Facultad de Ingeniería
UNLP

Matemática C

Módulo II

VI Transformaciones Lineales

2022

Temario:

1. Clase 1: Introducción. Definición. Ejemplos básicos. Proyección. Reflexión. Rotación. Casos especiales. Imagen y espacio nulo de una transformación lineal. Propiedades fundamentales. Isomorfismos.

2. Clase 2: Representación matricial. Ejemplos. Diferentes representaciones de acuerdo a las bases consideradas (cambio de base). Representación matricial y características de la transformación. Matrices Semejantes. Propiedades. Composición de transformaciones lineales. Ejemplos. Aplicaciones.

6.1. Introducción

Estudiaremos aquí una clase de funciones denominadas **transformaciones lineales**, que transforman un vector $\mathbf{v}$ de un espacio vectorial V en otro vector $\mathbf{w}$ de un espacio vectorial W, cumpliendo con ciertas condiciones. Es decir, son casos especiales de funciones $F : V \to W$ entre dos espacios vectoriales. Se las puede puede pensar como *funciones de "una sola variable"*, donde el argumento de la función es un vector del espacio vectorial V, y el "valor" de la función es un vector del espacio vectorial W.

Si $V = W$, de modo que transforma un vector $\boldsymbol{v}$ de V en otro vector $\boldsymbol{w}$ del **mismo espacio** V, la transformación lineal se denomina usualmente **operador lineal**.

Usaremos la notación $L : V \longrightarrow W$ para describir una transformación lineal:

$$L(\mathbf{v}) = \mathbf{w}, \qquad \mathbf{v} \in V, \;\; \mathbf{w} \in W$$

Veremos que una transformación lineal L de un espacio vectorial n-dimensional V en otro espacio vectorial m-dimensional W podrá representarse por una matriz A de $m \times n$. Esto permitirá trabajar con la matriz A para discutir las características y propiedades de la transformación L, como así también, en ciertos casos, determinar las propiedades de la matriz A a partir de las propiedades de la transformación lineal que representa.

6.1.1. Definición general

Definición.
Una función $L : V \longrightarrow W$ de un espacio vectorial V en otro espacio vectorial W (ambos sobre el mismo conjunto de escalares) es una **transformación lineal** si satisface

$$L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$$

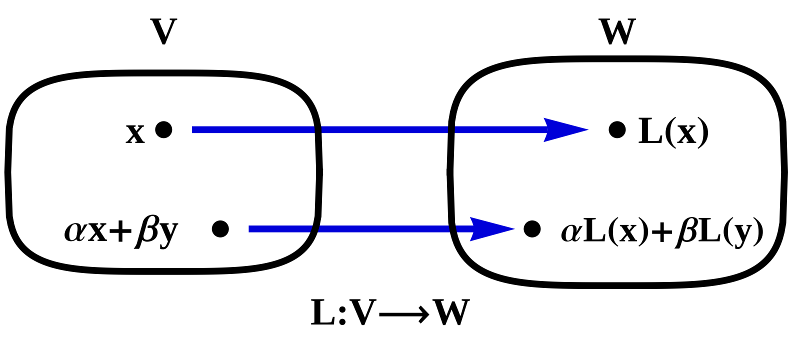
para todo par de vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ de V y todo par de escalares α, β. Es decir,

$$\begin{aligned} L(\mathbf{v}_1 + \mathbf{v}_2) &= L(\mathbf{v}_1) + L(\mathbf{v}_2) \quad \text{para todo } \mathbf{v}_1 \text{ y } \mathbf{v}_2 \text{ en } V \\ L(\alpha \mathbf{v}) &= \alpha L(\mathbf{v}) \quad \text{para todo } \mathbf{v} \text{ en } V \text{ y } \alpha \text{ escalar} \end{aligned}$$

En particular, para $\alpha = 0$ la última ecuación implica que **toda** transformación lineal $L : V \longrightarrow W$ satisface

$$L(\mathbf{0}_V) = \mathbf{0}_W$$

con $\mathbf{0}_V$ y $\mathbf{0}_W$ los vectores nulos de V y W, ya que $L(\mathbf{0}_V) = L(0\mathbf{v}) = 0L(\mathbf{v}) = \mathbf{0}_W$.



Ejemplos 6.1.1: Transformaciones lineales de $\Re^2$ en $\Re^2$
Comenzaremos con algunos ejemplos básicos:

1. **Transformación de dilatación (escalamiento):**

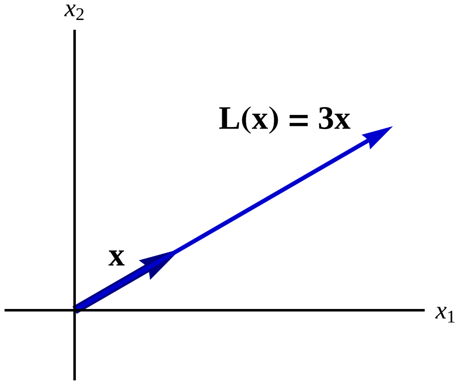
   Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ es un vector de $\mathbb{R}^2$, definimos, como primer ejemplo,

$$L(\mathbf{x}) = 3\mathbf{x} = \begin{pmatrix} 3x_1 \\ 3x_2 \end{pmatrix}$$

   L es una transformación lineal, ya que

$$\begin{aligned} L(\alpha\mathbf{x}) &= 3(\alpha\mathbf{x}) = \alpha(3\mathbf{x}) = \alpha L(\mathbf{x}) \\ L(\mathbf{x}+\mathbf{y}) &= 3(\mathbf{x}+\mathbf{y}) = 3\mathbf{x}+3\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y}) \end{aligned}$$

   verificándose que $L(\mathbf{0}) = \mathbf{0}$. Geométricamente, L tiene el efecto de “dilatar” el vector $\mathbf{x}$, multiplicando su longitud por un factor 3 y conservando su dirección y sentido:



   Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar!)

$$L(\mathbf{x}) = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

2. **Proyección ortogonal sobre el eje x_1:**

   Definimos ahora

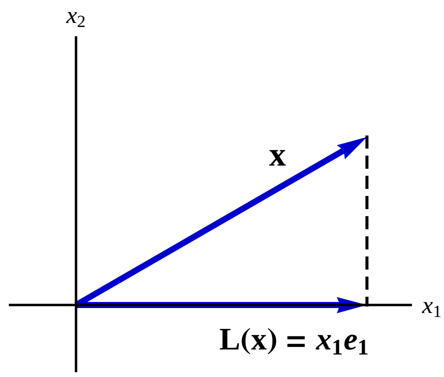
$$L(\mathbf{x}) = x_1\mathbf{e}_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

   Si $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, entonces $\alpha\mathbf{x}+\beta\mathbf{y} = \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \end{pmatrix}$. Luego

$$\begin{aligned} L(\alpha\mathbf{x}+\beta\mathbf{y}) &= (\alpha x_1 + \beta y_1)\mathbf{e}_1 = \alpha(x_1\mathbf{e}_1) + \beta(y_1\mathbf{e}_1) \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   lo que prueba que L es una transformación lineal.

   Geométricamente, $L(\mathbf{x})$ es la proyección del vector $\mathbf{x}$ sobre el eje x_1



Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar!)

$$L(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

3. **Reflexión respecto del eje $\boldsymbol{x}_1$:**
   Definimos

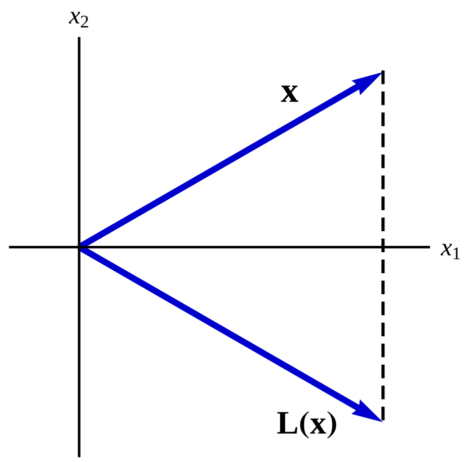
$$L(\mathbf{x}) = x_1\mathbf{e}_1 - x_2\mathbf{e}_2 = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

Esta tranformación satisface

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} \alpha x_1 + \beta y_1 \\ -(\alpha x_2 + \beta y_2) \end{pmatrix} \\ &= \alpha \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} + \beta \begin{pmatrix} y_1 \\ -y_2 \end{pmatrix} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

$\Rightarrow$ L es un operador lineal.

Geométricamente, $L(\mathbf{x})$ es la reflexión del vector $\mathbf{x}$ respecto (o a través) del eje x_1.



Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar)

$$L(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

4. **Rotación de ángulo $\pi/2$ antihorario:**

   Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ definimos

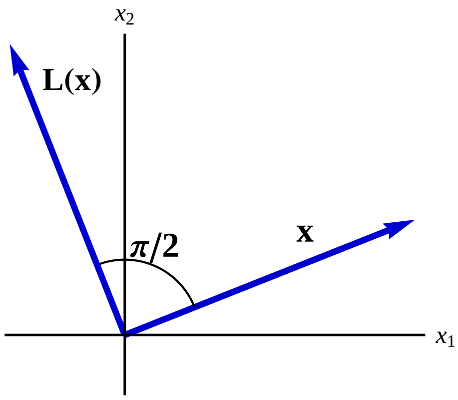
$$L(\mathbf{x}) = -x_2\mathbf{e}_1 + x_1\mathbf{e}_2 = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

   Vemos que cumple

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} -(\alpha x_2 + \beta y_2) \\ \alpha x_1 + \beta y_1 \end{pmatrix} = \alpha \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} + \beta \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   $\Rightarrow L$ es una transformación lineal.

   Geométricamente, $L(\mathbf{x})$ representa la rotación de ángulo $\theta = \pi/2$ (en sentido antihorario) del vector $\mathbf{x}$:



   Podemos expresar $L(\mathbf{x})$ en forma matricial como (verificar)

$$L(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5. **Transformación de escalamiento general**

   En general, el operador L_c definido por

$$L_c(\mathbf{x}) \quad = \quad c\mathbf{x} = \begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix}$$

   con c un escalar fijo, es una transformación lineal, como podrá el lector probar fácilmente.

   Si $c > 0$, L_c tiene el efecto de multiplicar la longitud del vector $\mathbf{x}$ por el factor de escala c, dilatando el vector un factor c si $c > 1$ y contrayendo el vector un factor c si $0 < c < 1$, pero siempre conservando su dirección y sentido.

Si $c = 1$, la transformación resultante

$$L_1(\mathbf{x}) = \mathbf{x}$$

se denomina **operador identidad** y se la denota como I: $I(\mathbf{x}) = \mathbf{x}\ \forall\ \mathbf{x} \in V$.
I no modifica ningún vector.

Si $c = 0$, la transformación resultante

$$L_0(\mathbf{x}) = \mathbf{0}$$

se denomina **operador nulo**. Envía a todos los vectores de V al vector nulo $\mathbf{0} \equiv \mathbf{0}_V$.

Si $c < 0$, L_c tendrá el efecto de invertir el sentido del vector, dilatándolo si $c < -1$, contrayéndolo si $-1 < c < 0$ y conservando su longitud si $c = -1$, en cuyo caso coincide con el operador de inversión.

Podemos expresar $L_c(\mathbf{x})$ en forma matricial como

$$L(\mathbf{x}) = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

6. **Inversión:**

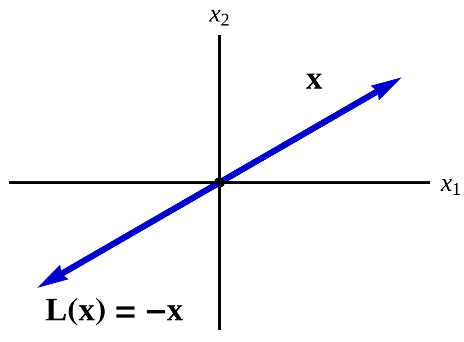
   Corresponde a

$$L(\mathbf{x}) = -\mathbf{x} = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

   La linealidad de L es inmediata:

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= \begin{pmatrix} -(\alpha x_1 + \beta y_1) \\ -(\alpha x_2 + \beta y_2) \end{pmatrix} = \alpha \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} + \beta \begin{pmatrix} -y_1 \\ -y_2 \end{pmatrix} \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

   Geométricamente, $L(\mathbf{x})$ es el vector opuesto a $\mathbf{x}$.



   Podemos expresar $L(\mathbf{x})$ en forma matricial como

$$L(\mathbf{x}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Observar que el operador de inversión puede obtenerse como caso particular de otras transformaciones. Por ejemplo,

1. La transformación de escala con $c = -1$

2. Una rotación de ángulo π (en sentido anti-horario o sentido horario)

Esta última puede también lograrse mediante dos rotaciones sucesivas de ángulo $\pi/2$ (por ejemplo, ambas en sentido antihorario): Si R rota al vector $\mathbf{x}$ en $\pi/2$ antihorario, entonces

$$R\left(R\left(\mathbf{x}\right)\right) = R\left(\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}\right) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} = -\mathbf{x}$$

Si definimos el cuadrado L^2 de un operador L (transformación de V en V) mediante

$$L^2\left(\mathbf{x}\right) \equiv L\left(L\left(\mathbf{x}\right)\right)$$

entonces el operador de inversión L puede expresarse en términos del operador de rotación previo como

$$L = R^2$$

Ejemplos de transformaciones no lineales:

1. Si

$$F\left(\mathbf{x}\right) = -x_1 x_2 \mathbf{e}_1 + x_2 \mathbf{e}_2 = \begin{pmatrix} -x_1 x_2 \\ x_2 \end{pmatrix}$$

obtenemos

$$\begin{aligned} F\left(\alpha\mathbf{x}\right) &= -\left(\alpha x_1\right)\left(\alpha x_2\right)\mathbf{e}_1 + \left(\alpha x_2\right)\mathbf{e}_2 = \alpha \begin{pmatrix} \alpha x_1 x_2 \\ x_2 \end{pmatrix} \\ &\neq \alpha F\left(\mathbf{x}\right) \quad \text{(excepto para } \alpha = 0 \text{ o } 1 \text{ o } x_1 x_2 = 0\text{)} \end{aligned}$$

$\Rightarrow F$ no es una transformación lineal.

2. **Traslación:**

La traslación $T_{\mathbf{a}}$ suma a todo vectir $\mathbf{x}$ un vector fijo $\mathbf{a}$:

$$T_{\mathbf{a}}\left(\mathbf{x}\right) = \mathbf{x} + \mathbf{a}$$

Si $\mathbf{a} \neq \mathbf{0}$, $T_{\mathbf{a}}$ no es una transformación lineal, ya que por ejemplo, $T_{\mathbf{a}}(\mathbf{0}) = \mathbf{a} \neq \mathbf{0}$ y $T_{\mathbf{a}}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{a} + \mathbf{x}_1 + \mathbf{x}_2 \neq T_{\mathbf{a}}(\mathbf{x}_1) + T_{\mathbf{a}}(\mathbf{x}_2)$.

Así, no podrá representarse directamente mediante una matriz aplicada a $\mathbf{x}$.

Problemas 6.1.1

1. (i) Definir la proyección ortogonal en $\mathbb{R}^3$ sobre el plano-xy y mostrar que es una transformación lineal. Graficar.
(ii) Definir la proyección ortogonal en $\mathbb{R}^3$ sobre el eje x y mostrar que es una transformación lineal. Graficar.
(iii) ¿Qué es la proyección al origen? ¿ Puede considerarse una transformación lineal?

2. Considerar la transformación $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x/2 \\ y/3 \end{pmatrix}$$

i) Verificar que es lineal. Expresarla en la forma matricial $L(\mathbf{x}) = A\mathbf{x}$.
ii) Hallar las imágenes $L(\boldsymbol{v})$ de los vectores $\binom{1}{0}$, $\binom{0}{1}$, $\binom{1}{1}$
iii) Dar una interpretación geométrica de L.
iv) La imagen por L de un conjunto de vectores C se define como $L(C) = \{L(\mathbf{v}),\ \mathbf{v} \in C\}$. Probar que la imagen $L(C)$ bajo esta aplicación de la elipse

$$C = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid (x^2/4) + (y^2/9) = 1 \right\}$$

es una circunferencia de radio 1.

Ejemplos de Transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

1. Sea $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ y $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^1$, definida por

$$L(\mathbf{x}) = x_1 + x_2$$

L es una transformación lineal, ya que

$$\begin{aligned} L(\alpha\mathbf{x} + \beta\mathbf{y}) &= (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \\ &= \alpha(x_1 + x_2) + \beta(y_1 + y_2) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \end{aligned}$$

L asocia a cada vector $\mathbf{x} \in \mathbb{R}^2$ un escalar dado por $x_1 + x_2$. Puede ser expresada en forma matricial como $L(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

2. Sea $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $L(\mathbf{x}) = \begin{pmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{pmatrix}$
Se verifica fácilmente que L es lineal (probar!) y que puede ser escrita también como

$$L(\mathbf{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

6.1.2. Ejemplos de transformaciones lineales en otros espacios

1. Dado un espacio vectorial V arbitrario, el operador **identidad** $I : V \longrightarrow V$ se define por

$$I(\mathbf{v}) = \mathbf{v} \quad \text{para todo } \mathbf{v} \in V$$

Es, obviamente, una transformación lineal (verificar!).
Notar que no existe $I : V \longrightarrow W$ si $W \neq V$, aun si V y W tienen la misma dimensión.

2. La transformación nula $0 : V \longrightarrow W$ se define por
$$0\,(\mathbf{v}) = \mathbf{0}_W \quad \text{para todo } \mathbf{v} \in V$$
Es, obviamente, una transformación lineal (verificar!), que generaliza el operador nulo L_0 visto anteriormente.

3. **Transformación definida por una matriz** A.
Dada una matriz A de $m \times n$, se puede definir una transformación lineal asociada $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por
$$L\,(\mathbf{x}) = A\,\mathbf{x}$$
Es fácil ver que L cumple las propiedades de linealidad:
$$\begin{aligned} L\,(\alpha\mathbf{x} + \beta\mathbf{y}) &= A\,(\alpha\mathbf{x} + \beta\mathbf{y}) \\ &= \alpha A\mathbf{x} + \beta A\mathbf{y} \\ &= \alpha L\,(\mathbf{x}) + \beta L\,(\mathbf{y}) \end{aligned}$$
Por lo tanto, cualquier matriz A de $m \times n$ puede verse como asociada a una transformación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Más aun, veremos luego que **toda** transformación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es de la forma anterior (para alguna matriz A de $m \times n$).

4. Sea $L : C_{[a,b]} \longrightarrow \mathbb{R}^1$ definida por
$$L\,(f) = \int_a^b f\,(x)\,dx$$
L es oviamente una transformación lineal, ya que si f y g son dos vectores cualesquiera de $C_{[a,b]}$, entonces
$$\begin{aligned} L\,(\alpha f + \beta g) &= \int_a^b (\alpha f + \beta g)\,(x)\,dx \\ &= \alpha \int_a^b f\,(x)\,dx + \beta \int_a^b g\,(x)\,dx \\ &= \alpha L\,(f) + \beta L\,(g) \end{aligned}$$
A diferencia de las anteriores, esta transformación lineal, cuyo dominio es un espacio vectorial de dimensión infinita, no puede representarse mediante una matriz.

5. Sea $D : C^\infty \longrightarrow C^\infty$ el operador derivada en el espacio C^∞ de funciones reales $f : \mathbb{R} \longrightarrow \mathbb{R}$ derivables a todo orden, definida por
$$D(f) = f'$$
es decir, $D(f)(x) = f'(x)$. Se la suele denotar directamente como $D = \frac{d}{dx}$.
D es obviamente un operador lineal, ya que si f y $g \in C^\infty$,
$$D(\alpha f + \beta q) = (\alpha f + \beta g)' = \alpha f' + \beta g' = \alpha D(f) + \beta D(g)$$
Nótese que D^2 es el operador derivada segunda $\frac{d^2}{dx^2}$:
$$D^2(f) = D(D(f)) = D(f') = f''$$

Dado que C^∞ tiene dimensión infinita, D no puede representarse mediante una matriz (pero si se restringe el dominio a un subespacio de C^∞ de dimensión finita, tal como el espacio P_n de polinomios de grado $\leq n$, D sí podrá representarse mediante una matriz, como veremos luego).

Importante: Si $L : V \to W$ es una transformación lineal, se cumplen siempre las siguientes reglas o propiedades:

1. $L\left(\mathbf{0}_V\right) = \mathbf{0}_W$
2. Si $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$, entonces
$$L\left(\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n\right) = \alpha_1 L\left(\mathbf{v}_1\right) + \cdots + \alpha_n L\left(\mathbf{v}_n\right)$$
3. $$L\left(-\mathbf{v}\right) = -L\left(\mathbf{v}\right) \;\; \forall \;\; \mathbf{v} \in V$$

Se dejan las demostraciones para el lector.

Problema 6.1.2

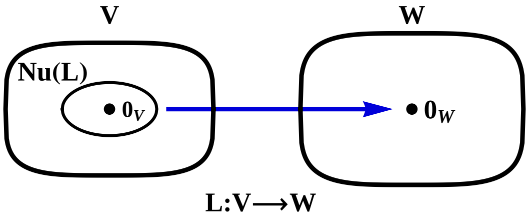
1. Sea $L : V \longrightarrow W$ una transformación lineal y sean $\boldsymbol{w}_1 = L(\boldsymbol{v}_1), \ldots, \boldsymbol{w}_k = L(\boldsymbol{v}_k)$ las imágenes de k vectores $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ de V.
a) Mostrar que si el conjunto de los vectores $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es linealmente dependiente $\Rightarrow \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ es linealmente dependiente.
b) Mostrar que si $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ es linealmente independiente $\Rightarrow$ el conjunto $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ no es necesariamente independiente. Dar un ejemplo (considere proyecciones ortogonales sobre un cierto eje o la transformación nula).

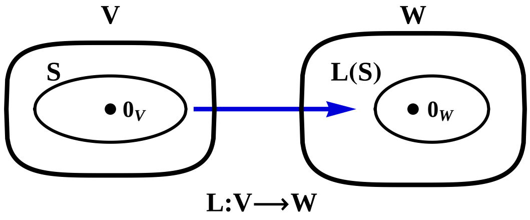
6.2. Imagen y núcleo de una transformación lineal

Sea $L : V \longrightarrow W$ una transformación lineal

1. **Núcleo** de L: es el conjunto de vectores $\mathbf{v}$ de V que son transformados o enviados al vector nulo $\mathbf{0}_W$ de W. Es decir,
$$\mathrm{Nu}\left(L\right) = \left\{\mathbf{v} \in V : L\left(\mathbf{v}\right) = \mathbf{0}_W\right\}$$
2. **Imagen** de un subespacio S de V: es el conjunto de vectores $\mathbf{w}$ de W que son imagen por L de vectores $\mathbf{v}$ de S, es decir,
$$L\left(S\right) = \left\{\mathbf{w} \in W : \mathbf{w} = L\left(\mathbf{v}\right) \text{para algún } \mathbf{v} \in S\right\} = \{L(\mathbf{v}),\ \mathbf{v} \in S\}$$
3. **Imagen de** L: Es la imagen $L\left(V\right)$ de todo el espacio vectorial V:
$$\mathrm{Im}(L) = L(V) = \{L(\mathbf{v}),\ \mathbf{v} \in V\} \subset W$$

Notar que la *imagen* por L del $\mathrm{Nu}(L)$ es el vector nulo $\mathbf{0}_W$ de W: $L(\mathrm{Nu}(L)) = \{\mathbf{0}_W\}$.
Cada uno de estos conjuntos de vectores es un subespacio en los respectivos espacios vectoriales:





Teorema.
Si $L : V \longrightarrow W$ es una transformación lineal, entonces

1. $\mathrm{Nu}\,(L)$ es un subespacio de V
2. Si S es un subespacio de V, $L\,(S)$ es un subespacio de W. Esto implica en particular que la imagen $\mathrm{Im}(L) = L(V)$ es un subespacio de W.
3. Si V es de dimensión finita, la suma de la dimensión de la imagen $\mathrm{Im}(L)$ y la dimensión del núcleo $\mathrm{Nu}(L)$ es la dimensión del espacio V:

$$\dim \mathrm{Im}\,(L) + \dim \mathrm{Nu}\,(L) = \dim V$$

Demostración de 1. En primer lugar, $L(\mathbf{0}_V) = \mathbf{0}_W$, por lo que $\mathbf{0}_V \in \mathrm{Nu}(L)$. Además, si $\mathbf{v}_1$ y $\mathbf{v}_2 \in \mathrm{Nu}\,(L)$,

$$\begin{aligned} L\,(\mathbf{v}_1 + \mathbf{v}_2) &= L\,(\mathbf{v}_1) + L\,(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W \\ L\,(\alpha \mathbf{v}_1) &= \alpha L\,(\mathbf{v}_1) = \mathbf{0}_W \end{aligned}$$

por lo que $\mathbf{v}_1 + \mathbf{v}_2$ y $\alpha \mathbf{v}_1 \in \mathrm{Nu}\,(L)$. El núcleo es pues cerrado por la suma y multiplicación por un escalar, siendo entonces un subespacio.

Demostración de 2. $L(S)$ contiene al $\mathbf{0}_W$ pues $L(\mathbf{0}_V) = \mathbf{0}_W$. Además, si $\mathbf{w}_1$ y $\mathbf{w}_2$ son vectores en $L\,(S)$, existen $\mathbf{v}_1$ y $\mathbf{v}_2$ en S tales que $\mathbf{w}_1 = L(\mathbf{v}_1)$, $\mathbf{w}_2 = L(\mathbf{v}_2)$. Entonces

$$\begin{aligned} \alpha \mathbf{w}_1 &= \alpha L\,(\mathbf{v}_1) = L\,(\alpha \mathbf{v}_1) \quad \text{para } \mathbf{v}_1 \in S \\ \mathbf{w}_1 + \mathbf{w}_2 &= L\,(\mathbf{v}_1) + L\,(\mathbf{v}_2) = L\,(\mathbf{v}_1 + \mathbf{v}_2) \quad \text{para } \mathbf{v}_1 \text{ y } \mathbf{v}_2 \in S \end{aligned}$$

Como S es un subespacio, ambos $\alpha \mathbf{v}_1$ y $\mathbf{v}_1 + \mathbf{v}_2$ pertenecen también a S, por lo que $\alpha \mathbf{w}_1 \in L(S)$ y $\mathbf{w}_1 + \mathbf{w}_2 \in L(S)$. Luego, $L\,(S)$ es cerrado por la suma y multiplicación por un escalar, siendo entonces un subespacio.

Demostración de 3. Partiendo de una base $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ de V tal que $\{\boldsymbol{v}_{k+1}, \ldots, \mathbf{v}_n\}$ es una base de $\mathrm{Nu}(L)$ ($L(\mathbf{v}_{k+1}) = \ldots = L(\mathbf{v}_n) = \mathbf{0}_W$), todo $\mathbf{v} \in V$ puede escribirse como $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k + \alpha_{k+1} \mathbf{v}_{k+1} + \ldots + \alpha_n \mathbf{v}_n$. Entonces,

$$\begin{aligned} L(\mathbf{v}) &= L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k + \alpha_{k+1} \mathbf{v}_{k+1} + \ldots + \alpha_n \mathbf{v}_n) \\ &= \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) + \alpha_{k+1} L(\mathbf{v}_{k+1}) + \ldots + \alpha_n L(\mathbf{v}_n) \\ &= \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) \end{aligned}$$

por lo que $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ **genera** $\mathrm{Im}(L)$. Además, $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es linealmente independiente, pues si $\mathbf{0}_W = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) = L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k) \Rightarrow$
$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k \in \mathrm{Nu}(L)$ y entonces $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k = \beta_1 \mathbf{v}_{k+1} + \ldots + \beta_n \mathbf{v}_n$. Pero por ser los $\mathbf{v}_i$ linealmente independientes, $\alpha_1 = \ldots = \alpha_k = \beta_{k+1} = \ldots = \beta_n = 0$. Por lo tanto,

$$\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = k + (n - k) = n = \dim V$$

Ejemplos 6.2

1. Sea $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

(operador de proyección). Es obvio que $L(\mathbf{x}) = 0$ si y sólo si $x_1 = 0$, es decir, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathrm{Nu}(L)$ si y sólo si $x_1 = 0$. Por lo tanto, $\mathrm{Nu}(L)$ es el subespacio 1-dimensional de $\mathbb{R}^2$ generado por el vector $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, es decir, es el eje x_2, como es obvio geométricamente (graficar!):

$$\mathrm{Nu}(L) = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$$

Por otra parte, dado que $L(\mathbf{x}) = x_1\mathbf{e}_1$, la imagen $\mathrm{Im}(L) = L(V)$ es el conjunto de vectores proporcionales a $\mathbf{e}_1$, es decir, el subespacio 1-dimensional de $\mathbb{R}^2$ generado por el vector $\mathbf{e}_1$, que geométricamente es el eje x_1:

$$\mathrm{Im}(L) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle$$

Se verifica que $\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = 1 + 1 = 2 = \dim V$ $(V = \mathbb{R}^2)$.
Nótese que $L(\mathbf{x}) = A\mathbf{x}$, con $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, y que $\mathrm{Nu}(L)$ coincide con el espacio nulo de $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, mientras que $\mathrm{Im}(L)$ coincide con el espacio columna de A.

2. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$$

Si $\mathbf{x} \in \mathrm{Nu}(L)$, entonces $\left\{ \begin{array}{rcl} x_1 + x_2 & = & 0 \\ x_2 + x_3 & = & 0 \end{array} \right.$. Resolviendo el sistema, si la variable independiente es $x_3 = t$, se tiene $x_2 = -x_3$, $x_1 = x_3$, es decir, $\mathbf{x} = \begin{pmatrix} t \\ -t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$:

$$\mathrm{Nu}(L) = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$$

Por otro lado, vemos que

$$L(\mathbf{x}) = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

con x_1, x_2, x_3 arbitrarios, por lo que la imagen será $\mathbb{R}^2$:

$$\mathrm{Im}(L) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

Se verifica que $\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = 2 + 1 = 3 = \dim V$ $(V = \mathbb{R}^3)$. Nótese también que $L(\mathbf{x}) = A\mathbf{x}$, con $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, coincidiendo $\mathrm{Nu}(L)$ con el espacio nulo de A y $\mathrm{Im}(L)$ con el espacio columna de A (ver problema 6.2.7).

Problemas 6.2

1. Verificar que las siguientes transformaciones $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ son lineales, y determinar $\text{Nu}(L)$, la imagen $\text{Im}(L)$ y sus dimensiones, junto con una base de los mismos:
   (a) $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x+y+z \end{pmatrix}$ **(b)** $L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+y \\ -4x-2y \end{pmatrix}$

2. Idem 1. para las siguientes transformaciones $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:
   (a) $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix}$ **(b)** $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$
   Interprételas geométricamente.

3. **Importante:** Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal definida por una matriz A de $m \times n$:
   $$L(\mathbf{x}) = A\mathbf{x}$$
   Probar que:
   a) El núcleo $\text{Nu}(L)$ es el **espacio nulo** de la matriz A.
   b) La imagen $\text{Im}(L)$ es el **espacio generado por las columnas** de la matriz A (o sea, el espacio columna de la matriz).
   c) La igualdad $\dim \text{Im}(L) + \dim \text{Nu}(L) = \dim V$ es equivalente a
   $$\text{rango } (A) + \text{nulidad } (A) = n$$
   d) Verifique estos resultados para las transformaciones del ejercicio 2., escribiéndolas en la forma $L(\mathbf{x}) = A\mathbf{x}$.

4. Determinar si son lineales las siguientes transformaciones $L : \mathbb{R}^{2\times 2} \longrightarrow \mathbb{R}$. En caso de que lo sean halle su núcleo e imagen.
   (a) $L(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = a + d$ (Traza de la matriz)
   (b) $L(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc$ (Determinante de la matriz)

5. a) Determine si la traza de una matriz cuadrada A de $n \times n$, definida por
   $$Tr(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + \ldots + a_{nn}$$
   es una transformación lineal $T : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$.
   b) Indique si el determinante $\det(A)$ de una matriz cuadrada A de $n \times n$, es una transformación lineal $\det : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$.

6. Halle el núcleo e imagen del operador identidad $I : V \to V$, y del operador nulo $0 : V \to V$.

7. Mostrar que una función $f : \Re \to \Re$ cuya gráfica es una recta no es necesariamente una transformación lineal $L : \Re \to \Re$ (considere por ejemplo la recta $y = mx + 1$).

8. Mostrar que $\forall\, k \geq 1$, la derivada k-ésima $D^k = d^k/dt^k$ en el espacio P de polinomios (de grado arbitrario) es una transformación lineal. ¿Cual es su núcleo e imagen?

9. **Importante: Propiedades geométricas de una transformación lineal**.

   a) Pobar que toda transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ con $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ transforma rectas $R = \{\mathbf{r}_0 + t\mathbf{n},\ t \in \mathbb{R}\}$, $\mathbf{n} \neq \mathbf{0}$, en rectas, y segmentos $Q = \{\mathbf{r}_0 + t\mathbf{n},\ t \in [t_0, t_1]\}$) en segmentos. ¿Qué puede suceder si $\dim \mathrm{Nu}(L) \geq 1$? ¿Siguen siendo válidos estos resultados en $\mathbb{R}^3$? ¿ y en $\mathbb{R}^n$?

   b) Probar que toda transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ con $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ transforma rectas y segmentos paralelos (caracterizados por un mismo vector $\mathbf{n} \neq \mathbf{0}$ pero distintos $\mathbf{r}_0$) en rectas y segmentos paralelos. Generalizar a $\mathbb{R}^3$ y $\mathbb{R}^n$. ¿Puede extenderse el resultado a planos paralelos en $\mathbb{R}^3$?

   c) Si $L(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$, determine la imagen por L de las rectas paralelas $y = 2x$ y $y = 2x + 1$. Grafique e interprete el resultado.

   d) Dados $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, i) probar que el segmento de recta que los conecta es el conjunto $Q = \{t\mathbf{v} + (1-t)\mathbf{u},\ t \in [0, 1]\}$. Verifíquelo para $\mathbf{u} = (1, 0)$, $\mathbf{v} = (0, 1)$. ii) Mostrar que su imagen $L(Q)$ bajo una transformación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ es el segmento de recta entre $L(\mathbf{u})$ y $L(\mathbf{v})$. Generalizar a $\mathbb{R}^3$ y $\mathbb{R}^n$.
   iii) Si $L(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$, determine la imagen por L del segmento de recta entre $\mathbf{u} = (1, 0)$ y $\mathbf{v} = (0, 1)$. Graficar.

   e) Un subconjunto $C \subset \mathbb{R}^n$ es **convexo** si para cualquier par de puntos de C el segmento de recta que los conecta yace enteramente en C, es decir,

   $$\mathbf{u} \in C, \mathbf{v} \in C \ \Rightarrow\ t\mathbf{v} + (1-t)\mathbf{u} \ \in\ C\ \forall\, t \in [0, 1]$$

   Por ejemplo, todo subespacio de $\mathbb{R}^n$ es un conjunto convexo (probar!). Pero un conjunto convexo no necesariamente es un subespacio. Por ejemplo, en $\mathbb{R}^2$ un círculo "lleno" $C = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$ es un conjunto convexo (probar!). En cambio, el círculo $\{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$ no es un conjunto convexo.

   Probar que toda transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ transforma un conjunto convexo en un conjunto convexo. Dé un ejemplo.

6.3. Propiedades fundamentales. Isomorfismos

1. Si $L : V \longrightarrow W$ es una transformación lineal, con V un espacio vectorial de dimensión finita n y $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base arbitraria de V, entonces L queda completamente determinada por los n vectores $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$, es decir, por las n imágenes de los vectores de la base.

   En efecto, si $\mathbf{v} \in V$, entonces

   $$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$$

   y por lo tanto

   $$L(\mathbf{v}) = L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_n L(\mathbf{v}_n)$$

es decir, $L(\mathbf{v})$ es una combinación lineal de los n vectores $L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)$.
La imagen $L(V)$ es entonces el espacio generado por estos n vectores:

$$L(V) = \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n) \rangle$$

Nótese, no obstante, que el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ puede ser linealmente dependiente (por ej., algunos vectores $L(\mathbf{v}_i)$ puden ser nulos), en cuyo caso no será una base de $L(V)$.

2. Una transformación lineal $L : V \longrightarrow W$ es inyectiva (o monomorfismo) si $L(\mathbf{v}_1) \neq L(\mathbf{v}_2)\ \forall\ \boldsymbol{v}_1 \neq \boldsymbol{v}_2$. Es fácil ver que es inyectiva **si y sólo si** $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$.

   En efecto, L inyectiva $\Rightarrow L(\mathbf{v}) \neq L(\mathbf{0_V}) = \mathbf{0}_W\ \forall\ \mathbf{v} \neq \mathbf{0}_V$, por lo que $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$. Y si $\mathrm{Nu}(L) = \{\mathbf{0}_V\} \Rightarrow L(\mathbf{v}_1) - L(\mathbf{v}_2) = L(\mathbf{v}_1 - \mathbf{v}_2) \neq \mathbf{0}_W\ \forall\ \mathbf{v}_1 \neq \mathbf{v}_2$, por lo que L es inyectiva.

3. Si $\mathrm{Nu}(L) = \{\mathbf{0}_V\}$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ es linealmente independiente, el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es **linealmente independiente**. En otras palabras, si la transformación lineal L es inyectiva, **conserva la independencia lineal**.
   En efecto, si

$$\mathbf{0}_W = \alpha_1 L(\mathbf{v}_1) + \ldots + \alpha_k L(\mathbf{v}_k) = L(\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k)$$

   entonces $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k \in \mathrm{Nu}(L)$, por lo que $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k = \mathbf{0}_V$. Pero esto implica $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$ por ser $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ linealmente independiente. Por lo tanto, $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_k)\}$ es **linealmente independiente**.

   En particular, si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de $V \Rightarrow \{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ es una **base** de la imagen $\mathrm{Im}(L) = L(V)$, pues por lo anterior, es linealmente independiente y por la propiedad 1, **genera** la imagen. Por lo tanto, para $L : V \to W$ inyectiva, $\dim \mathrm{Im}(L) = n$ y $\dim \mathrm{Nu}(L) = 0$, verificándose que

$$\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = n + 0 = n = \dim V$$

   Esto también implica que si $L : V \to W$ es inyectiva, necesariamente $\dim V \leq\ \dim W$ pues $\mathrm{Im}(V) \subset W$.

4. Si $L : V \longrightarrow W$ es una transformación lineal y $L(V) = W \Rightarrow L$ es sobreyectiva (epimorfismo). En este caso la imagen $L(V)$ es todo el codominio W, es decir, cubre todo W, por lo que el conjunto $\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$ **genera** W.

   Además, en este caso se cumple que

$$\dim V = \dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = \dim W + \dim \mathrm{Nu}(L)$$

   por lo que necesariamente $\dim V \geq \dim W$.

Isomorfismos:

5. Si $L : V \Rightarrow W$ es una transformación lineal **biyectiva**, es decir, que es a la vez inyectiva ($\text{Nu}(L) = \{\mathbf{0}_V\}$) y sobreyectiva ($L(V) = W$) $\Rightarrow$ se dice que L es un **isomorfismo**. Si $V = W$ al isomorfismo se lo denota **automorfismo**.

   Si L es un isomorfismo y $\dim V = n$, con $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base de $V \Rightarrow$

   $$\{L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\}$$

   es una **base** de W, pues son linealmente independientes (por ser L inyectiva) y generan W (por ser L sobreyectiva). Un isomorfismo transforma entonces cualquier **base** de V en una **base** de W.

   Por lo tanto, V y W **deben tener la misma dimensión** n (cuando son de dimensión finita). Para un isomorfismo se verifica entonces que

   $$\dim \text{Im}(L) + \dim \text{Nu}(L) = n + 0 = \dim V = \dim W$$

6. Una función $L : V \longrightarrow W$ tiene inversa $L^{-1} : W \longrightarrow V$ si y sólo si L es biyectiva. Por lo tanto, si $L : V \longrightarrow W$ es una transformación lineal, L tendrá inversa $L^{-1} : W \longrightarrow V$ **si y sólo si** L **es un isomorfismo:**

   $$L(\mathbf{v}) = \mathbf{w} \Rightarrow \mathbf{v} = L^{-1}(\mathbf{w})$$

   cumpliéndose que

   $$L(L^{-1}(\mathbf{w})) = L(\mathbf{v}) = \mathbf{w} \quad \forall\, \mathbf{w} \in W$$
   $$L^{-1}(L(\mathbf{v})) = L^{-1}(\mathbf{w}) = \mathbf{v} \quad \forall\, \mathbf{v} \in V$$

   es decir,

   $$LL^{-1} = I_W, \quad L^{-1}L = I_V$$

   donde I_W y I_V son los operadores identidad en W y V.
   La inversa $L^{-1} : W \longrightarrow V$ de un isomorfismo L **es también un isomorfismo** (probar!), es decir, una transformación lineal biyectiva.

7. Se dice que dos espacios vectoriales V, W son **isomorfos** si existe un isomorfismo $L : V \longrightarrow W$ entre ellos. Si V y W son de dimensión finita $\Rightarrow$ V y W son isomorfos **si y sólo si tienen la misma dimensión** (probar!).

Ejemplo 6.3 Sea $L : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ la transformación dada por

$$L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$$

Se verifica en primer lugar que L es lineal y que $\forall\, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \in \mathbb{R}^2$,

$$L(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_2 \\ -x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = x_1 L(\mathbf{e}_1) + x_2 L(\mathbf{e}_2)$$

con $L(\mathbf{e}_1) = \begin{pmatrix}1\\1\end{pmatrix}$, $L(\mathbf{e}_2) = \begin{pmatrix}1\\-1\end{pmatrix}$, por lo que basta conocer $L(\mathbf{e}_1)$ y $L(\mathbf{e}_2)$ para determinar $L(\mathbf{x})$ para cualquier $\mathbf{x} \in \mathbb{R}^2$.

Se comprueba también que $\mathrm{Nu}(L) = \mathbf{0}$, pues si $x_1+x_2 = 0$ y $x_1-x_2 = 0 \Rightarrow x_1 = x_2 = 0$ (verificar!). Por lo tanto L es **inyectiva**. Y finalmente, vemos que la imagen de L es

$$L(V) = \{x_1 \begin{pmatrix}1\\1\end{pmatrix} + x_2 \begin{pmatrix}1\\-1\end{pmatrix},\ x_1, x_2 \in \mathbb{R}\} = \left\langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix} \right\rangle = \mathbb{R}^2$$

por lo que L es también **sobreyectiva**. Este resultado puede también obtenerse directamente de

$$\dim \mathrm{Im}(L) = \dim V - \dim \mathrm{Nu}(L) = 2 - 0 = 2$$

L es entonces un **automorfismo**, es decir, un **isomorfismo** entre V y V, con $V = \mathbb{R}^2$.

L tendrá entonces inversa $L^{-1} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, dada por (verificar!)

$$L^{-1} \begin{pmatrix}x_1\\x_2\end{pmatrix} = \frac{1}{2} \begin{pmatrix}x_1 + x_2\\x_1 - x_2\end{pmatrix}$$

Notemos finalmente que podemos expresar $L(\mathbf{x})$ como

$$L(\mathbf{x}) = \begin{pmatrix}1 & 1\\1 & -1\end{pmatrix} \begin{pmatrix}x_1\\x_2\end{pmatrix}$$

y $L^{-1}(\mathbf{x})$ como

$$L^{-1}(\mathbf{x}) = \frac{1}{2} \begin{pmatrix}1 & 1\\1 & -1\end{pmatrix} \begin{pmatrix}x_1\\x_2\end{pmatrix}$$

siendo la matriz que representa a L^{-1} la **inversa** de la matriz que representa a L.

Problemas 6.3

1. Si $L : V \longrightarrow V$ es un operador lineal y $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de V, probar
   (a) Si $L(\mathbf{v}_i) = \mathbf{0}$ para cada elemento de la base entonces L es la transformación lineal nula.
   (b) Si $L(\mathbf{v}_i) = \mathbf{v}_i$ para cada vector de la base entonces L es la identidad.
   (c) Si existe un escalar r tal que $L(\mathbf{v}_i) = r\mathbf{v}_i$ para cada vector de la base entonces $L(\mathbf{v}) = r\mathbf{v}$ para todo los vectores en V.

2. Sea $L\colon V \to W$ una transformación lineal y supongamos que $L(\mathbf{v}_1) = \mathbf{w}_1, \ldots, L(\mathbf{v}_k) = \mathbf{w}_k$ para vectores $\mathbf{v}_1, \ldots, \mathbf{v}_k$ de V.
   (a) Si $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ genera W, ¿debe $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ generar V? Pensar, por ejemplo, en transformaciones de $\Re^3$, sobre $\Re^2$.
   (b) Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ genera V, ¿debe $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ generar W? Pensar por ejemplo en $L : \Re^2 \longrightarrow \Re^3$.
   (c) Si ahora L es un isomorfismo (que implica dim V = dim W en espacios de dimensión finita) ¿ cual es la respuesta a (a) y (b)?

3. Si $L : V \to W$ es inyectiva, mostrar que la imagen $L(S)$ de un subespacio de dimensión m de V es un subespacio de dimensión m de W.

