrar este polinomio, y a manera de ejemplo, lo expresaremos como una combinación lineal de polinomios de Legendre, una base ortogonal que hemos discutido con anterioridad en la secciones 2.3.4 y 2.8. Esto es:

$$f(x) = \sum_{k=0}^{n-1} C^k |P_k\rangle = \sum_{k=0}^{n-1} C^k P_k(x) \Rightarrow \begin{cases} y_1 = f(x_1) = C^0 P_0(x_1) + C^1 P_1(x_1) + \dots + C^{n-1} P_{n-1}(x_1) \\ y_2 = f(x_2) = C^0 P_0(x_2) + C^1 P_1(x_2) + \dots + C^{n-1} P_{n-1}(x_2) \\ \vdots \\ y_n = f(x_n) = C^0 P_0(x_n) + C^1 P_1(x_n) + \dots + C^{n-1} P_{n-1}(x_n) \end{cases}$$

que no es otra cosa que un sistema de n ecuaciones con n incógnitas: los coeficientes $\{C^0, C^1, \dots, C^{n-1}\}$.

Al resolver el sistema de ecuaciones y obtener los coeficientes, podremos obtener la función polinómica que interpola esos puntos. Una expansión equivalente se pudo haber logrado con cualquier otro conjunto de polinomios ortogonales, ya que ellos son base del espacio de funciones.

Ejemplo 2.17

Serie de Fourier Un ejemplo sencillo para aprender el mecanismo del cálculo de los coeficientes para la aproximación de funciones en series de Fourier lo podemos hacer para la siguiente función.

$$f(x) = x^2, \quad -\pi \leq x \leq \pi.$$

Los coeficientes serán:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{2\pi^2}{3}, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos(nx) dx = \frac{4 \cos(n\pi)}{n^2} = \frac{4(-1)^n}{n^2}, \quad b_n = 0 \quad \forall n.$$

Se puede verificar que si $f(x)$ es par ($f(-x) = f(x)$) sólo la parte que contiene $\cos(nx)$ contribuirá a la serie.

Si $f(x)$ es impar ($f(-x) = -f(x)$), lo hará sólo la parte que contiene $\sin(nx)$.

Por lo tanto:

$$f(x) = x^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx) = \frac{\pi^2}{3} - 4 \cos(x) + \cos(2x) - \frac{4}{9} \cos(3x) + \frac{1}{4} \cos(4x) - \dots$$



Ejemplo 2.18

Mínimos cuadrados:

Mostraremos como se puede utilizar el método de mínimos cuadrados para ajustar un conjunto de datos experimentales a un polinomio de cualquier grado. Veamos estos tres primeros casos:

1. La situación más simple será el conjunto de datos experimentales: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, y nos preguntamos ¿Cuál será la recta, $y = cx$, que ajusta más acertadamente a estos puntos? Es inmediato darnos cuenta que la pendiente de esa recta será:

$$|y\rangle = c|x\rangle \Rightarrow c = \frac{\langle x|y\rangle}{\langle x|x\rangle}.$$

Consideremos los puntos experimentales: $\{(1, 2), (3, 2), (4, 5), (6, 6)\}$, entonces:

$$|y\rangle = c|x\rangle \Rightarrow \begin{pmatrix} 2 \\ 2 \\ 5 \\ 6 \end{pmatrix} = c \begin{pmatrix} 1 \\ 3 \\ 4 \\ 6 \end{pmatrix} \Rightarrow c = \frac{\langle x|y\rangle}{\langle x|x\rangle} = \frac{2 + 6 + 20 + 36}{1 + 9 + 16 + 36} = \frac{64}{62} = 1,03226.$$

¿Cuál sería el ajuste para el caso de una recta $y = cx + b$? Queda como ejercicio para el lector utilizar el esquema descrito en las ecuaciones (2.4.3) y encontrar los parámetros c y b para este nuevo ajuste.

2. El segundo caso que consideraremos será la generalización a una contribución lineal de varios parámetros, i.e. un “ajuste” multilinear a la mejor función, $y = y(x_1, x_2, \dots, x_m) = c^1 x_1 + c^2 x_2 + \dots + c^m x_m$. El procedimiento anterior puede generalizarse de forma casi inmediata:

$$|y\rangle = \sum_{j=1}^m c^j |x_j\rangle \Rightarrow \begin{cases} \langle x^1|y\rangle = c^1 \langle x^1|x_1\rangle + c^2 \langle x^1|x_2\rangle + \dots + c^m \langle x^1|x_m\rangle \\ \langle x^2|y\rangle = c^1 \langle x^2|x_1\rangle + c^2 \langle x^2|x_2\rangle + \dots + c^m \langle x^2|x_m\rangle \\ \vdots \\ \langle x^m|y\rangle = c^1 \langle x^m|x_1\rangle + c^2 \langle x^m|x_2\rangle + \dots + c^m \langle x^m|x_m\rangle \end{cases}.$$

Es decir, un sistema de m ecuaciones con m incógnitas que corresponden a los “pesos”, c^1, c^2, \dots, c^m , de las contribuciones a la función multilinear.