4. **Importante:** Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal definida por una matriz A de $m \times n$:
$$L(\mathbf{x}) = A\mathbf{x}$$
probar que:
a) L es inyectiva si y sólo si rango$(A) = n$. Esto implica $n \leq m$.
b) L es sobreyectiva si y sólo si rango$(A) = m$. Esto implica $n \geq m$.
c) L es biyectiva (isomorfismo) si y sólo si rango$(A) = m = n$, es decir, si y sólo si A es una matriz cuadrada no singular.
d) Probar que en c), $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ está dada por
$$L^{-1}(\mathbf{x}) = A^{-1}(\mathbf{x})$$
e) Discuta las implicancias de los resultados anteriores para el sistema de ecuaciones lineales (de $m \times n$)
$$L(\mathbf{x}) = \mathbf{b}$$
En particular, muestre que
i) Es compatible si y sólo si $\mathbf{b} \in \text{Im}(L)$.
ii) Es compatible $\forall\, \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si L es sobreyectiva
iii) La solución, cuando existe, es única si y sólo si L es inyectiva
iv) La solución existe y es única $\forall\, \boldsymbol{b} \in \mathbb{R}^m$ si y sólo si L es biyectiva (isomorfismo), es decir si y sólo si $m = n$ y A es no singular.

5. **Importante:** Sea V un espacio vectorial de dimensión n, con $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ una base (ordenada) de V, tal que si $\mathbf{v} \in V$,
$$\mathbf{v} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$$
Sea $L : V \longrightarrow \mathbb{R}^n$ la transformación lineal definida por
$$L(\mathbf{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
Es decir, $L(\mathbf{v}) = [\mathbf{v}]_B$ es el vector columna de coordenadas de $\mathbf{v}$ en la base B.
a) Mostrar que L está bien definida, es decir, que $[\mathbf{v}]_B$ existe y es único $\forall\, \mathbf{v} \in V$.
b) Mostrar que L es una transformación lineal.
c) Mostrar que L es un isomorfismo (es decir, $\text{Nu}(L) = \{\mathbf{0}_V\}$, $\text{Im}(L) = \mathbb{R}^n$).
Este resultado muestra en forma explícita que todo espacio V de dimensión n (con escalares reales) es isomorfo a $\mathbb{R}^n$, es decir, que existe una correspondencia biyectiva entre ambos.
d) ¿Cual es la inversa L^{-1}?

6. Mostrar que dos espacios V, W son isomorfos si y sólo si tienen la misma dimensión.

7. Mostrar que la inversa de un isomorfismo $L : V \longrightarrow W$ es un isomorfismo, es decir, que L^{-1} es lineal y biyectiva.

6.4. Representación matricial de transformaciones lineales

Veremos ahora que cualquier transformación lineal L entre espacios vectoriales de dimensión finita V y W se puede representar mediante una matriz A, que dependerá de las bases que se consideren en V y W.

En primer lugar consideramos transformaciones de $\mathbb{R}^n$ en $\mathbb{R}^m$:

$$L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

Asumimos primero que la base en $V = \mathbb{R}^n$ es la base canónica $B_c = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. Dado $\mathbf{x} \in \mathbb{R}^n$, podemos representarlo como

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$$

Como L es lineal,

$$L(\mathbf{x}) \;=\; x_1 L(\mathbf{e}_1) + \ldots + x_n L(\mathbf{e}_n)$$

Luego si para cada $\mathbf{e}_j$, $j = 1, \ldots, n$, se tiene

$$L(\mathbf{e}_j) = \mathbf{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

entonces

$$\begin{aligned} L(\mathbf{x}) &= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

es decir,

$$L(\mathbf{x}) = A\mathbf{x}$$

con A la matriz de $\boldsymbol{m \times n}$

$$A = (L(\mathbf{e}_1), \ldots, L(\mathbf{e}_n)) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

La matriz A se denomina **representación matricial** de L en la bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. Queda completamente determinada por las n imágenes $L(\mathbf{e}_j)$ de los vectores de la base canónica. Se emplea también la notación

$$A = [L]_{B_c}^{B_c}$$

o simplemente $A = [L]_{B_c}$. La representación de L con respecto a otras bases será una matriz diferente, pero también de $m \times n$.

Ejemplos 6.4

1. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, definida por $L(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}$ (ejemplo anterior). Tenemos

$$\mathbf{a}_1 = L(\mathbf{e}_1) = L\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{a}_2 = L\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}, \quad \mathbf{a}_3 = L\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$$

Por lo tanto,

$$A = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}$$

Verificación:

$$A\mathbf{x} = \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_1 + x_2\\ x_2 + x_3\end{pmatrix} = L(\mathbf{x})$$

2. Dada A de $m \times n$, consideremos la transformación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por

$$L(\mathbf{x}) = A\mathbf{x}$$

Es fácil ver que

$$L(\mathbf{e}_j) = A\mathbf{e}_j = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mm} \end{pmatrix}\begin{pmatrix}0\\ \vdots \\ 1 \\ \vdots \\ 0\end{pmatrix} = \begin{pmatrix}a_{1j}\\ \vdots \\ a_{mj}\end{pmatrix} = \mathbf{a}_j$$

con $\mathbf{a}_j$ **la columna j-ésima de** A, por lo que la representación matricial de L en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ es precisamente A:

$$[L]_{B_c}^{B_c} = (L(\mathbf{e}_1), \dots, L(\mathbf{e}_2)) = A$$

Por ejemplo, si $A = \begin{pmatrix} a & b \\ c & d\end{pmatrix}$,

$$L(\mathbf{x}) = \begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}ax_1 + bx_2\\ cx_1 + dx_2\end{pmatrix}$$

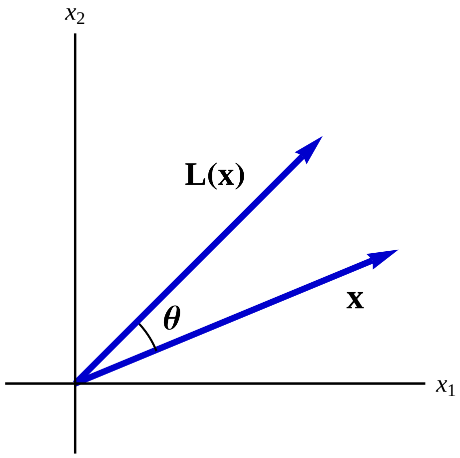
y entonces

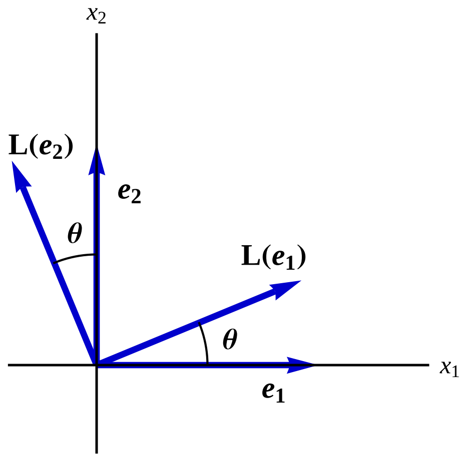
$$\begin{aligned} L(\mathbf{e}_1) &= \begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}a\\c\end{pmatrix} = \mathbf{a}_1 \\ L(\mathbf{e}_2) &= \begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}b\\d\end{pmatrix} = \mathbf{a}_2 \end{aligned}$$

por lo que $(\mathbf{a}_1, \mathbf{a}_2) = A$.

El problema 6.2.3 implica entonces que la imagen $\mathrm{Im}(L)$ de una transformación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ es el espacio columna de $A = [L]_{B_c}^{B_c}$ (es decir, el espacio generado por los vectores columna de A) mientras que el núcleo $\mathrm{Nu}(L)$ es espacio nulo de A.

3. Rotación general en $\mathbb{R}^2$. Tomemos la transformación $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, tal que a cada vector $\mathbf{x}$ lo hace rotar un ángulo θ, en sentido antihorario:





Tenemos

$$L\left(\mathbf{e}_1\right)=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \quad \text{y} \quad L\left(\mathbf{e}_2\right)=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$$

Por lo tanto

$$A=\left(L\left(\mathbf{e}_1\right),L\left(\mathbf{e}_2\right)\right)=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

Entonces, para rotar un ángulo θ un vector $\mathbf{x}$ de $\mathbb{R}^2$ en sentido antihorario, se debe multiplicar A por $\mathbf{x}$:

$$\mathbf{y}=L(\mathbf{x})=A\mathbf{x}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}x_1\cos\theta-x_2\sin\theta\\ x_1\sin\theta+x_2\cos\theta\end{pmatrix}$$

Problemas 6.4

1. Hallar la matriz que representa, respecto de las bases canónicas, las siguientes transformaciones lineales y escribir la forma explícita de $L(\mathbf{x})$:
   (a) La aplicación $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ que representa una dilatación con factor c_1 a lo largo del eje x y c_2 a lo largo del eje y.
   (b) La reflexión $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ respecto de la recta $y = x$.
   (c) La rotación $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ de ángulo $\pi/4$ antihorario.
   (d) La rotación $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ de ángulo θ antihorario alrededor del eje z.
   (e) Idem anterior pero alrededor del eje y.
   (f) La proyección $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ sobre el plano xy.
   (g) $L : \mathbb{R}^2 \to \mathbb{R}^2$ lineal, definida por $L\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$, $L\begin{pmatrix}0\\2\end{pmatrix}=\begin{pmatrix}-2\\2\end{pmatrix}$.
   (h) $L : \mathbb{R}^3 \to \mathbb{R}^2$ lineal, definida por $L\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$, $L\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}1\\-1\end{pmatrix}$, $L\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$.
   (i) Determinar la imagen y núcleo de las transformaciones anteriores.

2. Mostrar, determinando las imágenes $L(\mathbf{e}_1)$ y $L(\mathbf{e}_2)$, que la siguiente matriz

$$A=\begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}$$

representa en la base canónica la reflexión $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecto de la recta $y = mx$ que forma un ángulo θ con el eje x ($m = \tan\theta$). Verificar resultados previos.

6.4.1. Caso general:

Consideremos una transformación lineal general $L : V \longrightarrow W$ entre espacios vectoriales V y W de dimensión n y m respectivamente. Sean

$$B_V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)\,, \qquad B_W = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$$

bases ordenadas de V y W. Cualquier vector $\mathbf{v} \in V$ puede escribirse en términos de los vectores de la base B_V:

$$\mathbf{v} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$$

siendo $[\mathbf{v}]_{B_V} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{x}$ el vector de coordenadas de $\mathbf{v}$ con respecto a la base B_V.

Por lo tanto, para L lineal,

$$L(\mathbf{v}) = x_1 L(\mathbf{v}_1) + \ldots + x_n L(\mathbf{v}_n)$$

Si ahora escribimos $L(\mathbf{v})$ y $L(\mathbf{v}_j)$ en términos de los vectores de la base B_W,

$$\begin{aligned} L(\mathbf{v}) &= y_1\mathbf{w}_1 + \ldots + y_m\mathbf{w}_m \\ L(\mathbf{v}_j) &= a_{1j}\mathbf{w}_1 + \ldots + a_{mj}\mathbf{w}_m, \quad j = 1, \ldots, m \end{aligned}$$

tal que sus vectores de coordenadas en la base B_W son

$$[L(\mathbf{v})]_{B_W} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \mathbf{y}\,, \quad [L(\mathbf{v}_j)]_{B_W} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \mathbf{a}_j$$

obtenemos

$$\begin{aligned} \mathbf{y} = [L(\mathbf{v})]_{B_W} &= x_1[L(\mathbf{v}_1)]_{B_W} + \ldots + x_n[L(\mathbf{v}_n)]_{B_W} \\ &= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= A\mathbf{x} \end{aligned}$$

es decir,

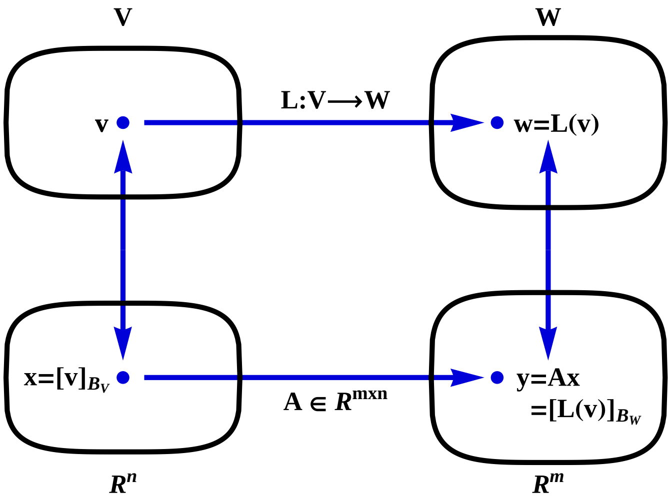
$$[L(\mathbf{v})]_{B_W} \;=\; A[\mathbf{v}]_{B_V}$$

donde A es la matriz de $m \times n$

$$A = ([L(\mathbf{v}_1)]_{B_W}, \ldots, [L(\mathbf{v}_n)]_{B_W}) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

Esta matriz A se denomina **representación matricial de L respecto de las bases B_V de V y B_W de W**. La denotaremos también como $A = [L]_{B_W}^{B_V}$.

El resultado se puede resumir en el siguiente esquema gráfico:



Ejemplos 6.4.1

1. Sea $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la transformación lineal definida por

$$L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\,\mathbf{b}_2$$

donde $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ y

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Tenemos

$$\begin{aligned} L(\mathbf{e}_1) &= \mathbf{b}_1 = 1\mathbf{b}_1 + 0\mathbf{b}_2 \\ L(\mathbf{e}_2) = L(\mathbf{e}_3) &= \mathbf{b}_2 = 0\mathbf{b}_1 + 1\mathbf{b}_2 \end{aligned}$$

Por lo tanto, la representación matricial de L con respecto a las bases $B_c = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de $V = \mathbb{R}^3$ y $B_W = (\mathbf{b}_1, \mathbf{b}_2)$ de $W = \mathbb{R}^2$ es

$$A = [L]_{B_W}^{B_c} = ([L(\mathbf{e}_1)]_{B_W}), [L(\mathbf{e}_2)]_{B_W}, [L(\mathbf{e}_3)]_{B_W}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Así,

$$[L(\mathbf{x})]_{B_W} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 + x_3 \end{pmatrix}$$

que implica justamente $L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$.
En cambio, en la base canónica de $\mathbb{R}^2$ obtenemos (probar!)

$$A_c = [L]_{B_c}^{B_c} = (L(\mathbf{e}_1), L(\mathbf{e}_2), L(\mathbf{e}_3)) = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

con

$$[L(\mathbf{x})]_{B_c} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 - x_3 \\ x_1 + x_2 + x_3 \end{pmatrix}$$

es decir, $L(\mathbf{x}) = (x_1 - x_2 - x_3)\mathbf{e}_1 + (x_1 + x_2 + x_3)\mathbf{e}_2$, que coincide con $x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$.

2. Sea $D : P_2 \to P_2$ el operador derivada $D = \frac{d}{dx}$ en el espacio de polinomios de grado ≤ 2. Como base ordenada de P_2 consideramos $B = (1, x, x^2)$. Tenemos

$$D(1) = 0, \quad D(x) = 1, \quad D(x^2) = 2x$$

Por lo tanto, su representación matricial en la base B es

$$A = [D]_B^B = ([D(1)]_B, [D(x)]_B, [D(x^2)]_B) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Para $p(x) = a_0 + a_1 x + a_2 x^2 \in P_2$ tenemos $[p]_B = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ y entonces

$$[D(p)]_B = A[p]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 0 \end{pmatrix}$$

que implica $D(p) = a_1 + 2a_2 x + 0x^2 = a_1 + 2a_2 x$. Al ser P_2 de dimensión 3, todo operador lineal en P_2 puede entonces ser representado por una matriz de 3×3, tal como un operador en $\mathbb{R}^3$.

Relación entre las propiedades de L y la matriz de representación

Sea $L : V \to W$ una transformación lineal, con dim $V = n$, dim $W = m$, y

$$A = [L]_{B_W}^{B_V}$$

la matriz que la representa con respecto a bases B_V de V y B_W de W. El rango y la nulidad de A **no dependen** de las bases elegidas, pues coinciden, respectivamente, con la dimensión de la imagen y del núcleo de L (los que no dependen de la representación).

En efecto, los n vectores columna $\mathbf{a}_j \in \mathbb{R}^m$ de A son las coordenadas de las n imágenes $L(\mathbf{v}_j) \in W$ en la base B_W. Como la correspondencia entre vectores y sus coordenadas en una base es un isomorfismo (Problema 6.3.5), la dimensión del subespacio de $\mathbb{R}^m$ generado por las n columnas de A (es decir, el rango de A) **es la misma** que la del subespacio de W generado por las n imágenes $L(\mathbf{v}_j)$, es decir, la misma que la dimensión de la imagen $\mathrm{Im}(L)$. Análogamente, los vectores $\mathbf{x} \in \mathbb{R}^n$ del espacio nulo de A ($A\mathbf{x} = \mathbf{0}$) son las coordenadas de los vectores $\mathbf{v} \in V$ del núcleo $\mathrm{Nu}(L)$ ($L(\mathbf{v}) = \mathbf{0}_W$), y por lo tanto la dimensión del espacio nulo de A (la nulidad de A) coincidirá con la del núcleo de L.

Se cumple entonces que

$$\mathrm{rango}(A) = \dim \mathrm{Im}(L)\,, \quad \mathrm{nulidad}(A) = \dim \mathrm{Nu}(L)$$

para **cualquier representación matricial** $A = [L]_{B_V}^{B_W}$ de L. Por lo tanto, la relación

$$\dim \mathrm{Im}(L) + \dim \mathrm{Nu}(L) = \dim V$$

resulta equivalente a

$$\mathrm{rango}\,(A) + \mathrm{nulidad}\,(A) = n$$

Los vectores $\in \mathbb{R}^m$ de una base del espacio columna de A son las coordenadas en la base B_W de los vectores de una base de la imagen $\text{Im}(L)$, mientras que los vectores $\in \mathbb{R}^n$ de una base del espacio nulo de A son las coordenadas en la base B_V de los vectores de una base del núcleo $\text{Nu}(L)$. Por lo tanto, pueden obtenerse bases de estos espacios mediante los métodos ya conocidos para obtener bases de los espacios columna y nulo de una matriz. Y la ecuación $L(\mathbf{v}) = \mathbf{w}$ es equivalente al sistema lineal $A[\mathbf{v}]_{B_V} = [\mathbf{w}]_{B_W}$.

Problemas 6.4 (continuación)

1. 3. A partir de consideraciones geométricas, encontrar la representación matricial $[L]_B^B$ respecto de la base $B = (\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix})$ de la transformación lineal $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que refleja todo vector respecto de la recta $y = x$. Mostrar que dicha matriz es diagonal. Comparar con la matriz que la representa en la base canónica $[L]_{B_c}^{B_c}$.
2. 4. Sea $D : P_3 \to P_3$ el operador derivada en el espacio de polinomios de grado ≤ 3.
   a) Determine su representación matricial A en la base $B = (1, x, x^2, x^3)$.
   b) Halle el núcleo y la imagen de D en este espacio, y su relación con el espacio columna y nulo de A.

6.5. Cambio de base

Determinaremos ahora en forma explícita la relación entre las representaciones matriciales de una transformación lineal L en distintas bases. La ventaja de "cambiar de base" es la posibilidad de encontrar una base en la que la matriz representativa de L sea simple (por ejemplo, diagonal) y por lo tanto sus propiedades esenciales (rango, imagen, etc.) resulten evidentes.

Consideremos primero un operador lineal $L : V \to V$, con V de dimensión finita n. Recordemos que si $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ y $B' = (\mathbf{v}'_1, \ldots, \mathbf{v}'_n)$ son dos bases de V, podemos escribir cualquier vector $\mathbf{v} \in V$ en las formas

$$\begin{aligned} \mathbf{v} &= x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n \\ &= x'_1\mathbf{v}'_1 + \ldots + x'_n\mathbf{v}'_n \end{aligned}$$

Las coordenadas $\mathbf{x} = [\mathbf{v}]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ y $\mathbf{x}' = [\mathbf{v}]_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ en estas bases se relacionan por

$$\mathbf{x}' = S^{-1}\mathbf{x}$$

o en forma equivalente,

$$\mathbf{x} = S\mathbf{x}'$$

donde

$$S = \big([\mathbf{v}'_1]_B, \quad \ldots, \quad [\mathbf{v}'_n]_B\big)$$

es la matriz de cambio de base (o matriz de transición), de $n \times n$ y no singular, formada por las coordenadas de los vectores de la base B' en la base B. Por lo tanto, escribiendo

$$\begin{aligned} L(\mathbf{v}) &= y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n \\ &= y'_1\mathbf{v}'_1 + \ldots + y'_n\mathbf{v}'_n \end{aligned}$$

con $\mathbf{y} = [L(\mathbf{v})]_B = A\mathbf{x} = A[\mathbf{v}]_B$, $\mathbf{y}' = [L(\mathbf{v})]_{B'} = A'\mathbf{x}' = A'[\mathbf{v}]_{B'}$, obtenemos

$$\begin{aligned}\mathbf{y}' &= S^{-1}\mathbf{y}\\ &= S^{-1}A\mathbf{x}\\ &= S^{-1}AS\mathbf{x}'\end{aligned}$$

es decir,

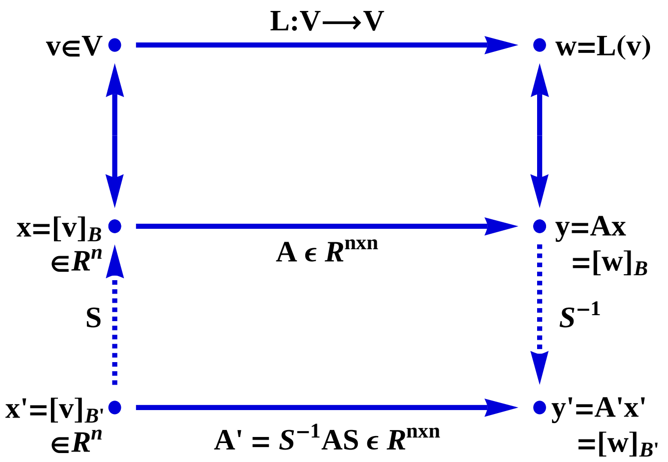
Si $A = [L]_B^B$ es la representación de L en la base B y $A' = [L]_{B'}^{B'}$ su representación en la base B', las matrices A y A' se relacionan por

$$A' = S^{-1}AS$$

donde $S = \big([\mathbf{v}_1']_B \quad \ldots \quad [\mathbf{v}_n']_B\big)$ es la matriz de cambio de base.
Notar que si se conoce A', la relación anterior permite obtener A como (probar!)

$$A = SA'S^{-1}$$

Podemos resumir el resultado en el siguiente esquema gráfico:



Ejemplo 6.5.1 Consideremos nuevamente la reflexión $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecto de la recta $y = x$. En la base $B' = (\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix})$, formada por el vector $\mathbf{v}_1' = \begin{pmatrix}1\\1\end{pmatrix}$ perteneciente a la recta y el vector $\mathbf{v}_2' = \begin{pmatrix}-1\\1\end{pmatrix}$ perpendicular a la recta, la representacion es obvia, ya que $L(\mathbf{v}_1') = \mathbf{v}_1'$ mientras que $L(\mathbf{v}_2') = -\mathbf{v}_2'$ (véase problema 6.4.3):

$$A' = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Para hallar la representación matricial A de L en la base canónica $B = (\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix})$, determinamos primero la correspondiente matriz de cambio de base,

$$S = ([\mathbf{v}_1']_B, [\mathbf{v}_2']_B) = \begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$$

y su inversa $S^{-1}=\frac{1}{2}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$. Luego

$$\begin{aligned} A=SA'S^{-1} &= \begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1/2 & 1/2\\ -1/2 & 1/2\end{pmatrix}\\ &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\end{aligned}$$

Este resultado coincide con el hallado previamente (problema 6.4.1 (b)).

6.5.1. Matrices semejantes

Dos matrices A y A' de $n\times n$ son **semejantes** si y sólo si existe una matriz no singular S tal que

$$A'=S^{-1}AS$$

Si A' es semejante a A, entonces, multiplicando la igualdad anterior por S a izquierda y por S^{-1} a derecha, obtenemos

$$\begin{aligned} SA'S^{-1} &= S(S^{-1}AS)S^{-1}=(SS^{-1})A(SS^{-1})\\ &= A\end{aligned}$$

es decir, $A=R^{-1}A'R$, con $R=S^{-1}$, por lo que A es también semejante a A'.

Por lo tanto, las matrices A y A' que representan a un operador lineal L en dos bases distintas son **semejantes**.

Dos propiedades fundamentales sobre matrices semejantes son las siguientes:

1. Las matrices semejantes tienen el mismo determinante:

$$\det(A')=\det(A)$$

En efecto,

$$\begin{aligned} \det(A') &= \det(S^{-1}AS)\\ &= \det(S^{-1})\det(A)\det(S)\\ &= (\det(S))^{-1}\det(A)\det(S)\\ &= \det(A)\end{aligned}$$

Esto implica que el determinante $\det(A)$ **es una propiedad del operador lineal** L **representado por** A, permaneciendo **invariante** frente a cambios de base.

Así, comprobamos en el ejemplo 6.5.1 que $\det(A')=\det(A)=-1$.

2. Las matrices semejantes tienen la misma traza:

$$\mathrm{Tr}\, A' = \mathrm{Tr}\, A$$

donde la traza $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$ es la suma de los elementos diagonales.

En efecto, probemos primero que para matrices A de $n \times m$ y B de $m \times n$, se cumple

$$\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$$

Demostración:

$$\mathrm{Tr}AB = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}(\sum_{j=1}^{m} a_{ij}b_{ji}) = \sum_{j=1}^{m}(\sum_{i=1}^{n} b_{ji}a_{ij}) = \sum_{j=1}^{m}(BA)_{jj} = \mathrm{Tr}BA$$

Luego

$$\mathrm{Tr}\, A' = \mathrm{Tr}\,(SAS^{-1}) = \mathrm{Tr}\,((AS^{-1})S) = \mathrm{Tr}\,(A(S^{-1}S)) = \mathrm{Tr}\, A$$

Este resultado implica que la traza es también una propiedad del operador L representado por A, permaneciendo **invariante** frente a cambios de base.

Así, comprobamos en el ejemplo 6.5.1 que $\mathrm{Tr}\,(A') = \mathrm{Tr}(A) = 0$.

Finalmente, mencionemos una tercera propiedad relacionada con 1.:

3. Si A y A' son matrices semejantes de $n \times n$ e I_n es la matriz identidad, entonces

$$\det\,(A' - \lambda I_n) = \det(A - \lambda I_n)$$

para cualquier escalar λ.

Este determinante (que es un polinomio de grado n en λ denominado **polinomio característico**) juega un rol central en la teoría de autovalores, como veremos luego.
Demostración:

$$\begin{aligned}\det(A' - \lambda I_n) &= \det(S^{-1}A'S - \lambda I_n) = \det[S^{-1}(A - \lambda I_n)S] \\ &= (\det(S^{-1})\det(A - \lambda I_n)\det(S) \\ &= \det(A - \lambda I_n)\end{aligned}$$

Así, comprobamos en el ejemplo 6.5.1 que

$$\begin{aligned}\det(A' - \lambda I_2) &= \begin{vmatrix} 1-\lambda & 0 \\ 0 & -1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) = \lambda^2 - 1 \\ \det(A - \lambda I_2) &= \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = \det(A' - \lambda I_2)\end{aligned}$$

Problemas 6.5

1. Dada $L:\mathbb{R}^3\longrightarrow\mathbb{R}^3$ definida por

$$L(\mathbf{x})=A\mathbf{x}\,,\quad A=\begin{pmatrix}2&2&0\\1&1&2\\1&1&2\end{pmatrix}$$

a) Obtener la expresión explícita de $L(\mathbf{x})$ para un vector $\mathbf{x}=\begin{pmatrix}x_1\\x_2\\x_2\end{pmatrix}$.

b) Hallar su representación A' en la base $B'=\left(\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}-2\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right)$.

Verificar que $L(\mathbf{v}_1')=0$, $L(\mathbf{v}_2')=\mathbf{v}_2'$, y $L(\mathbf{v}_3')=4\mathbf{v}_3'$, y que por lo tanto, A' es diagonal.
c) Usar esta representación diagonal para hallar su núcleo e imagen.
d) Usar esta representación diagonal para interpretar L geométricamente.

e) Indique si el vector $\mathbf{v}=\begin{pmatrix}0\\1\\1\end{pmatrix}$ pertenece a la imagen de L.

f) Verifique que la traza y determinante permanecen invariantes frente al cambio de base: $\mathrm{Tr}(A)=\mathrm{Tr}(A')$, $\det(A)=\det(A')$.

2. Si la matriz A del problema anterior representa ahora una transformación $L:P_2\to P_2$ en la base canónica de P_2, con P_2 el espacio de polinomios reales de grado ≤ 2 (y su base canónica dada por $(1,x,x^2)$), dé una expresión para $L(p)$ y determine la base B' de P_2 donde la representación A' es diagonal (utilice resultados previos).

3. Considere la reflexión $L:\mathbb{R}^2\to\mathbb{R}^2$ respecto de una recta que pasa por el origen y que forma un ángulo θ con el eje x, dada por $y=mx$, con $m=\tan\theta$ (graficar).
a) Determine, a partir de consideraciones geométricas, su representación matricial A' en la base $B'=\left(\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix},\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}\right)$ formada por un vector perteneciente a la recta y otro perpendicular a dicha recta (verificar!). Muestre que A' es diagonal.
b) Utilice a) y cambio de base para obtener la representación matricial A en la base canónica. Verifique que se obtiene el resultado del problema 6.4.2.

4. Sea $L:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ definida por $L\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}2\\2\end{pmatrix}$, $L\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$.

a) Encuentre su representación matricial en la base canónica de $\mathbb{R}^2$ y dé una expresión para $L(\mathbf{x})$.

b) Encuentre su representación matricial en la base $B'=(\begin{pmatrix}1\\-2\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix})$, y verifique que es diagonal. c) Determine su núcleo e imagen.

5. a) Muestre que la representación matricial $A=[I]_B^B$ del operador identidad $I:V\to V$ con respecto a una base arbitraria B de un espacio vectorial V de dimensión n, es siempre la matriz identidad I_n.
b) Muestre que la representación matricial del operador nulo $0:V\to V$ en cualquier base B de V es la matriz nula de $n\times n$.

6. Muestre que si una matriz B es semejante a una matriz A, entonces:
   i) B^2 es semejante a A^2.
   ii) Si A no singular, B es no singular y B^{-1} es semejante a A^{-1}.

7. Dada $L : V \to W$ y A su representación matricial respecto de bases B_V de V y B_W de W (con dim$V = n$, dim$W = m$) muestre que su representación matricial A' respecto de bases B'_V de V y B'_W de W es

$$A' = R^{-1}AS$$

   con $S = ([\mathbf{v}'_1]_{B_V}, \ldots, [\mathbf{v}'_n]_{B_V})$ la matriz de cambio de base en V (de $n \times n$, no singular) y $R = ([\mathbf{v}'_1]_{B_W}, \ldots, [\mathbf{v}'_n]_{B_W})$ la matriz de cambio de base en W (de $m \times m$, no singular).

6.6. Composición de transformaciones (operaciones sucesivas)

Consideremos dos transformaciones lineales $L : V \to W$, $G : W \to U$, con V, W, U espacios vectoriales. Supongamos que primero aplicamos L sobre un vector $\boldsymbol{v} \in V$, y luego aplicamos G sobre el vector resultante $\boldsymbol{w} = L(\boldsymbol{v}) \in W$. El resultado es el vector

$$\boldsymbol{u} = (G\,L)(\boldsymbol{v}) = G(L(\boldsymbol{v})) \tag{6.1}$$

que pertenece al espacio vectorial U:

$$\boldsymbol{v} \underset{L}{\to} \boldsymbol{w} \underset{G}{\to} \boldsymbol{u}$$

En (6.1), $G\,L : V \to U$ denota la *composición* de G con L (también escrita como $G \circ L$), que es también una transformación **lineal**.

Problemas 6.6.1

1. Demostrar que si $L : V \to W$ y $G : W \to U$ son transformaciones lineales, entonces $G\,L : V \to U$ definida por (6.1) es también una transformación lineal (probar que satisface $(G\,L)(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) = \alpha(G\,L)(\boldsymbol{v}_1) + \beta(G\,L)(\boldsymbol{v}_2))$.

2. Si $L : \mathbb{R}^3 \to \mathbb{R}^2$ está definida por $L(x, y, z) = (x + y, y + z)$,
   y $G : \mathbb{R}^2 \to \mathbb{R}^2$ está definida por $G(x, y) = (x + y, 2x - 2y)$, encontrar una expresión para $(G\,L)(x, y, z) = G(L(x, y, z))$.

6.6.1. Representación matricial de la composición

Consideremos primero que $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ y $U = \mathbb{R}^p$, y sean A_L (de $m \times n$) y A_G (de $p \times m$) las representaciones matriciales de L y G en las bases canónicas correspondientes, tal que $L(\boldsymbol{v}) = A_L(\boldsymbol{v})$, $G(\boldsymbol{w}) = A_G\boldsymbol{w}$. Entonces

$$GL(\boldsymbol{v}) = G(L(\boldsymbol{v})) = G(A_L\boldsymbol{v}) = A_G(A_L\boldsymbol{v}) = (A_GA_L)\boldsymbol{v}$$

por lo que la representación matricial de la composición $G\,L$ con respecto a las bases canónicas de V y U **es el producto** $A_G A_L$ de las matrices que representan a G y a L:

$$A_{GL} = A_G\, A_L \tag{6.2}$$

Observar que la matriz A_G se aplica a la izquierda de la matriz A_L. El producto está así bien definido.

Ejemplo 6.4: En el ejercicio 6.6.2, en las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$, obtenemos las representaciones matriciales (verificar!)

$$A_L = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad A_G = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$$

tal que

$$\begin{aligned} L\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y+z \end{pmatrix} \\ G\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ 2x-2y \end{pmatrix} \end{aligned}$$

Por lo tanto,

$$\begin{aligned} (GL)\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} x+2y+z \\ 2x-2z \end{pmatrix} \end{aligned} \tag{6.3}$$

o sea, $A_{GL} = A_G A_L = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \end{pmatrix}$, con $(G\,L)(x,y,z) = (x+2y+z, 2x-2z)$.

En el caso general, para representaciones matriciales en bases arbitrarias B_V, B_W y B_U de V, W, U, se obtiene también (se deja la demostración para el lector)

$$[GL]_{B_U}^{B_V} = A_G A_L \;\text{ con }\; A_L = [L]_{B_W}^{B_V}\,, \;\; A_G = [G]_{B_U}^{B_W}$$

Este resultado es válido para espacios vectoriales generales de dimensión finita.

6.6.2. Potencias de operadores lineales

Recordemos que si $W = V$, la transformación lineal $L : V \to V$ se denomina también operador lineal o endomorfismo. Queda en este caso definido el cuadrado $L^2 = L\,L$ como la composición de L con L:

$$L^2(\boldsymbol{v}) = L(L(\boldsymbol{v})) \tag{6.4}$$

Si A_L es la representación matricial de L en una cierta base de V, los resultadoa anteriores implican que la representación matricial de L^2 es el cuadrado de la matriz A_L:

$$A_{L^2} = A_L\, A_L = A_L^2 \tag{6.5}$$

En general, la potencia L^k $\forall\ k \geq 2$ queda definida por

$$L^k(\boldsymbol{v}) = L(L^{k-1}(\boldsymbol{v})),\quad k \geq 2$$

y su representación matricial es

$$A_{L^k} = A_L\, A_{L^{k-1}} = A_L^k$$

es decir, la potencia k de la matriz A_L. Por definición, $L^0 = I$, con I el operador identidad. Y si L es un operador lineal inversible (o sea un *automorfismo*: un isomorfismo de V en V), de forma que existe la transformación inversa L^{-1} (tal que $L^{-1}L = I_V$) entonces

$$A_{L^{-1}} = A_L^{-1} \tag{6.6}$$

ya que $A_{L^{-1}}A_L = I_n$. La representación matricial de la inversa L^{-1} es la inversa de la matriz A_L que representa a L, como se vió previamente.

Ejemplo: Si $L : \mathbb{R}^2 \to \mathbb{R}^2$ está dado por $L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ -x - 2y \end{pmatrix}$, su representación matricial en la base canónica es (probar!)

$$A_L = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}$$

La representación matricial de L^2 es entonces

$$A_{L^2} = A_L^2 = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

de forma que

$$L^2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix}$$

es decir, $L^2(\boldsymbol{x}) = 3\boldsymbol{x}$, lo que equivale a $L^2 = 3I$, con I el operador identidad.

Además, como A_L es no singular, L es inversible y la representación matricial de su inversa está dada por

$$A_{L^{-1}} = A_L^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}$$

de forma que

$$L^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2x + y \\ -x - 2y \end{pmatrix}$$

o sea, $L^{-1} = L/3$. Se verifica entonces $L^{-1}L = I$.

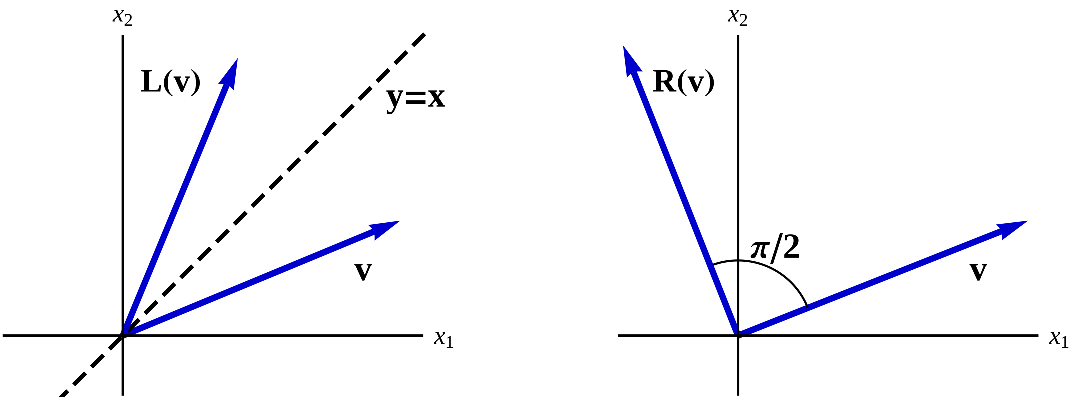
Problemas 6.6.2

1. Si $L : \mathbb{R}^3 \to \mathbb{R}^3$ queda definido por $L(x, y, z) = (y, x, -z)$,
   i) mostrar que en la base canónica,
$$A_L = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
   ii) Hallar A_L^2 y verificar que L^2 es el operador identidad.

2. Sea P_2 el subespacio de los polinomios reales de grado ≤ 2.
   Si $D : P_2 \to P_2$ denota el operador derivada ($D = \frac{d}{dx}$),
   i) Hallar las representaciones matriciales A_D y A_{D^2} de D y la derivada segunda D^2 en las base canónica $(1, x, x^2)$ y verificar que
$$A_{D^2} = A_D A_D = A_D^2$$

6.6.3. Composición de transformaciones lineales en $\mathbb{R}^2$

Consideremos ahora algunos ejemplos de operadores lineales $L : \mathbb{R}^2 \to \mathbb{R}^2$. La reflexión de un vector $\boldsymbol{v}$ respecto de la recta $y = x$ está dada por (recordar!)
$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$$
siendo su representación matricial en la base canónica $A_L = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
Es evidente de la definición que $L^2\binom{x}{y} = L(L\binom{x}{y}) = \binom{x}{y}$ $\forall$ x, y, o sea, $L^2 = I$ (operador identidad), verificándose que (probar!)
$$A_L^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
Es, decir, la inversa de una reflexión L es la misma reflexión L, como es obvio geométricamente.



La rotación de ángulo $\pi/2$ en sentido antihorario de un vector $\boldsymbol{v}$ está dada por
$$R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

siendo su representación matricial en la base canónica $A_R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Consideremos ahora la transformación lineal $R\,L$ que primero refleja un vector respecto de la recta $y = x$ y luego lo rota un ángulo de $\pi/2$ en sentido antihorario. Su representación matricial en la base canónica será

$$A_{(R\,L)} = A_R\,A_L = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

y por lo tanto,

$$R\,L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$$

Esto representa una **reflexión respecto del eje** y (mostrar!).

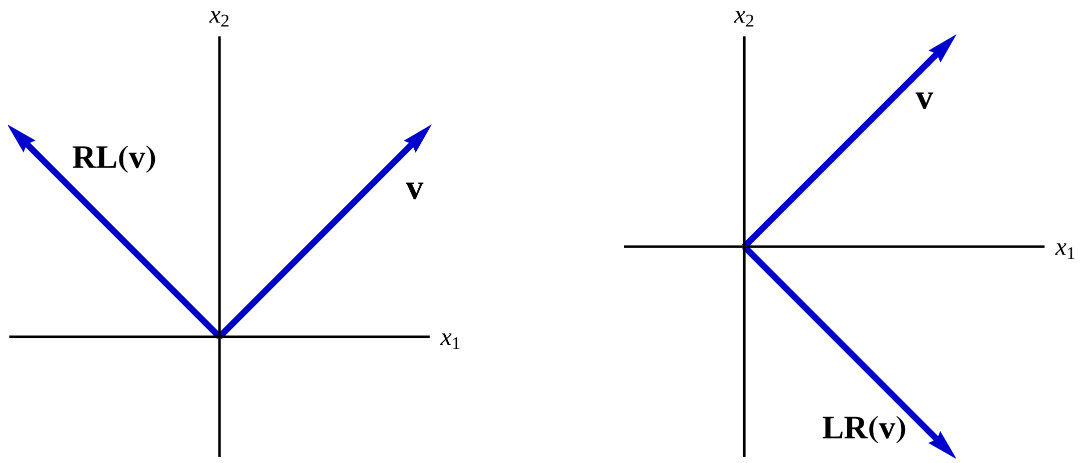
Por otro lado, la transformación lineal $L\,R$, que primero rota un vector un ángulo de $\pi/2$ en sentido antihorario y luego lo refleja respecto de la recta $y = x$, queda representada por la matriz

$$A_{(L\,R)} = A_L\,A_R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

y por lo tanto,

$$L\,R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$$

Esto representa una **reflexión respecto del eje** x (mostrar!).



Por lo tanto, vemos que el resultado final depende **del orden de la composición**, lo que se refleja en la **no conmutatividad** del producto matricial asociado.

Se define el **conmutador** de dos operadores $L : V \to V$, $G : V \to V$ como

$$[G, L] = GL - LG$$

que es en general un operador no nulo. La matriz que lo representa en una base de V es el conmutador de las matrices que representan a G y L en dicha base: $A_{[G,L]} = A_G\,A_L - A_L\,A_G$. En el ejemplo anterior se obtiene

$$A_R\,A_L - A_L\,A_R = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

que implica $RL - LR = 2RL$, es decir, $LR = -RL$.

Recordemos finalmente que la proyección ortogonal de un vector sobre el eje x está dada por (graficar!)

$$P_x \binom{x}{y} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \binom{x}{y} = \binom{x}{0}$$

y la proyección ortogonal de un vector sobre el eje y está dada por

$$P_y \binom{x}{y} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \binom{x}{y} = \binom{0}{y}$$

Problemas 6.6.3

1. a) Hallar la representación matricial en la base canónica de la inversa de los operadores R y L anteriores.
   b) ¿Tienen P_x y P_y inversa?
   c) Encuentre la representación matricial de P_xP_y y P_yP_x. Justificar el resultado.

2. Sea $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador lineal que primero rota todo vector un ángulo $\pi/2$ en sentido antihorario y luego lo proyecta sobre el eje x. Hallar su representación matricial en la base canónica y una expresión para $F\binom{x}{y}$.

3. Sea $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ el operador lineal que primero proyecta todo vector sobre el eje x y luego lo rota un ángulo $\pi/2$ en sentido antihorario. Hallar su representación matricial en la base canónica y una expresión para $G\binom{x}{y}$.

4. Recordando que la representación matricial en la base canónica de una reflexión respecto de la recta $y = mx$, con $m = \tan\theta$, es

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

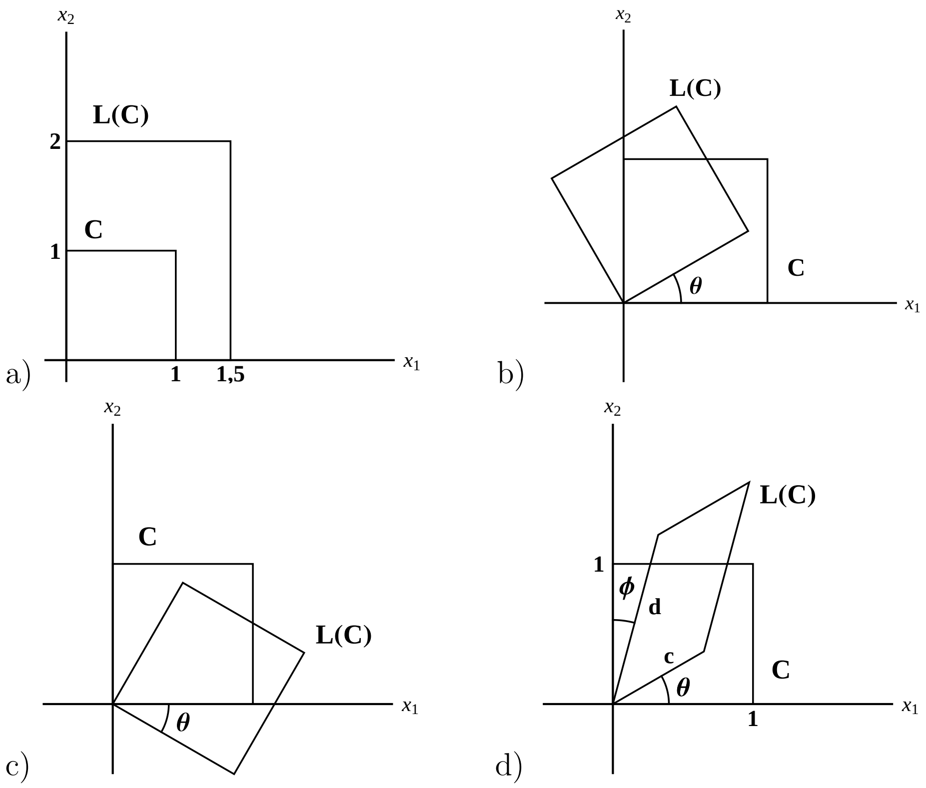
   a) Halle su determinante.
   b) Muestre que la composición de dos reflexiones es una rotación.

5. Recordando que la representación matricial en la base canónica de una rotación de ángulo θ antihorario es

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

   a) Halle su determinante.
   b) Muestre que la composición de dos rotaciones es otra rotación.
   c) Muestre que la composición de una reflexión con una rotación es otra reflexión.

6. Encuentre, a partir de las figuras, la representación matricial en la base canónica de las siguientes transformaciones lineales $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:



a)
x_2
x_1
L(C)
C
2
1
1
1,5
b)
x_2
x_1
L(C)
C
θ
c)
x_2
x_1
C
L(C)
θ
d)
x_2
x_1
L(C)
C
1
1
ϕ
d
c
θ

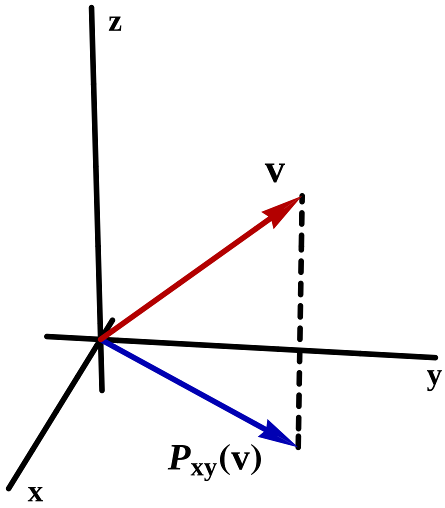
Facultad de Ingeniería
UNLP

Matemática C

VI Proyección ortogonal – Bases ortogonales

1 Proyección ortogonal

En secciones previas hemos usado la proyección ortogonal de vectores de $\mathbb{R}^3$ sobre el plano xy (subespacio). Asumiendo que el vector está por encima del plano xy, esta proyección es la sombra ("al mediodía") del vector sobre este plano:



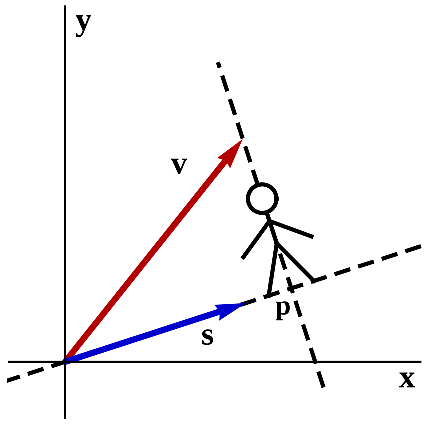
En otras palabras, la proyección de $\mathbf{v}$ sobre el plano xy es un vector $\mathbf{p} = P_{xy}(\mathbf{v})$ tal que si alguien camina sobre el plano y se detiene sobre el extremo de $\mathbf{p}$ mirando hacia arriba en línea recta, ve el extremo del vector $\mathbf{v}$.

En esta sección se generalizará este concepto a otras proyecciones.

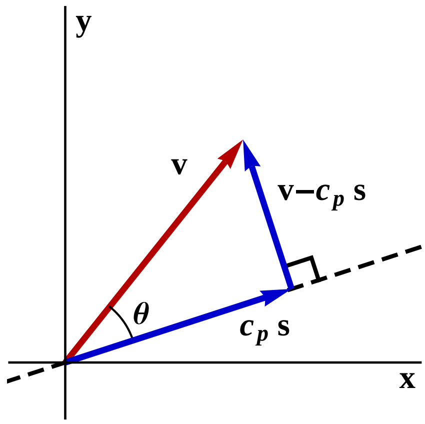
1.1 Proyección ortogonal sobre una recta

Primero consideramos de nuevo la proyección ortogonal sobre una recta que pasa por el origen, dirigida por un vector $\boldsymbol{s}$ no nulo.

Proyectar ortogonalmente un vector $\boldsymbol{v}$ sobre esta recta da un punto $\boldsymbol{p}$ sobre la recta tal que el segmento que une $\boldsymbol{p}$ con $\boldsymbol{v}$ es **perpendicular** a la recta. Un observador que camine sobre esa recta y se detenga en el punto $\boldsymbol{p}$, al mirar hacia arriba (o hacia abajo, según la posición del vector) en forma perpendicular a la recta, verá el extremo de $\boldsymbol{v}$:



Es decir, ha caminado sobre la recta $\{c\boldsymbol{s},\ c \in \mathbb{R}\}$ hasta hallar el punto $\boldsymbol{p} = c_p\boldsymbol{s}$ determinado por un coeficiente c_p, con la propiedad de que el vector $\boldsymbol{v}-c_p\boldsymbol{s}$ es ortogonal a $c_p\boldsymbol{s}$.



Para determinar c_p, basta con observar que la condición anterior implica que el producto escalar de estos dos vectores debe ser nulo:

$$\boldsymbol{s} \cdot (\boldsymbol{v} - c_p\boldsymbol{s}) = 0$$

De aquí obtenemos

$$c_p = \frac{\boldsymbol{s} \cdot \boldsymbol{v}}{\boldsymbol{s} \cdot \boldsymbol{s}}$$

es decir, $c_p = \frac{|\boldsymbol{v}|}{|\boldsymbol{s}|}\cos\theta$, donde θ es el ángulo entre $\boldsymbol{v}$ y $\boldsymbol{s}$. Estas consideraciones son aplicables tanto en $\mathbb{R}^2$ o $\mathbb{R}^3$ como en $\mathbb{R}^n$ (y también en todo espacio vectorial en el que se ha definido un producto escalar!), siendo el plano del dibujo el subespacio de dimensión 2 generado por $\boldsymbol{v}$ y $\boldsymbol{s}$.

Definición. Sea $\boldsymbol{v} \in \mathbb{R}^n$. La proyección ortogonal de $\boldsymbol{v}$ sobre la recta (que pasa por el origen) generada por el vector no nulo $\boldsymbol{s} \in \mathbb{R}^n$ es el vector

$$P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{\boldsymbol{s} \cdot \boldsymbol{v}}{\boldsymbol{s} \cdot \boldsymbol{s}}\,\boldsymbol{s}$$

Es decir, $P_{\boldsymbol{s}}(\boldsymbol{v}) = (|\boldsymbol{v}|\cos\theta)\frac{\boldsymbol{s}}{|\boldsymbol{s}|}$. El resultado depende sólo de la dirección de $\boldsymbol{s}$ y no de su longitud (o sentido): $P_{(\alpha\boldsymbol{s})}(\boldsymbol{v}) = P_{\boldsymbol{s}}(\boldsymbol{v})\ \forall\ \alpha \neq 0$. De esta forma, el vector $\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})$ es perpendicular a $\boldsymbol{s}$:

$$\boldsymbol{s} \cdot (\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})) = 0$$

Ejemplo 1.1.1 Para proyectar $\boldsymbol{v} = (x, y) \in \mathbb{R}^2$ ortogonalmente sobre la recta $y = 2x$, primero obtenemos un vector que indique la dirección de la recta. Por ejemplo,

$$\boldsymbol{s} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Luego

$$P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{(1,2) \cdot \begin{pmatrix} x \\ y \end{pmatrix}}{(1,2) \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{x + 2y}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Ejemplo 1.1.2 En $\mathbb{R}^3$ la proyección ortogonal de un vector $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ sobre el eje y es

$$P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}}{\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$$

Y la proyección de $\boldsymbol{v}$ sobre el vector $\boldsymbol{s} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ es

$$P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}}{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{x + 2y + 3z}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Puede verificarse que $\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{1}{14} \begin{pmatrix} 13x - 2y - 3z \\ -x + 5y - 3z \\ -3x - 6y + 5z \end{pmatrix}$ es ortogonal a $\boldsymbol{s}$.

Distancia mínima de $\boldsymbol{v}$ a la recta generada por $\boldsymbol{s}$.
Es importante destacar que la proyección ortogonal $P_{\boldsymbol{s}}(\boldsymbol{v})$ de $\boldsymbol{v}$ sobre la recta generada por $\boldsymbol{s}$ es, precisamente, el punto de la recta **más cercano** a $\boldsymbol{v}$: la distancia al cuadrado entre $\boldsymbol{v}$ y un punto arbitrario $\alpha\boldsymbol{s}$ de dicha recta está dada por el producto escalar

$$D^2 = |\boldsymbol{v} - \alpha\boldsymbol{s}|^2 = (\boldsymbol{v} - \alpha\boldsymbol{s}) \cdot (\boldsymbol{v} - \alpha\boldsymbol{s})$$

Para $c_p = \boldsymbol{s} \cdot \boldsymbol{v}/\boldsymbol{s} \cdot \boldsymbol{s}$ se obtiene, dado que $\boldsymbol{s} \cdot (\boldsymbol{v} - c_p\boldsymbol{s}) = 0$,

$$\begin{aligned} D^2 = |\boldsymbol{v} - \alpha\boldsymbol{s}|^2 &= |\boldsymbol{v} - c_p\boldsymbol{s} - (\alpha - c_p)\boldsymbol{s}|^2 \\ &= (\boldsymbol{v} - c_p\boldsymbol{s} - (\alpha - c_p)\boldsymbol{s}) \cdot (\boldsymbol{v} - c_p\boldsymbol{s} - (\alpha - c_p)\boldsymbol{s}) \\ &= |\boldsymbol{v} - c_p\boldsymbol{s}|^2 + (\alpha - c_p)^2|\boldsymbol{s}|^2 \\ &\geq |\boldsymbol{v} - c_p\boldsymbol{s}|^2 \end{aligned}$$

La cota inferior $D^2_{\min} = |\boldsymbol{v} - c_o\boldsymbol{s}|^2$ se alcanza pues para $\alpha = c_p$, es decir, para $\alpha\boldsymbol{s}$ igual a la proyección ortogonal $c_p\boldsymbol{s} = P_{\boldsymbol{s}}(\boldsymbol{v})$ de $\boldsymbol{v}$ sobre la recta. La distancia mínima a la recta es entonces

$$D_{\min} = |\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})|$$

Ejercicios 1.1

1.1 Encontrar la proyección ortogonal de $\boldsymbol{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ sobre las rectas generadas por los vectores i) $\boldsymbol{s} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y ii) $\boldsymbol{s} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Determinar también la distancia mínima $D_{\min}$ de $\boldsymbol{v}$ a la recta y verificar que $\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})$ es ortogonal a $\boldsymbol{s}$.

1.2 Idem anterior para:

i) $\boldsymbol{s} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\boldsymbol{v} = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}$ ii) $\boldsymbol{s} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 2 \end{pmatrix}$, $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

1.3 Encontrar la proyección ortogonal de un vector genérico $\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ sobre la recta $y = 2x$ (es decir, la recta generada por $\boldsymbol{s} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$) y determinar la distancia mínima de $\boldsymbol{v}$ a dicha recta.

1.4 Encontrar la proyección ortogonal de $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ sobre la recta $\left\{ c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, c \in \mathbb{R} \right\}$, y determinar la distancia mínima de $\boldsymbol{v}$ a dicha recta.

1.5 a) ¿Si $\boldsymbol{v}$ pertenece a la recta generada por $\boldsymbol{s}$, ¿cual es su proyección ortogonal $P_{\boldsymbol{s}}(\boldsymbol{v})$?
b) ¿Si $\boldsymbol{v}$ es ahora ortogonal a $\boldsymbol{s}$, ¿Cual es su proyección ortogonal $P_{\boldsymbol{s}}(\boldsymbol{v})$?

1.6 a) Mostrar que la proyección ortogonal $P_{\boldsymbol{s}} : \mathbb{R}^n \to \mathbb{R}^n$ dada por $P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{\boldsymbol{s}\cdot\boldsymbol{v}}{\boldsymbol{s}\cdot\boldsymbol{s}}\,\boldsymbol{s}$ (con $\boldsymbol{s} \neq \boldsymbol{0}$) es una **transformación lineal**.
b) ¿Cual es su imagen? ¿Cual es su núcleo?
c) Hallar explícitamente la matriz representativa de $P_{\boldsymbol{s}}$ en la base canónica en los ejercicios 1.3 y 1.4, junto con su núcleo e imagen. Interpretarlos geométricamente.

1.7 a) Mostrar que $|P_{\boldsymbol{s}}(\boldsymbol{v})| = |\boldsymbol{v}|\,|\cos\theta|$, con θ el ángulo entre $\boldsymbol{v}$ y $\boldsymbol{s}$.
b) Mostrar que $P_{\alpha\boldsymbol{s}}(\boldsymbol{v}) = P_{\boldsymbol{s}}(\boldsymbol{v})\ \forall\ \alpha \in \mathbb{R},\ \alpha \neq 0$ (válido $\forall\ \boldsymbol{v}, \boldsymbol{s} \in \mathbb{R}^n$, $\boldsymbol{s} \neq \boldsymbol{0}$).

1.8 Si $D^2(c) = |\boldsymbol{v} - c\boldsymbol{s}|^2 = |\boldsymbol{v}|^2 - 2c\boldsymbol{v}\cdot\boldsymbol{s} + c^2\boldsymbol{s}\cdot\boldsymbol{s}$, verificar, minimzando esta función respecto de c, que su valor mínimo se alcanza para $c = c_p = \boldsymbol{s}\cdot\boldsymbol{v}/\boldsymbol{s}\cdot\boldsymbol{s}$. Concluir que la distancia mínima de $\boldsymbol{v}$ a $\boldsymbol{s}$ es $D_{\min} = |\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})|$.

1.2 Bases ortogonales y método de ortogonalización de Gram-Schmidt

Al proyectar un vector $\boldsymbol{v}$ sobre una recta (que pasa por el origen) dirigida por un vector $\boldsymbol{s}$, se descompone al vector $\boldsymbol{v}$ en dos componentes ortogonales:

$$\boldsymbol{v} = P_{\boldsymbol{s}}(\boldsymbol{v})\ +\ (\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v}))$$

con $P_{\boldsymbol{s}}(\boldsymbol{v})\cdot(\boldsymbol{v}-P_{\boldsymbol{s}}(\boldsymbol{v})) = 0$. Extenderemos ahora esta idea a una base formada por vectores ortogonales.

Definición. Los n vectores no nulos $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ son mutuamente ortogonales si todo par de vectores son ortogonales:

$$\boldsymbol{b}_i \cdot \boldsymbol{b}_j = 0 \ \ \forall \ \ i \neq j$$

Es fácil ver que si son no nulos y ortogonales, el conjunto $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n\}$ es necesariamente **linealmente independiente**: Si consideramos la combinación lineal nula

$$c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \ldots + c_n\boldsymbol{b}_n = \mathbf{0}$$

entonces, multiplicando escalarmente por $\boldsymbol{b}_i$ ambos miembros y usando la ortogonalidad para $i \neq j$,

$$\begin{array}{rcl} \boldsymbol{b}_i \cdot (c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \ldots + c_n\boldsymbol{b}_n) & = & \boldsymbol{b}_i \cdot \mathbf{0} \\ \Rightarrow \qquad c_i(\boldsymbol{b}_i \cdot \boldsymbol{b}_i) & = & 0 \end{array}$$

lo que implica $c_i = 0$ ya que $\boldsymbol{b}_i \cdot \boldsymbol{b}_i = |\boldsymbol{b}_i|^2 \neq 0$. Esto vale para todo $i = 1, \ldots, n$.

En consecuencia, un conjunto de n vectores ortogonales en un espacio de dimensión n es necesariamente **una base del mismo**, ya que son linealmente independientes, y todo conjunto de n vectores linealmente independientes en un espacio de dimensión n es una base.

Definición. Una base **ortogonal** $B = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ de un espacio vectorial V es una base formada por vectores mutuamente ortogonales.

La gran ventaja de tener una base ortogonal B es que si $\boldsymbol{v} \in V$, entonces las coordenadas de $\boldsymbol{v}$ en esta base, es decir los coeficientes α_i en la expansión

$$\boldsymbol{v} = \alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \ldots + \alpha_n\boldsymbol{b}_n$$

pueden obtenerse directamente mediante productos escalares, sin necesidad de resolver un sistema de ecuaciones: Por ser la base ortogonal, se cumple que

$$\begin{array}{rcl} \boldsymbol{b}_i \cdot \boldsymbol{v} & = & \boldsymbol{b}_i \cdot (\alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \ldots \alpha_n\boldsymbol{b}_n) \\ & = & \alpha_i(\boldsymbol{b}_i \cdot \boldsymbol{b}_i) \end{array}$$

Por lo tanto, dado que $\boldsymbol{b}_i \cdot \boldsymbol{b}_i = |\boldsymbol{b}_i|^2 \neq 0$,

$$\alpha_i = \frac{\boldsymbol{b}_i \cdot \boldsymbol{v}}{\boldsymbol{b}_i \cdot \boldsymbol{b}_i}, \quad i = 1, \ldots, n$$

Se cumple entonces que

$$\alpha_i\boldsymbol{b_i} = \frac{\boldsymbol{b}_i \cdot \boldsymbol{v}}{\boldsymbol{b_i} \cdot \boldsymbol{b}_i}\,\boldsymbol{b}_i = P_{\boldsymbol{b_i}}(\boldsymbol{v})\,, \quad i = 1, \ldots, n$$

Es decir, cada término $\alpha_i\boldsymbol{v}$ en la expansión $\boldsymbol{v} = \alpha_1\boldsymbol{b}_1 + \ldots + \alpha_n\boldsymbol{b}_n$ es la **proyección ortogonal** de $\boldsymbol{v}$ sobre la recta generada por $\boldsymbol{b}_i$.
Además, por la ortogonalidad se cumple también que

$$\begin{array}{rcl} |\boldsymbol{v}|^2 & = & \boldsymbol{v} \cdot \boldsymbol{v} \\ & = & \alpha_1^2|\boldsymbol{b}_1|^2 + \alpha_2^2|\boldsymbol{b}_2|^2 + \ldots + \alpha_n^2|\boldsymbol{b}_n|^2 \end{array}$$

Definición. Una base $B=\{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n\}$ de un espacio vectorial V se dice **ortonormal** si es una base ortogonal y además los n vectores $\boldsymbol{b}_i$ tienen longitud 1:

$$\boldsymbol{b}_i\cdot\boldsymbol{b}_j=\left\{\begin{array}{ll}0 & i\neq j\\ 1 & i=j\end{array}\right.\quad,\quad i,j=1,\ldots,n$$

En este caso, las fórmulas anteriores se simplifican aun más: Si $\boldsymbol{v}\in V$, los coeficientes α_i en

$$\boldsymbol{v}=\alpha_1\boldsymbol{b}_1+\alpha_2\boldsymbol{b}_2+\ldots+\alpha_n\boldsymbol{b}_n$$

se reducen, dado que $|\boldsymbol{b}_i|^2=\boldsymbol{b}_i\cdot\boldsymbol{b}_i=1\ \forall\ i$, a

$$\alpha_i=\boldsymbol{b}_i\cdot\boldsymbol{v}\,,\quad i=1,\ldots,n$$

con $\alpha_i\boldsymbol{b_i}=P_{\boldsymbol{b}_i}(\boldsymbol{v})=(\boldsymbol{b}_i\cdot\boldsymbol{v})\,\boldsymbol{b}_i$. Además:

$$|\boldsymbol{v}|^2\ \ =\ \ \boldsymbol{v}\cdot\boldsymbol{v}=\alpha_1^2+\alpha_2^2+\ldots+\alpha_n^2$$

La base canónica de $\mathbb{R}^n$, $B_c=\left\{\boldsymbol{e}_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\ldots,\boldsymbol{e}_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right\}$ es un ejemplo de base ortonormal.

Ejemplo 1.2.1 La base $B=\left\{\boldsymbol{b}_1=\begin{pmatrix}1\\1\end{pmatrix},\ \boldsymbol{b}_2=\begin{pmatrix}-1\\1\end{pmatrix}\right\}$ es una base ortogonal de $\mathbb{R}^2$, ya que $\boldsymbol{b}_1\cdot\boldsymbol{b}_2=0$, y $\boldsymbol{b}_1\neq\boldsymbol{0}$, $\boldsymbol{b}_2\neq\boldsymbol{0}$. No es ortonormal, ya que $\boldsymbol{b}_1\cdot\boldsymbol{b}_1=\boldsymbol{b}_2\cdot\boldsymbol{b}_2=2\neq 1$. Si $\boldsymbol{v}=\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2$, podemos escribir $\boldsymbol{v}$ como

$$\boldsymbol{v}=\alpha_1\begin{pmatrix}1\\1\end{pmatrix}+\alpha_2\begin{pmatrix}-1\\1\end{pmatrix}$$

donde los coeficientes α_1,α_2 (las coordenadas de $\boldsymbol{v}$ en esta base) pueden obtenerse como

$$\alpha_1=\frac{\boldsymbol{b}_1\cdot\boldsymbol{v}}{\boldsymbol{b}_1\cdot\boldsymbol{b}_1}=\frac{\begin{pmatrix}1 & 1\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}}{2}=\frac{x+y}{2},\qquad\alpha_2=\frac{\boldsymbol{b}_2\cdot\boldsymbol{v}}{\boldsymbol{b}_2\cdot\boldsymbol{b}_2}=\frac{\begin{pmatrix}-1 & 1\end{pmatrix}\cdot\begin{pmatrix}x\\y\end{pmatrix}}{2}=\frac{-x+y}{2}$$

Por lo tanto,

$$\boldsymbol{v}=\frac{x+y}{2}\begin{pmatrix}1\\1\end{pmatrix}+\frac{y-x}{2}\begin{pmatrix}-1\\1\end{pmatrix}$$

Además, $|\boldsymbol{v}|^2=x^2+y^2=2(\alpha_1^2+\alpha_2^2)$.

Ejemplo 1.2.2 La base $B'=\left\{\boldsymbol{b}'_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},\ \boldsymbol{b}'_2=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}\right\}$ es una base **ortonormal** de $\mathbb{R}^2$, ya que $\boldsymbol{b}'_1\cdot\boldsymbol{b}'_2=0$, y además $\boldsymbol{b}'_1\cdot\boldsymbol{b}'_1=\boldsymbol{b}'_2\cdot\boldsymbol{b}'_2=1$. Es la base canónica de $\mathbb{R}^2$ rotada 45º en sentido antihorario. Si $\boldsymbol{v}=\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2$, podemos escribir $\boldsymbol{v}$ como

$$\boldsymbol{v}=\alpha'_1\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}+\alpha'_2\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$$

donde los coeficientes α'_1, α'_2 (las coordenadas de $\boldsymbol{v}$ en esta base) están dadas por

$$\alpha'_1 = \boldsymbol{b}'_1 \cdot \boldsymbol{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \frac{x+y}{\sqrt{2}}, \qquad \alpha'_2 = \boldsymbol{b}'_2 \cdot \boldsymbol{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \frac{-x+y}{\sqrt{2}}$$

Por lo tanto,

$$\boldsymbol{v} = \frac{x+y}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{y-x}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Además, $|\boldsymbol{v}|^2 = x^2 + y^2 = {\alpha'_1}^2 + {\alpha'_2}^2$.

Método de ortogonalización de Gram-Schmidt

Todo conjunto de n vectores mutuamente ortogonales y no nulos de $\mathbb{R}^n$ es una base de $\mathbb{R}^n$, pero no toda base de $\mathbb{R}^n$ es un conjunto de vectores mutuamente ortogonales. No obstante, es siempre posible **ortogonalizar** la base.

Método de Gram-Schmidt. Dado un conjunto $M = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ **linealmente independiente** de m vectores $\boldsymbol{v}_i \in V$, que generan un subespacio $S_m \subset V$ de dimensión m, los m vectores

$$\begin{aligned}
\boldsymbol{k}_1 &= \boldsymbol{v}_1 \\
\boldsymbol{k}_2 &= \boldsymbol{v}_2 - P_{\boldsymbol{k}_1}(\boldsymbol{v}_2) = \boldsymbol{v}_2 - \frac{\boldsymbol{k}_1 \cdot \boldsymbol{v}_2}{\boldsymbol{k}_1 \cdot \boldsymbol{k}_1} \boldsymbol{k}_1 \\
\boldsymbol{k}_3 &= \boldsymbol{v}_3 - P_{\boldsymbol{k}_1}(\boldsymbol{v}_3) - P_{\boldsymbol{k}_2}(\boldsymbol{v}_3) = \boldsymbol{v}_3 - \frac{\boldsymbol{k}_1 \cdot \boldsymbol{v}_3}{\boldsymbol{k}_1 \cdot \boldsymbol{k}_1} \boldsymbol{k}_1 - \frac{\boldsymbol{k}_2 \cdot \boldsymbol{v}_3}{\boldsymbol{k}_2 \cdot \boldsymbol{k}_2} \boldsymbol{k}_2 \\
\vdots & \quad \vdots \\
\boldsymbol{k}_m &= \boldsymbol{v}_m - \sum_{i=1}^{m-1} P_{\boldsymbol{k}_i}(\boldsymbol{v}_m) = \boldsymbol{v}_m - \sum_{i=1}^{m-1} \frac{\boldsymbol{k}_i \cdot \boldsymbol{v}_m}{\boldsymbol{k}_i \cdot \boldsymbol{k}_i} \boldsymbol{k}_i
\end{aligned}$$

forman una base **ortogonal** del mismo subespacio S_m (del cual M es una base no ortogonal):

$$\boldsymbol{k}_i \cdot \boldsymbol{k}_j = 0 \quad \forall\, i \neq j$$

Demostración: Partiendo de $\boldsymbol{k}_1 = \boldsymbol{v}_1$, se propone $\boldsymbol{k}_2 = \boldsymbol{v}_2 - \alpha \boldsymbol{k}_1$, con α determinado por la condición

$$\boldsymbol{k}_1 \cdot \boldsymbol{k}_2 = \boldsymbol{k}_1 \cdot \boldsymbol{v}_2 - \alpha \boldsymbol{k}_1 \cdot \boldsymbol{k}_1 = 0$$

De aqui se obtiene $\alpha = (\boldsymbol{k}_1 \cdot \boldsymbol{v}_2)/(\boldsymbol{k}_1 \cdot \boldsymbol{k}_1)$, y por lo tanto la expresión anterior para $\boldsymbol{k}_2$.
Se procede en forma análoga con los restantes vectores: proponiendo $\boldsymbol{k}_i = \boldsymbol{v}_i - \sum_{j=1}^{i-1} \alpha_j \boldsymbol{k}_i$ para $2 \leq i \leq m$, las condiciones $\boldsymbol{k}_j \cdot \boldsymbol{k}_i = 0$ para $j < i$ conducen a $\alpha_j = (\boldsymbol{k}_j \cdot \boldsymbol{v}_i)/(\boldsymbol{k}_j \cdot \boldsymbol{k}_j)$ y por lo tanto a $\boldsymbol{k}_i = \boldsymbol{v}_i - \sum_{j=1}^{i-1} P_{\boldsymbol{k}_j}(\boldsymbol{v}_i)$.
La expresión final es clara: a cada vector $\boldsymbol{v}_i$ se le restan sus proyecciones ortogonales sobre los vectores $\boldsymbol{k}_j$ anteriores, de forma que $\boldsymbol{k}_i$ contenga solo la componente de $\boldsymbol{v}_i$ ortogonal a todos ellos.
Notemos que $\boldsymbol{k}_i \neq \boldsymbol{0}$, ya que si $\boldsymbol{k}_i = \boldsymbol{0}$, $\boldsymbol{v}_i$ sería combinación lineal de los vectores anteriores, lo que es imposible por ser M linealmente independiente. Además, por ser ortogonales y no nulos, los m vectores $\boldsymbol{k}_i$ son también linealmente independientes. Y generan el mismo espacio S_m, ya que son todos combinaciones lineales de los vectores de M.
Si $m = n = \dim V$, se obtiene así una base **ortogonal** del espacio completo.
Para obtener una base ortonormal, basta normalizar luego los vectores: $\boldsymbol{k}_i \to \boldsymbol{k}'_i = \boldsymbol{k}_i / \sqrt{\boldsymbol{k}_i \cdot \boldsymbol{k}_i}$.

Ejemplo 1.2.3 Consideremos $M = \left\{ \boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\}$. Este conjunto es linealmente independiente pero no ortogonal, siendo la base de un subespacio S_2 de dimensión 2 de $\mathbb{R}^3$ (plano que pasa por el origen). Mediante el método anterior, podemos obtener una base ortogonal de S_2: tenemos $\boldsymbol{k}_1 = \boldsymbol{v}_1$ y

$$\boldsymbol{k}_2 = \boldsymbol{v}_2 - P_{\boldsymbol{k}_1}(\boldsymbol{v}_2) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

que resulta ortogonal a $\boldsymbol{k}_1 = \boldsymbol{v}_1$. Si se desea ahora una base ortogonal de $\mathbb{R}^3$ que contenga a $\boldsymbol{k}_1$ y $\boldsymbol{k}_2$, podemos considerar por ej. $\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, ya que $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ resulta linealmente independiente. Se obtiene (notar que con esta elección $\boldsymbol{k}_2 \cdot \boldsymbol{v}_3 = 0$)

$$\boldsymbol{k}_3 = \boldsymbol{v}_3 - P_{\boldsymbol{k}_1}(\boldsymbol{v}_3) - P_{\boldsymbol{k}_2}(\boldsymbol{v}_3) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 0 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$$

que es ortogonal a $\boldsymbol{k}_1$ y $\boldsymbol{k}_2$. Y el conjunto $\{\frac{\boldsymbol{k}_1}{\sqrt{3}}, \frac{\boldsymbol{k}_2}{\sqrt{2}}, \frac{\sqrt{3}\boldsymbol{k}_3}{\sqrt{2}}\}$ es una base ortonormal de $\mathbb{R}^3$.

Ejercicios 1.2

2.1 Determinar una base ortogonal del espacio generado por los vectores

a) $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$, b) $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ c) $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$

2.2 Determinar una base ortonormal de $\mathbb{R}^3$ que contenga al vector $(1, -1, 1)$.

2.3 a) Hallar una base ortogonal del plano $S = \{(x, y, z) \in \mathbb{R}^3, x + y - 2z = 0\}$.
b) Expresar un vector general $\boldsymbol{v} = (x, y, z) \in S$ como combinación lineal de los vectores de la base hallada en a).

2.4 Determinar una base ortogonal de $\mathbb{R}^3$ que contenga a la base de S hallada en 2.3.a).

2.5 Hallar un vector $\boldsymbol{v}$ que sea ortogonal a $\boldsymbol{v}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}^T$ y a $\boldsymbol{v}_2 = \begin{pmatrix} 1 & 2 & 3 & 2 \end{pmatrix}^T$.

2.6 Explicar lo que ocurre cuando se aplica el método de Gram-Schmidt a un conjunto linealmente dependiente. Como ejemplo, aplicarlo al conjunto $M = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \\ 2 \end{pmatrix} \right\}$.

2.7 Una gran ventaja de las bases ortonormales es que ellas simplifican la representación de un vector con respecto a esa base. Por ejemplo,

(a) Representar $\boldsymbol{v} = \begin{pmatrix}1\\3\end{pmatrix}$ respecto de la base no ortogonal $B = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}$ de $\mathbb{R}^2$. Mostrar que $\boldsymbol{v}$ NO es la suma de proyecciones ortogonales sobre cada vector de B.

(b) Representar ahora $\boldsymbol{v} = \begin{pmatrix}1\\3\end{pmatrix}$ respecto de la base ortogonal $B^o = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$. Mostrar que en este caso el vector SÍ es la suma de proyecciones ortogonales sobre cada vector de esta base.

(c) Mostrar que para una base ortogonal general $B^o = \{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$ de un subespacio S de $\mathbb{R}^n$ de dimensión no nula $m \leq n$, todo vector $\boldsymbol{v} \in S$ se escribe en la forma

$$\begin{aligned}\boldsymbol{v} &= P_{\boldsymbol{k}_1}(\boldsymbol{v}) + \ldots + P_{\boldsymbol{k}_m}(\boldsymbol{v})\\ &= \alpha_1 \boldsymbol{k}_1 + \ldots + \alpha_m \boldsymbol{k}_n\end{aligned}$$

donde $P_{\boldsymbol{k}_i}(\boldsymbol{v}) = \mathrm{proj}_{\boldsymbol{k}_i}(\boldsymbol{v}) = \alpha_i \boldsymbol{k}_i$ con

$$\alpha_i = \frac{\boldsymbol{k}_i \cdot \boldsymbol{v}}{\boldsymbol{k}_i \cdot \boldsymbol{k}_i}\,, \quad i = 1, \ldots, m$$

(Sugerencia: considerar el producto $\boldsymbol{k}_i \cdot \boldsymbol{v}$ y usar la ortogonalidad $\boldsymbol{k}_i \cdot \boldsymbol{k}_j = 0$ si $i \neq j$). El resultado vale también para $m = n$ ($S = \mathbb{R}^n$).

2.8 Mostrar que las columnas de una matriz A de $n \times n$ ortogonal forman una base ortonormal de $\mathbb{R}^n$ si y sólo si $A^{-1} = A^T$. Por ejemplo, la siguiente matriz es ortogonal:

$$A = \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2} & 0\\ 1/\sqrt{2} & 1/\sqrt{2} & 0\\ 0 & 0 & 1\end{pmatrix}$$

1.3. Subespacios ortogonales. Proyección sobre un subespacio

Al proyectar un vector sobre una recta, se lo descompone en dos partes: una que yace sobre la recta, $P_{\boldsymbol{s}}(\boldsymbol{v})$, y la restante $\boldsymbol{v} - P_{\boldsymbol{s}}(\boldsymbol{v})$ que es ortogonal a la recta. Para generalizar a subespacios arbitrarios, se sigue la misma idea.

Definición. Dos subespacios S_1 y S_2 de $\mathbb{R}^n$ son ortogonales si todo vector de S_1 es ortogonal a todo vector de S_2, es decir, si $\forall\ \boldsymbol{v} \in S_1$ y $\forall\ \boldsymbol{w} \in S_2$ se cumple $\boldsymbol{v} \cdot \boldsymbol{w} = 0$.

Por ejemplo, en $\mathbb{R}^2$ el eje x, $S_x = \{(x, 0), x \in \mathbb{R}\}$, es ortogonal al eje y, $S_y = \{(0, y), y \in \mathbb{R}\}$, ya que el primero contiene vectores de la forma $\boldsymbol{v} = (x, 0)$ y el segundo $\boldsymbol{w} = (0, y)$, cumpliéndose que $\boldsymbol{v} \cdot \boldsymbol{w} = 0\ \forall\ x, y$.

Y la recta $y = 2x$ es ortogonal a la recta $y = -x/2$, ya que la primera contiene vectores de la forma $\boldsymbol{v} = (x, 2x)$ y la segunda $\boldsymbol{w} = (-2y, y)$, cumpliéndose $\boldsymbol{v} \cdot \boldsymbol{w} = 0\ \forall\ x, y$.

Y en $\mathbb{R}^3$, el eje z es perpendicular al eje x y también al plano $x\,y$. En general, la recta $\boldsymbol{r} = t\boldsymbol{n}$, con $\boldsymbol{r} = (x, y, z)$, $t \in \mathbb{R}$ y $\boldsymbol{n} \neq \boldsymbol{0}$, es perpendicular al plano $\boldsymbol{r} \cdot \boldsymbol{n} = 0$ (probar!).

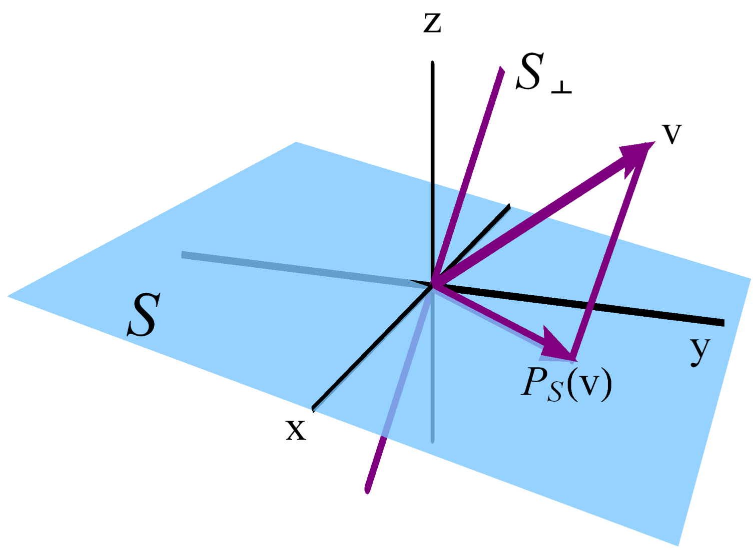
Definición. El complemento ortogonal de un subespacio $S \subset \mathbb{R}^n$ se define como el conjunto de vectores de $\mathbb{R}^n$ ortogonales a todo vector de S:

$$\begin{aligned} S_{\perp} &= \{\boldsymbol{w} \in \mathbb{R}^n \mid \boldsymbol{w} \text{ es perpendicular a todos los vectores de } S\ \} \\ &= \{\boldsymbol{w} \in \mathbb{R}^n \mid \boldsymbol{w} \cdot \boldsymbol{v} = 0 \ \ \forall \ \ \boldsymbol{v} \in S\ \} \end{aligned}$$

Es fácil ver que $S_{\perp}$ es siempre un subespacio de $\mathbb{R}^n$ (probar!).

Así, en los ejemplos anteriores se verifica que en $\mathbb{R}^2$, el eje y es el complemento ortogonal del eje x y la recta $y = -x/2$ el complemento ortogonal de la recta $y = 2x$.

Y en $\mathbb{R}^3$, el plano xy es el complemento ortogonal del eje z (y el eje z el complemento ortogonal del plano xy), y la recta $\boldsymbol{r} = t\boldsymbol{n}$ el complemento ortogonal del plano $\boldsymbol{r} \cdot \boldsymbol{n} = 0$ (y viceversa).



Definición. La proyección ortogonal $P_S(\boldsymbol{v})$ de un vector $\boldsymbol{v} \in \mathbb{R}^n$ sobre un subespacio $S \subset \mathbb{R}^n$, es la proyección sobre S a lo largo de $S_{\perp}$, de forma que se cumpla

$$P_S(\boldsymbol{v}) \in S, \quad \boldsymbol{v} - P_S(\boldsymbol{v}) \in S_{\perp}$$

Así, todo vector $\boldsymbol{v} \in \mathbb{R}^n$ puede escribirse como suma de un vector de S y otro de $S_{\perp}$:

$$\boldsymbol{v} \quad = \quad P_S(\boldsymbol{v}) + (\boldsymbol{v} - P_S(\boldsymbol{v}))$$

El vector $P_S(\boldsymbol{v})$ es entonces el vector de S **más cercano** a $\boldsymbol{v}$, y la **distancia mínima** de $\boldsymbol{v}$ a S es

$$D_{\min} = |\boldsymbol{v} - P_S(\boldsymbol{v})|$$

donde $|\boldsymbol{u}| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$ denota la longitud del vector $\boldsymbol{u}$.

En efecto, la distancia al cuadrado de $\boldsymbol{v}$ a un vector arbitrario $\boldsymbol{w}_S \in S$ es

$$\begin{array}{rcl} D^2 &=& |\boldsymbol{v} - \boldsymbol{w}_S|^2 \\ &=& |\boldsymbol{v} - P_S(\boldsymbol{v}) - (\boldsymbol{w}_S - P_S(\boldsymbol{v}))|^2 \\ &=& |\boldsymbol{v} - P_S(\boldsymbol{v})|^2 + |\boldsymbol{w}_S - P_S(\boldsymbol{v})|^2 \\ &\geq& |\boldsymbol{v} - P_S(\boldsymbol{v})|^2 \end{array}$$

donde hemos utilizado que $\boldsymbol{v} - P_S(\boldsymbol{v})$ es ortogonal a $\boldsymbol{w}_S - P_S(\boldsymbol{v})$ dado que $\boldsymbol{v} - P_S(\boldsymbol{v}) \in S_\perp$ y $\boldsymbol{w}_S - P_S(\boldsymbol{v}) \in S$. La distancia mínima se obtiene entonces para $\boldsymbol{w}_S = P_S(\boldsymbol{v})$, en cuyo caso $D = |\boldsymbol{v} - P_S(\boldsymbol{v})| = D_{\min}$. ∎

Veamos ahora como proyectar sobre S. Sea $B_S = \{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$ una base **ortogonal** de S. Entonces, como $P_S(\boldsymbol{v}) \in S$,

$$P_S(\boldsymbol{v}) = \alpha_1 \boldsymbol{k}_1 + \ldots + \alpha_m \boldsymbol{k}_m$$

con $\boldsymbol{k}_i \cdot P_S(\boldsymbol{v}) = \alpha_i \boldsymbol{k}_i \cdot \boldsymbol{k}_i$. La condición $\boldsymbol{v} - P_S(\boldsymbol{v}) \in S_\perp$ implica

$$\boldsymbol{k}_i \cdot (\boldsymbol{v} - P_S(\boldsymbol{v})) = 0\,, \quad i = 1, \ldots, m$$

es decir, $\boldsymbol{k}_i \cdot \boldsymbol{v} = \boldsymbol{k}_i \cdot P_S(\boldsymbol{v}) = \alpha_i\, \boldsymbol{k}_i \cdot \boldsymbol{k}_i$, de donde

$$\alpha_i = \frac{\boldsymbol{k}_i \cdot \boldsymbol{v}}{\boldsymbol{k}_i \cdot \boldsymbol{k}_i}\,, \quad i = 1, \ldots, m$$

Entonces $\alpha_i \boldsymbol{k}_i = P_{\boldsymbol{k}_i}(\boldsymbol{v})$, por lo que

$$P_S(\boldsymbol{v}) = P_{\boldsymbol{k}_1}(\boldsymbol{v}) + \ldots + P_{\boldsymbol{k}_m}(\boldsymbol{v}) = \frac{\boldsymbol{k}_1 \cdot \boldsymbol{v}_1}{\boldsymbol{k}_1 \cdot \boldsymbol{k}_1}\boldsymbol{k}_1 + \ldots + \frac{\boldsymbol{k}_m \cdot \boldsymbol{v}_m}{\boldsymbol{k}_m \cdot \boldsymbol{k}_m}\boldsymbol{k}_m$$

es decir, $P_S = P_{\boldsymbol{k}_1} + \ldots + P_{\boldsymbol{k}_m}$. La proyección ortogonal de $\boldsymbol{v}$ sobre S es pues *la suma de las proyecciones ortogonales de $\boldsymbol{v}$ sobre los vectores de una base ortogonal de S.*

Para obtener $P_S(\boldsymbol{v})$, basta entonces con disponer de una base ortogonal de S, y sumar las proyecciones sobre los vectores de esta base. Y para obtener una base ortogonal de S, puede aplicarse Gram-Schmidt a una base arbitraria de S.

Ejemplo 1.3.1. Sea $S = \{(x, y, z) \in \mathbb{R}^3,\ x+y+z = 0\}$. Este es un plano perpendicular a $\boldsymbol{n} = (1, 1, 1)$ que pasa por el origen. Como $x = -y - z$, podemos escribir

$$S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ \ y, z, \in \mathbb{R} \right\}$$

Por lo tanto, $B_S = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ es una base de S, no ortogonal. Mediante Gram-Schmidt, podemos llegar a la base ortogonal (probar!)

$$B_S^o = \left\{ \boldsymbol{k}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},\ \boldsymbol{k}_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\}$$

De esta forma

$$P_S(\boldsymbol{v}) = P_{\boldsymbol{k}_1}(\boldsymbol{v}) + P_{\boldsymbol{k}_2}(\boldsymbol{v}) = \frac{\boldsymbol{k}_1 \cdot \boldsymbol{v}}{\boldsymbol{k}_1 \cdot \boldsymbol{k}_1}\boldsymbol{k}_1 + \frac{\boldsymbol{k}_2 \cdot \boldsymbol{v}}{\boldsymbol{k}_2 \cdot \boldsymbol{k}_2}\boldsymbol{k}_2$$

Por ejemplo, para $\boldsymbol{v} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$,

$$P_S(\boldsymbol{v}) = \frac{2}{2}\boldsymbol{k}_1 + \frac{6}{6}\boldsymbol{k}_2 = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}$$

y la distancia mínima de $\boldsymbol{v}$ a S es

$$D_{\min} = |\boldsymbol{v} - P_S(\boldsymbol{v})| = \sqrt{12} = 2\sqrt{3}$$

Dimensión de $S_\perp$. Si ahora $B_{S_\perp} = \{\boldsymbol{k}_{m+1}, \ldots, \boldsymbol{k}_p\}$ es una base **ortogonal** de $S_\perp$, entonces

$$\boldsymbol{v} - P_S(\boldsymbol{v}) = \alpha_{m+1}\boldsymbol{k}_{m+1} + \ldots + \alpha_p \boldsymbol{k}_p$$

con $\alpha_j = \boldsymbol{k}_j \cdot \boldsymbol{v}/(\boldsymbol{k}_j \cdot \boldsymbol{k}_j)$ para $j = m+1, \ldots, p$ ya que $\boldsymbol{k}_j \cdot P_S(\boldsymbol{v}) = 0$. Podemos así escribir cualquier vector $\boldsymbol{v} \in \mathbb{R}^n$ como

$$\begin{aligned} \boldsymbol{v} &= P_S(\boldsymbol{v}) + (\boldsymbol{v} - P_S(\boldsymbol{v})) \\ &= \alpha_1 \boldsymbol{k}_1 + \ldots + \alpha_m \boldsymbol{k}_m + \alpha_{m+1}\boldsymbol{k}_{m+1} + \ldots + \alpha_p \boldsymbol{k}_p \qquad (*) \end{aligned}$$

Es fácil ver ahora que $B = B_S \cup B_{S_\perp} = \{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m, \boldsymbol{k}_{m+1}, \ldots, \boldsymbol{k}_p\}$ es una base de $\mathbb{R}^n$:

a) La expresión (*) muestra que el conjunto $\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m, \boldsymbol{k}_{m+1}, \ldots, \boldsymbol{k}_p\}$ *genera* $\mathbb{R}^n$, ya que todo $\boldsymbol{v} \in \mathbb{R}^n$ puede escribirse como combinación lineal de ellos.
b) El conjunto $\{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m, \boldsymbol{k}_{m+1}, \ldots, \boldsymbol{k}_p\}$ es *linealmente independiente*, ya que son todos no nulos y ortogonales entre si ($\boldsymbol{k}_i \cdot \boldsymbol{k}_j = 0$ si $\boldsymbol{k}_i \in B_S$ y $\boldsymbol{k}_j \in B_{S_\perp}$, por pertenecer a subespacios ortogonales, y además $\boldsymbol{k}_i \cdot \boldsymbol{k}_j = 0$ si $i \neq j$ y ambos pertenecen al mismo subespacio, ya que forman una base ortogonal del mismo).