Supongamos ahora un paso más en este afán de generalizar y consideremos que ejecutamos n experimentos con $n > m$. De esta forma nuestro problema adquiere un grado mayor de riqueza y las medidas experimentales estarán representadas ahora por:

$$(y_1, x_{11}, x_{12}, \dots, x_{1m}; y_2, x_{21}, x_{22}, \dots, x_{2m}; y_3, x_{31}, x_{32}, \dots, x_{3m}; \dots y_n, x_{n1}, x_{n2}, x_{n3}, \dots, x_{nm}),$$

por lo tanto, podremos construir m vectores:

$$|\tilde{x}_1\rangle = (x_{11}, \dots, x_{1n}); |\tilde{x}_2\rangle = (x_{21}, \dots, x_{2n}); \dots |\tilde{x}_m\rangle = (x_{m1}, \dots, x_{mn}); |y\rangle = (y_1, \dots, y_n),$$

y el conjunto $\{|\tilde{x}_1\rangle, |\tilde{x}_2\rangle, \dots, |\tilde{x}_m\rangle\}$ expande el subespacio $\mathcal{S}(|\tilde{x}_1\rangle, |\tilde{x}_2\rangle, \dots, |\tilde{x}_m\rangle)$, que alberga la aproximación de $|y\rangle$. Otra vez, la distancia de este subespacio al vector $|y\rangle$, será mínima, y por lo tanto, será la mejor aproximación:

$$[d(\mathcal{S}(\tilde{c}^i|\tilde{x}_i), |y\rangle)]^2 = \langle \mathcal{S}(\tilde{c}^i|\tilde{x}_i) - y | \mathcal{S}(\tilde{c}^i|\tilde{x}_i) - y \rangle,$$

y $|\mathcal{S}(\tilde{c}^i|\tilde{x}_i) - y\rangle$ será ortogonal a los $|\tilde{x}_i\rangle$:

$$\langle \tilde{x}^j | \mathcal{S}(\tilde{c}^i|\tilde{x}_i) - y \rangle \equiv \langle \tilde{x}^j | \sum_{i=1}^m \tilde{c}^i |\tilde{x}_i\rangle - y \rangle = 0 \quad \forall \quad i, j = 1, 2, 3, \dots, m.$$

Finalmente, el problema queda de la misma manera que en el caso multilineal (2.):

$$|y\rangle = \sum_{j=1}^m \tilde{c}^j |\tilde{x}_j\rangle \Rightarrow \begin{cases} \langle \tilde{x}^1 | y \rangle = \tilde{c}^1 \langle \tilde{x}^1 | \tilde{x}_1 \rangle + \tilde{c}^2 \langle \tilde{x}^1 | \tilde{x}_2 \rangle + \cdots + \tilde{c}^m \langle \tilde{x}^1 | \tilde{x}_j \rangle \\ \langle \tilde{x}^2 | y \rangle = \tilde{c}^1 \langle \tilde{x}^2 | \tilde{x}_1 \rangle + \tilde{c}^2 \langle \tilde{x}^2 | \tilde{x}_2 \rangle + \cdots + \tilde{c}^m \langle \tilde{x}^2 | \tilde{x}_j \rangle \\ \vdots \\ \langle \tilde{x}^m | y \rangle = \tilde{c}^1 \langle \tilde{x}^m | \tilde{x}_1 \rangle + \tilde{c}^2 \langle \tilde{x}^m | \tilde{x}_2 \rangle + \cdots + \tilde{c}^m \langle \tilde{x}^m | \tilde{x}_j \rangle \end{cases}$$

3. Se puede extender el razonamiento anterior y generar un ajuste “lineal no lineal”. Esto es: el ajuste lineal es en los coeficientes, pero la funcionalidad de la ley a la cual queremos ajustar los datos puede ser un polinomio de cualquier orden. Ese es el caso de una parábola que ajusta, por ejemplo, al siguiente conjunto de puntos experimentales:

$$\{(0, 1), (1, 3), (2, 7), (3, 15)\} \Leftrightarrow y = ax^2 + bx + c.$$

Las ecuaciones toman la siguiente forma:

$$\begin{aligned} 1 &= 0 + 0 + c \\ 3 &= a + b + c \\ 7 &= 4a + 2b + c \\ 15 &= 9a + 3b + c \end{aligned}$$

y los vectores construidos a partir de los datos experimentales serán:

$$|x_1\rangle = (0, 1, 4, 9), \quad |x_2\rangle = (0, 1, 2, 3), \quad |x_3\rangle = (1, 1, 1, 1), \quad |y\rangle = (1, 3, 7, 15).$$

Una vez más, la ecuación vectorial sería: $|y\rangle = a|x_1\rangle + b|x_2\rangle + c|x_3\rangle$, y las ecuaciones normales (2.) para este sistema se construyen como:

$$\begin{aligned} \left. \begin{aligned} 136 &= 98a + 36b + 14c \\ 62 &= 36a + 14b + 6c \\ 26 &= 14a + 6b + 4c \end{aligned} \right\} \Rightarrow \begin{cases} a = -6 \\ b = \frac{113}{5} \\ c = -\frac{32}{5} \end{cases} \Rightarrow y = -6x^2 + \frac{113}{5}x - \frac{32}{5}. \end{aligned}$$

4. Para finalizar analicemos el caso típico de aproximación por mínimos cuadrados. Se sospecha que una determinada propiedad de un material cumple con la ecuación $y = ax_1 + bx_2$ y realizamos 4 mediciones independientes obteniendo:

$$\begin{pmatrix} y_1 \\ x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 15 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} y_2 \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 12 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} y_3 \\ x_{31} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 10 \\ 1 \\ 1 \end{pmatrix}; \quad \begin{pmatrix} y_4 \\ x_{41} \\ x_{42} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Es claro que tenemos un subespacio de $m = 2$ dimensiones, que hemos hecho $n = 4$ veces el experimento y que los vectores considerados arriba serán:

$$|x_1\rangle = (1, 2, 1, 1), \quad |x_2\rangle = (2, 1, 1, -1), \quad |y\rangle = (15, 12, 10, 0).$$

Por lo tanto, vectorialmente tendremos: $|y\rangle = a|x_1\rangle + b|x_2\rangle$, es decir, las ecuaciones normales (2.) se escribirán de la siguiente manera:

$$\left. \begin{aligned} 7a + 4b &= 49 \\ 4a + 7b &= 52 \end{aligned} \right\} \Rightarrow \begin{cases} a = \frac{45}{11} \\ b = \frac{56}{11} \end{cases} \Rightarrow y = \frac{45}{11}x_1 + \frac{56}{11}x_2.$$



2.4.4.1 Aproximación polinómica

Consideremos los puntos experimentales representados en la figura 2.3:

$$\{(2, 8), (4, 10), (6, 11), (8, 18), (10, 20), (12, 34)\}.$$

Es importante hacer notar que debido a que los polinomios de Legendre están definidos en el intervalo $[-1, 1]$ los puntos experimentales deberán reescalarsse a ese intervalo para poder encontrar el polinomio de interpolación como combinación lineal de los polinomios de Legendre. Esto se puede hacer con la ayuda del siguiente cambio de variable:

$$x = \frac{(b-a)t + b + a}{2}, \quad dx = \frac{b-a}{2} dt,$$

donde a y b son los valores mínimos y máximos de los datos, respectivamente.

En este caso podemos ver que $a = 2$ y $b = 12$, y por lo tanto:

$$x = 5t + 7 \text{ y } dx = 5dt \Rightarrow t = (x - 7)/5 \text{ y } dt = dx/5.$$

Ahora los datos quedan de la siguiente forma:

$$\left\{(-1, 8), \left(-\frac{3}{5}, 10\right), \left(-\frac{1}{5}, 11\right), \left(\frac{1}{5}, 18\right), \left(\frac{3}{5}, 20\right), (1, 34)\right\}.$$

Ahora bien, necesitamos encontrar los coeficientes de la serie

$$f(x) = \sum_{k=0}^5 C^k P_k(x) = C^0 P_0(x) + C^1 P_1(x) + \cdots + C^5 P_5(x)$$

y el sistema de ecuaciones que resulta es el siguiente:

$$\left\{ \begin{array}{l} (-1, 8) \Rightarrow 8 = C^0 - C^1 + C^2 - C^3 + C^4 - C^5 \\ \left(-\frac{3}{5}, 10\right) \Rightarrow 10 = C^0 - \frac{3}{5} C^1 + \frac{1}{25} C^2 + \frac{9}{25} C^3 - \frac{51}{125} C^4 + \frac{477}{3125} C^5 \\ \left(-\frac{1}{5}, 11\right) \Rightarrow 11 = C^0 - \frac{1}{5} C^1 - \frac{11}{25} C^2 + \frac{7}{25} C^3 + \frac{29}{125} C^4 - \frac{961}{3125} C^5 \\ \left(\frac{1}{5}, 18\right) \Rightarrow 18 = C^0 + \frac{1}{5} C^1 - \frac{11}{25} C^2 - \frac{7}{25} C^3 + \frac{29}{125} C^4 + \frac{961}{3125} C^5 \\ \left(\frac{3}{5}, 20\right) \Rightarrow 20 = C^0 + \frac{3}{5} C^1 + \frac{1}{25} C^2 - \frac{9}{25} C^3 - \frac{51}{125} C^4 - \frac{477}{3125} C^5 \\ (1, 34) \Rightarrow 34 = C^0 + C^1 + C^2 + C^3 + C^4 + C^5 \end{array} \right.$$

Al resolver este sistema obtendremos los siguientes resultados:

$$C^0 = \frac{2249}{144}, \quad C^1 = \frac{3043}{336}, \quad C^2 = \frac{1775}{504}, \quad C^3 = -\frac{175}{216}, \quad C^4 = \frac{625}{336}, \quad C^5 = \frac{14375}{3024},$$

con lo cual:

$$\mathcal{P}(x) = f(x) = \frac{2249}{144} + \frac{3043}{336}x + \frac{1775}{504}P(2, x) - \frac{175}{216}P(3, x) + \frac{625}{336}P(4, x) + \frac{14375}{3024}P(5, x).$$

y la interpolación queda representada en la figura 2.3.

Nótese que mientras más puntos experimentales se incluyan para la interpolación, el polinomio resultante será de mayor grado y, por lo tanto incluirá oscilaciones que distorsionarán una aproximación más razonable. Por ello, la estrategia de hacer la interpolación a trozos, digamos de tres puntos en tres puntos, generará un mejor ajuste, pero será una función (un polinomio) continua a trozos.