Por lo tanto $p = n$ y entonces $\dim S_\perp = p - m = n - \dim S$, o sea,

$$\dim S + \dim S_\perp = n$$

La suma de las dimensiones de un subespacio S y de su complemento ortogonal $S_\perp$ es siempre la dimensión total del espacio. Además, se tiene

$$P_S + P_{S_\perp} = I_n$$

con I_n la identidad de $n \times n$ y $P_{S_\perp} = P_{\boldsymbol{k}_{m+1}} + \ldots + P_{\boldsymbol{k}_n} = I_n - P_S$ el proyector sobre $S_\perp$.

Ejemplo 1.3.2.. Siguiendo con el ej. **1.3.1**, se tiene $S_\perp = \{t(1,1,1), t \in \mathbb{R}\}$, ya que S es el plano ortogonal a $\boldsymbol{n} = (1,1,1)$. Por lo tanto, $B_{S_\perp} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ y $B = B_S^o \cup B_{S_\perp} = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ es una base ortogonal de $\mathbb{R}^3$, cumpliéndose que $\dim S = 2$, $\dim S_\perp = 1$ y $\dim S + \dim S_\perp = 3$.

Ejemplo 1.3.3. Espacio fila y espacio nulo de una matriz $A \in \mathbb{R}^{m\times n}$.

Si $S = EF(A) \subset \mathbb{R}^n$ es el espacio fila de A, su espacio nulo $Nu(A) = \{X \in \mathbb{R}^n,\ AX = 0\}$ (el conjunto solución del sistema homogéneo asociado) *es el espacio ortogonal* $S_\perp$.

En efecto, $AX = 0$ implica que X debe ser ortogonal a todas las filas de A, por lo que X será ortogonal a cualquier vector de $S = EF(A)$, es decir, del espacio generado por las filas. La relación dimensional $\dim S + \dim S_\perp = n$ implica entonces

$$\dim EF(A) + \dim Nu(A) = n$$

Pero esta igualdad es el teorema rango-nulidad, ya que $\dim EF(A) = r(A)$ (rango de A) y $\dim Nu(A) = nul(A)$ (nulidad de A).

Obtención de $S_\perp$. Sea $B_S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ una base arbitraria de un subespacio S de $\mathbb{R}^n$ de dimensión m. Para obtener $S_\perp$, deben resolverse las ecuaciones

$$\boldsymbol{v}_i \cdot \boldsymbol{w} = 0,\ \ i = 1, \ldots, m$$

ya que esto asegura que $\boldsymbol{w}$ será ortogonal a todo vector de S (justificar!).

Sea $A = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ la matriz de $n\times m$ formada por los m vectores de la base puestos en columna. Entonces A^T de $m \times n$ tendrá los m vectores $\boldsymbol{v}_i$ en forma de fila (tal que $S = EC(A) = EF(A^T)$). Si escribimos $\boldsymbol{w}$ como vector columna de $n\times 1$, las m ecuaciones anteriores equivalen al sistema de $m \times n$

$$A^T\boldsymbol{w} = \mathbf{0}$$

Su conjunto solución $Nu(A^T)$ (espacio nulo de A^T) es entonces $S_\perp$, y una base de $Nu(A^T)$ será una base de $S_\perp$.

Ejemplo 1.3.4. Consideremos nuevamente el subespacio S del ejemplo **1.3.1**.

Tenemos $B_S = \left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right\}$. Por lo tanto, una forma de hallar $S_\perp$ es resolver el sistema $A^T\boldsymbol{w} = \mathbf{0}$, con $A^T = \begin{pmatrix}-1 & 1 & 0\\-1 & 0 & 1\end{pmatrix}$ y $\boldsymbol{w} = \begin{pmatrix}x\\y\\z\end{pmatrix}$. Se obtiene

$$\begin{pmatrix}-1 & 1 & 0\\-1 & 0 & 1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \quad\Rightarrow\quad \left(\begin{array}{ccc|c}-1 & 1 & 0 & 0\\-1 & 0 & 1 & 0\end{array}\right) \to \left(\begin{array}{ccc|c}-1 & 1 & 0 & 0\\0 & -1 & 1 & 0\end{array}\right)$$

que implica $y = z$, $x = z$, con z libre. Por lo tanto, $\boldsymbol{w} = z\begin{pmatrix}1\\1\\1\end{pmatrix}$ y $B_{S_\perp} = \left\{\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$.

Este resultado coincide con el obtenido en **1.3.2** para $B_{S_\perp}$ a partir de la definición original de S. Se obtiene también al mismo resultado si se emplea la base ortogonal B_S^o en lugar de B_S (probar!).

Obtención directa de $P_S(\boldsymbol{v})$: Es también posible obtener la proyección ortogonal $P_S(\boldsymbol{v})$ sobre un subespacio $S \subset \mathbb{R}^n$ partiendo directamente de una base arbitraria del mismo, no necesariamente ortogonal. Si $B_S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ es una base de S, entonces

$$P_S(\boldsymbol{v}) = c_1\boldsymbol{v}_1 + \ldots + c_m\boldsymbol{v}_m = A\boldsymbol{c}$$

donde $A = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ es la matriz de $n \times m$ cuyas m columnas son los vectores de la base. Entonces la condición de que $\boldsymbol{v} - P_S(\boldsymbol{v})$ sea ortogonal a S equivale a que sea ortogonal a todo vector $\boldsymbol{v}_i$ de B_S, lo que implica

$$A^T(\boldsymbol{v} - A\boldsymbol{c}) = 0$$

Por lo tanto, $A^T\mathrm{v} - A^TA\boldsymbol{c} = \mathbf{0}$, o sea $A^TA\boldsymbol{c} = A^T\boldsymbol{v}$, de donde (ver problema 3.8 a)

$$\boldsymbol{c} = (A^TA)^{-1}A^T\boldsymbol{v}$$

Finalmente,

$$P_S(\boldsymbol{v}) = A(A^TA)^{-1}A^T\boldsymbol{v}$$

donde la matriz $A(A^TA)^{-1}A^T$, de $n \times n$, es la representación matricial en la base canónica del operador de proyección ortogonal P_S. Esta matriz no depende de la base de S elegida.

Ejemplo 1.3.5.

Consideremos de nuevo el subespacio $S \subset \mathbb{R}^3$ generado por la base $B_S = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

En este caso

$$A = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^T = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

con $A^TA = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $(A^TA)^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ y

$$A(A^TA)^{-1}A^T = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Para $\boldsymbol{v} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$ se obtiene

$$P_S(\boldsymbol{v}) = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}$$

que coincide con el resultado obtenido en el ejemplo **1.3.1**.

Si se emplea la base ortogonal B_S^o hallada en **1.3.1**, $A \to \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ 0 & 2 \end{pmatrix}$ y A^TA se torna diagonal, pero la matriz final $A(A^TA)^{-1}A^T$ (y por lo tanto $P_S(\boldsymbol{v})$) no cambia (probar!).

Problemas

3.1 Si $S=\{(x,y,z)\in\mathbb{R}^3,\ \ x+2y-z=0\}$,

a) Hallar una base de S.

b) Determinar una base ortogonal de S utilizando Gram-Schmidt.

c) Hallar la proyección ortogonal de $\boldsymbol{v}=(1,1,1)$ sobre S y la distancia mínima de $\boldsymbol{v}$ a S (usar las proyecciones sobre los vectores de la base ortogonal).

d) Verificar el resultado anterior aplicando la fórmula general matricial $P_S(\boldsymbol{v})=A(A^TA)^{-1}A^T\boldsymbol{v}$, usando en A la base hallada en a).

3.2 Idem anterior para $S=\{(x,y,z)\in\mathbb{R}^3,\ \ 2x+y=0\}$ y $\boldsymbol{v}=(1,1,1)$.

3.3 Determinar el subespacio ortogonal $S_\perp$ en los dos casos anteriores, y escribir $\boldsymbol{v}$ como suma de un vector de S y otro de $S_\perp$.

3.4 Encontrar el subespacio ortogonal $S_\perp$, si S es el subespacio generado por las bases

a) $B_S=\{(1,2)\}$ $(V=\mathbb{R}^2)$

b) $B_S=\{(1,2,3)\}$ $(V=\mathbb{R}^3)$

c) $B_S=\{(1,2,3),(1,1,1)\}$ $(V=\mathbb{R}^3)$

d) $B_S=\{(1,1,1,1),(1,0,1,0)\}$ $(V=\mathbb{R}^4)$

e) Encontrar en todos los casos una base de $B_{S_\perp}$ y verificar que $B_S\cup B_{S_\perp}$ es base del espacio V, con $\dim S+\dim S_\perp=\dim V$.

3.5 ¿Cual es el complemento ortogonal del subespacio nulo? ¿Qué es la proyección ortogonal sobre dicho subespacio?

3.6 Mostrar que si un vector es ortogonal a todos los vectores de una base de un subespacio S, entonces es ortogonal a todo vector de S.

3.7 Mostrar qué el único vector que puede pertenecer a S y a $S_\perp$ es el vector nulo.

3.8 *a*) Mostrar que las columnas de una matriz A de $m\times n$ forman un conjunto linealmente independiente si y sólo si A^TA es no singular.

b) Probar que si las columnas de una matriz A son mutuamente ortogonales y no nulas, entonces A^TA es diagonal y no singular.

c) Mostrar que la matriz de proyección ortogonal $P_S=A(A^TA)^{-1}A^T$ sobre un subespacio S es simétrica. ¿En qué caso se reduce a una matriz identidad?

d) Probar que dicha matriz de proyección es invariante frente a un cambio de base de S.

3.9 Probar que $D^2=|\boldsymbol{v}-A\boldsymbol{c}|^2$ es mínimo cuando $\boldsymbol{c}$ satisface la ecuación $A^T(\boldsymbol{v}-A\boldsymbol{c})=0$, o sea, $A^TA\boldsymbol{c}=A^T\boldsymbol{v}$, y que esto implica $A\boldsymbol{c}=P_S(\boldsymbol{v})$, con $S=EC(A)$ el espacio columna de A.

1.4 Proyección ortogonal y aproximación de cuadrados mínimos

La proyección ortogonal de un vector $\boldsymbol{v} \in \mathbb{R}^n$ sobre un subespacio $S \subset \mathbb{R}^n$ puede también verse como la mejor aproximación al vector $\boldsymbol{v}$ basada en un vector de S. Si se desea aproximar $\boldsymbol{v}$ mediante un vector $\boldsymbol{w}_S \in S$, la mejor aproximación se logra eligiendo $\boldsymbol{w}_S$ como el vector de S **más cercano** a $\boldsymbol{v}$, y tal vector es entonces la proyección ortogonal

$$\boldsymbol{v}_S = P_S(\boldsymbol{v})\,.$$

Un ejemplo es el sistema de n ecuaciones con m incógnitas

$$A\boldsymbol{c} = \boldsymbol{v}$$

donde A es de $n \times m$, $\boldsymbol{c}$ de $m \times 1$ y $\boldsymbol{v}$ de $n \times 1$, con $n \geq m$, tal que el número de ecuaciones es mayor que el número de incógnitas. Si las columnas de A son independientes, entonces la solución, si existe, es única, pero típicamente el sistema será incompatible, ya que las columnas de A generan sólo un subespacio $S = EC(A)$ de dimensión $m < n$ de $\mathbb{R}^n$.

En estos casos, se elige en general la "solución" **de cuadrados mínimos**, que es la que **minimiza** la cantidad

$$D^2 = |\boldsymbol{v} - A\boldsymbol{c}|^2$$

Pero D^2 es el cuadrado de la distancia de $\boldsymbol{v}$ a un vector $A\boldsymbol{c}$ de S, y por lo tanto la distancia mínima se logra para $\boldsymbol{c}$ tal que $A\boldsymbol{c}$ sea la **proyección ortogonal** de $\boldsymbol{v}$ sobre S:

$$A\boldsymbol{c} = P_S(\boldsymbol{v})$$

de forma que

$$D = |\boldsymbol{v} - P_S(\boldsymbol{v})| = D_{\min}$$

Utilizando las formulas de la sección anterior tenemos que tal $\boldsymbol{c}$ estará dado por

$$\boldsymbol{c} = (A^T A)^{-1} A^T \boldsymbol{v}$$

es decir, $\boldsymbol{c}$ es la solución del sistema cuadrado de $m \times m$

$$A^T A\boldsymbol{c} = A^T \mathbf{v}$$

donde $A^T A$ es de $m \times m$ y no singular.

Un ejemplo corriente de aplicación es el ajuste de los valores $f(x_i)$ de una función $f : \mathbb{R} \to \mathbb{R}$ para un conjunto de puntos $(x_1, \ldots, x_n)$, por un polinomio de grado k,

$$P(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_k x^k \tag{I}$$

con k normalmente menor o mucho menor que n, y $c_0, \ldots, c_k$ los coeficientes a ajustar. Lo que se desea es elegir el polinomio que minimiza la suma de la diferencia de cuadrados

$$D^2 = \sum_{i=1}^{n} [f(x_i) - P(x_i)]^2 \tag{II}$$

ya que en general no existirá un polinomio de grado $k < n-1$ que reproduzca exactamente los n datos $f(x_i)$, es decir, que satisfaga

$$P(x_i) = f(x_i) \quad i = 1, \ldots, n \tag{III}$$

Las ecuaciónes (III) forman un sistema de ecuaciones **lineales** en los coeficientes c_i de $P(x)$:

$$c_0 + c_1 x_i + c_2 x_i^2 + \ldots + c_k x_i^k = f(x_i), \;\; i = 1, \ldots, n$$

que puede escribirse en forma matricial como

$$A\boldsymbol{c} = \boldsymbol{f} \tag{IV}$$

con

$$A = \begin{pmatrix} 1 & x_1 & \ldots & x_1^k \\ 1 & x_2 & \ldots & x_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \ldots & x_n^k \end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_k \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix}$$

siendo A de $n \times m$, $\boldsymbol{c}$ de $m \times 1$ y $\boldsymbol{f}$ de $n \times 1$, con $m = k+1$. El sistema (IV) es normalmente incompatible ya que en general $k < n$ y típicamente $k \ll n$.

La cantidad (II) resulta así la distancia al cuadrado entre los vectores $\boldsymbol{f}$ y $A\boldsymbol{c}$:

$$D^2 = |\boldsymbol{f} - A\boldsymbol{c}|^2$$

El vector $\boldsymbol{c}$ que minimiza D^2 es pues aquel tal que $A\boldsymbol{c} = P_S(\boldsymbol{f})$, donde S es el subespacio de $\mathbb{R}^n$ generado por las m columnas de A. Está dado entonces por

$$\boldsymbol{c} = (A^T A)^{-1} A^T \boldsymbol{f} \tag{V}$$

y el valor mínimo de D por

$$D_{\min} = |\boldsymbol{f} - P_S(\boldsymbol{f})|$$

con $P_S(\boldsymbol{f}) = A\boldsymbol{c} = A(A^T A)^{-1} A^T \boldsymbol{f}$.

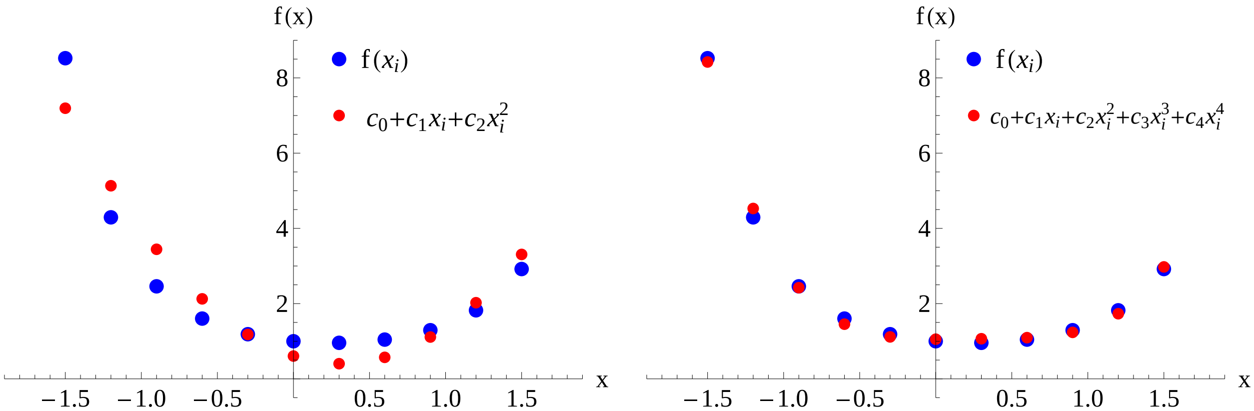
Notemos que $A^T A$ es una matriz simétrica de $m \times m$, y $A^T \boldsymbol{f}$ un vector de $m \times 1$, de elementos

$$(A^T A)_{jl} = \sum_{i=1}^{n} x_i^{j+l-2}, \quad (A^T \boldsymbol{f})_j = \sum_{i=1}^{n} x_i^{j-1} f(x_i)$$

con $j, l, = 1, \ldots, m$ y $m = k + 1$.

Nótese que la dimensión de $A^T A$ depende sólo del grado del polinomio y no del número de puntos a ajustar. Por ejemplo, si $k = 2$, $P(x) = c_0 + c_1 x + c_2 x^2$ es un polinonio de grado 2 y $A^T A$ es de 3×3, independientemente del número n de valores a ajustar.

Se muestra en la figura la aproximación de 11 datos $f(x_i)$ por medio de un polinomio de grado 2 y uno de grado 4. No obstante, debe tenerse en cuenta que si los datos no exhiben un comportamiento polinomial, el ajuste puede presentar oscilaciones, cuya magnitud puede aumentar al incrementarse el grado k del polinomio. La matriz $A^T A$ se torna también mal condicionada al aumentar k.



El formalismo anterior puede extenderse a un ajuste basado en la combinación lineal de funciones linealmente independientes $g_j(x)$, reemplazando $P(x)$ por

$$G(x) = c_1 g_1(x) + \ldots + c_m g_m(x)$$

En este caso $A_{ij} = g_j(x_i)$, y el formalismo puede aplicarse si $A^T A$ es no singular, con

$$(A^T A)_{jl} = \sum_{i=1}^{n} g_j(x_i) g_l(x_i)\,, \quad (A^T \boldsymbol{f})_j = \sum_{i=1}^{n} g_j(x_i) f(x_i)\,, \quad j, l = 1, \ldots, m$$

Ejercicios

4.1 Verifique que la minimización de D^2 (II) respecto de los coeficientes c_j conduce a la solución (V).
4.2 Determine el polinomio $P_2(x) = c_0 + c_1 x + c_2 x^2$ que minimiza D^2 si los datos son $f(-1{,}5) = 8{,}5$, $f(-0{,}9) = 2{,}5$, $f(-0{,}3) = 1$, $f(0) = 0{,}5$, $f(0{,}3) = 0{,}6$, $f(0{,}9) = 1{,}5$, $f(1{,}5) = 3$ (Utilice PC). Estos corresponden a la posición (en una cierta escala) en distintos tiempos de un móvil con aceleración variable. Si la aceleración fuese constante, ¿como sería el ajuste con este polinomio?

Facultad de Ingeniería
UNLP

Matemática C

VII Números Complejos

7.1. Conceptos básicos

7.1.1. Introducción

Los números complejos constituyen una extensión de los números reales, que permite obtener todas las raíces de cualquier polinomio. El conjunto de números complejos, denotado por la letra $\mathbb{C}$, forma así lo que se denomina un conjunto algebraicamente cerrado. Geométricamente, un número complejo puede representarse mediante un punto en un plano. El concepto de número compejo extiende pues la recta real a un espacio bidimensional, denominado *plano complejo*.

Los números complejos son además utilizados, como veremos, en la representación de funciones trigonométricas y por ende, de funciones *periódicas*, siendo por lo tanto su uso muy extendido en Ingeniería y Física (representación de corriente alterna, análisis de Fourier, etc.). Asimismo, proporcionan una representación alternativa de vectores en dos dimensiones, la cual resulta muy conveniente en ciertos problemas bidimensionales (utilizada por ejemplo en dinámica de fluidos, electromagnetismo, etc.). La Mecánica Cuántica está también formulada en términos de una función de onda compleja. En este curso los utilizaremos en los próximos dos capítulos (autovalores y ecuaciones diferenciales). Históricamente, su introducción se atribuye al matemático italiano Girolamo Cardano (siglo XVI), en el marco de su solución de ecuaciones cúbicas y cuárticas.

7.1.2. Definición

Comenzamos por definir el número i (unidad imaginaria), que satisface

$$i^2 = -1$$

De esta forma, la ecuación

$$z^2 = -1$$

tendrá las raíces $z = i$ y $z = -i$. Mencionemos que en ciertos contextos se utilizan otras notaciones para la unidad imaginaria (por ejemplo, en ingeniería eléctrica se utiliza j en lugar de i, ya que i denota la corriente eléctrica).

Un número complejo es de la forma

$$z = x + y\,i, \qquad x, y \in \mathbb{R}$$

donde tanto x como y son números reales. La **parte real** de z es x, y la **parte imaginaria** de z es y (también un número real!):

$$\mathrm{Re}(z) = x, \quad \mathrm{Im}(z) = y$$

Ejercicio: Hallar la parte real e imaginaria de a) $z = -2 - 3i$, b) $z = -i$.

Un número complejo cuya parte imaginaria es 0 es identificado como un número real. Por ejemplo, $5 = 5 + 0i$, $0 = 0 + 0i$.
También, $0 + yi = yi$. Un número complejo no nulo cuya parte real es 0, tal como $5i$, se

dice que es **imaginario** o imaginario puro.
Además, $yi = iy$, y $x + yi = yi + x$. Por ejemplo, $1 + 2i = 1 + i2 = 2i + 1$.

Dos números complejos $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, son iguales si y sólo si tanto sus partes reales como imaginarias son respectivamente iguales:

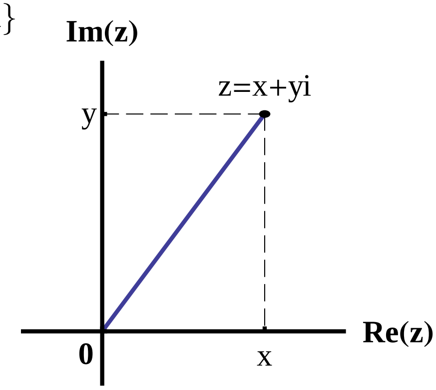
$$x_1 + y_1 i = x_2 + y_2 i \Leftrightarrow x_1 = x_2\,, \quad y_1 = y_2 \qquad (x_1, x_2,\ y_1, y_2 \in \mathbb{R})$$

Por ejemplo, $x + yi = 2 + 3i$ implica $x = 2$, $y = 3$ si x e y son reales.

Como hemos mencionado, el conjunto de todos los números complejos se denota con la letra $\mathbb{C}$:

$$\mathbb{C} = \{x + iy, \ \ x \in \mathbb{R},\ y \in \mathbb{R}\}$$

Podemos representar un número complejo $z = x + yi$ mediante un par ordenado (x, y). Este par corresponde a un punto o vector en el plano cartesiano, que en este contexto se denomina **plano complejo**.



7.1.3. Conjugación

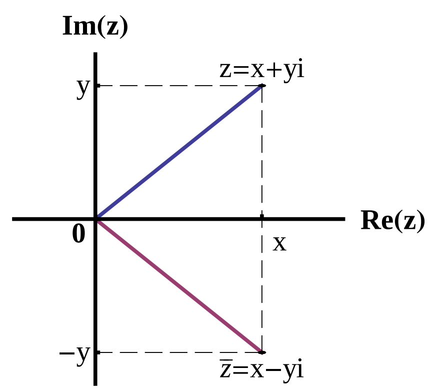
El **conjugado** de un número complejo $z = x + iy$ se lo denota como $\bar{z}$ o z^*, y se lo define como

$$\bar{z} = x - iy$$

de forma que $\mathrm{Re}(\bar{z}) = \mathrm{Re}(z)$, $\mathrm{Im}(\bar{z}) = -\mathrm{Im}(z)$.

Por ejemplo, $\bar{i} = -i$, $\overline{1 + 2i} = 1 - 2i$ y $\overline{1 - 2i} = 1 + 2i$ y .

Geométricamente, conjugar corresponde a reflejar z respecto del eje real.

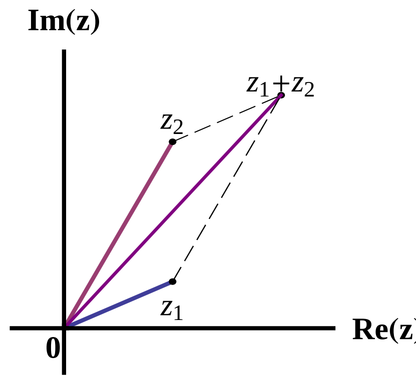


Obviamente, $\bar{\bar{z}} = z$ (o sea, $(z^*)^* = z$).

7.1.4. Suma de números complejos

Dados dos números complejos $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$, su suma se define como

$$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i$$



Por ejemplo, $(1 + 2i) + (3 - 4i) = 4 - 2i$.

Geométricamente, $z_1 + z_2$ corresponde al vector suma de los vectores que representan a z_1 y z_2.

La resta $z_1 - z_2$ es la suma de z_1 y el opuesto $-z_2 = -x_2 - y_2 i$ de z_2:

$$z_1 - z_2 = (x_1 - x_2) + (y_1 - y_2)i$$

Por ejemplo, $(1 + 2i) - (3 - 4i) = -2 + 6i$, $(1 + 2i) - (1 + 2i) = 0$.
Ejercicio: Indique el significado geométrico de $z_1 - z_2$.

Problema 1: Mostrar que el conjugado de una suma es la suma de los conjugados: y que el conjugado de una resta es la resta de los conjugados:

$$\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$$

Problema 2: Mostrar que es posible expresar la parte real y la parte imaginaria de un número complejo z en términos de z y su conjugado $\bar{z}$:

$$\mathrm{Re}(z) = \frac{z + \bar{z}}{2}, \quad \mathrm{Im}(z) = \frac{z - \bar{z}}{2i}$$

7.1.5. Producto de dos números complejos

A partir de la propiedad básica $i^2 = -1$, el producto de dos números complejos $z_1 z_2 = (x_1 + y_1 i)(x_2 + y_2 i)$ se realiza aplicando la propiedad distributiva (respecto de la suma) y asociativa (respecto del producto):

$$\begin{aligned} (x_1 + y_1 i)(x_2 + y_2 i) &= x_1(x_2 + y_2 i) + y_1 i(x_2 + y_2 i) \\ &= x_1 x_2 + x_1 y_2 i + y_1 x_2 i - y_1 y_2 \\ &= (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i \end{aligned} \tag{1.1}$$

Por ejemplo,

$$\begin{aligned} (1 + 2i)(3 - 4i) &= 3 - 4i + 2i(3 - 4i) = 3 - 4i + 6i + 8 \\ &= 11 + 2i \end{aligned} \tag{1.2}$$

Problema 3: Probar que a) $z_1 z_2 = z_2 z_1$ b) $z_1(z_2 z_3) = (z_1 z_2) z_3$ c) $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$

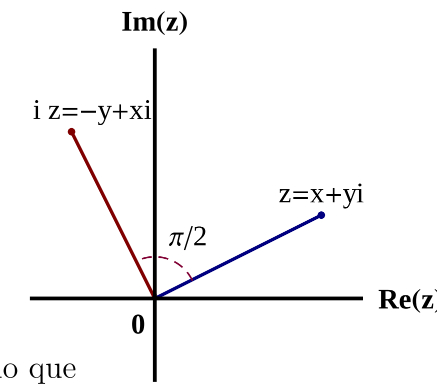
Problema 4: Probar que el conjugado de un producto es el producto de los conjugados:

$$\overline{z_1 z_2} = \overline{z}_1 \, \overline{z}_2$$

En particular, el producto de un número complejo z por un número real α,

$$\alpha(x + yi) = \alpha x + \alpha yi$$

corresponde, geométricamente, a la multiplicación del vector que representa a z por el escalar real α. Por ejemplo, $3(1 + i) = 3 + 3i$.



Por otro lado, el producto de un número complejo z por la unidad imaginaria i,

$$i(x + yi) = -y + xi$$

corresponde, geométricamente, a **rotar** el vector que representa a z un ángulo de 90^o en sentido antihorario. Por ejemplo, $i1 = i$, $ii = -1$, $i(1 + i) = -1 + i$.

Problema 5: Probar el enunciado general anterior, recordando que la matriz que representa a dicha rotación en la base canónica de $\mathbb{R}^2$ es

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{1.3}$$

de forma que $R\binom{x}{y} = \binom{-y}{x}$.
En consecuencia, multiplicar un número complejo z por un número imaginario αi (α real) corresponde, geométricamente, a rotar z un ángulo de 90^o en sentido antihorario y luego multiplicar el vector resultante por el escalar real α.

El significado geométrico de multiplicar a z por un número complejo arbitrario, $(\alpha + i\beta)z = \alpha z + i\beta z$, puede obtenerse sumando αz y el vector rotado $i\beta z$, pero es más fácil visualizar el resultado final por medio de la representación polar de un número complejo, que discutiremos luego.

Potencias de i: Las potencias pares son **reales:** $i^2 = -1$, $i^4 = (i^2)^2 = 1$ y en general,

$$i^{2n} = (i^2)^n = (-1)^n \tag{1.4}$$

Las potencias impares son en cambio **imaginarias:** $i^1 = i$, $i^3 = (i^2)i = -i$ y en general,

$$i^{2n+1} = i^{2n} i = (-1)^n i \tag{1.5}$$

Ejercicio: Graficar las potencias de i en el plano complejo.

7.1.6. Valor absoluto de un número complejo

El valor absoluto (o módulo) de un número complejo z se define como

$$|z| = \sqrt{x^2 + y^2} \tag{1.6}$$

y es la longitud del vector que lo representa en el plano complejo. El módulo $|z|$ es siempre un número real no negativo. Por ejemplo: $|3+4i| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$, $|i| = 1$, $|-2| = 2$.

Es posible expresar $|z|$ mediante z y su conjugado $\bar{z}$ como

$$|z| = \sqrt{z\bar{z}}$$

En efecto, si $z = x+yi$, con x, y reales, $z\bar{z} = (x+yi)(x-iy) = x^2 - (iy)^2 = x^2+y^2 = |z|^2$.

Problema 6: Probar que $|z| = 0$ si y sólo si $z = 0$.
Problema 7: Probar que $|\bar{z}| = |z|$.
Problema 8: Para z_1, z_2 complejos generales, probar que

$$|z_1 z_2| = |z_1||z_2|$$

$$||z_1| - |z_2||| \leq |z_1 + z_2| \leq |z_1| + |z_2|$$

Problema 9: Probar que $|z_1 - z_2|$ es, geométricamente, la **distancia** entre z_1 y z_2.

7.1.7. Inverso de un número complejo

El inverso (o recíproco) z^{-1} de un número complejo $z \neq 0$ es el número complejo que satisface

$$zz^{-1} = 1$$

Para obtenerlo, vimos en el punto anterior que $z\bar{z} = |z|^2 \neq 0$ si $z \neq 0$. Por lo tanto,

$$z(\bar{z}/|z|^2) = |z|^2/|z|^2 = 1$$

de donde

$$z^{-1} \;=\; \frac{\bar{z}}{|z|^2} = \frac{x - yi}{x^2 + y^2}, \quad z \neq 0 \tag{1.7}$$

Notamos aquí que el cociente z/α de un número complejo $z = x + yi$ por un número real $\alpha \neq 0$ se define como $(1/\alpha)z$, es decir, $z/\alpha = x/\alpha + (y/\alpha)i$. Por ejemplo, $(2+4i)/2 = 1+2i$.

Problema 10: Si $z = x + yi \neq 0$, probar, escribiendo $z^{-1} = a + ib$, que la ecuación $zz^{-1} = 1$ (o sea, $xa - yb + (xb + ya)i = 1 + 0i$) implica necesariamente $a = x/|z|^2$, $b = -y/|z|^2$.

Podemos ahora definir

$$\frac{1}{z} = z^{-1} = \frac{\bar{z}}{|z|^2}$$

resultado que puede obtenerse multiplicando numerador y denominador de $\frac{1}{z}$ por $\bar{z}$. Por ejemplo, $\frac{1}{i} = \frac{-i}{-i^2} = -i$:

$$\frac{1}{i} = -i$$

verificándose que $i(-i) = -i^2 = 1$. Como segundo ejemplo,

$$\frac{1}{1+i} = \frac{1}{1+i}\frac{1-i}{1-i} = \frac{1-i}{2} = \frac{1}{2} - \frac{1}{2}i$$

verificándose que $(1+i)(1-i)/2 = 1$.

Problema 11: Probar que para $z \neq 0$,

$$\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$$

$$\left|\frac{1}{z}\right| = \frac{1}{|z|}$$

7.1.8. Cociente de dos números complejos

Dados ahora dos números complejos z_1 y z_2, con $z_2 \neq 0$, definimos el cociente como

$$\begin{aligned} \frac{z_1}{z_2} &= z_1 z_2^{-1} = \frac{z_1 \bar{z}_2}{|z_2|^2} \\ &= \frac{(x_1 + iy_1)(x_2 - iy_2)}{x_2^2 + y_2^2} \end{aligned} \tag{1.8}$$

En la práctica, se obtiene este resultado multiplicando numerador y denominador por $\bar{z}_2$. Por ejemplo:

$$\frac{1+i}{3+4i} = \frac{(1+i)(3-4i)}{(3+4i)(3-4i)} = \frac{7-i}{25}$$

Problema 12: Probar que para $z_2 \neq 0$,

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}, \qquad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$$

7.1.9. Representaciones reales y propiedades algebraicas

El conjunto $\mathbb{C}$ de los números complejos puede ser considerado como el conjunto $\mathbb{R}^2$ de pares ordenados (x, y) de números reales, con la suma y producto definidos por

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \tag{1.9}$$

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2) \tag{1.10}$$

Puede mostrarse que el producto así definido resulta conmutativo y asociativo ($z_1 z_2 = z_2 z_1$, $(z_1 z_2) z_3 = z_1 (z_2 z_3)$, para $z_j = (x_j, y_j)$, $j = 1, 2, 3$), y distributivo con la suma:

$z_1(z_2+z_3) = z_1z_2 + z_1z_3$. De esta forma, denotando a los elementos de la base canónica como $1 = (1,0)$ y $i = (0,1)$, tenemos

$$i \cdot i = (0,1)\cdot(0,1) = (-1,0) = -(1,0) = -1$$

con $1 \cdot 1 = 1$ y $1 \cdot i = i$. Además, podemos escribir

$$(x,y) = x(1,0) + y(0,1) = x1 + yi$$

con lo cual recuperamos la notación estándar $z = x + yi$ si identificamos x con $x1$.

Los números complejos pueden ser también representados por matrices reales de 2×2 de la forma

$$\begin{aligned} z &= \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} && (1.11) \\ &= xI + yR\,, \quad x, y \in \mathbb{R} && (1.12) \end{aligned}$$

donde $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ es la matriz identidad y $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ la matriz de rotación definida en (1.3). Obtenemos, con el producto matricial usual, $I.I = I$, $I.R = R.I = R$, y

$$R.R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I \qquad (1.13)$$

por lo que el par I, R constituye una representación matricial real de 1 e i, jugando R el rol de la unidad imaginaria i. Si $z_1 = x_1I + y_1R$, $z_2 = x_2I + y_2R$, obtenemos, con la suma y producto matricial usual,

$$z_1 + z_2 = (x_1 + x_2)I + (y_1 + y_2)R$$

$$z_1.z_2 = z_2.z_1 = (x_1x_2 - y_1y_2)I + (x_1y_2 + y_1x_2)R$$

con lo cual, identificando $1 = I$ y $i = R$, recuperamos las reglas de suma y producto de los números complejos. Además, el determinante de la matriz (1.11) es el módulo de z:

$$\begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2 = |z|$$

Como estructura algebraica, el conjunto de números complejos es, al igual que los números reales o racionales, un *cuerpo*. Una estructura algebraica $\{F, +, \cdot\}$, donde F es un cierto conjunto de números y $+$, $\cdot$ son operaciones binarias cerradas entre ellos, es un cuerpo (o campo) si:

i) La suma satisface las propiedades de asociatividad, conmutatividad, existencia de elemento neutro (0, tal que $z+0 = z\ \forall\ z \in F$) y opuesto ($\forall\ z \in F\ \exists\, -z$ tal que $z + (-z) = 0$).

ii) El producto es también asociativo y conmutativo, con existencia de elemento neutro 1 ($z \cdot 1 = z\ \forall\ z \in F$) e inverso $z^{-1}\ \forall\ z \neq 0$ (tal que $z \cdot z^{-1} = 1$)

iii) El producto es distributivo: $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$.

Finalmente, considerando escalares **reales**, el conjunto de los números complejos $\mathbb{C}_\mathbb{R}$ forma un **espacio vectorial de dimensión 2**: Todo complejo $z = x1 + yi$ es la combinación lineal de 1 e i, con coeficientes reales x, y, siendo $\{1,\ i\}$ linealmente independientes $(x1 + yi = 0 \Rightarrow x = 0,\, y = 0)$.

Si consideramos en cambio escalares **complejos**, $\mathbb{C}_\mathbb{C}$ es un espacio vectorial de dimensión 1 (Probar!).

7.1.10. Raíces de polinomios

Mediante los números complejos, es posible encontrar todas las raíces de un polinomio de grado arbitrario $n \geq 1$, y así escribir el mismo como el producto de n polinomios elementales de grado 1.

Consideremos primero un polinomio de grado 2, $P_2(z) = az^2 + bz + c$, con $a \neq 0$, y la correspondiente ecuación cuadrática

$$az^2 + bz + c = 0, \quad (a \neq 0) \tag{1.14}$$

Sabemos que las raíces están dadas por la muy conocida fórmula

$$z_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{1.15}$$

Problema 13: Obtener la fórmula (1.15), completando cuadrados.
Problema 14: Probar que para $a \neq 0$,

$$az^2 + bz + c = a(z - z_+)(z - z_-) \tag{1.16}$$

Para a, b y c reales, las raíces $z_\pm$ serán entonces:
i) Reales y distintas si $b^2 > 4ac$
ii) Reales e iguales si $b^2 = 4ac$
iii) **Complejas conjugadas** si $b^2 < 4ac$:
En este caso $b^2 - 4ac = -(4ac - b^2)$, con $4ac - b^2 > 0$, y por lo tanto,

$$z_\pm = \frac{-b \pm i\sqrt{4ac - b^2}}{2}$$

con $z_- = \bar{z}_+$. Mediante la introducción de los números complejos, podemos entonces obtener siempre raíces de la ecuación cuadrática (1.15), aún si $b^2 > 4ac$, y así escribir el polinomio $az^2 + bz + c$ en la forma factorizada (1.16).

Por ejemplo, la ecuación

$$z^2 + 2z + 2 = 0$$

posee las raíces complejas conjugadas

$$z_\pm = \frac{-2 \pm \sqrt{4 - 8}}{2} = \frac{-2 \pm 2i}{2} = -1 \pm i$$

Ejercicio: Verificar que $z_\pm = -1 \pm i$ satisfacen $z^2 + 2z + 2 = 0$.
Ejercicio: Verificar que $z^2 + 2z + 2 = (z + 1 - i)(z + 1 + i)$

La fórmula (1.15) es válida también para a, b, c complejos si $a \neq 0$ (veremos luego como obtener raíces de números complejos), aunque en este caso las raíces complejas **no** aparecerán necesariamente como pares conjugados.

Los resultados anteriores se generalizan a polinomios de grado arbitrario $n \geq 1$. Si

$$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0, \quad a_n \neq 0 \tag{1.17}$$

es un polinomio de grado $n \geq 1$, donde los coeficientes (constantes) a_i, $i = 0, \ldots, n$, pueden ser reales o complejos, la ecuación

$$P_n(z) = 0$$

posee siempre al menos una raíz (real o compleja). Este enunciado constituye el **teorema fundamental del algebra** e indica que el cuerpo de los números complejos es algebraicamente **cerrado**.
Más aun, existirán siempre n raíces $z_1, z_2, \ldots, z_n$, en general complejas y no necesariamente distintas, tales que $P_n(z_i) = 0$ para $i = 1, \ldots, n$ y

$$P_n(z) = a_n (z - z_1)(z - z_2) \ldots (z - z_n) \tag{1.18}$$

Como las raíces pueden repetirse, esta igualdad suele escribirse como

$$P_n(z) = a_n (z - z_1)^{m_1} (z - z_2)^{m_2} \ldots (z - z_k)^{m_k} \tag{1.19}$$

donde $z_1, \ldots, z_k$ denotan ahora raíces distintas ($z_i \neq z_j$ si $i \neq j$), con $1 \leq k \leq n$, y m_i, $i = 1, \ldots, k$, denota la multiplicidad de la raíz iésima ($1 \leq m_i \leq n$, $m_1 + \ldots + m_k = n$).

Si los coeficientes a_i del polinomio son todos **reales**, entonces las raíces complejas aparecen siempre en **pares conjugados**.

Problema 15: Probar el enunciado anterior, mostrando primero que si los coeficientes son reales $\Rightarrow \overline{P_n(z)} = P_n(\bar{z})$.

7.1.11. Problemas

1) Realizar las siguientes operaciones, expresando el resultado como $z = x + yi$, con $x, y \in \mathbb{R}$:

$$a)\ (1-2i)(3+4i) \qquad b)\ -i + (2+4i)(2-\frac{1}{2i}) \qquad c)\ 5 + \frac{25}{3+4i}$$

$$d)\ \left(\frac{1+i}{1-i}\right)^* \qquad e)\ \frac{3+4i}{3-4i} \qquad f)\ \frac{i}{3-i} \qquad g)\ \frac{2+i}{1+i} + \frac{1+i}{1+1/i}$$

2) Determinar la parte real, la parte imaginaria y el módulo de los resultados de las operaciones anteriores.

3) Hallar todas las raíces de las siguientes ecuaciones, y escribir el polinomio correspondiente en la forma factorizada (1.16) o (1.18)–(1.19):

$$\begin{array}{lll} a)\ z^2-2z+2=0 & b)\ 2z^2+4=0 & c)\ 2z^2+2z+1=0 \\ d)\ z^2+2iz=0 & e)\ z^2+2iz-1=0 & f)\ z^4-1=0 \end{array}$$

4) Resolver el sistema $\begin{cases} ix + y = i \\ x + iy = -2 \end{cases}$

5) Indicar el error en a) $-1 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1$
y b) $\frac{1}{i} = \frac{\sqrt{1}}{\sqrt{-1}} = \sqrt{\frac{1}{-1}} = \sqrt{-1} = i$.

6) Mostrar que la ecuación

$$|z| = r, \quad r > 0$$

corresponde geométricamente a un círculo de radio r centrado en el origen, y

$$|z - z_0| = r, \quad r > 0$$

a un círculo de radio r centrado en z_0.

7) Mostrar que si $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$,

$$\mathrm{Re}(z_1\bar{z}_2) = x_1x_2 + y_1y_2$$

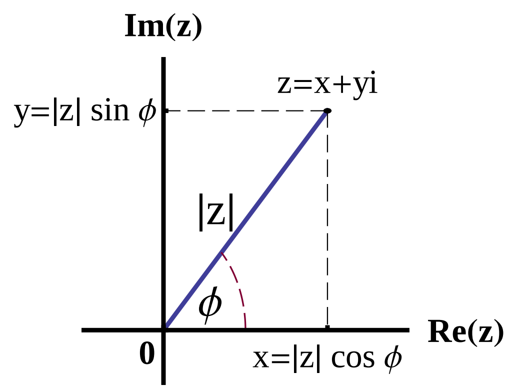
por lo que $\mathrm{Re}(z_1\bar{z}_2)$ es el producto escalar de (x_1, y_1) y (x_2, y_2).
Dar también una interpretación geométrica a $\mathrm{Im}(z_1\bar{z}_2)$.