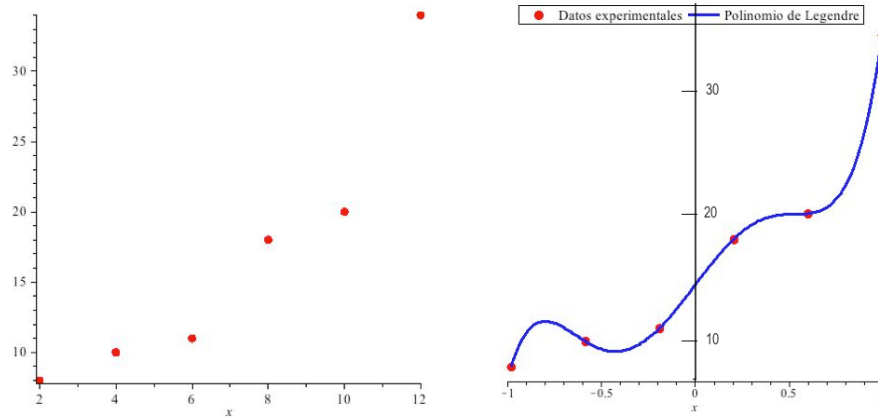


Figura 2.3: A la izquierda los puntos experimentales: $\{(2, 8), (4, 10), (6, 11), (8, 18), (10, 20), (12, 34)\}$ y a la derecha la función polinómica que los interpola.

2.4.5 Ejercicios

1. Demuestre que con las identidades siguientes:

$$\cos(kx) = \frac{e^{ikx} + e^{-ikx}}{2}, \quad \text{sen}(kx) = \frac{e^{ikx} - e^{-ikx}}{2i},$$

la serie de Fourier definida para funciones en el intervalo $(t, t + 2\pi)$ se escribe en su forma compleja como:

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \text{sen}(kx)] \Rightarrow f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

con:

$$c_k = \frac{1}{2\pi} \int_t^{t+2\pi} f(x) e^{-ikx} dx.$$

Y que para funciones definidas en el intervalo $(l, l + 2L)$ como:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{ik\pi x}{L}}, \quad \text{con } c_k = \frac{1}{2L} \int_l^{l+2L} f(x) e^{-\frac{ik\pi x}{L}} dx.$$

Nota: Para los siguientes ejercicios supondremos la utilización del programa **Maxima**.

2. Para las siguientes funciones determine la serie de Fourier calculando los coeficientes como en la sección 2.4 y compare los resultados con los cálculos hechos en el ambiente de manipulación simbólica.

- (a). $f(x) = x \text{ sen}(x)$, si $-\pi < x < \pi$.
- (b). $f(x) = e^x$, si $-\pi < x < \pi$.
- (c). $f(x) = x$, si $0 < x < 2$.
- (d). $f(x) = 2x - 1$, si $0 < x < 1$.

3. Considere el espacio vectorial, $\mathcal{C}_{[-1,1]}^{\infty}$, de funciones reales, continuas y continuamente diferenciables definidas en el intervalo $[-1, 1]$. Es claro que una posible base de este espacio de funciones la constituye el conjunto de monomios $\{1, x, x^2, x^3, x^4, \dots\}$ por cuanto estas funciones son linealmente independientes.

- (a). Si suponemos que este espacio vectorial está equipado con un producto interno definido por $\langle f|g \rangle = \int_{-1}^1 dx f(x)g(x)$, muestre que esa base de funciones no es ortogonal.
- (b). Utilizando la definición de producto interno $\langle f|g \rangle = \int_{-1}^1 dx f(x)g(x)$ ortogonalice la base

$\{1, x, x^2, x^3, x^4, \dots\}$ y encuentre los 10 primeros vectores ortogonales, base para $C_{[-1,1]}^\infty$. Esta nueva base de polinomios ortogonales se conoce como los polinomios de Legendre.

- (c). Modificando un poco la definición de producto interno para que ahora sea

$$\langle f|g \rangle = \int_{-1}^1 dx f(x)g(x)\sqrt{1-x^2},$$

ortogonalice la base $\{1, x, x^2, x^3, x^4, \dots\}$ y encuentre otros 10 primeros vectores ortogonales base para el mismo $C_{[-1,1]}^\infty$. Esta nueva base de polinomios ortogonales se conoce como los polinomios de Chebyshev.

- (d). Suponga la función $h(x) = \sin(3x)(1-x^2)$:

- I. Expanda la función $h(x)$ en términos de la base de monomios y de polinomios de Legendre, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- II. Expanda la función $h(x)$ en términos de la base de monomios y de polinomios de Chebyshev, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- III. Expanda la función $h(x)$ en términos de la base de polinomios de Legendre y de Chebyshev, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- IV. Estime en cada caso el error que se comete como función del grado del polinomio (o monomio) de la expansión.

¿Qué puede concluir respecto a la expansión en una u otra base?

4. Parecido al ejercicio anterior, considere el espacio vectorial, $C_{[0,1]}^\infty$, de funciones reales, continuas y continuamente diferenciables definidas en el intervalo $[0, 1]$. Es claro que otra posible base de este espacio de funciones la constituye el conjunto de funciones exponenciales $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ por cuanto, al igual que el caso anterior estas funciones también son linealmente independientes. Adicionalmente, considere aquí la función $g(x) = \cos(3x^3)(1-x^2)$,

- (a). Suponga que este espacio vectorial está equipado con un producto interno definido por

$$\langle f|g \rangle = \int_0^1 dx f(x)g(x),$$

y muestre que esa base de funciones $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ no es ortogonal.

- (b). Utilizando la definición de producto interno ortogonalice la base $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ y encuentre los 7 primeros vectores ortogonales, base para $C_{[0,1]}^\infty$, i.e. $\{1, E_1(x), E_2(x), E_3(x), E_4(x), \dots\}$.
- (c). Encuentre el valor de los coeficientes, C_i de la expansión

$$g_M(x) \approx \sum_{i=0}^7 C_i x^i,$$

utilizando la definición de producto interno anterior y estime también el error ϵ_M en esta aproximación

$$g_M(x) = \sum_{i=0}^7 C_i x^i + \epsilon_M.$$

- (d). Expanda en serie de Taylor la función anterior

$$g_T(x) \approx \sum_{n=0}^7 \left. \frac{dg(x)}{dx} \right|_{x=0} \frac{x^n}{n!},$$

estime el error ϵ_T y compárelo con el caso anterior.

- (e). Ahora encuentre el valor de los coeficientes, \tilde{C}_i de la expansión

$$g_E(x) \approx \sum_{m=0}^7 \tilde{C}_m e^{mx},$$

utilizando la definición de producto interno anterior y, del mismo modo, estime también el error ϵ_E en esta aproximación

$$g_E(x) = \sum_{m=0}^7 \tilde{C}_m e^{mx} + \epsilon_E.$$

¿Qué puede concluir de la comparación de los errores ϵ_M , ϵ_T y ϵ_E ?

- (f). A continuación encuentre el valor de los coeficientes, \tilde{C}_i de la expansión

$$g_{Eo}(x) \approx \sum_{m=0}^7 \tilde{C}_m E_m(x),$$

y estime también el error ϵ_{Eo} en esta aproximación

$$g_{Eo}(x) = \sum_{m=0}^7 \tilde{C}_m E_m(x) + \epsilon_{Eo}.$$

¿Otra vez, qué puede concluir de la comparación de los errores ϵ_M , ϵ_T , ϵ_E y ϵ_{Eo} ?

- (g). Grafique las funciones $g(x)$, $g_M(x)$, $g_T(x)$, $g_E(x)$ y $g_{Eo}(x)$ y compare con los errores.

5. Al medir la temperatura a lo largo de una barra material obtenemos los siguientes valores:

x_i (cm)	1,0	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0
T_i (°C)	14,6	18,5	36,6	30,8	59,2	60,1	62,2	79,4	99,9

Encuentre, mediante el método de los mínimos cuadrados los coeficientes que mejor ajustan a la recta $T = ax + b$.

6. Los precios de un determinado producto varían como se muestra a continuación:

Año	2007	2008	2009	2010	2011	2012	2013	2014
Precio	133,5	132,2	138,7	141,5	144,2	144,5	148,6	153,8

Realice una interpolación polinomial que permita modelar estos datos.

Practicando con Maxima

1. Series de Fourier

Existe una biblioteca llamada **fourie** en **Maxima** que contiene instrucciones para el cálculo simbólico de series de Fourier de una función $f(x)$ en el intervalo $[-l, l]$: **fourier** (f, x, l). También encontraremos un conjunto de comandos en el paquete para calcular los coeficientes y para manipular las expresiones resultantes.

(%i1) `load(fourie)$`

En el ejemplo anterior aproximamos la función:

$$f(x) = x^2.$$

Veamos como se trabaja con el programa para calcular la serie de Fourier. Los resultados aparecerán en la forma de listas temporales y entre ellas los coeficientes. Las listas temporales serán indicadas con la notación (%t).

```
(%i2) f:x^2;
(%o2) x^2

(%i3) fourier(f,x,%pi);
(%t3) a0 =  $\frac{\pi^2}{3}$ 
      2  $\left( \frac{\pi^2 \sin(\pi n)}{n} - \frac{2 \sin(\pi n)}{n^3} + \frac{2 \pi \cos(\pi n)}{n^2} \right)$ 
(%t4) an =  $\frac{\pi^2 \sin(\pi n)}{n} - \frac{2 \sin(\pi n)}{n^3} + \frac{2 \pi \cos(\pi n)}{n^2}$ 
(%t5) bn = 0
(%o5) [%t3, %t4, %t5]
```

Lo anterior se puede simplificar con el comando **foursimp**:

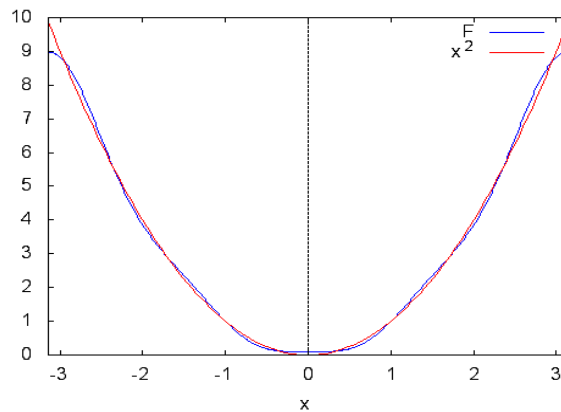
```
(%i6) foursimp(%);
(%t6) a0 =  $\frac{\pi^2}{3}$ 
(%t7) an =  $\frac{4(-1)^n}{n^2}$ 
(%t8) bn = 0
(%o8) [%t6, %t7, %t8]
```

Podemos evaluar la lista de los coeficiente hasta el término k . Aquí lo haremos hasta $k = 4$ y el resultado lo lo asignaremos a la variable F . Por otro lado, usaremos (%o8), la última salida, como entrada para siguiente comando.

```
(%i9) F:fourexexpand(%o8,x,%pi,4);
(%o9)  $\frac{\cos(4x)}{4} - \frac{4 \cos(3x)}{9} + \cos(2x) - 4 \cos x + \frac{\pi^2}{3}$ 
```

Construiremos ahora en un mismo gráfico la función original y los primeros 5 términos de la serie, de esta manera podremos comparar el resultado de la aproximación. Las opciones para realizar los diferentes gráficos en **Maxima** se pueden consultar en el manual de programa.

```
(%i10) wxplot2d([F,f],[x,-%pi,%pi],[legend,"F","x^2"]);
(%o10)
```



Veamos que sucede si escribimos:

```
(%i11) totalfourier(f,x,%pi);
```

```
(%t11) a0 = 
$$\frac{\pi^2}{3}$$

```

```
(%t12) a_n = 
$$\frac{2 \left( \frac{\pi^2 \sin(\pi n)}{n} - \frac{2 \sin(\pi n)}{n^3} + \frac{2 \pi \cos(\pi n)}{n^2} \right)}{\pi}$$

```

```
(%t13) b_n = 0
```

```
(%t14) a0 = 
$$\frac{\pi^2}{3}$$

```

```
(%t15) a_n = 
$$\frac{4 (-1)^n}{n^2}$$

```

```
(%t16) b_n = 0
```

```
(%o16) 4 \sum_{n=1}^{\infty} \frac{(-1)^n \cos(nx)}{n^2} + \frac{\pi^2}{3}
```

En este caso fueron aplicados de manera simultánea los comandos **fourier** y **foursimp** para finalmente presentar la serie en forma de una sumatoria.

```
(%i17) kill(all)$
```

2. Mínimos Cuadrados

Maxima puede estimar los parámetros que mejor se ajusten a una función $f = (x, y)$ para un conjunto de datos, utilizando el método de mínimos cuadrados. El programa buscará primero una solución exacta, si no la encuentra buscará una aproximada. El resultado lo presentará como una lista de ecuaciones. La función a utilizar será **lsquares**.

Vamos a considerar los ejemplos estudiados con anterioridad:

(a). En el primer ejemplo los datos eran los siguientes:

$$(x, y) = (1, 2), (3, 2), (4, 5), (6, 6) .$$

Es necesario hacer uso de la biblioteca **lsquares** y los datos deben introducirse en forma de matriz.

```
(%i1) load(lsquares)$
```

```
(%i2) datos: matrix([1,2],[3,2],[4,5],[6,6])$
```

Por conveniencia, para más adelante hacer un gráfico, convertimos la matrix “datos” en una lista.

Esto es sencillo si utilizamos el comando **args**:

```
(%i3) datosL: args(datos);
```

```
(%o3) [[1, 2], [3, 2], [4, 5], [6, 6]]
```

Supondremos entonces que los puntos se ajustan a un polinomio lineal del tipo: $y = ax$. El parámetro a se calcula con la función **lsquares_estimates**. Es importante prestar bastante atención a la sintaxis del siguiente comando.

```
(%i4) param: lsquares_estimates(datos,[x,y],y=a*x,[a]), numer;
```

```
(%o4) [[a = 1.032258064516129]]
```

Este será entonces el valor del parámetro a de la ecuación de la recta $y = ax$ que pasa por el origen. Notemos también que le hemos asignado el valor del parámetro a a la variable `param`.

Lo que haremos ahora es escribir la ecuación de dicha recta. Podemos hacer uso de la instrucción `ev` que nos permite evaluar una expresión.

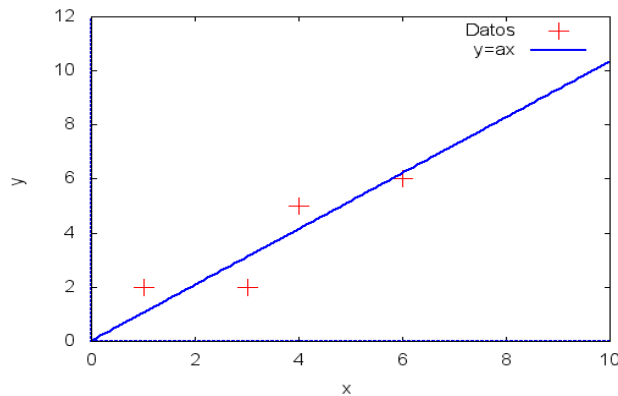
```
(%i5) y:ev(a*x,first(param));
(%o5) 1,032258064516129 x
```

Procederemos ahora a graficar los datos experimentales *vs* el ajuste por mínimos cuadrados en un mismo gráfico.