7.2. Representación polar y fórmula de Euler

7.2.1. Forma polar de un número complejo

A partir del gráfico, vemos que podemos expresar la parte real e imaginaria de $z = x + iy$ como

$$x = |z| \cos \phi, \quad y = |z| \operatorname{sen} \phi$$



donde

$$|z| = \sqrt{x^2 + y^2}$$

es el valor absoluto de z y ϕ es el ángulo que forma z con el eje real en el plano complejo, denominado **argumento** de z o $\arg(z)$.
Este ángulo satisface

$$\tan \phi = y/x$$

y el cuadrante al que pertenece queda determinado por los signos de x e y. Normalmente se toma $\phi \in (-\pi, \pi]$ (valor principal), pero $\phi + 2n\pi$ resulta equivalente a ϕ $\forall$ n entero.

Podemos entonces expresar $z = x + iy$ en términos de $|z|$ y ϕ como

$$z = |z|(\cos \phi + i \operatorname{sen} \phi) \tag{2.20}$$

Esta expresión se denomina **representación polar** del número complejo.

Por ejemplo, si $z = 1 + i \Rightarrow |z| = \sqrt{2}$ y $\tan \phi = 1$, con $x > 0$, $y > 0$, por lo que $\phi = \pi/4$. Por lo tanto

$$1 + i = \sqrt{2}(\cos \pi/4 + i \operatorname{sen} \pi/4)$$

Ejercicio: Verificar esta igualdad.

Problema 1: Probar que $arg(\bar{z}) = -arg(\bar{z})$

La forma polar resulta muy conveniente par el **producto y cociente**: Si

$$z_1 = |z_1|(\cos\phi_1 + i\ \mathrm{sen}\ \phi_1),\ z_2 = |z_2|(\cos\phi_2 + i\ \mathrm{sen}\ \phi_2)$$

entonces

$$z_1 z_2 = |z_1||z_2|\left[\cos(\phi_1+\phi_2) + i\,\mathrm{sen}(\phi_1+\phi_2)\right] \tag{2.21}$$

o sea, $|z_1 z_2| = |z_1||z_2|$, $arg(z_1 z_2) = arg(z_1) + arg(z_2)$.

Y si $z_2 \neq 0$,

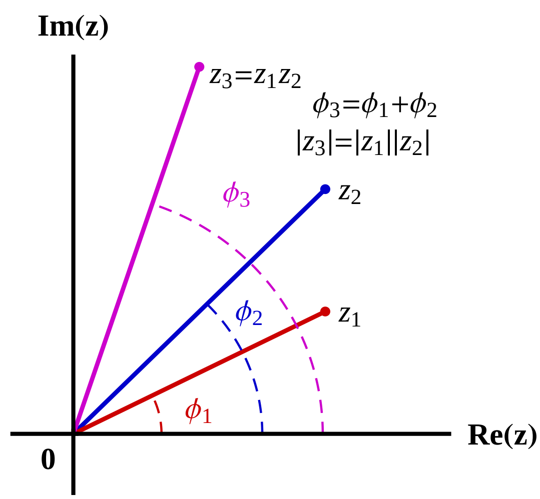
$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}\left[\cos(\phi_1-\phi_2) + i\,\mathrm{sen}(\phi_1-\phi_2)\right] \tag{2.22}$$

o sea, $|z_1/z_2| = |z_1|/|z_2|$ y $arg(z_1 z_2) = arg(z_1) - arg(z_2)$.

Problema 2: Probar (2.21), utilizando $\cos(\phi_1+\phi_2) = \cos\phi_1\cos\phi_2 - \mathrm{sen}\,\phi_1\,\mathrm{sen}\,\phi_2$,

$\mathrm{sen}(\phi_1+\phi_2) = \mathrm{sen}(\phi_1)\cos\phi_2 + \cos\phi_1\,\mathrm{sen}\,\phi_2$.
Problema 3: Probar (2.22), partiendo de $z_1/z_2 = z_1\bar{z}_2/|z_2|^2$ y el resultado de los dos problemas anteriores.

Podemos ver así el significado geométrico del producto $z_1 z_2$: Equivale a rotar z_2 un ángulo ϕ_1 en sentido antihorario, y multiplicar el vector resultante por el escalar real $|z_1|$:



7.2.2. Fórmula de Euler:

La forma polar resulta en realidad más cómoda cuando se utiliza con la famosa fórmula de Euler para la exponencial de un número imaginario:

$$e^{i\phi} = \cos\phi + i \operatorname{sen}\phi \tag{2.23}$$

Demostremos primero esta igualdad: Definiendo $e^{i\phi}$ por medio de la serie exponencial, y separando en términos pares e impares, obtenemos, utilizando (1.4) y (1.5) ($i^{2n} = (-1)^n$, $i^{2n+1} = (-1)^n i$),

$$\begin{aligned} e^{i\phi} &= \sum_{n=0}^{\infty} \frac{(i\phi)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{i^{2n}\phi^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{i^{2n+1}\phi^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n\phi^{2n}}{(2n)!} + i\sum_{n=0}^{\infty} \frac{(-1)^n\phi^{2n+1}}{(2n+1)!} \\ &= \cos\phi + i\operatorname{sen}\phi \end{aligned} \tag{2.24}$$

Con este resultado podemos también evaluar la exponencial de un número complejo general como

$$e^{x+yi} = e^x e^{yi} = e^x(\cos y + i \operatorname{sen} y) \tag{2.25}$$

donde x e y son reales. Por ejemplo, $e^{-2+i} = e^{-2}[\cos(1) + i \operatorname{sen}(1)]$.

Podemos ahora expresar z en la froma polar como

$$z = |z|\, e^{i\phi} \tag{2.26}$$

donde $\phi = arg(z)$.

Por ejemplo,

$$i = e^{i\pi/2}$$

ya que $|i| = 1$ y $arg(i) = \pi/2$. Se comprueba que $e^{i\pi/2} = \cos\pi/2 + i\operatorname{sen}\pi/2 = 0 + i = i$. También,

$$-1 = e^{i\pi}$$

ya que $|-1| = 1$, $arg(-1) = \pi$. Se comprueba que $e^{i\pi} = \cos\pi + i\operatorname{sen}\pi = -1 + 0i = -1$.

Ejercicio: Mostrar que

$$\frac{1+i}{\sqrt{2}} = e^{i\pi/4}$$

Problema 4: Probar que si $z = |z|e^{i\phi} \Rightarrow \bar{z} = |z|e^{-i\phi}$ y, si $z \neq 0$,

$$\frac{1}{z} = \frac{1}{|z|}e^{-i\phi}$$

Problema 5: Probar que $|e^{i\phi}| = 1$, y $|\alpha e^{i\phi}| = |\alpha|$ (α, ϕ reales).

Dado que

$$e^{i\phi_1}e^{i\phi_2} = e^{i(\phi_1+\phi_2)}, \quad \frac{e^{i\phi_1}}{e^{i\phi_2}} = e^{i(\phi_1-\phi_2)}$$

resulta entonces obvio, utilizando la forma polar (2.26), que

$$z_1z_2 = |z|_1e^{i\phi_1}|z_2|e^{i\phi_2} = |z_1||z_2|\,e^{i(\phi_1+\phi_2)} \tag{2.27}$$

$$\frac{z_1}{z_2} = \frac{|z|_1e^{i\phi_1}}{|z|_2e^{i\phi_2}} = \frac{|z_1|}{|z_2|}\,e^{i(\phi_1-\phi_2)}\,, \quad z_2 \neq 0 \tag{2.28}$$

y por lo tanto obtener las fórmulas (2.21)–(2.22). Resulta también obvio que $|z_1z_2| = |z_1||z_2|$, $arg(z_1z_2) = arg(z_1)+arg(z_2)$, y $|z_1/z_2| = |z_1|/|z_2|$, $arg(z_1/z_2) = arg(z_1)-arg(z_2)$.

7.2.3. Potencias de un número complejo

Mediante la forma polar (2.23), resulta muy sencillo calcular cualquier potencia z^n para todo n natural o entero: Dado que $(e^{i\phi})^n = e^{in\phi}$, obtenemos

$$\begin{aligned} z^n &= (|z|e^{i\phi})^n \\ &= |z|^ne^{in\phi} && (2.29) \\ &= |z|^n[\cos(n\phi) + i\,\mathrm{sen}(n\phi)] && (2.30) \end{aligned}$$

La última expresión se denota normalmente fórmula de De Moivre.
Estas expresiones son válidas para $n = 0, 1, 2, \ldots$, y si $z \neq 0$, también para $n = -1, -2, \ldots$.

Ejercicio: Probar que

$$(1+i)^n = 2^{n/2}[\cos(n\pi/4) + i\,\mathrm{sen}(n\pi/4)]$$

y que por ejemplo, $(1+i)^2 = 2i$, $(1+i)^{20} = -2^{10}$, $(1+i)^{-4} = -\frac{1}{4}$.

Ejercicio: Representar gráficamente las primeras cuatro potencias de $1+i$.

7.2.4. Raíces enésimas de la unidad

Consideremos ahora la ecuación

$$z^n = 1$$

para $n \geq 1$. Como $1 = e^{i0} = e^{i2k\pi}$ para cualquier k entero, vemos que todo número complejo $z = e^{i2k\pi/n}$ satisface $z^n = 1$:

$$(e^{i2k\pi/n})^n = e^{i2k\pi} = \cos(2k\pi) + i\,\text{sen}(2k\pi) = 1 \qquad k = 0, \pm 1, \pm 2, \ldots$$

Por lo tanto, obtenemos así n **raíces distintas**

$$z_0 = 1, \;\; z_1 = e^{i2\pi/n}, \;\; z_2 = e^{i4\pi/n}, \;\ldots, \; z_{n-1} = e^{i(n-1)2\pi/n}$$

ya que $z_n = e^{in2\pi/n} = e^{i2\pi} = 1 = z_0$.
Las n raíces de $z^n = 1$ pueden entonces escribirse como

$$z_k \;=\; e^{i2k\pi/n} = \cos(2k\pi/n) + i\,\text{sen}(2k\pi/n)\,, \quad k = 0, \ldots, n-1, \qquad (2.31)$$

Todas las raíces satisfacen $|z_k| = 1$, por lo que están sobre el círculo unidad $|z| = 1$. Obviamente, $z_0 = 1 \;\forall\; n$.
Las raíces aparecen en pares conjugados: $\bar{z}_k = z_{n-k}$, pues $e^{-i2\pi k/n} = e^{i2\pi(n-k)/n}$.

La suma de todas las raíces es 0:

$$z_0 + z_1 + \ldots + z_{n-1} = 0$$

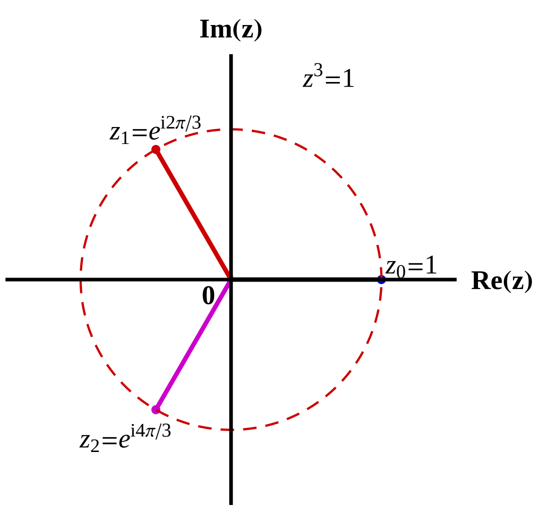
Para $n = 2$, obtenemos obviamente las dos raíces cuadradas de 1, $z_0 = 1$, $z_1 = e^{i\pi} = -1$.
Las tres raíces cúbicas de 1 son en cambio

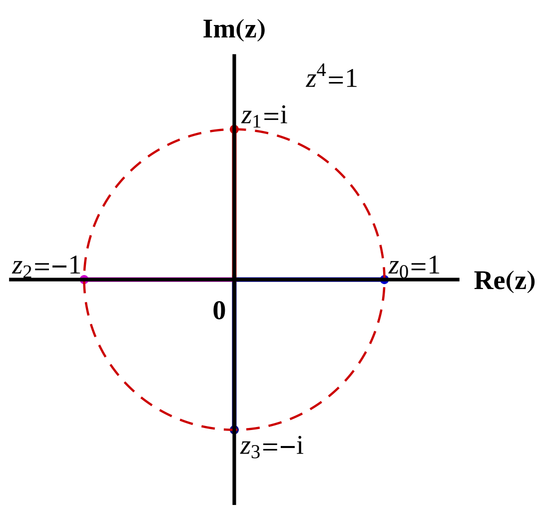
$$\begin{aligned} z_0 &= 1 \\ z_1 &= e^{i2\pi/3} = \cos(2\pi/3) + i\sin(2\pi/3) = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ z_2 &= e^{i4\pi/3} = \cos(4\pi/3) + i\sin(4\pi/3) = -\frac{1}{2} - i\frac{\sqrt{3}}{2} = \bar{z}_1 \end{aligned}$$

Finalmente, las 4 raíces de $z^4 = 1$ son

$$z_0 = 1, \;\; z_1 = e^{i2\pi/4} = i, \;\; z_2 = e^{i4\pi/4} = -1, \;\; z_3 = e^{i6\pi/4} = -i = \bar{z}_1$$

Se muestran en la figura las raíces cúbicas y cuárticas de 1.





7.2.5. Raíces de un número complejo

Podemos ahora fácilmente obtener las n raíces de la ecuación

$$z^n = a + ib$$

para a y b reales arbitrarios. Escribiendo $a + ib$ en la forma polar,

$$a + ib = re^{i\phi}, \quad r = |a + ib| = \sqrt{a^2 + b^2}, \quad \phi = arg(a + ib)$$

vemos que

$$z_k = r^{1/n} e^{i(\phi+2k\pi)/n}$$

satisface

$$z_k^n = re^{i(\phi+2k\pi)} = re^{i\phi} = a + ib$$

para todo k entero. Por lo tanto, las n raíces distintas (asumiendo $r \neq 0$) son

$$z_k = \sqrt[n]{r}\, e^{i(\phi+2k\pi)/n} = \sqrt[n]{r}\, [\cos(\frac{\phi + 2k\pi}{n}) + i \,\mathrm{sen}(\frac{\phi + 2k\pi}{n})]\,, \qquad k = 0, \ldots, n-1$$

Las n raíces son pues de la forma

$$z_k = \sqrt[n]{r}\, e^{i\phi/n} e^{i2k\pi/n}$$

donde $e^{i2k\pi/n}$ son las raíces enésimas de la unidad (ecuación (2.31)).

Por ejemplo, la ecuación

$$z^2 = a + ib$$

escribimos $a + ib = re^{i\phi}$, con $r = \sqrt{a^2 + b^2}$, $\phi = arg(z)$, y entonces, las tiene las dos raíces

$$z_0 = \sqrt{r} e^{i\phi/2}, \quad z_1 = \sqrt{r} e^{i(\phi+2\pi)/2} = -z_0$$

o sea, $z = \pm\sqrt{a + ib}$, con $\sqrt{a + ib} = \sqrt{r} e^{i\phi/2}$ y $\phi \in (-\pi, \pi]$.

Por ejemplo, para resolver la ecuación

$$z^2 = i$$

escribimos i en la forma polar, $i = e^{i\pi/2}$, y entonces

$$z = \pm e^{i\pi/4} = \pm[\cos(\pi/4) + i \,\mathrm{sen}(\pi/4)] = \pm\frac{1 + i}{\sqrt{2}}$$

Análogamente, las 4 raíces de

$$z^4 = -3$$

son, escribiendo $-3 = 3e^{i\pi}$,

$$z_0 = \sqrt[4]{3} e^{i\pi/4} = \sqrt[4]{3}\, \frac{1 + i}{\sqrt{2}}, \quad z_1 = iz_0, \quad z_2 = -z_0, \quad z_3 = -iz_0$$

7.2.6. Funciones complejas

Consideraremos aquí funciones complejas de dominio real $f : \mathbb{R} \to \mathbb{C}$. Estas tienen la forma general

$$f(t) = u(t) + iv(t) \tag{2.32}$$

donde $t \in \mathbb{R}$ y $u(t) = \mathrm{Re}[f(t)]$, $v(t) = \mathrm{Im}[f(t)]$ son funciones reales. Un ejemplo de especial interés es la función exponencial compleja

$$f(t) \;=\; e^{\lambda t}, \quad \lambda = \gamma + i\omega \;\in\; \mathbb{C} \tag{2.33}$$

con γ, ω reales. Aplicando la fórmula general (2.25), obtenemos

$$\begin{aligned} e^{\lambda t} &= e^{\gamma t} e^{i\omega t} \\ &= e^{\gamma t}[\cos(\omega t) + i\,\mathrm{sen}(\omega t)] && (2.34) \\ &= e^{\gamma t}\cos(\omega t) + i e^{\gamma t}\,\mathrm{sen}(\omega t) && (2.35) \end{aligned}$$

o sea, $u(t) = e^{\gamma t}\cos(\omega t)$, $v(t) = e^{\gamma t}\,\mathrm{sen}(\omega t)$.

Asumiendo $u(t)$ y $v(t)$ derivables, la derivada de f respecto de t es la función compleja

$$f'(t) = u'(t) + iv'(t)$$

En el caso (2.33) obtenemos, partiendo de (2.35),

$$\begin{aligned} (e^{\lambda t})' &= e^{\gamma t}[\gamma\cos(\omega t) - \omega\,\mathrm{sen}(\omega t) + i\gamma\,\mathrm{sen}(\omega t) + i\omega\cos(\omega t)] \\ &= (\gamma + i\omega)e^{\gamma t}[\cos(\omega t) + i\,\mathrm{sen}(\omega t)] \\ &= \lambda e^{\lambda t} \quad !! && (2.36) \end{aligned}$$

Por lo tanto, la expresión $(e^{\lambda t})' = \lambda e^{\lambda t}$ resulta válida también para λ complejo. Este resultado será esencial a la hora de resolver ecuaciones diferenciales, y puede obtenerse de forma más general por medio de la teoría de funciones analíticas de variable compleja, que no trataremos en este curso.

De la misma forma, la integral de f se define como

$$\int f(t)dt = \int u(t)dt + i\int v(t)dt$$

Se deja como ejercicio mostrar que entonces,

$$\int e^{\lambda t}dt = \frac{e^{\lambda t}}{\lambda} + c$$

para $\lambda \in \mathbb{C}$, $\lambda \neq 0$, donde $c \in \mathbb{C}$ es en general una constante compleja.

Problema: La potencia t^λ para $\lambda = \gamma + i\omega \in \mathbb{C}$ y $t \in \mathbb{R}_+$ se define como

$$\begin{aligned} t^\lambda &= e^{\lambda\ln t} = e^{\gamma\ln t}e^{i\omega\ln t} \\ &= t^\gamma\,[\cos(\omega\ln t) + i\,\mathrm{sen}(\omega\ln t)] && (2.37) \end{aligned}$$

donde $\ln t$ es el logaritmo natural. Por ejemplo $t^i = \cos(\ln t) + i\,\mathrm{sen}(\ln t)$. Se deja como ejercicio probar que para $\lambda \in \mathbb{C}$, también se obtiene

$$(t^\lambda)' = \lambda t^{\lambda-1}$$

La aplicación más simple y común de las funciones complejas es la representación de funciones periódicas, por ejemplo de la forma

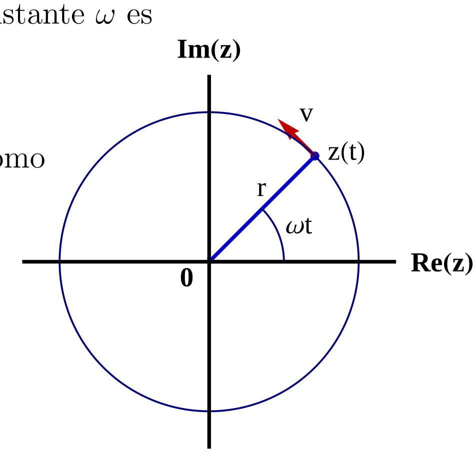
$$u(t) = A\cos(\omega t + \phi) = A[\cos\omega t\cos\phi - \mathrm{sen}\,\omega t\,\mathrm{sen}\,\phi]$$

tales como una corriente alterna o la posición de una partícula en movimiento oscilatorio armónico. Resulta en general más conveniente expresar $u(t)$ como

$$u(t) = \mathrm{Re}[Ae^{i(\omega t+\phi)}]$$

donde hemos asumido A, ω, ϕ, t reales. Se verán ejemplos en la parte de Ecuaciones Diferenciales.

Ejemplo: Movimiento circular. El formalismo complejo es también útil para tratar problemas en dos dimensiones. Por ejemplo, el vector posición de una partícula que se mueve en un círculo de radio r con velocidad angular constante ω es



$$(x(t), y(t)) = r(\cos\omega t, \mathrm{sen}\,\omega t)$$

Podemos escribir este par ordenado en forma compleja como

$$z(t) = re^{i\omega t}$$

Si r y ω son constantes, la velocidad de la partícula será

$$v(t) = z'(t) = i\omega r\, e^{i\omega t} = i\omega z(t)$$

donde i indica que la velocidad será tangencial, es decir, perpendicular a $z(t)$ (recordar el significado de multiplicar por i).
El módulo de la velocidad es el valor absoluto $|v(t)| = \omega|z(t)| = \omega r$. La aceleración será

$$a(t) = z''(t) = (i\omega)^2 re^{i\omega t} = -\omega^2 z(t)$$

es decir, antiparalela a $z(t)$ y de módulo $|a(t)| = \omega^2 r$. Esta es la aceleración centrípeta.

Problema: Generalizar los resultados anteriores al caso de velocidad angular variable $\omega(t)$, tal que $z(t) = re^{i\phi(t)}$, con $\phi'(t) = \omega(t)$. Mostrar que

$$v(t) = z'(t) = i\omega(t)z(t)$$

$$a(t) = z''(t) = -\omega^2(t)z(t) + i\omega'(t)z(t)$$

Interpretar el resultado e identificar la aceleración centrípeta y la aceleración tangencial.

7.2.7. Problemas

1) Encontrar la representación polar $z = |z|e^{i\phi}$ de

$$a)\ z = 1 - i \quad b)\ z = \frac{\sqrt{3}+i}{1+i} \quad c)\ z = -3i \quad d)\ z = \frac{1}{1+\sqrt{3}i} \quad e)\ z = -i$$

2) Mediante la forma polar, evaluar z^n en los casos anteriores, y hallar el valor explícito para $n = 6$ y $n = -3$.

3) Dar una expresión de la forma $z = x + yi$, con $x, y \in \mathbb{R}$, para

$$a)\ z = e^{i\pi/6} \quad b)\ z = \frac{e^{i\pi/3}}{e^{i\pi/2}} \quad c)\ z = e^{-i2\pi/3}e^{i\pi/3}$$

4) Determinar todos los z que satisfacen

$$a)\ z^2 = -i \quad b)\ z^3 = -1 \quad c)\ z^4 = i \quad d)\ z^6 = 1 \quad e)\ z^2 = 1 - i$$

5) Resolver las ecuaciones

$$a)\ z^2 + 2iz - 1 + i = 0 \quad b)\ z^2 + 2i = 0 \quad c)\ z^4 + 2z^2 + 2 = 0$$

6) Escribir en la forma $u(t) + iv(t)$, con $u(t)$, $v(t)$ funciones reales, las funciones

$$a)\ \ f(t) = e^{-(1+2i)t} \quad b)\ f(t) = t^2 e^{3it} \quad c)\ f(t) = t^{1+i}$$

7) Determinar las derivadas de las funciones anteriores, y escribir su parte real y su parte imaginaria.

Facultad de Ingeniería
UNLP

Matemática C

VIII Autovalores y Autovectores

Temario

Clase 1: Introducción. Definiciones. Cálculo de autovalores. Polinomio característico. Autoespacios asociados a los autovalores: cálculo de autovectores. Estimación de autovalores y localización (Círculos de Gershgorin). Operadores geométricos y autovalores. Casos particulares.

Clase 2: Multiplicidad algebraica y geométrica. Autovectores independientes. Bases de autovectores. Matrices semejantes. Diagonalización de matrices. Matrices simétricas y resultados relacionados, etc..

Clase 3: Algunas Aplicaciones.

1 Introducción

Motivaremos este capítulo sobre autovalores discutiendo la ecuación

$$ax^2 + 2hxy + by^2 = 1$$

donde los coeficientes a, h, y b cumplen que al menos uno es no nulo.

La expresión $ax^2 + 2hxy + by^2$ *se llama* forma cuadrática *en* x *e* y. *Se puede escribir como*

$$ax^2 + 2hxy + by^2 = \begin{pmatrix} x & y \end{pmatrix} . \begin{pmatrix} a & h \\ h & b \end{pmatrix} . \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{x}^T . A . \mathbf{x}$$

donde $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ *y* $A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$. *La matriz A se llama matriz de la forma cuadrática.*

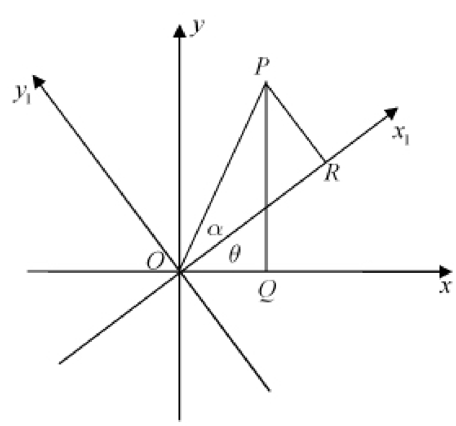
Objetivo *... Queremos aplicar un cambio de variables apropiado, la rotación necesaria* θ *de los ejes* x *e* y, *para que las coordenadas de los puntos de esa figura geométrica respecto de los nuevos ejes* x_1 *e* y_1 *satisfagan una ecuación reducida a la forma canónica:* $\lambda_1 x_1^2 + \lambda_2 y_1^2 = 1$.

La forma canónica es fácil de analizar. Reconocemos a una elipse si los coeficientes son positivos, o una hipérbola si tienen distinto signo, etc.

... Ahora rotamos los ejes x e y en sentido antihorario mediante el ángulo θ hasta los nuevos ejes x_1 e y_1.

La ecuación que describe la rotación de los ejes se deduce de lo siguiente:

Sea P un punto arbitrario con coordenadas (x, y) respecto a los ejes x e y, y con coordenadas (x_1, y_1) respecto a los nuevos ejes x_1 y y_1.



La relación entre ambas coordenadas se obtiene geométricamente considerando ...

$$\begin{aligned} x &= OQ = OP\cos(\theta + \alpha) \\ &= OP(\cos\theta\cos\alpha - \sin\theta\sin\alpha) \\ &= (OP\cos\alpha)\cos\theta - (OP\sin\alpha)\sin\theta \\ &= OR\cos\theta - PR\sin\theta \\ &= x_1\cos\theta - y_1\sin\theta \end{aligned}$$

De igual manera se obtiene que $y = x_1 \sin\theta + y_1 \cos\theta$.

Así la matriz de transición de las coordenadas (x_1, y_1), coordenadas respecto de la base nueva (los vectores que dirigen los ejes nuevos), a las coordenadas (x, y) respecto de la base canónica sabemos que es la matriz S:

$$S = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Así, para cada (x_1, y_1) se tiene que ...

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

... si denotamos $\tilde{x}$ al vector de coordenadas x_1, y_1

$$\mathbf{x} = S.\tilde{\mathbf{x}}$$

... Notamos que las columnas de S dan las direcciones de los semiejes positivos de los ejes x_1 e y_1. La matriz S de rotación tiene la propiedad especial que $\det(S) = 1$ *(verficarlo).*

Sabemos, que para obtener las nuevas coordenadas en términos de las viejas ...

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \tilde{\mathbf{x}} = S^{-1}.\mathbf{x} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} .\mathbf{x}$$

... Por tanto

$$\begin{aligned} x_1 &= x\cos\theta + y\sin\theta \\ y_1 &= -x\sin\theta + y\cos\theta \end{aligned}$$

Además ...

... si reemplazamos en la ecuación dada

$$\mathbf{x}^T.A.\mathbf{x} = (S.\tilde{\mathbf{x}})^T.A.(S.\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}^T.\left(S^T.A.S\right).\tilde{\mathbf{x}}$$

... Supongamos ahora, como veremos más adelante, que es posible elegir un ángulo θ tal que $S^T.A.S$ sea una matriz diagonal, entonces ...

$$\begin{aligned} x^T.A.x &= \tilde{x}^T.\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.\tilde{x} \\ &= \lambda_1 x_1^2 + \lambda_2 y_1^2 \end{aligned}$$

... Así encontrando esa rotación adecuada, la ecuación original $ax^2 + 2hxy + by^2 = 1$ se escribiría, en términos de los nuevos ejes como

$$\lambda_1 x_1^2 + \lambda_2 y_1^2 = 1$$

.

Las soluciones de esta ecuación representan una curva simétrica respecto a los ejes x_1 e y_1.

Las columnas de S, las columnas s_1 y s_2, dan las direcciones de los ejes de simetría. Esta matriz es una matriz ortogonal ($S^T = S^{-1}$, verificarlo).

... Además, para esa rotación apropiada se cumple $S^T.A.S = S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Desde esa igualdad se obtiene que ...

$A.S = S.\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Así si S es la matriz adecuada para cumplir con nuestro objetivo, sus columnas s_1 y s_2 deben satisfacer las ecuaciones siguientes ...

$$A.s_1 = \lambda_1.s_1 \quad y \quad A.s_2 = \lambda_2.s_2$$

... Estas ecuaciones establecen restricciones sobre λ_1 y λ_2, y determinan los vectores s_1 y s_2, no nulos.

Para determinar el vector $s_1 = \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix}$, la primer ecuación se escribe como

$$\begin{pmatrix} a & h \\ h & b \end{pmatrix} \cdot \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix} = \lambda_1. \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix} \tag{1}$$

o

$$\begin{pmatrix} a-\lambda_1 & h \\ h & b-\lambda_1 \end{pmatrix} \cdot \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

... Hemos planteado un sistema homogéneo de 2×2, que debe tener solución no trivial (s_{11}, s_{21}) (es el nuevo eje). Por tanto, el determinante de la matriz de ese sistema debe ...

$$\det \begin{pmatrix} a-\lambda_1 & h \\ h & b-\lambda_1 \end{pmatrix} = 0$$

... Similarmente, λ_2 debe satisfacer la misma ecuación para hallar el $s_2 \neq 0$, y también cumplirá ...:

$$\det \begin{pmatrix} a-\lambda_2 & h \\ h & b-\lambda_2 \end{pmatrix} = 0$$

... Por tanto, ambas λ_1 y λ_2 satisfacen ...

$$\det \begin{pmatrix} a-\lambda & h \\ h & b-\lambda \end{pmatrix} = 0$$

... Expandiendo el determinante se obtiene ...

$$\lambda^2 - (a+b)\,\lambda + ab - h^2 = 0$$

Esta ecuación tiene las raíces

$$\lambda = \frac{a+b \pm \sqrt{(a+b)^2 - 4\left(ab - h^2\right)}}{2}$$
$$= \frac{a+b \pm \sqrt{(a-b)^2 + 4h^2}}{2}$$

Las raíces son reales, y *son distintas si $a \neq b$ o si $h \neq 0$.*

El caso en que $a = b$ y $h = 0$ no necesita investigación porque representa la ecuación de una circunferencia.

La ecuación encontrada $\lambda^2 - (a+b)\lambda + ab - h^2 = 0$ se llama la ecuación de los autovalores *de la matriz A.*

... Si conocemos los valores λ_1 y λ_2, podremos encontrar s_1 y s_2, que son las columnas de la matriz de rotación que se necesita, resolviendo el sistema (1) después de reemplazar el valor de λ_1 en la ecuación, y resolviendo la ecuación similar reemplazando el valor de λ_2.

...

2 El "problema de los autovalores"

Ese problema consiste en el planteo ...

... de la ecuación

$$A.\mathbf{x} = \lambda.\mathbf{x} \tag{2}$$

para determinar el valor λ para el cual existe algún vector x "no nulo" que satisfaga la igualdad.

... Esta ecuación puede mirarse como un sistema lineal de ecuaciones si el vector $\mathbf{x}$ pertenece a $\mathbb{R}^n$ y A es una matriz de $n \times n$.

También $\mathbf{x}$ puede ser un vector en algún otro espacio vectorial V, y A un operador lineal sobre ese espacio. En tal caso ...

... la ecuación (2) también se aplica a la matriz $\hat{A}$, que representa a A con respecto a una base $[u_i]$, y al vector de coordenadas $\hat{\mathbf{x}}$ del vector $\mathbf{x}$ respecto a la misma base.

Definición *Si la ecuación (2) tiene solución $\mathbf{x}$ no nula, se dice que λ es un "autovalor" de A (del operador o la matriz), y el vector $\mathbf{x}$ se llama "autovector" de A, correspondiente al autovalor λ.*

... Otros nombres dados a λ son: valor característico, valor propio o valor principal del operador (o matriz) A.

El conocimiento de los autovalores de un operador lineal (o matriz) ayuda a comprender la acción del operador sobre un autovector y sobre un vector cualquiera.

... Dado un operador (o matriz de representación de $n \times n$) A , si podemos construir una base del espacio compuesta enteramente por autovectores de A, ...

... $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, entonces cualquier vector $\mathbf{x}$ puede escribirse como una combinación lineal de estos autovectores ...

$$\mathbf{x} = c_1.\mathbf{u}_1 + c_2.\mathbf{u}_2 + \cdots + c_n.\mathbf{u}_n = \sum_{i=1}^{n} c_i.\mathbf{u}_i$$

... La acción de A sobre $\mathbf{x}$ es simplemente

$$\begin{aligned} A.\mathbf{x} &= c_1.A.\mathbf{u}_1 + c_2.A.\mathbf{u}_2 + \cdots + c_n.A.\mathbf{u}_n \\ &= c_1\lambda_1.\mathbf{u}_1 + c_2\lambda_2.\mathbf{u}_2 + \cdots + c_n\lambda_n.\mathbf{u}_n \\ &= \sum_{i=1}^{n} (c_i\lambda_i)\,.\mathbf{u}_i \end{aligned}$$

... Por tanto, si $(c_1, c_2, \ldots, c_n)^T$ es el vector de coordenadas de $\mathbf{x}$, con respecto a esa base de autovectores del operador lineal (o de su matriz de representación) A, entonces el vector de las coordenadas de $A.\mathbf{x}$ con respecto a la misma base es

$$\begin{pmatrix} \lambda_1 c_1 \\ \lambda_2 c_2 \\ \vdots \\ \lambda_n c_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} . \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Conclusión *...(I) La matriz de representación de un operador A con respecto a una base de sus autovectores (si existe esa base) es una "matriz diagonal", cuyos coeficientes en la diagonal son "los autovalores" de A.*

Además ...

...(II) el efecto de un operador A sobre cualquiera de sus autovectores es simplemente "escalar al autovector" por el correspondiente autovalor ($\mathbf{u}_i \longrightarrow \lambda_i.\mathbf{u}_i$), expandiendo o contrayendo $\|\mathbf{u}_i\|$ por el factor λ_i (que también puede ser igual a cero).

Observación: *En una sección próxima se verá que "no siempre" se puede tener una base de autovectores (el lindo resultado (I)).*

Ejemplo

$$\begin{aligned} A &= \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \quad y \quad \mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ A.\mathbf{x} &= \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} . \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} \\ &= 3. \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 3.\mathbf{x} \end{aligned}$$

Por tanto, $\lambda = 3$ es un autovalor de A y $\mathbf{x} = (2,1)^T$ es un autovector de A asociado a $\lambda = 3$.

Observación: *$(4,2)^T$ es también otro autovector para el mismo autovalor $\lambda = 3$, por tanto*

$$\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} . \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \end{pmatrix} = 3. \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

... Así cualquier múltiplo escalar de $(2,1)^T$ también será un autovector (asociado al mismo λ), ya que

$$A.(c.\mathbf{x}) = c.A.\mathbf{x} = c.\lambda.\mathbf{x} = \lambda.(c.\mathbf{x})$$

Además ...

Los autovectores asociados a un particular autovalor λ de A, con el agregado del vector 0, forman un subespacio de $\Re^n$ que se denomina " autoespacio" de A asociado a λ.
Para ver eso ...

Dado un λ que es autovalor ...
Si $A.\mathbf{x} = \lambda.\mathbf{x}$, y si y también satisface
$A.\mathbf{y} = \lambda.\mathbf{y}$ entonces

$$A.(\alpha.\mathbf{x} + \beta.\mathbf{y}) = \alpha.A.\mathbf{x} + \beta.A.\mathbf{y} = \alpha\lambda.\mathbf{x} + \beta\lambda.\mathbf{y} = \lambda.(\alpha.\mathbf{x} + \beta.\mathbf{y})$$

esto es ...

$$(\alpha.\mathbf{x} + \beta.\mathbf{y}) \text{ es también un autovector de} A \text{con el mismo autovalor } \lambda$$

... Así el conjunto de los autovectores correspondientes a un mismo λ, con el agregado del vector 0 es un subespacio.

Por ahora, consideramos que A es una matriz real $n \times n$, y x un vector de $\Re^n$

....¿ Cómo se obtiene el conjunto de "autovectores" asociados a un λ ?
La ecuación de autovalores $A.\mathbf{x} = \lambda.\mathbf{x}$ puede escribirse ...

$$\begin{aligned} A.\mathbf{x} &= \lambda.I_n.\mathbf{x} && (3) \\ (A - \lambda.I_n).\mathbf{x} &= 0 && (4) \end{aligned}$$

Así estamos buscando vectores "no nulos" $\mathbf{x}$ que satisfagan la ecuación (3). Es decir,
... buscamos soluciones "no triviales" del sistema lineal homogéneo (4). El conjunto de todas estas soluciones "no nulas" está en ...
... el subespacio o espacio nulo de la matriz $A - \lambda.I_n$: $N(A - \lambda.I_n)$. Este subespacio es un subespacio de $\mathbb{R}^n$).

El espacio nulo $N(A - \lambda.I_n)$ se llama el *autoespacio* de A correspondiente al autovalor λ.
Además ...
... si λ es un autovalor de A implica que existe algún x no nulo en $N(A - \lambda.I)$. Cualquier vector no nulo en $N(A - \lambda.I)$ es un autovector asociado a λ. La dimensión $N(A - \lambda.I) \geq 1$ (por lo menos hay un autovector para ese valor λ).

La ecuación (4) tiene solución x no nula " si y sólo si" la matriz $A - \lambda.I_n$ es singular. Por tanto, debe ser ...

$$\det(A - \lambda.I_n) = 0$$

... Esta ecuación se denomina la ecuación característica *de la matriz A.*

3 Cálculo de los autovalores

Si expandimos el determinante $\det(A - \lambda.I_n) = 0$ *(recordando la definición del determinante) tendremos un polinomio de grado n en la variable λ*

$$\det(A - \lambda.I) = p(\lambda)$$

3.1 Polinomio característico

Definición *El $p(\lambda)$ se llama el* polinomio característico *de A, y sus raíces (las soluciones de la ecuación característica) son los autovalores de A.*

... Como cualquier polinomio de grado n...

$$p(\lambda) = (-1)^n(\lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0)$$

... tiene n raíces (contando las raíces múltiples) entonces A tiene exactamente n autovalores.

Sin embargo, hay que tener en cuenta que ...

- *Algunas raíces pueden repetirse, por tanto A tendrá a lo sumo n autovalores "distintos";*
- *Algunas raíces pueden ser números complejos, por tanto A puede tener autovalores complejos (y consecuentemente,* autovectores *complejos).*

Las siguientes proposiciones, correspondientes a una matriz A de $n \times n$, son equivalentes:

- λ *es un autovalor de A.*
- $(A - \lambda.I)\,.\mathbf{x} = 0$ *tiene solución no trivial.*
- $N(A - \lambda.I) \neq \{0\}$.
- $A - \lambda.I$ *es singular.*
- $\det(A - \lambda.I) = 0$.

... En la práctica, la última condición es la que se usa para determinar los autovalores de una matriz de $n \times n$ (al menos para matrices pequeñas y manejables).

Ejemplo *Encontrar los autovalores y autovectores de* $A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$

La ecuación característica es

$$\begin{aligned} \det \begin{pmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{pmatrix} &= 0 \\ (3-\lambda)(-2-\lambda) - 6 &= 0 \\ \lambda^2 - \lambda - 12 &= 0 \\ (\lambda - 4)(\lambda + 3) &= 0 \end{aligned}$$

$$\lambda_1 = 4 \quad y \quad \lambda_2 = -3$$

Ahora, buscamos el espacio nulo de $A - \lambda_1.I$ *y de* $A - \lambda_2.I$ *para determinar los correspondientes* autovectores.

Para hacer esto, resolvemos los correspondientes sistemas homogénos

$$(A - \lambda_i.I).x = 0 \qquad (i = 1, 2)$$

Si tomamos $\lambda_1 = 4$*, y resolvemos por eliminación el sistema*

$$A - \lambda_1.I = \begin{pmatrix} -1 & 2 \\ 3 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}$$

se obtiene ...

$$\mathbf{u}_1 = \beta. \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

esto es ...

... cualquier múltiplo no nulo de este vector es un autovector asociado a $\lambda_1 = 4$*; y este vector es una base para* $N(A - 4.I)$*, que es el autoespacio de ese autovalor.*

Si consideramos $\lambda_2 = -3$

$$A - \lambda_2.I = \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 6 & 2 \\ 0 & 0 \end{pmatrix}$$

al resolver el sistema homogéneo ...

$$\mathbf{u}_2 = \beta. \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

como autovector asociado a λ_2*. Así este vector forma una base de* $N(A + 3.I)$*.*

En este ejemplo ambos "autoespacios" tienen dimensión uno.

Para practicar ...

Ejercicios

3.1 *Para cada matriz, encontrar el polinomio característico y los autovalores.*

(a) $\begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$ ***(b)*** $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ ***(c)*** $\begin{pmatrix} 0 & 3 \\ 7 & 0 \end{pmatrix}$ ***(d)*** $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

✓ ***3.2*** *Para cada matriz, encontrar la ecuación característica, los autovalores y los autovectores asociados a cada uno.*

(a) $\begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ ***(b)*** $\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$

Otro ejemplo ...

Ejemplo $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \longrightarrow \det \begin{pmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{pmatrix} = 0 \longrightarrow (1-\lambda)^2 + 4 = 0$

Por tanto ...

$$\lambda_1 = 1 + 2i \quad y \quad \lambda_2 = 1 - 2i \qquad \text{¡oh-oh!}$$

... Esto tenía que pasar tarde o temprano. *No nos habituaremos a los autovalores o autovectores complejos, aunque haremos algunos comentarios:*

Si λ es un autovalor complejo de una matriz real A, y su correspondiente autovector es $\mathbf{z}$, *entonces si $\bar{\lambda}$ es el complejo conjugado de λ y $\bar{\mathbf{z}}$ es el complejo conjugado de z ...*

$$A.\bar{z} = \bar{A}.\bar{z} = \overline{A.z} = \overline{\lambda.z} = \bar{\lambda}.\bar{z}$$

... Por tanto, $\bar{z}$ es también un autovector de A con autovalor $\bar{\lambda}$.

Así autovalores y autovectores "complejos de las matrices reales" de $n \times n$ siempre aparecen con su complejo conjugado, es decir aparecen en "pares".

Problema *(1)Dada una matriz de dimensión $n \times n$, con n impar, verificar que tiene por lo menos un autovalor real.*

Problema *(2) Usando los comandos del MATLAB. Los comandos del MATLAB para calcular autovalores (después de haber entrado la matriz $A = [fila1; fila2; fila3]$) es $E = eig(A)$, dando en E los autovalores.*

Para obtener los autovectores: $[U, D] = eig(A)$, da en las columnas de U los autovectores (no garantiza que las columnas de U sean linealmente independientes), y en D los autovalores.

Aplicar a la matriz A: $\begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$

(3) Con la ayuda del MATLAB, calcular los autovalores de la matriz de 5×5 cuyos términos son $a_{ij} = \frac{1}{i+j-1}$.

(4) Dada la matriz A: $\begin{pmatrix} 10 & -9 & 1 \\ 4 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *(a) Usando el MATLAB hallar los autovalores. Hallar una base del autoespacio de cada autovalor λ_i encontrado, para eso usar : $null(A - \lambda_i * I_n)$. En general, el comando $null(B)$, da una base del subespacio nulo de la matriz B.*

(b) Explicar porqué puede conocer todos los autovectores asociados a cada λ_i, conociendo las bases halladas en (a).

3.2 Localización de los autovalores. Círculos de Gershgorin

Objetivo *Despues de haber calculado los autovalores de una matriz, y haber visto las dificultades en calcular las raíces de los polinomios característicos, sabiendo además lo dificil de ese tema cuando las matrices son de dimensión $n > 4$, es importante conocer el Teorema de Gershgorin. Este teorema nos provee un procedimiento para estimar la localización de los autovalores de cada matriz, y en muchas oportunidades permite sacar conclusiones sobre los autovalores sin calcularlos.*

Teorema *Sea A una matriz arbitraria de $n \times n$. Entonces los autovalores λ de A están localizados en la unión de los n discos (en el plano complejo, ya que los autovalores pueden ser complejos) que satisfacen para cada fila $i = 1, \dots, n$:*

$$|\lambda - a_{i,i}| \leq \sum_{j \neq i} |a_{i,j}| \tag{5}$$

siendo $a_{i,i}$ el elemento i-ésimo de la diagonal de A, y $a_{i,j}$ con $j \neq i$ los elementos de la fila i-ésima, diferentes del $a_{i,i}$.

Demostración. *Dado λ y un correspondiente autovector u, $u \neq 0$, tal que $Au = \lambda u$. Tal u se puede normalizar dividiéndolo por la mayor componente en valor absoluto de u ($u = (u_1, u_2, \dots, u_j, \dots, u_n)$, se divide por la componente u_i con mayor $|u_j|$, $j = 1, \dots, n$), así llamamos x a ese vector tal que $Ax = \lambda x$, y tal que ahora sus componentes satisfacen que $|x_j| \leq 1$ (equivale a decir que la $\|x\|_\infty = 1 = x_i$). Entonces,*

$$\sum_{j=1}^{n} a_{i,j} x_j = \lambda x_i = \lambda.$$

Por tanto,

$$\lambda - a_{i,i} = \sum_{j \neq i}^{n} a_{i,j} x_j,$$

lo que implica que

$$|\lambda - a_{i,i}| \leq \sum_{j \neq i}^{n} |a_{i,j} x_j| \leq \sum_{j \neq i}^{n} |a_{i,j}|. \tag{6}$$

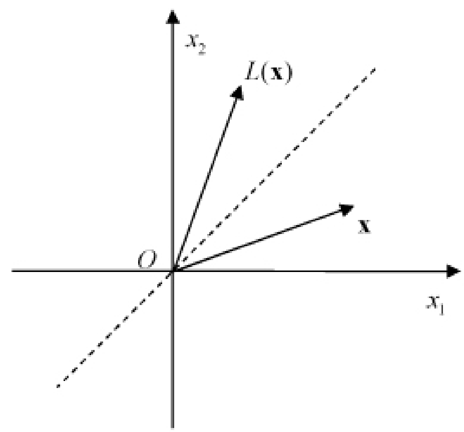
Ejemplo *Estimar la localización de los autovalores de* $A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$

$|\lambda - 3| \leq 2$, *desde la fila 1, y desde la fila 2 se obtiene que* $|\lambda - (-2)| \leq 3$. *Si fuesen reales, estarían (dibujar la recta) entre* $[-5, 5]$.

3.3 Operadores geométricos y autovalores (para visualizar).

Este ejemplo explica visualmente el significado de los autovectores de esta transformación y visualiza también porqué los restantes vectores de $\Re^2$ no pueden ser autovectores.

Consideremos el operador $L(\mathbf{x})$ *que refleja un vector (en* $\mathbb{R}^2$*) a través de la recta* $y = x$*, diagonal que corta el primer y tercer cuadrante:*



La matriz que representa a L *con respecto a la base canónica* $[\mathbf{e}_1, \mathbf{e}_2]$ *es*

$$A = (L(\mathbf{e}_1), L(\mathbf{e}_2)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

por tanto ...

$$L(\mathbf{x}) = L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

¿Cuáles son los autovalores y los autovectores de A *(o, equivalentemente de* L*)?*

Planteamos ...

$$\det(A - \lambda.I) = \det\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$$

Tenemos entonces que $\lambda_1 = 1$ *y* $\lambda_2 = -1$.
Resolvemos $(A - \lambda_1.I).\mathbf{x} = 0$
La matriz de este sistema homogéneo ...

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \textit{ que } \longrightarrow \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \textit{ por eliminación for filas } \longrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

Entonces ...
el autoespacio asociado a $\lambda_1 = 1$ *es el conjunto:*
$\{\mathbf{u}_1 = \alpha.\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ *para todo* $\alpha \in \Re\}$

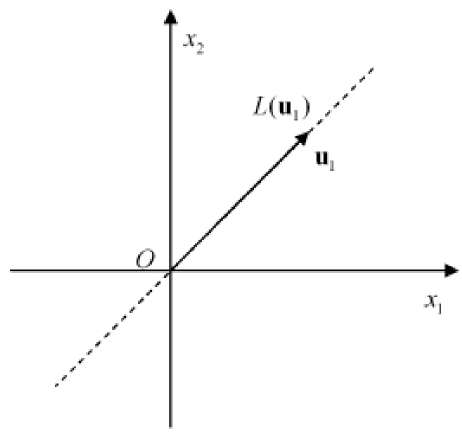
Para el otro autovalor, resolvemos $(A - \lambda_2.I)\,.\mathbf{x} = 0$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

y ...

... el autoespacio asociado ... $\mathbf{u}_2 = \alpha.\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

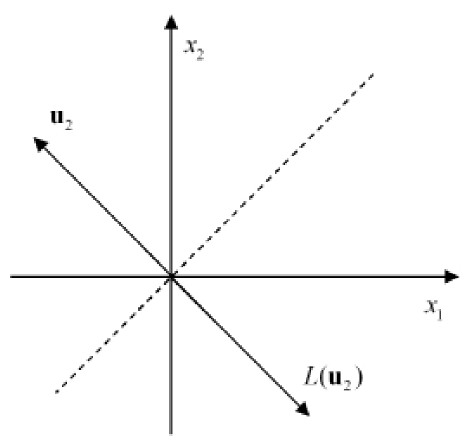
Lo que sigue explica visualmente el significado de los autovectores hallados y explica también porque no hay otros autovectores en esta aplicación.



Para $\mathbf{u}_1 = (1,1)^T$*, el vector permanece sobre el eje de reflección, y al reflejarlo a través del eje retorna al mismo vector* $\mathbf{u}_1$*.*

$$L(\mathbf{u}_1) = \mathbf{u}_1 = 1.\mathbf{u}_1$$

Esto es, $\mathbf{u}_1$ *es un autovector de* L *con autovalor* 1*.*



Para $\mathbf{u}_2 = (-1,1)^T$*, el vector permanece perpendicular al eje de reflección, y al reflejarlo a través del eje produce el opuesto de* $\mathbf{u}_2$*.*

$$L(\mathbf{u}_2) = -\mathbf{u}_2 = -1.\mathbf{u}_2$$

Esto es, $\mathbf{u}_2$ *es un autovector de* L *con autovalor* -1*.*

Cualquier otra orientación del vector $\mathbf{x}$ *al reflejarse a través del eje "no produce" un vector que sea múltiplo escalar de* $\mathbf{x}$*:* $L(\mathbf{x}) \neq \lambda.\mathbf{x}$

Sólo los dos tipos de vectores $\mathbf{u}_1$ *y* $\mathbf{u}_2$ *tienen la propiedad de conservar la dirección invariante (pudiendo cambiar sólo el sentido y longitud).*

3.4 Singularidad y No-singularidad

¿Qué papel juega la singularidad o no-singularidad de la matriz A *en el problema de los autovalores?.*

(i) Primero, supongamos que A *es singular. Entonces* $\det A = 0$*, y el sistema homogéneo*

$$A.\mathbf{x} = 0$$

tiene solución no trivial $\mathbf{x}$*. Entonces existe un vector no nulo* x *en* $\mathbb{R}^n$ *tal que ...*

$$A.\mathbf{x} = 0 = 0.\mathbf{x}$$

Así $\lambda = 0$ *es un autovalor de* A *si* A *es singular.*

Recíprocamente, si $\lambda = 0$ *es un autovalor de* A*, entonces el correspondiente autovector (no nulo)* $\mathbf{x}$ *debe satisfacer*

$$A.\mathbf{x} = \lambda.x = 0, \quad \textit{ya que } \lambda = 0$$

Esto es, $\mathbf{x}$ *es una solución no trivial del sistema lineal* $A.\mathbf{x} = 0$*, y por tanto* A *es "singular".*

... Hemos encontrado que: " A es una matriz de $n \times n$ singular *si y sólo si* al menos uno de sus autovalores es cero".

Entonces, lo anterior se puede expresar de otra manera ...

(ii)Si una matriz A *es no-singular, todos sus autovalores deben ser no nulos. Recíprocamente, si* A *tiene todos sus autovalores no nulos, es no-singular.*

Casos especiales...

4 Matrices diagonales

Si A *es una matriz diagonal, entonces*

$$\begin{aligned} 0 &= \det(A - \lambda.I) = \det \begin{pmatrix} a_{11}-\lambda & 0 & \cdots & 0 \\ 0 & a_{22}-\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}-\lambda \end{pmatrix} \\ &= (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) \end{aligned}$$

Por tanto ...

$$\lambda_1 = a_{11}; \;\; \lambda_2 = a_{22}; \quad \cdots \;\; ; \lambda_n = a_{nn}$$

Así los autovalores de una matriz diagonal son los coeficientes de la diagonal $\{a_{ii}\}$ *(eso muestra también que son todos reales).*

Además, en este caso los autovectores también tienen una forma particularmente simple. En este caso: son precisamente los vectores de la base canónica $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$*, ya que*

$$A.\mathbf{e}_1 = \lambda_1.\mathbf{e}_1 \;\; ; \;\; A.\mathbf{e}_2 = \lambda_2.\mathbf{e}_2 \;\; ; \; etc$$

En otros casos, se podrán encontrar bases que se puedan usar para obtener una representación con una matriz diagonal ...?

Si A *no es diagonal... pero podemos determinar cuales son sus autovalores y autovectores* $\{\lambda_i\}$ *y* $\{\mathbf{x}_i\}$*.*

... Y supogamos que los "autovectores de A *son linealmente independientes", con lo cual forman una base de* $\mathbb{R}^n$ *(porque son* n *vectores).*

Entonces considerando esa base, los vectores coordenadas $\{\hat{\mathbf{x}}_i\}$ *de los autovectores* $\{\mathbf{x}_i\}$ *con respecto a esta " autobase" serán* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$*.*

... La matriz de representación de A *con respecto a esta " autobase " será*

$$\hat{A} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

> *... Veremos más adelante, que cualquier matriz* A *de* $n \times n$ *con "n autovectores linealmente independientes" puede ser diagonalizada en este sentido. Puede ser representada con respecto a una base de sus autovectores mediante una " matriz diagonal", con sus respectivos autovalores en la diagonal.*

4.1 Otras propiedades de las matrices diagonales

- *Dado que los autovalores de una matriz diagonal* A *son los coeficientes de la diagonal* $\{a_{ii}\}$*, tenemos que*

$$\det A = a_{11}a_{22}\cdots a_{nn} = \lambda_1\lambda_2\cdots\lambda_n = \prod_{i=1}^{n} \lambda_i$$

Además,

- *la traza de* A *o* $\operatorname{tr} A$*, definida como la suma de los coeficientes de la diagonal de* A*, está dada por*

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$$

> *Estos dos resultado son válidos también para matrices que no son diagonales.*
> *...*

5 En general - Producto y suma de autovalores

Si A es una matriz cualquiera de $n \times n$, entonces

> $\det A$ *es igual al producto de los autovalores de A:* $\det A = \prod_{i=1}^{n} \lambda_i$

Demostración. *Recordar que $p(\lambda) = \det(A - \lambda.I)$ tiene n raí ces $\{\lambda_i\}$. Por lo tanto, cuando expandimos el determinante y factorizamos el polinomio resultante obtendremos*

$$p(\lambda) = \pm(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

... Una mirada detallada a la estructura del determinante muestra que de hecho

$$\begin{aligned} p(\lambda) &= (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda) \end{aligned}$$

y por tanto ...

$$\det A = p(0) = \lambda_1 \lambda_2 \cdots \lambda_n$$

■

> $\operatorname{tr} A$ *es igual a la suma de los autovalores de A:* $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$

Demostración. *Esta demostración no es tan fácil de verificar, y no la hacemos aquí˙Ver la bibliografía.* ■

Observación *El punto (1) es coherente con el resultado previo :* $\det A = 0$ *(A singular) si y sólo si al menos uno de los autovalores de A es nulo.*

Observación *Como el determinante de una matriz real debe ser un número real, se sigue que el producto de los autovalores λ_i debe ser también real.*

... ¿Qué pasa si alguno de los autovalores es complejo?

Como los autovalores complejos sólo aparecen con su pareja conjugada, y dado que

$$\lambda.\bar{\lambda} = |\lambda|^2$$

es siempre un número real, el producto de todos los autovalores será un número real.

Observación *... Similarmente, dado que* la traza de una matriz real es real, *la suma de los autovalores también debe ser real. Y esto es cierto, aún cuando haya autovalores complejos, dado que ...*

$$\lambda + \bar{\lambda} = 2\operatorname{Re}\lambda \quad \textit{(es también real)}$$

En un ejemplo previo ...

Ejemplo *la matriz dada es* $A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$. *Vimos que los autovalores son* $\lambda_1 = 4$ *y* $\lambda_2 = -3$.

Por tanto ...

$$\begin{aligned} \operatorname{tr} A &= 3 + (-2) = 4 + (-3) = 1 \\ \det A &= -6 - 6 = 4.(-3) = -12 \end{aligned}$$

En otro ejemplo previo ...

Ejemplo *la matriz* $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$ *tiene los autovalores* $\lambda_1 = 1 + 2i$ *y* $\lambda_2 = 1 - 2i$

$$\begin{aligned} \operatorname{tr} A &= 1 + 1 = (1 + 2i) + (1 - 2i) = 2 \\ \det A &= 1 + 4 = (1 + 2i)(1 - 2i) = 5 \end{aligned}$$

5.1 Matrices triangulares

Si A es una matriz triangular ...

... para calcular sus autovalores consideramos

$$\det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - \lambda \end{pmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$$

Conclusión *Los autovalores de una matriz triangular (superior o inferior) son los coeficientes de la diagonal (como en el caso de las matrices diagonales).*

5.2 Matrices inversas

Si A es no singular...

... ¿cuáles son los autovalores de A^{-1}?

Si los $\lambda's$ son los autovalores de A, satisfacen ...

$$A.\mathbf{x} = \lambda.\mathbf{x}$$

entonces multiplicando por A^{-1}

$$A^{-1}.A.\mathbf{x} = \lambda.A^{-1}.\mathbf{x}$$

o

$$\frac{1}{\lambda}.\mathbf{x} = A^{-1}.\mathbf{x}$$

Así si $\mathbf{x}$ *es un autovector de* A *con autovalor* λ,

... entonces $\mathbf{x}$ *también es un autovector de* A^{-1} *correspondiente al autovalor* $1/\lambda$ *(como* A *es no singular los* $\lambda's$ *no pueden ser nulos).*

6 Potencia de matrices

¿Cómo son los autovalores y autovectores de A^k?

Si se conocen los autovalores de A

$$A.\mathbf{x} = \lambda.\mathbf{x}$$

entonces

$$A^2.\mathbf{x} = A.\left(A.\mathbf{x}\right) = A.\left(\lambda.\mathbf{x}\right) = \lambda.A.\mathbf{x} = \lambda^2.\mathbf{x}$$

y por inducción ...

$$A^k.\mathbf{x} = \lambda^k.\mathbf{x}$$

Así si $\mathbf{x}$ *es un autovector de* A *con autovalor* λ*, entonces...*
... $\mathbf{x}$ *es también un autovector de* A^k*, con autovalor* λ^k*.*

Esto también es cierto para potencias negativas de A*, con* $A^{-k} = \left(A^{-1}\right)^k$*.*

Ejercicios para seguir practicando en su casa ...

Ejercicios

✓ **6.1** Para cada matriz, encontrar la ecuación característica, los autovalores y los autovectores asociados a cada uno.

(a) $\begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$

6.2 Encontrar la ecuación característica, autovalores (son complejos) y los asociados autovectores.

$$\begin{pmatrix} -2 & -1 \\ 5 & 2 \end{pmatrix}$$

6.3 Para esta matriz, encontrar la ecuación característica, los autovalores y los autovectores asociados a cada uno.

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

✓ **6.4** Lo mismo del previo ejercicio.

(a) $\begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ **(b)** $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}$

✓ **6.5** Encontrar los autovalores y asociados autovectores del operador de diferenciación $d/dx\colon \mathcal{P}_3 \to \mathcal{P}_3$. La base natural es $B = \langle 1, x, x^2, x^3 \rangle$. La acción de la aplicación es $1 \mapsto 0$, $x \mapsto 1$, $x^2 \mapsto 2x$, y $x^3 \mapsto 3x^2$, y verificar que la matriz de representación es :

$$T = \mathrm{Rep}_{B,B}(d/dx) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{B,B}$$

Hallar los autovalores, planteando

$$0 = detT - \lambda I = \begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ 0 & 0 & -\lambda & 3 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} = \lambda^4$$

La aplicación tiene el único autovalor $\lambda = 0$. Para hallar los asociados autovectores, se resuelve

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{B,B} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}_B = 0 \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}_B \qquad \Longrightarrow \qquad u_2 = 0,\, u_3 = 0,\, u_4 = 0$$

Así para alcanzar su autoespacio

$$\{ \begin{pmatrix} u_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B \quad u_1 \in \Re\} = \{u_1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \mid u_1 \in \Re\} = \{u_1 \mid u_1 \in \Re\}$$

6.6 Probar que los autovalores de una matriz triangular (superior o inferior) son las entradas de la diagonal.

✓ **6.7** Encontrar la fórmula del polinomio característico de una matriz de 2×2.

? **6.8** Mostrar que si A es una n- matriz cuadrada y cada fila suma c entonces c es un autovalor de A.

✓ **6.9** **(a)** Puede cualquier vector no-$\vec{0}$ en un espacio vectorial ser un autovector ?

✓ **6.10** Mostrar que λ es un autovalor de T si y sólo si la aplicación representada por $T - \lambda * I$ es singular.

6.11 **(a)** Mostrar que si λ es un autovalor de A entonces λ^k es un autovalor de A^k.

(b) ¿ Qué es lo erróneo en la demostración siguiente ? "Si λ es un autovalor de A y si μ es un autovalor de B, entonces $\lambda\mu$ es un autovalor de AB, ya que si $A\vec{x} = \lambda\vec{x}$ y $B\vec{x} = \mu\vec{x}$ entonces $AB\vec{x} = A\mu\vec{x} = \mu A\vec{x} = \mu\lambda\vec{x}$"? Cuidado !, con la obtención de resultados falsos. Observar si se puede afirmar que A y B tienen igual autovector? (consultar).

7 Propiedad de Independencia lineal de los autovectores. Bases de autovectores

Un resultado importante ...

> *Autovectores correspondientes a autovalores "distintos" son linealmente independientes.*

Demostración. Supongamos que $\mathbf{x}$ e $\mathbf{y}$ son autovectores de A correspondientes a λ_1 y λ_2 distintos. Luego , ...

$$A.\mathbf{x} = \lambda_1.\mathbf{x} \quad \text{y} \quad A.\mathbf{y} = \lambda_2.\mathbf{y} \quad (\text{con } \lambda_1 \neq \lambda_2)$$

Veamos que $\mathbf{x}$ e $\mathbf{y}$ son linealmente independientes.
Consideremos la combinación ...

$$c_1.\mathbf{x} + c_2.\mathbf{y} = 0$$

entonces ...

$$\begin{aligned} A.(c_1.\mathbf{x} + c_2.\mathbf{y}) &= 0 \\ &= c_1.A.\mathbf{x} + c_2.A.\mathbf{y} \\ &= c_1\lambda_1.\mathbf{x} + c_2\lambda_2.\mathbf{y} \end{aligned}$$

Tenemos dos ecuaciones vectoriales ...

$$\begin{aligned} c_1.\mathbf{x} + c_2.\mathbf{y} &= 0 \\ c_1\lambda_1.\mathbf{x} + c_2\lambda_2.\mathbf{y} &= 0 \end{aligned}$$

Despejando de la primera queda

$$c_1.\mathbf{x} = -c_2.\mathbf{y} \tag{7}$$

y reemplazando en la segunda ...

$$\begin{aligned} -\lambda_1 c_2.\mathbf{y} + c_2.\lambda_2\mathbf{y} &= 0 \\ c_2\left(\lambda_2 - \lambda_1\right).\mathbf{y} &= 0 \end{aligned}$$

Como $\mathbf{y} \neq 0$ (¿por qué?), deberá ser $c_2\left(\lambda_2 - \lambda_1\right) = 0$. Pero $\lambda_2 \neq \lambda_1$. Luego, deberá ser $c_2 = 0$. Entonces, de acuerdo a (7), también $c_1 = 0$.

... Por tanto, $\mathbf{x}$ e $\mathbf{y}$ son linealmente independientes. ■

> **Conclusión** ... *Así si una matriz tiene n autovalores distintos* $\Rightarrow$ *tiene n autovectores linealmente independientes.*

7.1 Autovalores múltiples

...¿ Qué ocurre cuando los autovalores son repetidos?

En este caso, ¿ se puede afirmar que existen n autovectores independientes ?.

... La respuesta es ... " no siempre existen n autovectores independientes", como se puede ver en los dos ejemplos que siguen.

Así nos interesará saber cuándo existen y cuándo no existen n autovectores linealmente independientes de A.

Consideremos el ejemplo ...

$$\begin{aligned} A &= \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix} \\ \det(A-\lambda.I) &= \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} \\ &= -\lambda^3-\lambda^2-21\lambda+45 \\ &= (\lambda-5)(\lambda+3)^2 \end{aligned}$$

Por tanto ...

$$\det(A-\lambda.I)=0$$

implica

$$\lambda_1 = 5 \quad ; \quad \lambda_2=\lambda_3=-3$$

Para $\lambda_1 = 5$, tenemos

$$\mathbf{x}_1 = (1,2,-1)^T \qquad\qquad (\blacksquare \text{ Verificar!})$$

Para los autovalores repetidos ...
$\lambda_2=\lambda_3=-3$, la matriz característica

$$(A-\lambda.I) = A+3.I = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{pmatrix}$$

se reduce por filas a

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

que tiene rango igual a 1. De la ecuación obtenida $x_1+2x_2-3x_3=0$, o $x_1=-2x_2+3x_3$, vemos que tenemos dos variables independientes (en este caso, x_2 y x_3). La solución general será

$$\mathbf{x} = \begin{pmatrix} -2\alpha+3\beta \\ \alpha \\ \beta \end{pmatrix}$$

Así podemos generar dos soluciones linealmente independientes eligiendo $(\alpha, \beta) = (1, 0)$ y $(\alpha, \beta) = (0, 1)$

$$\mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \mathbf{x}_3 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

... Esto es coherente, ya que si el rango de la matriz $A + 3.I$ es igual 1 y $n = 3$, entonces la dimensión del espacio nulo $N((A + 3.I))$ es la diferencia $n - 1 = 2$.

... Hay 3 autovectores linealmente independientes ya que para los repetidos hay 2 autovectores linealmente independientes pués $dim(N(A + 3.I)) = 2$

Definición *Sea λ un autovalor de una matriz A.*
Se denomina ...

Multiplicidad algebraica de λ: *al número M_λ que indica el orden de la multiplicidad de λ como raíz del polinomio característico $p(\lambda)$.*

Multiplicidad geométrica de λ: *al número m_λ de autovectores linealmente independientes asociados a λ. Este número indica la dimensión del "autoespacio" correspondiente a λ (que es $dimN(A - \lambda.I)$).*

Observación *Como el polinomio característico tiene grado n se tiene que ...*
... $\sum(multiplicidades\ algebraicas) = n$

Observación *En general, sucede que ...*
... $m_\lambda \leq M_\lambda$. *La diferencia se llama defecto:* $\Delta_\lambda = M_\lambda - m_\lambda$.

Ejemplo *Sea la matriz ...*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \textit{para calcular sus autovalores...se plantea} \det(A - \lambda.I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = 0$$

Por tanto, $\lambda = 0$ es un autovalor de multiplicidad algebraica 2.

... Pero la multiplicidad geométrica es 1, ya que los autovectores x correspondientes a $\lambda = 0$ son resolviendo

... $(A - 0.I)x = 0$, las soluciones de $Ax = 0$ (hacerla).

Resolviendo se obtiene que los autovectores son todos de la forma

$$x = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$$

Por tanto, $M_0 = 2$; $m_0 = 1$; $\Delta_0 = 1$.

... *Este es un ejemplo dónde la suma de las multiplidades geométricas es " menor " que n.* Entonces, no existen n autovectores independientes para la matriz dada.

...

Conclusión *Si hay autovalores repetidos ... la existencia o no de n autovectores independientes...*

... depende de la suma de las multiplicidades geométricas de los autovalores. Si la multiplicidad geométrica es $n \Rightarrow$ hay n autovectores linealmente independientes. Sino, hay tantos independientes como el valor de la suma $\sum m_\lambda$.

Problema *¿Cuántos autovectores linealmente independientes tiene esa matriz?.*

$$\begin{pmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{pmatrix}$$

(a) Usando el Matlab calcular los autovalores. (b)Hallar los autovectores resolviendo $(A - \lambda_i * I_3)x = 0$, para cada autovalor.

(c) Cúal es la multiplicidad geométrica de los autovalores repetidos (si hay)?.

(d) Verificar si es correcta la resolución de (b) usando el MATLAB y el comando: $null(A - \lambda_i * eye(3)$ (observando cuántos elementos tiene una base de ese espacio nulo).

(e) Finalmente, cuántos autovectores linealmente independientes tiene A ?.

8 Matrices semejantes

Ya las hemos definido ...

... Recordemos que una matriz B de $n \times n$ es semejante a una matriz A de $n \times n$ si existe una matriz no singular S tal que

$$B = S^{-1}.A.S$$

También recordemos que las matrices representativas de operadores lineales respecto de diferentes bases "son semejantes".

Los autovalores de B están determinados por las raíces de su *polinomio característico:*

$$\begin{aligned} p_B(\lambda) &= \det(B - \lambda.I) = \det\left(S^{-1}.A.S - \lambda.I\right) \\ &= \det\left(S^{-1}(A - \lambda.I).S\right) \\ &= \det S^{-1}.\det(A - \lambda.I).\det S \\ &= \det\left(S^{-1}.S\right).\det(A - \lambda.I) \\ &= \det(A - \lambda.I) \\ &= p_A(\lambda) \end{aligned}$$

... A y B semejantes, tienen precisamente el mismo polinomio característico ...

... y por tanto el mismo conjunto de autovalores $\{\lambda_i\}$.

¿Qué pasa con los autovectores de B?

Si $\mathbf{x}$ es una autovector de A con autovalor λ, entonces

$$\begin{aligned} A.\mathbf{x} &= \lambda.\mathbf{x} \\ S.B.S^{-1}. &= \lambda.\mathbf{x} \\ S^{-1}\left(S.B.S^{-1}\right).\mathbf{x} &= S^{-1}.(\lambda.\mathbf{x}) \\ B.\left(S^{-1}.\mathbf{x}\right) &= \lambda.\left(S^{-1}.\mathbf{x}\right) \end{aligned}$$

... Así el autovector de B correspondiente al autovalor λ es justamente $S^{-1}.\mathbf{x}$, donde $\mathbf{x}$ es el autovector de A asociado a λ.

Conclusión *... Si A y B son matrices semejantes de $n \times n$, con $B = S^{-1}.A.S$, entonces tienen el mismo conjunto de autovalores $\{\lambda_i\}$; y si $\mathbf{u}_i$ es un autovector de A correspondiente a λ_i, entonces $S^{-1}.\mathbf{u}_i$ es un autovector de B, correspondiente a λ_i.*

9 Diagonalización

Definición *Una matriz A de $n \times n$ es diagonalizable si existe una matriz no singular X tal que*

$$X^{-1}.A.X = D$$

donde D es una matriz diagonal. En este caso, se dice que X diagonaliza a A.

Teorema *Una matriz A de $n \times n$ es diagonalizable si y sólo si A tiene n autovectores linealmente independientes.*

Demostración. Supongamos que A tienen n autovectores linealmente independientes $\{\mathbf{x}_i\}$ correpondientes a los autovalores $\{\lambda_i\}$ (no necesariamente todos distintos). Definimos X como la matriz de $n \times n$ cuyas columnas son los autovectores $\{\mathbf{x}_i\}$

$$X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$$

Entonces

$$\begin{aligned} A.X &= (A.\mathbf{x}_1, A.\mathbf{x}_2, \ldots, A.\mathbf{x}_n) \\ &= (\lambda_1.\mathbf{x}_1, \lambda_2.\mathbf{x}_2, \ldots, \lambda_n.\mathbf{x}_n) \\ &= (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) . \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ &= X.D \end{aligned}$$

Dado que los $\{\mathbf{x}_i\}$ son linealmente independientes, X es no singular; y por tanto ...

$$A.X = X.D \Longrightarrow D = X^{-1}.A.X$$

La recíproca se demuestra usando la misma idea, comenzando desde el conocimiento que existe una X no singular tal que $X^{-1}AX = D$, D diagonal. Luego, vale que $A.X = X.D$...

... desde ahí se llega a que las columnas de X son autovectores de A (linealmente independientes, ya que X es no singular). ∎

Conclusión ***Observación*** *Sea A un matriz de $n \times n$*

- *Si A es diagonalizable, entonces los vectores columnas de la matriz de diagonalización X son autovectores de A, y los elementos de la diagonal de D son los autovalores correspondientes.*
- *La matriz de diagonalización X no es única. Reordenando su columnas, o multiplicandolas por un escalar no nulo, se obtiene una nueva matriz de diagonalización.*
- *Si A tiene n autovalores "distintos", los autovectores correspondientes son linealmente independientes, y por tanto A es diagonalizable.*
- *Si los autovalores de A no son todos distintos, A puede ser o no diagonalizable.*
- *Si A tiene menos de n autovectores linealmente independientes, no es diagonalizable.*

  Así...
- *... si hay autovalores repetidos se debe hallar la multiplicidad geométrica de los autovalores repetidos (que coincide con la dimensión del autoespacio, o sea con la $dim(N(A - \lambda_i * I))$, si el autovalor es λ_i).*

  ... Si la multiplicidad geométrica coincide con la multiplicidad algebraica de los autovalores repetidos, entonces hay n autovectores linealmente independientes.

  ... Si la multiplicidad geométrica para un λ_i repetido, es menor que su multiplicidad algebraica, entonces A no es diagonalizable (no hay n autovectores independientes, sino sólo $r = \sum m_\lambda$).

Ejercicios

9.1 Diagonalizar la matriz (hallar autovalores, y una base de autovectores si es posible):

$$\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$$

9.2 Para la misma matriz anterior, usando el MATLAB, si es diagonalizable hallar la descomposición $A = XDX^{-1}$.

Usar comandos: $E = eig(A)$, obtiene los autovalores.

$[U, D] = eig(A)$, muestra los autovectores de A en las columnas de U (estas columnas pueden ser dependientes si hay autovalores repetidos), y en D los autovalores.

Otros comandos: $null(A - \lambda_i * eye(n))$, muestra una base del autoespacio asociado a λ_i (se usa si el autovalor es repetido). Recordar que $eye(n)$ es I_n.

9.3 ¿Es posible diagonalizar la matriz A siguiente?. Hallar los autovalores. El autosistema para cada autovalor es el subespacio $N(A - \lambda_i * I_n)$. Hallar la multiplicidad geométrica de cada autovalor.

¿Cuántos autovectores linealmente independientes hay ?.

$$\begin{pmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{pmatrix}$$

....**Para visualizar y repasar**

Otros resultados importantes ...

- *Si una matriz A es diagonalizable, puede ser factoreada en el producto $X.D.X^{-1}$. Por tanto ...*

$$A^2 = \left(X.D.X^{-1}\right).\left(X.D.X^{-1}\right) = X.D^2.X^{-1} \tag{8}$$

y en general,

$$\begin{aligned} A^k &= X.D^k.X^{-1} \\ &= X.\begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.X^{-1} \end{aligned}$$

Ejemplo *Sea*

$$A = \begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}$$

Entonces ...

$$\det(A - \lambda.I) = (2-\lambda)(-5-\lambda) + 6 = \lambda^2 + 3\lambda - 4 = (\lambda - 1)(\lambda + 4)$$

... entonces $\lambda_1 = 1$ y $\lambda_2 = -4$ son los autovalores de A (que son distintos). Si buscamos los autovectores, obtendremos $\mathbf{u}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Obtenemos ...

$$X = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$$

Entonces ...

$$\begin{aligned} X^{-1}.A.X &= \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}.\begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}.\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = D \end{aligned}$$

Observar, que si reordenamos $\mathbf{u}_1$ y $\mathbf{u}_2$ en X

$$\hat{X} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$$

Entonces ...

$$\begin{aligned} \hat{D} &= \hat{X}^{-1}.A.\hat{X} = -\frac{1}{5}\begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix}.\begin{pmatrix} 2 & -3 \\ 2 & -5 \end{pmatrix}.\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -4 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} \end{aligned}$$

Conclusión *... Esto es, si A es diagonalizable, puede factorearse $A = XDX^{-1}$, y los autovalores de A simplemente se reordenan en la diagonal de D, y X está formada por los autovectores independientes.*

Además, al multiplicar X por un escalar $\alpha \neq 0$ también implica multiplicar X^{-1} por $1/\alpha$, lo cual deja invariante a D.

Otra consecuencia interesante para matrices diagonalizables:

Si A es diagonalizable, $A = XDX^{-1}$, y para toda potencia de A ...
... $A^k = XD^kX^{-1}$. Así la función exponencial de una matriz A de $n \times n$ se define (como ocurre con e^x):

$$e^A = exp(A) = I + A + A^2 + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

siendo $A^0 = I_n$.
... Usando $A = XDX^{-1}$, se obtiene

$$e^A = exp(A) = X(I + D + D^2 + \frac{D^2}{2!} + \frac{D^3}{3!} + \ldots)X^{-1} = X(\sum_{k=0}^{\infty} \frac{D^k}{k!})X^{-1}$$

Luego, considerando que

$$e^D = lim_{n->\infty} \left(I + D + D^2 + \frac{D^2}{2!} + \ldots + \frac{D^n}{n!} \right)$$

.. . y usando (8)

$$e^D \quad = \quad lim_{m->\infty} \begin{pmatrix} \sum_{k=0}^{m} \frac{\lambda_1^k}{k!} & & & \\ & \sum_{k=0}^{m} \frac{\lambda_2^k}{k!} & & \\ & & \ddots & \\ & & & \sum_{k=0}^{m} \frac{\lambda_n^k}{k!} \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{pmatrix}$$

$\Rightarrow$

$$e^A = Xe^DX^{-1}.$$

... es calculable si se conocen los autovalores de A, y los n autovectores de A.
... *Esto es muy útil cuando se resuelven sistemas de ecuaciones diferenciales.*

Para practicar con diagonalización de matrices...

Ejercicios

✓ **9.1** S i T es una matriz $n\times n$ y c, d son escalares.

(a) Probar que si T tiene un autovalor λ con un asociado autovector $\vec{v}$ entonces $\vec{v}$ es un autovector de $cT + dI_n$ asociado con el autovalor $c\lambda + d$.

(b) Probar que si T es diagonalizable entonces lo es $cT + dI_n$.

9.2 Matrices equivalentes por filas tienen iguales sus autovalores?.

La respuesta es no!. Verificar que las siguientes matrices son equivalentes y no tienen los mismos autovalores. Estas matrices equivalentes tienen la misma magnitud, igual rango y distintos autovalores.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

9.3 Mostrar que una matriz cuadrada con coeficientes reales y un número impar de filas tiene siempre al menos un autovalor real.

9.4 Diagonalizar la matriz (hallar los autovalores, y autovectores, multiplicidad algebraica y multiplicidad geométrica):

$$\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$$

9.5 Es posible diagonalizar la matriz A siguiente?. Hallar autovalores. El autoespacio para cada autovalor (es $N(A - \lambda_i * I_n)$). Hallar la multiplicidad geométrica para cada autovalor. ¿Cántos autovectores linealmente independientes hay ?.

$$\begin{pmatrix} -1 & -3 & -9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{pmatrix}$$

9.6 Diagonalizar la matriz (hallar autovalores, y una base de autovectores (si es posible):

$$\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$$

9.7 Repetir el ejercicio anterior resolviendo con el MATLAB, usando: E=eig(A). Luego con el comando : [U,D]= eig(A) (obtiene los autovectores de A en las columnas de U (pueden ser dependientes), y en D los autovalores. Hallar el $N(A - lambda.I)$, y una base correspondiente a cada autoespacio de cada autovalor.

9.8 Es posible diagonalizar la matriz siguiente?. (a)Hallar autovalores y los autovectores linealmente independientes (observar que es simétrica):

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

10 Matrices Simétricas.Matrices hermitianas.

- **En el caso de matrices *reales* especialmente tenemos:**

  ...

Matrices simétricas: $A^T = A$

Matrices ortogonales: las columnas de A forman un conjunto ortonormal de $\mathbb{R}^n$ (y por lo tanto, una base).

- **En cambio en el caso de las matrices *complejas* tenemos lo equivalente ... :**

  Matrices hermitianas: Son las que satisfacen $M^H \ \left(= \bar{M}^T\right) = M$.

  Matrices unitarias: las columnas de M forman un conjunto ortonormal de $\mathbb{C}^n$ (y por tanto, una base).

Luego ...

- Una matriz *real* hermitiana es una matriz simétrica
- Una matriz *real* unitaria es una matriz ortogonal

Resultados relacionados importantes:

Estos resultados se dan aquí sin demostración (se recomienda ver la bibliografía S. Grossman para conocer los fundamentos que conducen a ellos).

Los resultados siguientes son de gran importancia y de ellos depende la resolución de muchas aplicaciones importantes.

En el caso de matrices reales...

- *Una matriz* real simétrica *tiene "todos" sus autovalores* reales.

  Además ...

- *Si A es simétrica, existe una matriz* ortogonal U *que diagonaliza a A*

$$U^{-1}.A.U = U^T.A.U = D$$

Conclusión ... *Esto quiere decir que si A es una matriz simétrica " existe siempre" una base de autovectores ortogonales.*

... Proviene esa conclusión de los resultados ...

(1) Si A es simétrica, los autovectores asociados a autovalores distintos son ortogonales.

(2) Si A es simétrica, los autovalores repetidos tienen la multiplicidad geométrica igual a la multiplicidad algebraica. Es decir, si la multiplicidad algebraica de un λ_i es r, entonces la dimensión $dim(N(A - \lambda_i.I)) = r$. Observar que, por ejemplo con el comando del MATLAB $null(A - \lambda_i.I)$, puede siempre obtener una base ortogonal de este autoespacio de λ_i. Así...

(3) Existe una base ortogonal de autovectores si A es simétrica.

Ejercicios

10.1 Rehacer un ejercicio previo: ¿ Es posible diagonalizar la matriz siguiente?. Hallar autovalores y los autovectores ortogonales (Observar que es simétrica).

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

10.2 Es posible diagonalizar la matriz ?

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

10.3 Idem para

$$\begin{pmatrix} 4 & -1 & 0 & -1 \\ -1 & 4 & 0 & 0 \\ 0 & 0 & 4 & -1 \\ -1 & 0 & -1 & 4 \end{pmatrix}$$

OBSERVACIONES:

Hacemos una *pequeña* incursión en el mundo de los escalares complejos, los vectores y matrices complejas ...

El conjunto $\mathbb{C}^n$ es el espacio vectorial formado por las n-uplas de números complejos. Serán los escalares a considerar α:

$$\alpha = a + ib \quad (\text{con } a \text{ y } \ b \text{ números reales})$$

Recordemos que...

Conjugado de α:

$$\bar{\alpha} = a - ib$$

Longitud o módulo de α :

$$|\alpha| = \sqrt{\alpha.\bar{\alpha}} = \sqrt{a^2 + b^2}$$

Para los vectores:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = (z_1, z_2, \dots, z_n)^T \quad \text{y} \quad \bar{\mathbf{z}} = \begin{pmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_n \end{pmatrix} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_n)^T$$

Longitud o módulo de un vector:

$$\begin{aligned} \|\mathbf{z}\| &= \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} \\ &= \left(\bar{\mathbf{z}}^T.\mathbf{z}\right)^{1/2} \end{aligned}$$

Notación:

$$\bar{\mathbf{z}}^T \equiv \mathbf{z}^H \quad \text{y} \quad \|\mathbf{z}\| = \left(\mathbf{z}^H.\mathbf{z}\right)^{1/2} \quad (\text{conjugado transpuesto de } \mathbf{z})$$

Producto interno:

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H.\mathbf{z}$$

Propiedades:

1. $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$ (”=” si y sólo si $\mathbf{z} = 0$)

2. $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ para todo $\mathbf{z}$ y $\mathbf{w}$ en V

3. $\langle \alpha.\mathbf{z} + \beta.\mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{z}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle$

Matrices:

$$\begin{aligned} M &= \{m_{ij}\} \quad \text{con } m_{ij} = a_{ij} + ib_{ij} \quad \text{o,} \\ M &= A + iB \quad \text{con } A = \{a_{ij}\} \text{ y } B = \{b_{ij}\} \\ \text{y} \quad M^H &= \bar{M}^T \end{aligned}$$

Para matrices complejas ...

- los autovalores de una matriz *hermitiana* son *todos reales*.
- La matriz conjugada transpuesta de una matriz unitaria U es su inversa: $U^{-1} = U^H$ (análogamente ocurres para matrices reales ortogonales: $A^{-1} = A^T$)
- Si A es una matriz hermitiana, entonces existe una matriz unitaria U que diagonaliza a A

$$U^{-1}.A.U = U^H.A.U = D$$

11 Aplicaciones

Nos limitaremos a ver dos casos, más adelante cuando veamos Ecuaciones Diferenciales veremos otras aplicaciones interesantes.

11.1 Funciones de varias variables.

Ahora, consideramos una aplicación de los autovalores de matrices simétricas ... para analizar la curvatura de las gráficas de funciones de varias variables. Para abreviar cálculos lo hacemos para dos variables.

- Sea $z = f(x, y)$ una función con derivadas segundas contínuas. Sea $P_0 = (x_0, y_0)$ un punto de su dominio.

  Sea un vector $u = (u_1, u_2)^T$, de longitud $\|u\| = 1$, que indica una dirección.

- Si restringimos el estudio del comportamiento de $f(x, y)$ en el punto P_0, en la dirección de u, podemos considerar la función compuesta $g(t) = f(P_0 + tu)$. Así en particular cuando $t = 0$, $g(0) = f(P_0)$.

- La derivada $g'(t) = \nabla f(P_0 + tu)^T.u$ usando derivada de funciones compuestas.

  Además, si calculamos

- la derivada segunda de $g(t)$ en $t = 0$, nos da la curvatura de $g(t)$ en $t = 0$, que coincide con la de la función f en P_0, en la dirección indicada por u.

  Tal derivada $g''(t) = (g'(t))' = (\frac{\partial f(P_0+tu)}{\partial x}u_1 + \frac{\partial f(P_0+tu)}{\partial y}u_2)'$. Usando regla de la cadena, se obtiene:

  $g''(t) == \frac{\partial^2 f(P_0+tu)}{\partial x^2}(u_1)^2 + 2\frac{\partial^2 f(P_0+tu)}{\partial xy}u_1.u_2 + \frac{\partial^2 f(P_0+tu))}{\partial y^2}(u_2)^2$.

  Luego, reemplazando en $t = 0$, se obtiene $g''(0) = \frac{\partial^2 f(P_0)}{\partial x^2}(u_1)^2 + 2\frac{\partial^2 f(P_0)}{\partial xy}u_1.u_2 + \frac{\partial^2 f(P_0)}{\partial y^2}(u_2)^2$.

  Si se considera la matriz Hessiana en P_0 ...

  $H = \begin{pmatrix} \frac{\partial^2 f(P_0)}{\partial x^2} & \frac{\partial^2 f(P_0)}{\partial xy} \\ \frac{\partial^2 f(P_0)}{\partial yx} & \frac{\partial^2 f(P_0)}{\partial y^2} \end{pmatrix}$,

  ... la curvatura de $f(x, y)$ en P_0, en la dirección de u, coincide con ...

  $g''(0) = (u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$

Conclusión *Así la curvatura de f en P_0, en una dirección u está dada por el valor de forma cuadrática: $(u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, siendo H la matriz simétrica de las derivadas segundas de f en P_0.*

... Así para saber si la función en el punto P_0 es convexa o cóncava hay que conocer el signo de la forma cuadrática $(u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ para toda dirección u.

Se puede demostrar que ...

(i) La matriz H (simétrica) tiene todos sus autovalores positivos $\Leftrightarrow$ la forma cuadrática $(u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ es "positiva" en toda dirección. Eso dice que si los autovalores del Hessiano H en P_0 son positivos entonces f es convexa en P_0, pués la curvatura es positiva en todas las direcciones u.

(ii) La matriz H (simétrica) tiene todos sus autovalores negativos $\Leftrightarrow$ la forma cuadrática $(u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ es negativa en toda dirección. Eso dice que si los autovalores del Hessiano H en P_0 son negativos, entonces f es cóncava en P_0, pués la curvatura es negativa en todas las direcciones u.

(iii) Si la matriz H (simétrica) tiene al menos un autovalor negativo y otro positivo entonces la forma cuadrática $(u_1, u_2)H\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ es negativa al menos en una dirección y positiva en otra dirección. Así $f(x, y)$ tiene en P_0, curvatura negativa en alguna dirección y curvatura positiva en otra.

Problema *(i) Dada una matriz real simétrica A $(n \times n)$. Verificar que si tiene n autovalores positivos equivale a que la forma cuadrática $u^T Au > 0$, para todo $u \in \Re^n$, con $u \neq 0$.*

Ayuda: considerar que existe una base de autovectores ortonormales $\{u_1, u_2, ..., u_n\}$ tales que $Au_i = \lambda_i u_i$. Luego usar esa base de autovectores para escribir: $u = \sum \alpha_i u_i$. Reemplazar en $u^T Au$, para ver que se obtiene un valor positivo para todo u, si cada $\lambda_i > 0$ ya que $u_i^T Au_i = \lambda_i > 0$.