Recordemos que el conjunto de puntos lo tenemos en la forma de una lista, que hemos denominada más arriba como `datosL`. Mientras que al ajuste que hemos calculado, es decir, la recta: $1,032258064516129 x$ le hemos asignado la variable denominada y .

Es recomendable consultar el manual del programa, en la parte que tiene que ver con gráficos, **plot2d**, **plot3d**, para identificar la sintaxis que aparecerá en la siguiente instrucción.

```
(%i6) wxplot2d([[discrete,datosL], y], [x,0,10],[style, [points,5,2], [lines,2,1]],
[point_type, plus], [legend,"Datos","y=ax"],[xlabel, "x"],[ylabel, " y"])$
(%o6)
```



Nota: Se deja como ejercicio repetir éste mismo cálculo pero usando un ajuste para los datos de la forma: $y = ax + b$.

- (b). Consideremos el conjunto de datos: $|x_1\rangle = (1, 2, 1, 1)$, $|x_2\rangle = (2, 1, 1, -1)$, $|y\rangle = (15, 12, 10, 0)$. Vamos a suponer que ajustan de la manera siguiente: $|y\rangle = a |x_1\rangle + b |x_2\rangle$.

```
(%i7) datos2: matrix([1,2,15],[2,1,12],[1,1,10],[1,-1,0])$
```

Cambiamos ligeramente la notación por: $z = ax + by$ y calculemos los parámetros a y b .

```
(%i8) param: lsquares_estimates(datos2,[x,y,z], z=a*x+b*y,[a,b]), numer;
(%o8) [[a = 4,090909090909091, b = 5,090909090909091]]
```

- (c). Para el tercer ejemplo se consideraron los siguientes datos:

$$\{(0, 1), (1, 3), (2, 7), (3, 15)\} \Leftrightarrow y = ax^2 + bx + c.$$

Haremos con **Maxima** el cálculo directo usando un ajuste cuadrático para los datos suministrados.

```
(%i9) datos3: matrix([0,1],[1,3],[2,7],[3,15])$
```

Como hicimos anteriormente, hacemos una lista:

```
(%i10) datosL3: args(datos3)$
```

Procedemos a calcular los parámetros:

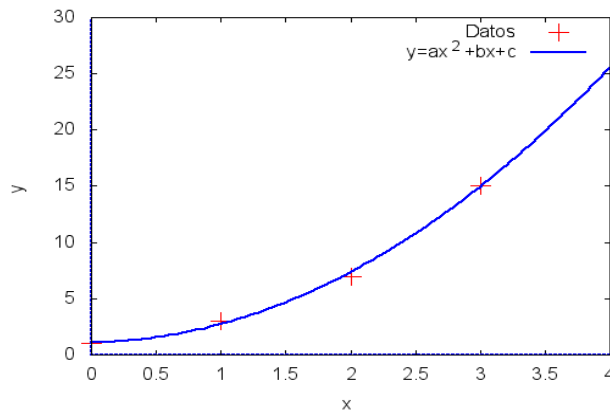
```
(%i11) param: lsquares_estimates(datos3,[x,y], y=a*x^2+b*x+c,[a,b,c]), numer;
(%o11) [[a = 1,5, b = 0,1, c = 1,1]]
```

Con los parámetros determinados podemos asignarlos a una nueva variable, en este caso y_2 ,

```
(%i12) y2:ev(a*x^2+b*x+c,first(param))$
```

Como hicimos anteriormente, graficamos los datos y el ajuste cuadrático en una misma figura.

```
(%i13) wxplot2d([[discrete,datosL3], y2], [x,0,4],[style,[points,5,2],[lines,2,1]],
[point_type, plus], [legend,"Datos", "y=ax^2+bx+c"],[xlabel,"x"],[ylabel,"y"])$
(%o13)
```



```
(%i14) kill(all)$
```

3. Polinomios ortogonales

Maxima contiene la librería **orthopoly** que nos permite acceder a la evaluación simbólica y numérica de los diferentes tipos de polinomios ortogonales: Chebyshev, Laguerre, Hermite, Jacobi, Legendre...

```
(%i1) load(orthopoly)$
```

Por ejemplo, para obtener los primeros 6 polinomios de Legendre escribimos los siguientes comandos:

```
(%i2) [legendre_p(0,x),legendre_p(1,x),legendre_p(2,x),
legendre_p(3,x),legendre_p(4,x),legendre_p(5,x)]$
```