Recíprocamente, si para todo $u \neq 0$, $u^T Au > 0$, entonces también para todo autovector $u_i^T Au_i = \lambda_i > 0$. Así tiene autovalores positivos.

(ii) Lo mismo se puede hacer cuando todos los autovalores son negativos.

Problema *(i) Sea la función $z = e^{xy}$, en el punto $(1,2)$. Ver cual es la curvatura en $(1,2)$ en cualquier dirección u, de acuerdo a los autovalores del Hessiano en ese punto.*

(ii) Sea la función $z = -x^3 + 4xy - 2y^2 + 1$, analizar la curvatura en los puntos dónde el gradiente es cero. Puede obtener alguna conclusión sobre el tipo de puntos estacionarios que tiene?.

11.2 Relación de recurrencia

En 1202 Leonardo de Pisa, conocido como Fibonacci, expuso el siguiente problema.

Problema *Un hombre tiene una pareja de conejos en un corral. ¿Cuántas parejas de conejos se pueden generar a partir de esa pareja en un año, si se sabe que cada mes se reproduce generando una nueva pareja, la que a partir de su segundo mes de vida también se reproduce de la misma forma?.*

Hay supuestos simplificadores, como que no hay períodos de gestación y que no hay mortalidad.

- Dado un cierto mes, el número de pares nuevos que aparecerán en el próximo mes es simplemente el número de pares que vivían en el mes previo pasado, desde que todos estos serán fértiles después de dos meses.
- Dado un cierto mes, el número de pares, $f(n+1)$, que vivirán en el mes siguiente es la suma del número de los que viven en este mes, $f(n)$, más el número de los nacimientos (igual a $f(n-1)$, correspondiente a los que vivían en el mes $n-1$).

Así se tiene la siguiente relación

$$f(n+1) = f(n) + f(n-1) \qquad \text{donde } f(0) = 1,\ f(1) = 1$$

Este es un ejemplo de una relación de recurrencia, llamada así porque los valores de f se calculan mirando los previos valores de f.

Podemos responder la pregunta de Fibonacci, despues de doce meses, sabiendo que $f(n+1) = f(n) + f(n-1)$, para cada mes n, y conociendo las condiciones iniciales $f(0)$ y $f(1)$. A partir de la fórmula, se obtiene $f(2) = f(1) + f(0)$, $f(3) = f(2) + f(1)$,y así siguiendo ...

mes	0	1	2	3	4	5	6	7	8	9	10	11	12
parejas	1	1	2	3	5	8	13	21	34	55	89	144	233

La sucesión $f(n)$ de números definida por la ecuación de arriba (se listan los primeros) se llama la sucesión de *Fibonacci* .

Lo que hemos visto en las clases previas es útil para calcular $f(n+1)$ sin tener que encontrar cada uno de los previos $f(n)$, $f(n-1)$, etc. Por ejemplo, para tal sucesión nos puede interesar $f(100)$, o cualquier otro más grande.

Cómo hacer para conocer $f(100)$, por ejemplo ?

Observar que la relacin de recurrencia es lineal y asípodemos dar una formulación matricial para ella.

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f(n) \\ f(n-1) \end{pmatrix} = \begin{pmatrix} f(n+1) \\ f(n) \end{pmatrix} \qquad \textit{donde} \begin{pmatrix} f(1) \\ f(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Entonces, si denominamos
A a la matriz y ...
... $\vec{v}_n$ al vector de componentes $f(n+1)$ y $f(n)$,
tenemos que

$$\vec{v}_n = A\vec{v}_{n-1},$$

así reemplazando $\vec{v}_{n-1}$ usando $\vec{v}_{n-2}$, se tendría

$$\vec{v}_n = AA\vec{v}_{n-2},$$

..., hasta reemplazar $\vec{v}_1$ mediante $A\vec{v}_0$, se obtiene que

$$\vec{v}_n = A^n\vec{v}_0.$$

La ventaja de esta formulación es que si diagonalizamos A alcanzamos una forma rápida para calcular sus potencias: donde $A = SDS^{-1}$, tenemos $A^n = SD^nS^{-1}$, y la n-ésima potencia de la matriz diagonal D es la matriz diagonal cuyos elementos diagonales son la potencia-n de los elememtos de D.

La ecuación característica de A is $\lambda^2 - \lambda - 1$, que tiene las raiíces $(1+\sqrt{5})/2$ y $(1-\sqrt{5})/2$. Diagonalizando se tiene:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{pmatrix}$$

Introduciendo los vectores y tomando la potencia-n, tenemos

$$\begin{aligned} \begin{pmatrix} f(n+1) \\ f(n) \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} f(1) \\ f(0) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2}^n & 0 \\ 0 & \frac{1-\sqrt{5}}{2}^n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{pmatrix} \begin{pmatrix} f(1) \\ f(0) \end{pmatrix} \end{aligned}$$

Podemos calcular $f(n)$ desde la segunda componente de la ecuación,

$$f(n) = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$$

Notamos que f es dominada por su primer término porque $(1-\sqrt{5})/2$ es menor que 1, así sus potencias tienden a cero.

En general, una *relación de recurrencia lineal* tiene la forma

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \cdots + a_k f(n-k)$$

(también se llama una ecuación en diferencias). Esta relación de recurrencia es *homogénea* porque no hay un término constante.

Se puede poner en la forma:

$$0 = -f(n) + a_1 f(n-1) + a_2 f(n-2) + \cdots + a_k f(n-k). \tag{9}$$

Esta es una relación de recurrencia de *orden k*. Esa relación, con las *condiciones iniciales*

$$f(0), \ldots, f(k-1)$$

determina completamente la sucesión, para $n \geq k$.

Por ejemplo, la relación de Fibonacci es de orden 2, y con las condiciones inicales $f(0) = 1$ y $f(1) = 1$, determina la secuencia de Fibonacci, para $f(2)$, $f(3)$....

Veremos como se puede usar el Algebra Lineal para resolver relaciones de recurrencia lineales.

- Primero, definimos el espacio vectorial en el cual trabajamos.

  Sea V el conjunto de funciones f definidas para los números naturales $N = \{0, 1, 2, \ldots\}$ y con valores $f(n)$ en los números reales. (Tendremos también funciones con dominio $\{1, 2, \ldots\}$, eso es , sin el 0, pero eso no cambia lo principal.)

  Dejando de lado "las condiciones iniciales" por un momento, para cualquier relación de recurrencia, podemos considerar el

  ... subconjunto S (del espacio V) de las funciones soluciones $f(n)$ de una relación de recurrencia.

- El subconjunto *S de soluciones* es un subespacio de V.

  Es no vaciío, ya que la función $f(n) = 0$, constante, satisface la relación (9) (aunque no las específicas condiciones iniciales).

  Además, si dos funciones f_1 y f_2 son soluciones, también la suma de ellas $f_1 + f_2$ lo es:

$$\begin{aligned}
&-(f_1+f_2)(n)+a_1(f_1+f_2)(n-1)+\cdots+a_k(f_1+f_2)(n-k) \\
&= (-f_1(n)+a_1 f_1(n-1)+\cdots+a_k f_1(n-k)) \\
&\quad +(-f_2(n) a_1 f_2(n-1)+\cdots+a_k f_2(n-k)) \\
&= 0.
\end{aligned}$$

También, si se multiplica por un escalar rf_1, lo es :

$$
\begin{aligned}
&(-rf_1)(n)+a_1(rf_1)(n-1)+\cdots+a_k(rf_1)(n-k)\\
&= r(-f_1(n)+a_1(f_1)(n-1)\cdots+a_kf_1(n-k))\\
&= r\cdot 0\\
&= 0.
\end{aligned}
$$

¿Cuál es la dimensión del subespacio de soluciones S ?.

Consideremos la aplicación: $T : S \to \Re^k$. desde el conjunto de las funciones de S al conjunto de vectores de $\Re^k$.

$$
f \mapsto \begin{pmatrix} f(0) \\ f(1) \\ \vdots \\ f(k-1) \end{pmatrix}
$$

Como cualquier solución de la relación de recurrencia está univocamente determinada por las k **condiciones iniciales**, *esta aplicación es "inyectiva" y "suryectiva" sobre $\Re^k$.*

Así S tiene la misma dimensión k que $\Re^k$, que es el orden de la relación de recurrencia.

Así (dejando de lado todavía las condiciones iniciales), podemos describir el conjunto de soluciones de cualquier relación de recurrencia lineal homogenéa de orden k a partir de conocer una base de S, o sea un conjunto de k funciones linealmente independientes, y luego combinando a estas funciones k se obtiene cualquier otra solución.

... Cómo obtener estas k funciones linealmente independientes?, y obtener las restantes...

Para eso, expresamos en forma matricial la relación de recurrencia,

$$
f(n) = a_1 f(n-1) + \cdots + a_k f(n-k)
$$

$$
\begin{pmatrix}
a_1 & a_2 & a_3 & \dots & a_{k-1} & a_k \\
1 & 0 & 0 & \dots & 0 & 0 \\
0 & 1 & 0 & & & \\
0 & 0 & 1 & & & \\
\vdots & \vdots & & \ddots & & \vdots \\
0 & 0 & 0 & \dots & 1 & 0
\end{pmatrix}
\begin{pmatrix} f(n-1) \\ f(n-2) \\ \vdots \\ f(n-k) \end{pmatrix}
=
\begin{pmatrix} f(n) \\ f(n-1) \\ \vdots \\ f(n-k+1) \end{pmatrix}
$$

Para encontrar la ecuación característica de la matriz, lo vemos primero para el caso de una matriz de 2×2*:*

$$\begin{pmatrix} a_1-\lambda & a_2 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - a_1\lambda - a_2$$

y para el caso de una matriz 3×3*.*

$$\begin{pmatrix} a_1-\lambda & a_2 & a_3 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{pmatrix} = -\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$$

que muestra que la ecuación característica es la siguiente.

$$\begin{vmatrix} a_1-\lambda & a_2 & a_3 & \dots & a_{k-1} & a_k \\ 1 & -\lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & -\lambda & & & \\ 0 & 0 & 1 & & & \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\lambda \end{vmatrix}$$

$$= \pm(-\lambda^k + a_1\lambda^{k-1} + a_2\lambda^{k-2} + \cdots + a_{k-1}\lambda + a_k)$$

Este es el polinomio asociado con la relación de recurrencia.

Si el polinomio

$$-\lambda^k + a_1\lambda^{k-1} + a_2\lambda^{k-2} + \cdots + a_{k-1}\lambda + a_k$$

no tiene raíces repetidas entonces la matriz es diagonizable y podemos, en teoría, tener la fórmula para $f(n)$ como en el caso de Fibonacci.

Pero, como sabemos que el subespacio S *de las soluciones tiene dimension* k*, " no necesitamos diagonalizar", si conocemos k funciones linealmente independientes satisfaciendo la relación.*

Si $r_1,\ r_2,\ \dots,\ r_k$ *son las raíces distintas, consideramos las funciones* $f_{r_1}(n) = r_1^n$*,* $f_{r_2}(n) = r_2^n$*... hasta* $f_{r_k}(n) = r_k^n$*, potencias de las raíces del polinomio.*

Problema *(i) Mostrar que cada una de esas funciones ees una solución de la relación (9).*

(ii) Ver que son linealmente independientes.

Síntesis *Dada una relación de recurrencia lineal homogenea* $f(n) = a_1 f(n-1) + \cdots + a_k f(n-k)$ *(es decir,* $0 = -f(n) + a_1 f(n-1) + \cdots + a_k f(n-k)$*) consideramos la ecuación asociada* $0 = -\lambda^k + a_1 \lambda^{k-1} + \cdots + a_{k-1}\lambda + a_k$.

Encontramos las raíces $r_1, \ldots, r_k$*, y si son distintas entonces toda solución de la relación de recurrencia tiene la forma*

$$f(n) = c_1 r_1^n + c_2 r_2^n + \cdots + c_k r_k^n$$

para ciertas constantes $c_1, \ldots, c_n \in \Re$*, que dependerán de las condiciones iniciales.*

El caso de raíces repetidas también es fácil de hacer, pero no lo cubriremos aquí .

Solamente diremos que en el caso de orden $k = 2$ *(si las raíces son repetidas* $r_1 = r_2$*), las soluciones independientes son :*

$$f_{r_1}(n) = r_1^n, \text{ y la otra se construye } \quad f_2(n) = n r_1^n$$

Ahora, dadas algunas condiciones iniciales:

Si se conocen las condiciones iniciales, vamos a determinar los coeficiente c_i que combinan adecuadamente a las funciones de la base.

Para hallar $c_1, \ldots, c_n$. Por ejemplo, sea el polinomio asociado con la relación de Fibonacci $-\lambda^2 + \lambda + 1$, cuyas raíces son $(1 \pm \sqrt{5})/2$ y toda solución satisface:

$f(n) = c_1((1+\sqrt{5})/2)^n + c_2((1-\sqrt{5})/2)^n$.

Incluyendo las condiciones iniciales para $n = 0$ ($f(0) = 1$) y $n = 1$ ($f(1) = 1$), tenemos el sistema lineal de ecuaciones en c_1 y c_2:

$$\begin{aligned} c_1 + \qquad\qquad c_2 &= 1 \\ (1+\sqrt{5}/2)c_1 + (1-\sqrt{5}/2)c_2 &= 1 \end{aligned}$$

Resolviendo ese sistema, conduce a $c_1 = 1/\sqrt{5}$ y $c_2 = -1/\sqrt{5}$, como se habia calculado antes.

Ejercicios

Resolver la siguientes relaciones lineales de recurrencia:

(a) $f(n) = f(n-1) + f(n-2), \quad f(0) = 1/2;\ f(1) = 1;$

(b) $f(n) = f(n-2)/2, \quad f(0) = 1;\ f(1) = 0;$

(c) $f(n) = f(n-1) + f(n-2)/2, \quad f(0) = 0;\ f(1) = 1;$

(d) Escribir la solución de las relaciones anteriores, para condiciones iniciales generales $f(0) = f_0$, $f(1) = f_1$.

Apéndices

1. Autovalores de matrices definidas por bloques

Consideremos la matriz

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

donde A es una matriz de $n \times n$, B de $m \times m$ y 0 denota matrices nulas, tales que M es de $(n+m) \times (n+m)$. Estas matrices aparecen frecuentemente en diversos problemas, representando por ejemplo dos sistemas A y B independientes.
Es fácil ver que para hallar los autovalores y autovectores de M, basta con obtener los autovalores de A y de B por separado, y luego unir ambos conjuntos, completando con ceros los autovectores correspondientes:

$$AX = \lambda^A X \quad \Rightarrow \quad \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} X \\ 0 \end{pmatrix} = \lambda^A \begin{pmatrix} X \\ 0 \end{pmatrix}$$

$$BY = \lambda^B Y \quad \Rightarrow \quad \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} 0 \\ Y \end{pmatrix} = \lambda^B \begin{pmatrix} 0 \\ Y \end{pmatrix}$$

Estas expresiones muestran que todo autovalor y autovector de A o B origina un autovalor y autovector de M. Además, hemos visto antes que $\det M = \det A \det B$ y por lo tanto,

$$\det[M - \lambda I] = \det\,[A - \lambda I]\det[B - \lambda I] \tag{1}$$

La ecuación característica $\det[M - \lambda I] = 0$ es entonces $\det\,[A - \lambda I]\det[B - \lambda I] = 0$, por lo que toda raíz λ debe ser necesariamente raíz del primer o segundo factor, es decir autovalor de A o B.

Y si λ es autovalor de M, el sistema homogéneo $(M - \lambda I)\binom{X}{Y} = 0$ para obtener los autovectores se reduce a dos sistemas independientes: $(A - \lambda I)X = 0$ y $(B - \lambda I)Y = 0$. Si λ es autovalor de A y no de B, el segundo sistema tendrá sólo la solución trivial $Y = 0$, por lo que los autovectores serán de la forma $\binom{X}{0}$, y si λ es autovalor de B y no de A, los autovectores serán de la forma $\binom{0}{Y}$. Y si λ es autovalor de A y de B, sigue siendo posible elegir autovectores de la forma $\binom{X}{0}$ y $\binom{0}{Y}$, que formen una base del espacio propio.

Ejemplo: Consideremos la matriz

$$M = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Para hallar sus autovalores y autovectores, obtenemos primero los de $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, que son $\lambda_1^A = 0$, $\lambda_2^A = 2$, con autovectores $X_1 \propto \binom{1}{1}$ y $X_2 \propto \binom{1}{-1}$, (probar!) y luego los de $B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, que son $\lambda_1^B = 3$, $\lambda_2^B = 1$, con autovectores $Y_1 \propto \binom{1}{1}$, $Y_2 \propto \binom{1}{-1}$.

Los autovalores y autovectores de M serán entonces

$$\lambda_1 = 0,\ V_1 \propto \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ \ \lambda_2 = 2,\ V_2 \propto \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},\ \ \lambda_3 = 3,\ V_3 \propto \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \lambda_4 = 1,\ V_4 \propto \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

Si algún autovalor de B de multiplicidad 1 coincidiese con alguno de A de multiplicidad 1, ese autovalor tendrá en M multiplicidad algebraica y geométrica 2. Una base del espacio propio asociado estará formada por la unión de los autovectores asociados $\binom{X}{0}$ y $\binom{0}{Y}$.
En general, la multiplicidad algebraica de un autovalor λ en M será la suma de las multiplicidades algebraicas en A y B, y lo mismo rige para la multiplicidad geométrica.

Problemas:

1. En base a los resultados anteriores, determinar los autovalores y una base de los espacios propios asociados de las matrices

$$a)\ M = \begin{pmatrix} 2 & -2 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \qquad b)\ M = \begin{pmatrix} 2 & -2 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 1 \end{pmatrix}$$

$$c)\ M = \begin{pmatrix} 2 & -2 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 2 \end{pmatrix} \quad d)\ M = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 3 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

2. Para pensar: Discuta el caso en que

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

con A de $n \times n$, B de $m \times m$ y C de $n \times m$. Los sistemas A y B ya no son independientes. Muestre sin embargo que la ecuación (1) sigue siendo válida, por lo que los autovalores de M siguen siendo los de A y B: Para hallar los autovalores basta nuevamente con hallar los de A y B.
No obstante, muestre que si bien los autovalores de A darán origen a autovectores $\binom{X}{0}$ de M que dependen sólo de A, los autovectores de M asociados a autovalores de B dependerán en general también de A y C, y pueden incluso no existir si el autovalor asociado es también autovalor de A.

3. Determine los autovalores y autovectores de

$$M = \begin{pmatrix} 2 & -2 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

2. Polinomio de Taylor en varias variables. Forma diagonal

Consideremos un campo escalar $F : \mathbb{R}^n \to \mathbb{R}$, es decir, una función que asigna a cada vector $\boldsymbol{r} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ un número real $F(\boldsymbol{r}) = F(x_1, \ldots, x_n)$. Asumiendo que poseea derivadas parciales de todo orden continuas en una región $|\boldsymbol{r} - \boldsymbol{a}| < R$, con $\boldsymbol{a} = (a_1, \ldots, a_n)$, se define el polinomio de Taylor de grado k como

$$\begin{aligned} P_k(\boldsymbol{r}) &= F(\boldsymbol{a}) + \sum_i \frac{\partial F}{\partial x_i}(\boldsymbol{a})(x_i - a_i) + \frac{1}{2!}\sum_{i,j}\frac{\partial^2 F}{\partial x_i \partial x_j}(\boldsymbol{a})(x_i - a_i)(x_j - a_j) \quad (1)\\ &+ \ldots + \frac{1}{k!}\sum_{i_1, i_2, \ldots, i_k} \frac{\partial^k F}{\partial x_{i_1}\partial x_{i_2}\ldots \partial x_{i_k}}(\boldsymbol{a})(x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2})\ldots(x_{i_k} - a_{i_k}) \quad (2)\end{aligned}$$

donde $\frac{\partial F}{\partial x_i}(\boldsymbol{a}) = \frac{\partial F}{\partial x_i}|_{\boldsymbol{r}=\boldsymbol{a}}$ y las sumas sobre i, j, $i_1, \ldots i_k$ corren entre 1 y n. Este polinomio ajusta todas las derivadas parciales de F hasta orden k en $\boldsymbol{r} = \boldsymbol{a}$, reduciéndose en el caso de una dimensión ($n = 1$) al polinomio Taylor ya estudiado. Para $\boldsymbol{r}$ cercano a $\boldsymbol{a}$, la diferencia $F(\boldsymbol{r}) - P_k(\boldsymbol{r})$ es pequeña, con $|F(\boldsymbol{r}) - P_k(\boldsymbol{r})| = O(|\boldsymbol{r} - \boldsymbol{a}|^{k+1})$.

Los tres primeros términos (1) del desarrollo constituyen el polinomio de Taylor de segundo orden $P_2(\boldsymbol{r})$. Mediante el gradiente $\nabla F(\boldsymbol{r}) = (\frac{\partial F}{\partial x_1}, \ldots, \frac{\partial F}{\partial x_n})$ y la matriz Hessiana

$$H(\boldsymbol{r}) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{pmatrix} \quad (3)$$

de $n \times n$, este polinomio puede escribirse en la forma compacta

$$P_2(\boldsymbol{r}) = F(\boldsymbol{a}) + \boldsymbol{\nabla} F(\boldsymbol{a}) \cdot (\boldsymbol{r} - \boldsymbol{a}) + \frac{1}{2}(\boldsymbol{r} - \boldsymbol{a})^T H(\boldsymbol{a})(\boldsymbol{r} - \boldsymbol{a}) \quad (4)$$

donde en el último término hemos considerado a $(\boldsymbol{r} - \boldsymbol{a})$ como matriz columna $(n \times 1)$ y a $(\boldsymbol{r} - \boldsymbol{a})^T$ matriz fila $(1 \times n)$. Por ejemplo, para $n = 2$ y $\boldsymbol{r} = (x, y)$, $\boldsymbol{a} = (a_x, a_y)$, este término resulta

$$\begin{aligned} \tfrac{1}{2}(\boldsymbol{r} - \boldsymbol{a})^T H(\boldsymbol{a})(\boldsymbol{r} - \boldsymbol{a}) &= \tfrac{1}{2}(x - a_x, y - a_y) \begin{pmatrix} \frac{\partial^2 F}{\partial x^2}(\boldsymbol{a}) & \frac{\partial^2 F}{\partial x \partial y}(\boldsymbol{a}) \\ \frac{\partial^2 F}{\partial y \partial x}(\boldsymbol{a}) & \frac{\partial^2 F}{\partial y^2}(\boldsymbol{a}) \end{pmatrix} \begin{pmatrix} x - a_x \\ y - a_y \end{pmatrix} \\ &= \tfrac{1}{2}\left[\tfrac{\partial^2 F}{\partial x^2}(\boldsymbol{a})(x - a_x)^2 + 2\tfrac{\partial^2 F}{\partial x \partial y}(\boldsymbol{a})(x - a_x)(y - a_y) + \tfrac{\partial^2 F}{\partial y^2}(\boldsymbol{a})(y - a_y)^2\right] \\ &= \tfrac{1}{2!}\textstyle\sum_{i,j} \tfrac{\partial^2 F}{\partial x_i \partial x_j}(\boldsymbol{a})(x_i - a_i)(x_j - a_j) \end{aligned}$$

en acuerdo con (1). Hemos asumido $\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x}$ en $\boldsymbol{r} = \boldsymbol{a}$, propiedad válida cuando ambas derivadas cruzadas existen y son continuas.

En las condiciones anteriores ($\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i}\ \forall\, i, j$) la matriz Hessiana $H(\boldsymbol{a})$ es real y **simétrica**, por lo que puede ser diagonalizada mediante una matriz **ortogonal** R formada por autovectores $\boldsymbol{v}_i$ ortogonales y **normalizados** de $H(\boldsymbol{a})$: $R = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, con

$$\begin{aligned} H(\boldsymbol{a})\boldsymbol{v}_i &= \lambda_i \boldsymbol{v}_i, \quad i = 1, \ldots, n \\ \boldsymbol{v}_i \cdot \boldsymbol{v}_j &= \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \end{aligned}$$

De esta forma, R satisface

$$R^{-1} = R^T, \qquad R^T H(\boldsymbol{a}) R = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \tag{5}$$

donde $\lambda_1, \dots, \lambda_n$ son los autovalores de $H(\boldsymbol{a})$. Reemplazando ahora $\boldsymbol{r} - \boldsymbol{a} = R(\boldsymbol{r'} - \boldsymbol{a'})$, $(\boldsymbol{r} - \boldsymbol{a})^T = (R(\boldsymbol{r'} - \boldsymbol{a'}))^T = (\boldsymbol{r'} - \boldsymbol{a'})^T R^T$, se obtiene la forma **diagonal**

$$\begin{aligned} \frac{1}{2}(\boldsymbol{r} - \boldsymbol{a})^T H(\boldsymbol{a})(\boldsymbol{r} - \boldsymbol{a}) &= \frac{1}{2}(\boldsymbol{r'} - \boldsymbol{a'})^T (R^T H(\boldsymbol{a}) R)(\boldsymbol{r'} - \boldsymbol{a'}) \\ &= \frac{1}{2}\sum_{i=1}^{n} \lambda_i \, (x'_i - a'_i)^2 \end{aligned} \tag{6}$$

donde $\boldsymbol{r'} = R^T \boldsymbol{r}$ y $\boldsymbol{a'} = R^T \boldsymbol{a}$ denotan las coordenadas de $\boldsymbol{r}$ y $\boldsymbol{a}$ en la base de autovectores:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = R^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{7}$$

Se puede siempre elegir el signo de los autovectores tal que $\text{Det}(R) = +1$, en cuyo caso la matriz R representa una **rotación** y $(x'_1, \dots, x'_n)$ serán las coordenadas respecto de un sistema de ejes ortogonales rotado respecto del original. En estas coordenadas, la matriz Hessiana es entonces **diagonal**:

$$\frac{\partial^2 F}{\partial x'_i \partial x'_j}(\boldsymbol{a}) = (R^T H(\boldsymbol{a}) R)_{ij} = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$$

como puede verificarse utilizando la forma diagonal (6) o aplicando (7) y regla de la cadena. Hemos así probado que **existen siempre n ejes ortogonales para los que la matriz Hessiana $H(\boldsymbol{a})$ respecto de las coordenadas asociadas resulta diagonal**. Estos ejes se denominan usualmente ejes principales (de la forma cuadrática $\boldsymbol{r}^T H(\boldsymbol{a}) \boldsymbol{r}$).

La forma diagonal (7) permite ver que si $\lambda_i > 0 \; \forall \; i$, F será cóncava "hacia arriba" en $\boldsymbol{a}$ en todas las direcciones, como en el caso en que $\boldsymbol{a}$ es un mínimo local, mientras que si $\lambda_i < 0 \; \forall \; i$, F será cóncava "hacia abajo" en $\boldsymbol{a}$ en todas las direcciones, como en el caso en que $\boldsymbol{a}$ es un máximo local. En general, F será cóncava hacia arriba en las variables x'_i asociadas a autovalores $\lambda_i > 0$, y cóncava hacia abajo en las variables x'_i asociadas a autovalores $\lambda_i < 0$. Si algun autovalor λ_i es nulo, se deben utilizar derivadas de orden superior u otros métodos para determinar la concavidad respecto de x'_i.

Ejercicio: Considere la función

$$F(x, y) = e^{-x^2 - y^2}(1 - x^2 - y^2 - 2\alpha x y)$$

1) Muestre que para $\boldsymbol{a} = \boldsymbol{0}$,

$$F(\boldsymbol{0}) = 1, \quad \boldsymbol{\nabla} F(\boldsymbol{0}) = \boldsymbol{0}, \quad H(\boldsymbol{0}) = \begin{pmatrix} -4 & -2\alpha \\ -2\alpha & -4 \end{pmatrix}$$

2) Muestre que $H(\mathbf{0})$ tiene autovalores

$$\lambda_1 = -2(2+\alpha), \quad \lambda_2 = -2(2-\alpha)$$

con autovectores normalizados $\boldsymbol{v}_1 = \binom{1}{1}/\sqrt{2}$, $\boldsymbol{v}_2 = \binom{-1}{1}/\sqrt{2}$.

3) Pruebe que si

$$R = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

entonces

$$R^T R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R^T H(\mathbf{0}) R = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

4) A partir de los resultados anteriores, muestre que el polinomio de Taylor de F de grado 2 alrededor de $\boldsymbol{a} = \mathbf{0}$ es

$$\begin{aligned} P_2(x,y) &= 1 - 2(x^2 + y^2 + \alpha x y) \\ &= 1 + \tfrac{1}{2}(x,y)\left(\begin{smallmatrix} -4 & -2\alpha \\ -2\alpha & -4 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = 1 + \tfrac{1}{2}(x', y')\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right) \\ &= 1 - (2+\alpha){x'}^2 - (2-\alpha){y'}^2 \end{aligned} \tag{8}$$

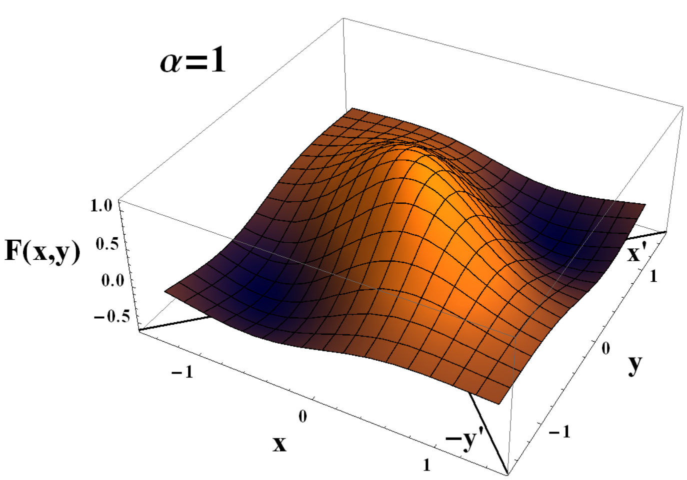
donde

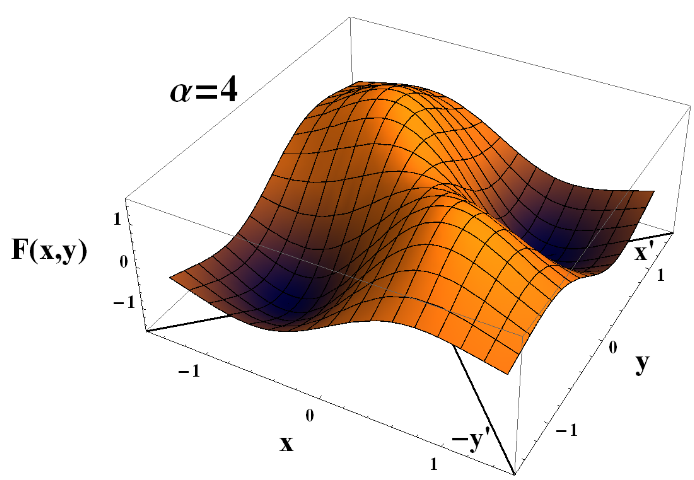
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R^T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ -x + y \end{pmatrix} / \sqrt{2}$$

5) Muestre que R representa una rotación de 45^o en sentido antihorario, y que por lo tanto, x', y', son las coordenadas respecto de ejes rotados 45^o en sentido antihorario respecto de los ejes originales.

6) A partir de la forma diagonal (8) y la anulación del gradiente, concluya que $\mathbf{0} = (0,0)$ es un **máximo local** de F si $|\alpha| < 2$ y un **punto silla** de F si $|\alpha| > 2$. En particular, si $\alpha > 2$, F es en el origen cóncava hacia abajo en x' y hacia arriba en y'. ¿Que sucede en el caso $\alpha = 2$?

7) Analice de la misma forma los otros puntos críticos de F. La figura muestra el gráfico de F cerca del origen para dos valores de α.





3. Autovalores de operadores lineales

Un operador lineal o endomorfismo en un espacio vectorial V es una transformación lineal $L : V \to V$. Se dice que un vector no nulo $\boldsymbol{v} \in V$ es *autovector* de L si existe un escalar λ, denominado *autovalor*, tal que

$$L(\boldsymbol{v}) = \lambda \boldsymbol{v}$$

Está definición es general, siendo independiente de la representación de L y válida tanto para espacios de dimensión finita o infinita.

Si $\boldsymbol{v}$ es autovector con autovalor λ, $\alpha\boldsymbol{v}$ es también autovector de L con el mismo autovalor $\lambda \; \forall \; \alpha \neq 0$, debido a que L es lineal:

$$L(\alpha\boldsymbol{v}) = \alpha L(\boldsymbol{v}) = \alpha\lambda\boldsymbol{v} = \lambda(\alpha\boldsymbol{v})$$

Y si $\boldsymbol{v}$ y $\boldsymbol{w}$ son autovectores con el mismo autovalor λ, $\boldsymbol{v} + \boldsymbol{w}$ también satisface (1):

$$L(\boldsymbol{v} + \boldsymbol{w}) = L(\boldsymbol{v}) + L(\boldsymbol{w}) = \lambda\boldsymbol{v} + \lambda\boldsymbol{w} = \lambda(\boldsymbol{v} + \boldsymbol{w})$$

Por lo tanto, el conjunto de todos los autovectores con autovalor λ, junto con el vector nulo $\mathbf{0}$ (que satisface $L(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0} \; \forall \; \lambda$) forma un subespacio de V, denominado espacio propio o autoespacio asociado al autovalor λ:

$$S_\lambda = \{\boldsymbol{v} \in V | L(\boldsymbol{v}) = \lambda\boldsymbol{v}\}$$

Por lo tanto, $S_\lambda = \{\boldsymbol{v} \in V | L(\boldsymbol{v}) - \lambda\boldsymbol{v} = \mathbf{0}\} = \{\boldsymbol{v} \in V | (L - \lambda I)\boldsymbol{v} = \mathbf{0}\} = \mathrm{Nu}[L - \lambda I]$, donde I denota el operador identidad y $\mathrm{Nu}[L - \lambda I]$ el núcleo del operador lineal $L - \lambda I$.

En espacios V de dimensión finita n, L queda completamente determinado por los vectores que asigna a los elementos de una base (ordenada) $B = (\boldsymbol{b}_1 \ldots, \boldsymbol{b}_n)$ de V. Si

$$\begin{aligned} L(\boldsymbol{b}_i) &= m_{1i}\boldsymbol{b}_1 + \ldots + m_{ni}\boldsymbol{b}_n \quad i = 1, \ldots, n\,, \\ \boldsymbol{v} &= x_1\boldsymbol{b}_1 + \ldots + x_n\boldsymbol{b}_n \end{aligned}$$

la ecuación (1) resulta equivalente a

$$M_L X = \lambda X$$

donde

$$M_L = \begin{pmatrix} m_{11} & \ldots & m_{1n} \\ & \ddots & \\ m_{n1} & \ldots & m_{nn} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

son, respectivamente, la matriz que representa a L en esta base y el vector columna de coordenadas de $\boldsymbol{v}$ en esta base. Esto implica que λ es autovalor de L si y sólo si λ es autovalor de la matriz representativa M_L, es decir, si y sólo si $\mathrm{Det}[M_L - \lambda I_n] = 0$.

Esto es válido para *cualquier* base elegida, es decir, para cualquier representación matricial M_L de L, por lo que todas las matrices que representan a L en alguna base deben tener los mismos autovalores. Podemos comprobar esto recordando que si M_L y M'_L representan a L en bases B y B' respectivamente, entonces son matrices *semejantes*:

$$M'_L = S^{-1} M_L S$$

donde S es la matriz no singular de cambio de base, cuyas columnas son las coordenadas de los vectores de la base B' en la base original B:

$$\boldsymbol{b'_i} = s_{1i}\boldsymbol{b}_1 + \ldots + s_{ni}\boldsymbol{b}_n, \quad i = 1, \ldots, n \;\Rightarrow\; S = \begin{pmatrix} s_{11} & \ldots & s_{1n} \\ & \ldots & \\ s_{n1} & \ldots & s_{nn} \end{pmatrix}$$

Por lo tanto, los polinomios característicos asociados a M_L y M'_L *son los mismos*:

$$\mathrm{Det}[M'_L - \lambda I_n] = \mathrm{Det}[S^{-1} M_L S - \lambda I_n] = \mathrm{Det}[S^{-1}(M_L - \lambda I_n) S] = \mathrm{Det}[M_L - \lambda I_n]$$

ya que $\mathrm{Det}[S^{-1}AS] = \mathrm{Det}(S)\mathrm{Det}(A)\mathrm{Det}(S) = \mathrm{Det}(A)$. Los autovalores de M_L y M'_L serán entonces **idénticos**.

Recordemos también que las coordenadas de $\boldsymbol{v}$ en estas bases se relacionan por

$$X' = S^{-1} X$$

por lo que si $M_L X = \lambda X \Rightarrow M'_L X' = S^{-1} M_L S S^{-1} X = S^{-1} M_L X = \lambda S^{-1} X = \lambda X'$. Esto muestra que si X es autovector de M_L con autovalor λ, $X' = S^{-1} X$ es autovector de M'_L con el mismo autovalor λ.

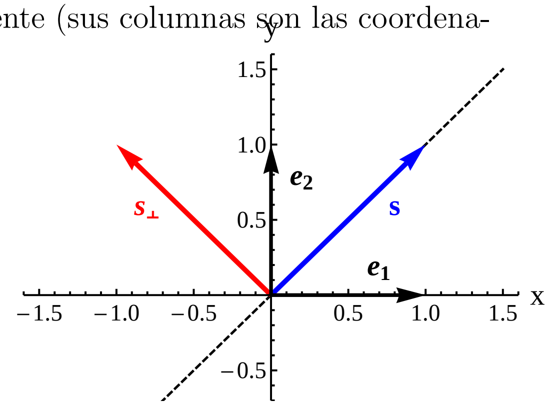
L es diagonalizable si existe una base $B^d = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ de V formada por autovectores de L:

$$L(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i, \quad i = 1, \ldots, n$$

La matriz representativa M_L^d de L en esta base es entonces diagonal:

$$M_L^d = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots \\ & & \ddots & \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} = S^{-1} M_L S$$

donde S es la matriz de cambio de base correspondiente (sus columnas son las coordenadas de los autovectores $\boldsymbol{v}_i$ en la base original).



Ejemplo: Sea $P_{\boldsymbol{s}} : \mathbb{R}^2 \to \mathbb{R}^2$ el proyector ortogonal sobre la recta generada por un vector $\boldsymbol{s} = \boldsymbol{e}_1 + \boldsymbol{e}_2$, donde $B = (\boldsymbol{e}_1, \boldsymbol{e}_2)$ es la base canónica. Tenemos

$$P_{\boldsymbol{s}}(\boldsymbol{v}) = \frac{\boldsymbol{s} \cdot \boldsymbol{v}}{\boldsymbol{s} \cdot \boldsymbol{s}}\, \boldsymbol{s}$$

$P_{\boldsymbol{s}}$ es lineal, y por su definición cumple

$$P_{\boldsymbol{s}}(\boldsymbol{s}) = \boldsymbol{s} = 1\, \boldsymbol{s}, \quad P_{\boldsymbol{s}}(\boldsymbol{s}_\perp) = \boldsymbol{0} = 0\, \boldsymbol{s}_\perp$$

donde $\boldsymbol{s}_\perp = -\boldsymbol{e}_1 + \boldsymbol{e}_2$ es un vector ortogonal a $\boldsymbol{s}$ ($\boldsymbol{s} \cdot \boldsymbol{s}_\perp = 0$). Por lo tanto,

1) $\boldsymbol{s}$ es autovector de $P_{\boldsymbol{s}}$ con autovalor $\lambda_1 = 1$
2) $\boldsymbol{s}_\perp$ es autovector de $P_{\boldsymbol{s}}$ con autovalor $\lambda_2 = 0$
Estos resultados fueron obtenidos de la definición de $P_{\boldsymbol{s}}$, sin utilizar ninguna representación matricial determinada. En la base $B^d = (\boldsymbol{s}, \boldsymbol{s}_\perp)$ de $\mathbb{R}^2$, la matriz representativa de $P_{\boldsymbol{s}}$ será entonces **diagonal**:

$$M_{P_s}^d = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Por otro lado, en la base canónica se obtiene

$$P_{\boldsymbol{s}}(\boldsymbol{e}_1) = P_{\boldsymbol{s}}(\boldsymbol{e}_2) = \tfrac{1}{2}\boldsymbol{s} = \tfrac{1}{2}(\boldsymbol{e}_1 + \boldsymbol{e}_2)$$

por lo que la matriz que representa a $P_{\boldsymbol{s}}$ en esta base es

$$M_{P_s} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Se verifica (probar!) que sus autovalores son $\lambda_1 = 1$ y $\lambda_2 = 0$, con espacios propios generados por los autovectores $X_1 = \binom{1}{1}$ y $X_2 = \binom{-1}{1}$ respectivamente:

$$M_{P_s}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad M_{P_s}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

siendo X_1 y X_2 las coordenadas de $\boldsymbol{s}$ y $\boldsymbol{s}_\perp$ en la base canónica. Se comprueba que

$$M_{P_s}^d = S^{-1}M_{P_s}S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{con} \quad S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Ejercicios:

1. Hallar las matrices que representan a la reflexión $R : \mathbb{R}^2 \to \mathbb{R}^2$ respecto de la recta $y = x$ en i) la base canónica, ii) la base $B' = ((1,1),(-1,1))$ y iii) la base $B'' = ((1,0),(1,1))$. Mostrar luego que poseen los mismos autovalores, y que B' es una base de autovectores de R.

2. Verificar que en el espacio vectorial V de las funciones $f : [0, L] \to \mathbb{R}$ derivables a segundo orden que satisfacen $f(0) = f(L) = 0$, la transformación lineal L definida por

$$L(f) = -\frac{d^2 f}{dx^2}$$

posee autovectores (denominados en este contexto autofunciones)

$$f_n(x) = \text{sen}(\frac{n\pi x}{L}), \quad n = 1, 2, 3, 4\ldots$$

con autovalores

$$\lambda_n = \frac{n^2\pi^2}{L^2}, \quad n = 1, 2, 3, 4\ldots$$

Estas son las únicas autofunciones de L que se anulan en los bordes del intervalo. Este conjunto de autofunciones es completo: toda función $f(x)$ perteneciente a este espacio puede escribirse como una serie de estas autofunciones: $f(x) = \sum_{n=1}^{\infty} c_n \text{sen}(\frac{n\pi x}{L})$ (serie de medio rango). En mecánica cuántica, estos autovalores λ_n determinan las energías posibles de una partícula confinada en un intervalo de longitud L, siendo $f_n(x)$ las funciones de onda asociadas.

Apéndice

La descomposición en valores singulares (DVS)

Referencias

[1] R.L. Burden, J.Douglas Faires, *Análisis Numérico* , Grupo Editorial Iberoamericana, México, 1985.

[2] J. Demmel, *Applied Numerical Linear Algebra*, SIAM,Philadelphia, 1997.

Hemos visto previamente que toda matriz simétrica A se puede descomponer $A = PDP^T$, siendo P una matriz ortogonal ($P^T = P^{-1}$), D una matriz diagonal que contiene los autovalores de A. Cuando A no es simétrica, pero si cuadrada, si A es diagonizable existe una descomposición de $A = SDS^{-1}$, siendo S no singular aunque no necesariamente ortogonal. Pero, cualquier matriz no es diagonizable ! !.

.... Ahora veremos que toda matriz (cuadrada o no, simétrica o no) tiene una factorización de la forma:

$$A = PDQ^T,$$

donde P y Q son matrices ortogonales y D una matriz diagonal. Este resultado se llama " descomposición en valores singulares"(DVS), y es una de las mas importantes entre las descomposiciones de matrices.

... Explicaremos como obtenerla y haremos consideraciones sobre sus aplicaciones.

1 Los valores singulares de una matriz

Para cualquier matriz A de $m \times n$, la matriz A^TA de $n \times n$ es simétrica y por tanto puede ser diagonalizada ortogonalmente. Los autovalores de A^TA son reales y no negativos (≥ 0). Eso surge ya que para cualquier autovector u, se tiene que

$$A^TAu = \lambda u$$

si se multiplica a ambos lados por u, se obtiene la igualdad

$$u^TA^TAu = \lambda u^Tu$$

que indica que $\|Au\|^2 = \lambda\|u\|^2$, por lo tanto $\lambda \geq 0$.

... Por tanto tiene sentido tomar las $\sqrt{\lambda_i}$, si λ_i, $i = 1, \ldots, n$, son los autovalores de A^TA.

Definición *Si A es una matriz $m\times n$, los valores singulares de A son las raíces cuadradas de los autovalores de A^TA, y se denotan mediante $\sigma_1, \ldots, \sigma_n$. Es convencional acomodar los valores singulares de modo que $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.*

Ejemplo 1: Encontrar los valores singulares de A:

Ejemplo

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

La matriz A^TA

$$A^TA = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

tiene autovalores : $\lambda_1 = 3$, $\lambda_2 = 1$. En consecuencia los valores singulares de A son : $\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{1}$.

Para comprender el significado lde los valores singulares de A consideremos los autovectores de A^TA. Por ser simétrica sabemos que hay n autovectores ortogonales, que se pueden considerar de longitud 1. Así sean $v_1, v_2, \ldots, v_n$ esa base de autovectores ordenados y correspondientes a los autovalores $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$. Estos autovectores satisfacen:

$$A^TAv_i = \lambda_i v_i,$$

o equivalentemente

$$\|Av_i\|^2 = \lambda_i,$$

por consiguiente

$$\|Av_i\| = \sqrt{\lambda_i}.$$

Así los valores singulares de A son las longitudes de los vectores $Av_1, \ldots, Av_n$. Geometricamente esto tiene una importante interpretación. Si consideramos el ejemplo 1, y si consideramos $x \in \{x : \|x\|^2 = 1\}$, entonces $\|Ax\|^2 = Ax \cdot Ax = x^TA^TAx =$

$$\mathbf{x^TA^TAx} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1^2 + 2x_1x_2 + 2x_2^2$$

lo cual es una forma cuadrática.