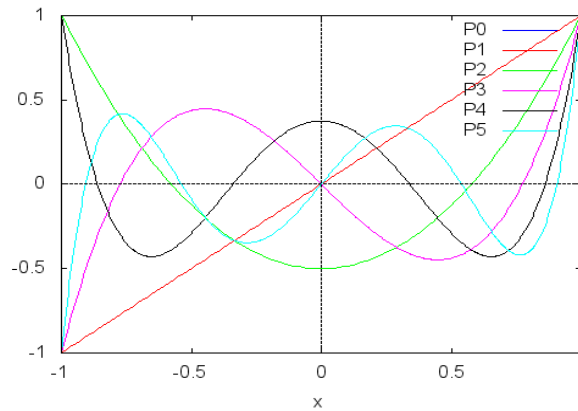
Simplificamos con **ratsimp**:

```
(%i3) ratsimp(%);
```

$$(\%o3) \left[1, x, \frac{3x^2 - 1}{2}, \frac{5x^3 - 3x}{2}, \frac{35x^4 - 30x^2 + 3}{8}, \frac{63x^5 - 70x^3 + 15x}{8} \right]$$

Los diferentes polinomios de Legendre se pueden visualizar de la manera siguiente:

```
(%i4) wxplot2d(%, [x, -1, 1], [legend, "P0", "P1", "P2", "P3", "P4", "P5"])$
(%o4)
```



Ahora bien, con los datos de la figura 2.3 planteamos un sistema de ecuaciones lineales que ahora introduciremos en el programa:

```
(%i4) ecu1:C0-C1+C2-C3+C4-C5=8$
(%i5) ecu2:C0-3/5*C1+1/25*C2+9/25*C3-51/125*C4+477/3125*C5=10$
(%i6) ecu3:C0-1/5*C1-11/25*C2+7/25*C3+29/125*C4-961/3125*C5=11$
(%i7) ecu4:C0+1/5*C1-11/25*C2-7/25*C3+29/125*C4+961/3125*C5=18$
(%i8) ecu5:C0+3/5*C1+1/25*C2-9/25*C3-51/125*C4-477/3125*C5=20$
(%i9) ecu6:C0+C1+C2+C3+C4+C5=34$
```

Procedemos a resolver el sistema de ecuaciones anterior:

```
(%i10) linsolve([ecu1,ecu2,ecu3,ecu4,ecu5,ecu6], [C0,C1,C2,C3,C4,C5]);
(%o10) [C0 = 2249/144, C1 = 3043/336, C2 = 1775/504, C3 = -175/216, C4 = 625/336, C5 = 14375/3024]
```

Para asignar cada solución a la variable correspondiente podemos hacer lo siguiente:

```
(%i11) [C0,C1,C2,C3,C4,C5]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4]),rhs(%[5]),rhs(%[6])];
(%o11) [2249/144, 3043/336, 1775/504, -175/216, 625/336, 14375/3024]
```

Por lo tanto, la función aproximada será la siguiente:

```
(%i13) f:C0+C1*legendre_p(1,x)+C2*legendre_p(2,x)+C3*legendre_p(3,x)
+C4*legendre_p(4,x)+legendre_p(5,x)*C5$
(%i14) f:expand(%);
```

$$(\%o14) \frac{14375x^5}{384} + \frac{3125x^4}{384} - \frac{8375x^3}{192} - \frac{325x^2}{192} + \frac{7367x}{384} + \frac{1863}{128}$$

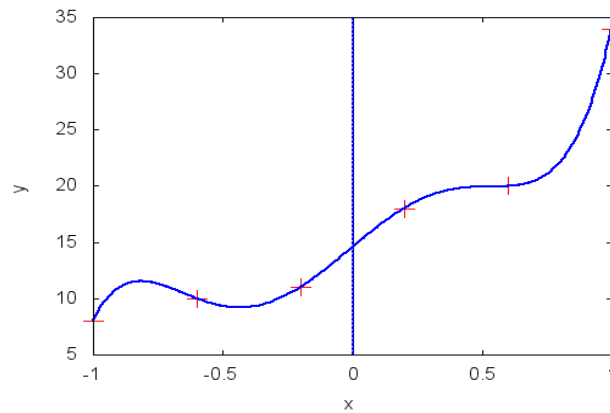
Procedemos a introducir los datos:

```
(%i15) datos:[[-1,8],[-3/5,10],[-1/5,11],[1/5,18],[3/5,20],[1,34]];
```

```
(%o15) [[-1, 8], [-3/5, 10], [-1/5, 11], [1/5, 18], [3/5, 20], [1, 34]]
```

Haremos ahora la gráfica con los datos y con la interpolación:

```
(%i16) wxplot2d([[discrete,datos],f], [x,-1,1],[style, [points,5,2], [lines,2,1]], [point_type, plus],
[legend, false],[xlabel, "x"],[ylabel, " y"])$
(%o16)
```



Bibliografía

- [1] Arfken, G. B., Weber, H., Weber, H.J. (2000) **Mathematical Methods for Physicists** 5ta Edición (Academic Press, Nueva York)
- [2] Scharnhorst, K. (2009) **Angles in complex vector spaces** Acta Applicandae Mathematica), **69**(1), 95-103
- [3] Dennery, P. y Krzywicki, A. (1995) **Mathematics for Physicists** (Dover Publications Inc, Nueva York)
- [4] Harper, C. (1971) **Introduction to Mathematical Physics** (*Prentice Hall, Englewood Cliff, N.J.:*)
- [5] Hassani, S. (1991) **Foundations of Mathematical Physics** (*Prentice Hall, International Edition, London:*
- [6] Riley, K.F., Hobson, M.P. y Bence, S.J. (2002) **Mathematical Methods for Physics and Engineering** (*Cambridge University Press*)