... Es facil ver que los valores mínimos y máximos que toma una forma cuadrática en los vectores x con $\|x\| = 1$, es

$$\lambda_{min} \leq x^TA^TAx \leq \lambda_{max}$$

si λ_{min} es el menor autovalor de A^TA, y si λ_{max} es el mayor autovalor de esa matriz. En el caso del ejemplo de arriba, el autovector correspondiente a $\lambda_{min} = 1$ es $\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$, y el autovector de $\lambda_{max} = 3$ es $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

Por tanto, $\|Av_i\|^2 = \lambda_i$, se tiene que $\sigma_1 = \sqrt{3} = \|Av_1\|$, $\sigma_2 = \sqrt{1} = \|Av_2\|$, son los valores máximo y mínimo de las longitudes $\|Ax\|$, cuando x recorre la circunferencia unitaria dada por : $\|x\| = 1$.

La transformación lineal T con matriz A, $T : \Re^2 - -- > \Re^3$ transforma el círculo unidad $\|x\| = 1$ en una "elipse" que está sobre el plano de ecuación: $x - y - z = 0$ (verificar que la imagen de la transformación es ese plano en $\Re^3$). Las longitudes σ_1 , y σ_2 son las longitudes de los semiejes mayor y menor respectivamente de esa elipse.

2 Descomposición de valor singular

Queremos demostrar que una matriz A de $m \times n$ se puede factorizar como

$$A = U\Sigma V^T$$

donde U es una matriz ortogonal de $m \times m$, V es una matriz ortogonal de $n \times n$, y Σ una matriz "diagonal(seudo)" de $m \times n$. Si los valores singulares NO NULOS de A son

$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$, y si $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$, entonces Σ de $m \times n$ tendrá una forma en bloques (un bloque D de $r \times r$, a su derecha una matriz de ceros $\mathcal{O}$ de $r \times n - r$, abajo a la izquierda otra matriz de ceros $\mathcal{O}$ de $m - r \times r$, y el último bloque en la diagonal de ceros $\mathcal{O}$ de $m - r \times n - r$:

$$\Sigma = \begin{pmatrix} D & \mathcal{O} \\ \mathcal{O} & \mathcal{O} \end{pmatrix}$$

donde

$$D = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{pmatrix}$$

Si r coincide con n o con m alguna de esas matrices nulas $\mathcal{O}$ no aparecerán. En el caso del ejemplo 1, previo,la matriz Σ tiene la forma:

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{1} \\ 0 & 0 \end{pmatrix}$$

.... Ahora, cómo hallar los factores U y V^T?....

Para hallar la matriz ortogonal V primero determinamos una base ortogonal $\{v_1, \ldots, v_n\}$ de vectores de $\Re^n$ compuesta por autovectores de la matriz A^TA de $n \times n$.

$$V = [v_1, v_2, \ldots, v_n]$$

es una matriz ortogonal de $n \times n$.

Con respecto a la determinación de la matriz U de $m \times m$, primero observamos que los vectores de $\Re^m$:

$$Av_1, Av_2, \ldots, Av_n$$

es un conjunto de n vectores ortogonales. Eso se obtiene de observar que para dos autovectores de A^TA, v_i y v_j con $i \neq j$, por ser ortogonales se cumple que

$$v_j^T A^T A v_i = v_j^T . \lambda_i v_i = 0$$

Así Av_i es ortogonal a Av_j.

Ahora recordemos que los valores singulares $\sigma_i = \|Av_i\|$, y que los primeros r son no nulos. Por tanto, podemos normalizar los Av_i, con $i = 1, \ldots, r$, considerando:

$$u_i = \frac{Av_i}{\sigma_i}$$

para cada $i = 1, \ldots, r$. Eso garantiza que $[u_1, \ldots, u_r]$ son ortonormales en $\Re^m$. Si $r < m$ ese conjunto no es una base de $\Re^m$. En ese caso hay que extender ese conjunto para tener m vectores ortonormales de $\Re^m$.

... Esa es la parte más dificil de la factorización que queremos obtener: $A = U\Sigma V^T$.

La matriz U estará formada por las columnas ortonormales $U = [u_1, \dots, u_r, u_{r+1}, ...u_m]$.
...Veamos que con tales matrices V, U, y Σ, se verifica que

$$A = U\Sigma V^T$$

Como $V^T = V^{-1}$ por ser una matriz ortogonal, verificar que vale $A = U\Sigma V^T$, es igual a ver que (multiplicando por V):

$$AV = U\Sigma$$

Sabemos que $Av_i = \sigma_i u_i$, para $i = 1, 2, \dots, r$, y que

$$\|Av_i\| = \sigma_i = 0, \text{para} \;\; i = r+1, \dots, n.$$

Por tanto,

$$Av_i = 0, \;\text{para} \;\; i = r+1, \dots, n.$$

Por consiguiente,

$$AV = A[v_1 \dots v_n] = [Av_1 \dots Av_n]$$

entonces

$$AV = [Av_1 \dots Av_r, 0 \dots 0] = [\sigma_1 u_1 \dots \sigma_r u_r, 0 \dots 0]$$

$$\mathbf{AV} = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \begin{pmatrix} \sigma_1 & \dots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \dots & \sigma_r & \\ & O & & O \end{pmatrix} = U\Sigma$$

como se requiere para verificar que $A = U\Sigma V^T$.

Observación *Los vectores de las columnas de U se denominan vectores singulares por la izquierda de A, mientras los vectores de las columnas de V se llaman vectores singulares por la derecha de A. Las matrices U y V no estan determinadas en forma única por A, en cambio Σ si, porque contiene los valores singulares de A.*

Ejemplo *Encontrar una descomposición de valor singular de las matrices*

(i)$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *(ii)*$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

Solución: (i) Consideramos $A^TA = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *y hallamos los autovalores:* $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$, *con sus autovectores:*

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Estos vectores son ortogonales, de manera que los normalizamos para obtener:

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \; v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \; v_3 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$$

Los valores singulares de A son $\sigma_1 = \sqrt{2}$,$\sigma_2 = \sqrt{1} = 1$,$\sigma_3 = \sqrt{0} = 0$.

Así

$$V = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \; y \; \Sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Para determinar U, *calculamos*

$$u_1 = \frac{Av_1}{\sigma_1} = 1/\sqrt{2} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

y

$$u_2 = \frac{Av_2}{\sigma_2} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Estos vectores ya forman una base ortonormal de $\Re^2$, *de manera que tenemos la matriz* U: $U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. *Esto produce la descomposición DVS de la matriz A:*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = U\Sigma V^T$$

Advierta que $\sigma_3 = 0$ *no aparece en* Σ, *y que la matriz que se necesita es la matriz* V^T *en lugar de la* V.

En el ejemplo(ii) se necesitará completar la matriz U(observar que en el ejemplo previo eso no fue necesario).

La matriz corresponde al Ejemplo 1: $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ Como antes consideramos $A^T A$, y sus autovalores y autovectores. Sabemos que : $\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$, y los correspondientes autovectores de $A^T A$ son ....

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \; v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Así

$$V = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \text{ y } \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Para U, calculamos

$$u_1 = \frac{Av_1}{\sigma_1} = 1/\sqrt{3}\begin{pmatrix}1 & 1\\ 1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix} = \begin{pmatrix}2/\sqrt{6}\\ 1/\sqrt{6}\\ 1/\sqrt{6}\end{pmatrix}$$

y

$$u_2 = \frac{Av_2}{\sigma_2} = \begin{pmatrix}1 & 1\\ 1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix} = \begin{pmatrix}0\\ -1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}$$

Esta vez necesitamos extender $\{u_1, u_2\}$ a una base ortonormal de $\Re^3$. Para eso, usamos el procedimiento de Gram-Schmidt. Necesitamos tener un vector de $\Re^3$ que sea linealmente independiente con respecto a los vectores $\{u_1, u_2\}$. Por ejemplo, $e_3 = (0, 0, 1)$ es tal que $\{u_1, u_2, e_3\}$ forman una base de $\Re^3$. Así usando el procedimiento de Gram-Schmidt sólo el tercer paso es necesario para obtener un vectoru_3 ortogonal a ambos $\{u_1, u_2\}$, se encuentra que este vector$u_3 = \begin{pmatrix}-1/\sqrt{3}\\ 1/\sqrt{3}\\ 1/\sqrt{3}\end{pmatrix}$.

...Otra forma de expresar la DVS de una matriz A.

Para eso consideramos

$$\mathbf{A} = U\Sigma V^T = \begin{pmatrix}u_1 & u_2 & \ldots & u_m\end{pmatrix}\begin{pmatrix}\sigma_1 & \ldots & 0 & \\ \vdots & \ddots & \vdots & O\\ 0 & \ldots & \sigma_r & \\ & O & & O\end{pmatrix}\begin{pmatrix}v_1^T\\ v_2^T\\ \vdots\\ v_n^T\end{pmatrix}$$

que es igual a

$$= \begin{pmatrix}u_1 & \ldots & u_r & |u_{r+1} & \ldots & u_m\end{pmatrix}\begin{pmatrix}\sigma_1 & \ldots & 0 & \\ \vdots & \ddots & \vdots & O\\ 0 & \ldots & \sigma_r & \\ & O & & O\end{pmatrix}\begin{pmatrix}v_1^T\\ \vdots\\ v_r^T\\ v_{r+1}^T\\ \vdots\\ v_n^T\end{pmatrix}$$

obteniéndose ...

$$= \begin{pmatrix} u_1 & u_2 & \dots & u_r \end{pmatrix} \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_r \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{pmatrix} + \begin{pmatrix} u_{r+1} & \dots & u_m \end{pmatrix} (O) \begin{pmatrix} v_{r+1}^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & u_2 & \dots & u_r \end{pmatrix} \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_r \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 u_1 & \dots & \sigma_r u_r \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_r^T \end{pmatrix} = \sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T$$

Se ha justificado que...

Lema *Sea A una matriz $m \times n$con valores singulares $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ y $\sigma_{r+1} = \dots = \sigma_n = 0$. Sean $u_1, \dots, u_r$ vectores singulares por la izquierda, y sean $v_1, \dots, v_r$ vectores singulares por la derecha de A correspondientes aesos valores singulares. entonces*

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T$$

Observación *La DVS de una matriz A da mucha información acerca de A como se resalta en el siguiente teorema:*

Teorema *Sea $A = U\Sigma V^T$ una descomposición de valor singular de una matriz A de $m \times n$. Sean $\sigma_1, \dots, \sigma_r$ todos los valores singulares NO NULOS de A. Entonces*

(i) el rango de A es r.

(ii)$\{u_1, u_2, \dots, u_r\}$ es una base ortonormal de $R(A)$

(iii)$\{u_{r+1}, u_{r+2}, \dots, u_m\}$ es una base ortonormal de $N(A^T)$

(iv)$\{v_1, v_2, \dots, v_r\}$ es una base ortonormal de $R(A^T)$

(v)$\{v_{r+1}, v_{r+2}, \dots, v_n\}$ es una base ortonormal de $N(A)$

Demostración.

(i)Como U y V^T son matrices no singulares(además, ortogonales) se sabe que $rango(A) = rango(U\Sigma V^T)$ coincide con $= rango(\Sigma V^T) = rango(\Sigma) = r$.
(ii) Como Av_i, con $i = 1, \dots, r$ son linealmente independientes (ortogonales), y como $u_i = (1/\sigma_i)Av_i$, $i = 1, \dots, r$, entonces $\{u_1, \dots, u_r$ forman una base ortonormal del $rango(A)$.
(iii) Como $\{u_1, \dots, u_r, \dots, u_m$ forman una base ortonormal de $\Re^m$, luego por (ii), y considerando que $\{u_{r+1}, \dots, u_m$ es una base del espacio complementario a $R(A)$, se obtiene que ese conjunto es una base ortonormal del subespacio $N(A^T)$ o del $R(A)^\perp$.
(v) Como $Av_{r+1} = Av_{r+2} = \dots Av_n = 0$, se tiene que $\{v_{r+1}, \dots, v_n$, que es un conjunto

ortonormal de vectores contenido en el anulador de A, de dimensión $n-r$, por lo que es una base ortonormal del $N(A)$.
(iv) Esta propiedad se desprende de considerar (v) y que $\{v_1, v_2, \ldots, v_r\}$ es un conjunto ortonormal complementario del de (v), por lo que es una base de $R(A^T)$.

Otro resultado importante que se obtiene desde la descomposición DVS ....

Lema *Sea $A = U\Sigma V^T$ la descomposición de A de $m \times n$ con rango r. Por consiguiente, la imagen de la esfera unitaria en $\Re^n$, bajo la transformación matricial que aplica $x \in \Re^n$ en $Ax \in \Re^m$, es*

(a)la superficie de un elipsoide en $\Re^m$ si $r = n$,

(b) un elipsoide sólido en $\Re^m$ si $r < n$.

Demostración: Sean $u_1, \ldots, u_m$ y $v_1, v_2, \ldots, v_n$ los vectores singulares por la izquierda y por la derecha de A, respectivamente. En razón que $rango(A) = r$, los valores singulares de A satisfacen $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r > 0$, y $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$.

Sea $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ un vector unitario de $\Re^n$. Ahora, como V es una matriz ortogonal, también lo es V^T, por tanto V^Tx es un vector unitario(conserva longitudes), así

$$V^Tx = \begin{pmatrix} v_1^Tx \\ v_2^Tx \\ \vdots \\ v_n^Tx \end{pmatrix}$$

de manera que $(v_1^Tx)^2 + (v_2^Tx)^2 + \ldots + (v_n^Tx)^2 = 1$.

Como $A = \sigma_1 u_1 v_1^T + \ldots + \sigma_r u_r v_r^T$. Por tanto,

$$Ax = \sigma_1 u_1 v_1^T x + \ldots + \sigma_r u_r v_r^T x = (\sigma_1 v_1^T x)u_1 + \ldots + (\sigma_r v_r^T x)u_r$$

$$Ax = (\sigma_1 v_1^T x)u_1 + \ldots + (\sigma_r v_r^T x)u_r = y_1u_1 + \ldots + y_ru_r$$

donde denotamos con $y_i = (\sigma_i v_i^T x)$.

(a) Si $r = n$, entonces corresponde al caso $n \leq m$, y

$$Ax = y_1u_1 + \ldots + y_nu_n = Uy$$

donde $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.Por consiguiente, como U es ortogonal, $\|Ax\| = \|Uy\| = \|y\|$. Como

$$(\frac{y_1}{\sigma_1})^2 + \ldots + (\frac{y_n}{\sigma_n})^2 = (v_1^Tx)^2 + \ldots + (v_n^Tx)^2 = 1$$

lo que muestra que los vectores Ax forman la superficie de un elipsoide en $\Re^m$.

(b) Si $r < n$, la única diferencia en los pasos anteriores es que la ecuación se convierte en

$$(\frac{y_1}{\sigma_1})^2 + \ldots + (\frac{y_r}{\sigma_r})^2 \leq 1$$

puesto que despreciamos algunos términos. Esto corresponde a un elipsoide sólido de $\Re^m$.

Facultad de Ingeniería
UNLP

Matemática C

IX Ecuaciones Diferenciales Lineales Ordinarias

Temario por clase:

Clase 1: Ecuaciones diferenciales lineales ordinarias de orden n. Generalidades. Propiedades fundamentales. Caso homogéneo. La ecuación lineal homogénea de segundo orden. El caso de coeficientes constantes. Aplicaciones.

Clase 2: El caso no homogéneo. Métodos de resolución generales y para el caso de coeficientes constantes. Aplicaciones.

Clase 3: Sistemas de ecuaciones diferenciales lineales ordinarias. Generalidades. Propiedades fundamentales. El caso homogéneo. El caso homogéneo de coeficientes constantes. Caso diagonalizable.

Clase 4: El caso no diagonalizable. Matriz fundamental. Sistemas lineales no homogéneos. Métodos de resolución generales y para el caso de coeficientes constantes.

Clase 5: Aplicaciones. Problemas.

Bibliografía:

1. E. Kreyszig, *Matemáticas Avanzadas para Ingeniería (2do.vol)*, Limusa.
2. R.K. Nagle, E.B. Saff, *Fundamentos de Ecuaciones Diferenciales*, Addison Wesley Iberoamericana.
3. M. R. Simmons, *Ecuaciones Diferenciales con aplicaciones* , McGrawHill.
4. Dennis G. Zill , *Ecuaciones Diferenciales con Aplicaciones*, Grupo Editorial Iberoamericana, México.

1. Ecuaciones diferenciales lineales ordinarias

1.1. Introducción

Estudiaremos en este capítulo ecuaciones diferenciales **lineales.** Recordemos primero que una ecuación diferencial es una ecuación donde la incógnita es una **función**, y donde esta aparece vinculada con sus **derivadas**. Si la función incógnita depende de una sóla variable, que llamaremos t, la ecuación diferencial se denomina **ordinaria**. Si la misma contiene derivadas hasta de orden n de la función incógnita, se dice que la ecuación diferencial es **de orden n**.

Una ecuación diferencial ordinaria de orden n puede escribirse en la forma

$$y^{(n)} = f(t, y, y', y'', \ldots, y^{(n-1)}) \tag{1.1}$$

donde $y(t)$ es la función incógnita y $y'(t) = \frac{dy}{dt}$, $y''(t) = \frac{d^2y}{dt^2}$, . . ., $y^{(n)} = \frac{d^ny}{dt^n}$, sus derivadas.

Las ecuaciones diferenciales juegan un rol fundamental en Ingeniería, Física y muchas otras áreas de la ciencia y tecnología. La razón es que en general, no resulta fácil hallar leyes que vinculen directamente las magnitudes que caracterizan un cierto fenómeno o sistema, pero sí resulta posible en muchos casos relacionar esas magnitudes con sus derivadas, lo que origina una ecuación diferencial o en general, un sistema de ecuaciones diferenciales. Por ejemplo, esto ocurre cuando la evolución de un sistema depende de su estado.

Un ejemplo típico es la segunda ley de Newton para el movimiento de una partícula,

$$\boldsymbol{F} = m\boldsymbol{a} \tag{1.2}$$

donde $\boldsymbol{F}$ es la fuerza neta que actúa sobre la partícula, m la masa de la misma y $\boldsymbol{a}$ su aceleración. Si $\boldsymbol{r}(t)$ denota el vector posición de la partícula, entonces $\boldsymbol{a} = \frac{d^2\boldsymbol{r}}{dt^2}$, y si la fuerza $\boldsymbol{F}$ depende de $\boldsymbol{r}$, $d\boldsymbol{r}/dt$ y t, esta ley conduce a la ecuación diferencial

$$m\frac{d^2\boldsymbol{r}}{dt^2} = \boldsymbol{F}(t, \boldsymbol{r}, \frac{d\boldsymbol{r}}{dt}) \tag{1.3}$$

que es en realidad un sistema de tres ecuaciones diferenciales ordinarias de segundo orden para las coordenadas del vector $\boldsymbol{r}(t) = (x(t), y(t), z(t))$. Con esta ecuación podemos determinar desde el movimiento de una masa unida a un resorte hasta el movimiento de un planeta alrededor del sol !

Si la partícula se mueve en una dimensión, la ecuación anterior se reduce a

$$m\frac{d^2y}{dt^2} = F(t, y, \frac{dy}{dt}) \tag{1.4}$$

donde $y(t)$ denota la posición de la partícula. Esta es una ecuación diferencial ordinaria de 2^o orden para $y(t)$, que corresponde a $n = 2$ y $f(t, y, y') = F(t, y, y')/m$ en (1.1).

Problema 1: Mostrar que la ecuación diferencial que describe el movimiento (en una dimensión) de un objeto de masa m unido a un resorte de constante k es

$$y'' + \frac{k}{m}y = 0 \tag{1.5}$$

donde y es la posición del objeto medida a partir del punto de equilibrio.

1.2. Ecuación diferencial lineal ordinaria de orden n

Una ecuación diferencial es **lineal** si f en (1.1) es una función lineal de y y sus derivadas (aunque no necesariamente de la variable independiente t). Una ecuación diferencial lineal ordinaria de orden n puede expresarse en la forma

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = f(t) \tag{1.6}$$

donde $a_i(t)$, $i = 0, \ldots, n-1$ y $f(t)$, son funciones definidas en un cierto intervalo $I \subset \mathbb{R}$. Así, una ecuación diferencial lineal ordinaria de primer orden es de la forma

$$y' + p(t)y = f(t) \tag{1.7}$$

y una ecuación diferencial lineal de segundo orden puede escribirse como

$$y'' + p(t)y' + q(t)y = f(t) \tag{1.8}$$

En todos los casos, si $f(t) = 0 \ \forall\ t \in I$, la ecuación diferencial se denomina **homogénea**.

Las ecuaciones diferenciales lineales surgen en numerosos problemas corrientes. Por ejemplo, la ecuación (1.5), $y'' + \frac{k}{m}y = 0$, que describe el movimiento de una masa unida a un resorte, es una ecuación diferencial lineal homogénea de 2^o orden.

Pero la importancia de las ecuaciones diferenciales lineales proviene además del hecho que resultan más fáciles de resolver y comprender que las ecuaciones diferenciales no lineales. A diferencia de estas últimas, en las lineales es válido, como veremos, el principio de superposición, el cual permite obtener soluciones de problemas complejos mediante la superposición de soluciones de problemas más sencillos.

Una ecuación diferencial lineal puede también escribirse como

$$L[y] = f(t) \tag{1.9}$$

donde L es el **operador lineal** definido por

$$L[y] = y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y \tag{1.10}$$

Si y_1 e y_2 son dos funciones derivables hasta orden n y c_1, c_2 son constantes, L satisface

$$L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2] \tag{1.11}$$

En efecto, para cualquier derivada de orden k, con $k = 0, \ldots, n$,

$$(c_1y_1 + c_2y_2)^{(k)} = (c_1y_1)^{(k)} + (c_2y_2)^{(k)} = c_1y_1^{(k)} + c_2y_2^{(k)}$$

lo que implica $L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2]$ (Justificar!). Por ejemplo, para $n = 2$,

$$\begin{aligned} L[c_1y_1 + c_2y_2] &= (c_1y_1 + c_2y_2)'' + a_1(t)(c_1y_1 + c_2y_2)' + a_0(t)(c_1y_1 + c_2y_2) \\ &= c_1(y_1'' + a_1(t)y_1' + a_0(t)y_1) + c_2(y_2'' + a_1(t)y_2' + a_0(t)y_2) = c_1L[y_1] + c_2L[y_2] \end{aligned}$$

Notar que $L = D^n + a_{n-1}(t)D^{n-1} + \ldots + a_1(t)D + a_0(t)$, donde D es el operador derivada: $D[y] = y'$, $D^2[y] = D[D[y]] = y''$ y $D^{(k)}[y] = y^{(k)}$. La linealidad de L es consecuencia de la linealidad de D.

1.3. Propiedades fundamentales de ecuaciones diferenciales lineales homogéneas

Las ecuaciones diferenciales lineales poseen propiedades especiales, que permiten establecer las propiedades fundamentales de sus soluciones aun sin conocerlas explícitamente.

1. El conjunto S de todas las soluciones de una ecuación diferencial lineal **homogénea**

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = 0\,, \tag{1.12}$$

es un espacio vectorial.

Demostración: Una solución de (1.12) es una función $y(t)$ que la satisface. El conjunto de soluciones $S = \{y(t) \mid L[y] = 0\}$, es el núcleo del operador lineal L y la propiedad **1** es entonces consecuencia directa de la linealidad de L:
a) La función nula $y(t) = 0 \ \forall \ t \in I$ es siempre una solución de (1.12), denominada **solución trivial**, como puede verse fácilmente ($L[0] = 0$).
b) Además, si $y_1(t)$ e $y_2(t)$ son soluciones de (1.12), la **combinación lineal**

$$y(t) = c_1 y_1(t) + c_2 y_2(t)$$

es también solución de (1.12), para cualquier valor de las constantes c_1 y c_2, ya que

$$L[c_1y_1 + c_2y_2] = c_1L[y_1] + c_2L[y_2] = c_1 0 + c_2 0 = 0$$

Es decir, si $y(t)$ es solución, $cy(t)$ también lo es, y si $y_1(t)$ e $y_2(t)$ son soluciones, $y_1(t)+y_2(t)$ también lo es. Por lo tanto, S es no vacío y cerrado bajo las operaciones usuales de suma de funciones y multiplicación por un escalar. Esto implica que S es un subespacio del espacio vectorial de funciones derivables hasta orden n en I, y por ende un espacio vectorial.

La propiedad anterior se conoce como **propiedad de superposición**: Si $y_1(t)$ e $y_2(t)$ son dos soluciones de una ecuación diferencial lineal homogénea, **cualquier combinación lineal de ellas es también una solución.**

Problema 2: Mostrar que el conjunto de soluciones de una ecuación diferencial lineal no homogénea **no es un espacio vectorial.**

Consideraremos en lo sucesivo que $a_i(t)$, $i = 0, \ldots, n-1$, son funciones **continuas** en un intervalo abierto I.

2. La dimensión del espacio S de soluciones de la ecuación (1.12) **es $\boldsymbol{n}$**. Esto significa que existen $\boldsymbol{n}$ soluciones linealmente independientes $y_1(t), \ldots, y_n(t)$ tales que cualquier solución de (1.12) puede escribirse como

$$y(t) = c_1 y_1(t) + \ldots + c_n y_n(t) \tag{1.13}$$

El conjunto $\{y_1(t), \ldots, y_n(t)\}$ es pues una base de S, y la expresión (1.13), con n constantes de integración arbitrarias $c_1, \ldots, c_n$, es la **solución general** de (1.12).

Probaremos este teorema en las páginas siguientes. Como ejemplo, consideremos la ecuación lineal de segundo orden ($n = 2$)

$$y'' + y = 0 \tag{1.14}$$

que corresponde a $k/m = 1$ en (1.5) (en unidades apropiadas!). Podemos ver que

$$y_1(t) = \cos t, \quad y_2(t) = \operatorname{sen} t$$

son soluciones de (1.14), ya que $(\cos t)'' = -\cos t$, $(\operatorname{sen} t)'' = -\operatorname{sen} t$. Por lo tanto, como el conjunto $\{\cos t\,,\ \operatorname{sen} t\}$ es linealmente independiente y la ecuación (1.14) es de orden 2, toda solución de (1.14) es necesariamente de la forma

$$y(t) = c_1 \cos t + c_2 \operatorname{sen} t \tag{1.15}$$

con c_1, c_2 constantes. El conjunto $\{\cos t, \operatorname{sen} t\}$ es pues una base del espacio S de soluciones.

3. Existencia y Unicidad de la solución para el problema de condiciones iniciales. Consideremos la ecuación diferencial homogénea (1.12) con las n condiciones iniciales

$$\left\{\begin{array}{ccc} y(t_0) & = & b_0 \\ y'(t_0) & = & b_1 \\ & \vdots & \\ y^{(n-1)}(t_0) & = & b_{n-1} \end{array}\right. \tag{1.16}$$

con $t_0 \in I$. Si todas las $a_i(t)$ son continuas para $t \in I$, **existe una única solución** $y(t)$ para $t \in I$ que satisface (1.12) junto con las n condiciones iniciales (1.16), $\forall\, b_0, b_1, \ldots, b_{n-1}$.

Este es un caso particular del teorema general de existencia y unicidad que se aplica, bajo ciertas condiciones, también al caso no homogéneo e incluso al caso no lineal, como veremos más adelante. La solución que satisface (1.12) junto con (1.16) es una **solución particular** de (1.12).

En el caso homogéneo, esto implica, utilizando la solución general (1.13), que **las n constantes** $c_1, c_2, \ldots, c_n$ **quedan completamente determinadas por las n condiciones iniciales (1.16),** para cualquier valor de $b_0, b_1, \ldots, b_{n-1}$. Las condiciones iniciales conducen a un sistema de n ecuaciones lineales para $c_1, \ldots, c_n$:

$$\left\{\begin{array}{ccccc} c_1 y_1(t_0)+ & \ldots & +c_n y_n(t_0) & = & b_0 \\ c_1 y_1'(t_0)+ & \ldots & +c_n y_n'(t_0) & = & b_1 \\ & \vdots & & \vdots & \vdots \\ c_1 y_1^{(n-1)}(t_0)+ & \ldots & +c_n y_n^{(n-1)}(t_0) & = & b_{n-1} \end{array}\right.$$

El presente teorema asegura que el sistema anterior posee solución única (es compatible determinado) $\forall\ t_0 \in I$. Por lo tanto, el determinante de la matriz de coeficientes debe ser no nulo $\forall\ t_0 \in I$, es decir,

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) & \ldots & y_n(t) \\ y_1'(t) & y_2'(t) & \ldots & y_n'(t) \\ & & \ddots & \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \ldots & y_n^{(n-1)}(t) \end{vmatrix} \neq 0\ \forall t \in I \tag{1.17}$$

Este determinante se denomina *Wronskiano.*

Problema 3: Probar que si las funciones $\{y_1(t), \ldots, y_n(t)\}$ fuesen linealmente dependientes para $t \in I$, entonces $W(t) = 0 \ \forall\ t \in I$ (Sug.: Usar que una de las $y_i(t)$ será en tal caso combinación lineal de las restantes!).

Observación 1: La condición $W(t) \neq 0 \ \forall\ t \in I$ es más fuerte que la independencia lineal. Por ejemplo, las funciones $y_1(t) = t$ y $y_2(t) = t^2$ son L.I. para $t \in \mathbb{R}$, pero $W(0) = 0$ (probar!). Y las funciones $y_1(t) = t^2$, $y_2(t) = |t|t$ son también L.I. si $t \in \mathbb{R}$ pero $W(t) = 0$ $\forall\ t$ (probar!).

Observación 2: El teorema implica que la única solución de la ecuación homogénea que satisface $y(t_0) = y'(t_0) = \ldots = y^{(n-1)}(t_0) = 0$ es la solución trivial $y(t) = 0 \ \forall\ t \in I$.

Observación 3: El presente teorema puede utilizarse para demostrar que la dimensión del espacio de soluciones debe ser n, pues si fuese menor, el sistema para determinar los coeficientes c_i sería sobredeterminado y no podría garantizarse la existencia de solución para cualquier valor de las n condiciones iniciales, mientras que si fuese mayor, el sistema sería subdeterminado y no habría unicidad. Veremos luego otra demostración más general.

Como ejemplo, si en la ecuación (1.14), $y'' + y = 0$, tenemos las condiciones iniciales

$$\left\{\begin{array}{rcl} y(t_0) & = & y_0 \\ y'(t_0) & = & v_0 \end{array}\right. \tag{1.18}$$

donde en el problema del resorte y_0 representa la posición inicial y v_0 la velocidad inicial del objeto, el presente teorema asegura que existe una única solución que las satisface. De (1.15) tenemos $y(t) = c_1 \cos t + c_2 \operatorname{sen} t$, $y'(t) = -c_1 \operatorname{sen} t + c_2 \cos t$, y las condiciones iniciales (1.18) conducen al sistema

$$\left\{\begin{array}{rcl} c_1 \cos t_0 + c_2 \operatorname{sen} t_0 & = & y_0 \\ -c_1 \operatorname{sen} t_0 + c_2 \cos t_0 & = & v_0 \end{array}\right. \tag{1.19}$$

Este sistema posee solución única para c_1, c_2 $\forall\ t_0 \in \mathbb{R}$, ya que

$$W(t_0) = \left| \begin{array}{cc} \cos t_0 & \operatorname{sen} t_0 \\ -\operatorname{sen} t_0 & \cos t_0 \end{array} \right| = \cos^2 t_0 + \operatorname{sen}^2 t_0 = 1$$

verificándose que $W(t_0) \neq 0 \ \forall\ t_0 \in \mathbb{R}$. Por ejemplo, si $t_0 = 0$, la única solución de (1.19) es $c_1 = y_0$, $c_2 = v_0$, y entonces la única solución de (1.14) que satisface $y(0) = y_0$, $y'(0) = v_0$ es

$$y(t) = y_0 \cos t + v_0 \operatorname{sen} t$$

Problema 4: a) Mostrar que $y_1(t) = e^t$, $y_2(t) = e^{-t}$ son dos soluciones lin. indep. de

$$y'' - y = 0$$

y que por lo tanto, la solución general de esta ecuación es

$$y(t) = c_1 e^t + c_2 e^{-t}$$

b) Encontrar la solución que satisface $y(0) = y_0$, $y'(0) = v_0$.

4. Soluciones complejas. Supongamos que t y las n funciones $a_i(t)$, $i = 0, \ldots, n-1$, son **reales**. Si $y(t)$ es una solución **compleja** de la ecuación homogénea (1.12),

$$L[y] = 0, \quad y(t) = y_1(t) + iy_2(t) \tag{1.20}$$

donde $y_1(t)$ e $y_2(t)$ son funciones reales y $i^2 = -1$, tanto la parte real

$$y_1(t) = \mathrm{Re}[y(t)]$$

como la parte imaginaria

$$y_2(t) = \mathrm{Im}[y(t)]$$

son también soluciones (reales!) de la ecuación homogénea (1.12):

$$L[y_1] = 0, \quad L[y_2] = 0$$

Esto es muy fácil de probar, dado que L es **lineal y real**. La linealidad implica

$$0 = L[y] = L[y_1 + iy_2] = L[y_1] + iL[y_2]$$

y como L es real, $L[y_1]$ y $L[y_2]$ son ambos reales, por lo que la igualdad anterior implica

$$L[y_1] = 0, \quad L[y_2] = 0$$

Además, la función **conjugada** $\overline{y}(t) = y_1(t) - iy_2(t)$ es también solución de (1.12), pues $L[\overline{y}] = \overline{L[y]} = \overline{0} = 0$.

Fórmula de Euler. El resultado anterior resulta útil en conjunto con la fórmula de Euler para la exponencial de un número complejo, que se obtiene del desarrollo en serie de la exponencial:

$$e^{ibt} = \cos(bt) + i\sin(bt) \tag{1.21}$$

y en general,

$$e^{(a+ib)t} = e^{at}e^{ibt} = e^{at}\cos(bt) + ie^{at}\sin(bt) \tag{1.22}$$

Por lo tanto, para a, b y t reales,

$$\mathrm{Re}[e^{(a+ib)t}] = e^{at}\cos(bt), \quad \mathrm{Im}[e^{(a+ib)t}] = e^{at}\sin(bt) \tag{1.23}$$

Como ejemplo, consideremos nuevamente la ecuación (1.14), $y'' + y = 0$, que es lineal y real. Es fácil ver que $y(t) = e^{it}$ es solución, pues $y'(t) = ie^{it}$ y $y''(t) = i^2e^{it} = -e^{it}$. Como

$$e^{it} = \cos(t) + i\sin(t)$$

entonces

$$y_1(t) = \mathrm{Re}[e^{it}] = \cos(t), \quad y_2(t) = \mathrm{Im}[e^{it}] = \sin(t)$$

son soluciones reales linealmente independientes de (1.14), en acuerdo con (1.15).
La función conjugada $\overline{y}(t) = e^{-it} = \cos t - i \,\mathrm{sen}\, t$ es también solución de (1.14), y forma con e^{it} un par de soluciones complejas linealmente independientes de (1.14) (probar!). Tanto $\{\cos t, \mathrm{sen}\, t\}$ como $\{e^{it}, e^{-it}\}$ son bases del mismo espacio S de soluciones.

1.4. La ecuación lineal homogénea de primer orden

Es el caso $n = 1$. Si $a_0(t) = p(t)$, la ecuación homogénea (1.12) toma la forma

$$y' + p(t)y = 0 \tag{1.24}$$

o sea, $y' = -p(t)y$. Su solución general es

$$y(t) = ce^{-\int p(t)dt} \tag{1.25}$$

como es fácil verificar, donde c es una constante y $\int p(t)dt$ una primitiva de $p(t)$ en el intervalo I donde es continua. La dimensión del espacio S de soluciones es 1, siendo generado por $y_1(t) = e^{-\int p(t)dt}$, que claramente satisface $y_1'(t) = -p(t)y_1(t)$. Cualquier constante aditiva C en la primitiva puede absorberse en c ($c \to ce^{-C}$).

La única solución del problema de condiciones iniciales

$$\left\{ \begin{array}{rcl} y' + p(t)y & = & 0 \\ y(t_0) & = & y_0 \end{array} \right. \tag{1.26}$$

es entonces

$$y(t) = y_0 e^{-\int_{t_0}^{t} p(t)dt} \tag{1.27}$$

ya que $y(t_0) = y_0 e^{\int_{t_0}^{t_0} p(t)dt} = y_0$. Aquí $\int_{t_0}^{t} p(t)dt$ es la primitiva que se anula en $t = t_0$.

Problema 5: Probar que la solución general de

$$y' + yt = 0$$

es $y(t) = ce^{-t^2/2}$, y que la única solución que satisface $y(0) = 1$ es $y(t) = e^{-t^2/2}$.

1.5. La ecuación lineal homogénea de segundo orden

Definiendo $p(t) = a_1(t)$, $q(t) = a_0(t)$, la forma general de esta ecuación es

$$y'' + p(t)y' + q(t)y = 0 \tag{1.28}$$

Esta ecuación juega un rol muy importante en diversas áreas de la física e ingeniería. Por ejemplo, la 2^a ley de Newton (1.4) conduce a una ecuación de este tipo cuando la fuerza F en (1.4) es una función **lineal** de la posición y velocidad: $F(t) = \alpha y + \beta y'$.
En tal caso $p(t) = -\alpha/m$, $q(t) = -\beta/m$.

Los resultados generales anteriores implican que si $p(t)$ y $q(t)$ son continuas en un intervalo I, la ecuación (1.28) posee siempre **dos soluciones linealmente independientes** $y_1(t)$ y $y_2(t)$, tales que toda solución de (1.28) puede escribirse en la forma

$$y(t) = c_1\, y_1(t) + c_2\, y_2(t) \tag{1.29}$$

Esta es la **solución general** de (1.28). El espacio S de soluciones tiene dimensión 2 y $\{y_1(t), y_2(t)\}$ forman una base de S.

Además, el problema de condiciones iniciales

$$\left\{ \begin{array}{rcl} y'' + p(t)y' + q(t)y & = & 0 \\ y(t_0) & = & y_0 \\ y'(t_0) & = & v_0 \end{array} \right. \tag{1.30}$$

con $t_0 \in I$, posee **solución única**. Estas condiciones indiciales conducen al sistema

$$\left\{ \begin{array}{rcl} c_1 y_1(t_0) + c_2 y_2(t_0) & = & y_0 \\ c_1 y_1'(t_0) + c_2 y_2'(t_0) & = & v_0 \end{array} \right. \tag{1.31}$$

que entonces posee $\forall\ t_0 \in I$ una solución única para c_1 y c_2, para cualquier valor de y_0 y v_0 (sistema compatible determinado). Esto implica que el determinante (Wronskiano)

$$W(t) = \left| \begin{array}{cc} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{array} \right| = y_1(t)y_2'(t) - y_2(t)y_1'(t) \tag{1.32}$$

debe ser no nulo $\forall\ t \in I$.

Cuando $y(t)$ denota la posición de un objeto, y_0 y v_0 representan su **posición inicial y velocidad incial.** Con estos dos datos podemos entonces determinar $y(t)$.

La unicidad implica en particular que si $y_0 = v_0 = 0$, entonces $y(t) = 0\ \forall\ t \in I$. Es decir, si una solución satisface $y(t) = y'(t) = 0$ para algún $t \in I$, debe coincidir por unicidad con la solución trivial.

Problema 6: Muestre que la única solución de (1.31) es

$$c_1 = \frac{y_2'(t_0)y_0 - y_2(t_0)v_0}{W(t_0)}, \quad c_2 = \frac{-y_1'(t_0)y_0 + y_1(t_0)v_0}{W(t_0)}$$

Problema 7: Muestre que si $W(t) = 0\ \forall\ t \in I = (a, b)$ y y_1, y_2 son derivables en I, con $y_1(t) \neq 0$, $y_2(t) \neq 0\ \forall\ t \in I$, entonces $y_1(t)$ e $y_2(t)$ son linealmente dependientes en este intervalo.

Problema 8: a) Pruebe que $W'(t) = y_1(t)y_2''(t) - y_2(t)y_1''(t)$.
b) Utilizando a), muestre que si $y_1(t)$ e $y_2(t)$ son dos soluciones de (1.28), $W'(t) = -p(t)W(t)$ y por lo tanto

$$W(t) = W(t_0)e^{-\int_{t_0}^{t} p(t)dt}$$

Esto muestra que si $t_0 \in I$ y $W(t_0) \neq 0$, entonces $W(t) \neq 0\ \forall\ t \in I$. También implica $W(t) = W(t_0)$ (constante) si $p(t) = 0\ \forall\ t \in I$.

A diferencia de la ecuación lineal de primer orden, para el caso general no es posible dar una expresión de y_1 e y_2 en términos de integrales de $p(t)$ y $q(t)$. Sin embargo, sí es posible encontrar la segunda solución $y_2(t)$ si se conoce una de las soluciones $y_1(t)$:

Método general para hallar la segunda solución:
Si $y_1(t)$ es una solución no nula de (1.28), una segunda solución linealmente independiente de $y_1(t)$ es

$$y_2(t) = v(t)y_1(t), \quad \text{con} \quad v(t) = \int \frac{e^{-\int p(t)dt}}{y_1^2(t)} dt \tag{1.33}$$

Demostración: Planteando $y_2(t) = v(t)y_1(t)$, con $v(t)$ a determinar, se obtiene $y_2' = v'y_1 + vy_1'$, $y_2'' = v''y_1 + 2v'y_1' + vy_1''$ y reemplazando en (1.28),

$$v''y_1 + v'[2y_1' + p(t)y_1] + v[y_1'' + p(t)y_1' + q(t)y_1] = 0$$

Como y_1 es solución, el último término se anula. Definiendo $w = v'$ y dividiendo por y_1, obtenemos entonces una ecuación diferencial lineal homogénea de 1^o orden para w,

$$w' + [2y_1'/y_1 + p(t)]w = 0$$

cuya solución es, aplicando (1.25),

$$w = ce^{-\int [\frac{2y_1'}{y_1} + p(t)]dt} = ce^{-\ln y_1^2 - \int p(t)dt} = c\frac{e^{-\int p(t)dt}}{y_1^2(t)} \tag{1.34}$$

Dado que $v' = w$, integrando la expresión anterior y fijando $c = 1$ obtenemos el resultado (1.33). La constante aditiva C en la integral de (1.33) puede descartarse, ya que añade un término $Cy_1(t)$ en $y_2(t)$ linealmente dependiente de $y_1(t)$. Como $v'(t) = w$ es no nulo si $c \neq 0$, $v(t)$ no es constante, y entonces $y_2(t)$ es linealmente independiente de $y_1(t)$.

Ejemplo: Es fácil ver que $y_1(t) = e^{-t}$ es solución de

$$y'' + 2y' + y = 0 \tag{1.35}$$

ya que $y' = -y$ y $y'' = y$. Utilizando (1.33) obtenemos, dado que $p(t) = 2$ y $y_1^2(t) = e^{-2t}$,

$$v(t) = \int \frac{e^{-2t}}{e^{-2t}} dt = \int 1 dt = t + C$$

Descartando C, una segunda solución linealmente independiente de $y_1(t)$ es entonces

$$y_2(t) = ty_1(t) = te^{-t}$$

y la solución general de (1.35) es

$$y(t) = c_1 e^{-t} + c_2 te^{-t}$$

Problema 9: Sabiendo que $y_1(t) = e^t$ es una solución de $\quad y'' - 2y' + y = 0$ encontrar la solución general de esta ecuación para $t \in \mathbb{R}$.

Problema 10: Sabiendo que $y_1(t) = t$ es una solución de $\quad y'' + y'/t - y/t^2 = 0$ encontrar la solución general de esta ecuación para $t \in (0, \infty)$.

1.6. Ecuación lineal homogénea de segundo orden con coeficientes constantes

Consideremos ahora el caso muy importante en el que $p(t)$ y $q(t)$ son **constantes** $\forall t \in \mathbb{R}$, es decir, $p(t) = a$, $q(t) = b$:

$$y'' + ay' + by = 0 \tag{1.36}$$

Este tipo de ecuación diferencial, en el que no aparece t explícitamente, se denomina **autónoma**. Planteamos en este caso una solución del tipo

$$y(t) = e^{\lambda t} \tag{1.37}$$

con λ a determinar. Reemplazando en (1.36), se obtiene

$$e^{\lambda t}(\lambda^2 + a\lambda + b) = 0 \tag{1.38}$$

Como $e^{\lambda t} \neq 0$ (tanto si λ es real como complejo), $e^{\lambda t}$ será solución de (1.36) si y sólo si

$$\lambda^2 + a\lambda + b = 0 \tag{1.39}$$

Esta ecuación se denomina **ecuación característica**. Tiene a lo sumo dos raíces distintas:

$$\lambda_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}, \quad \lambda_2 = \frac{-a - \sqrt{a^2 - 4b}}{2}$$

Caso I: $\lambda_1 \neq \lambda_2$ ($a^2 \neq 4b$): Cada raíz origina una solución distinta,

$$y_1(t) = e^{\lambda_1 t}, \qquad y_2(t) = e^{\lambda_2 t} \tag{1.40}$$

que son linealmente independientes. La solución general de (1.36) es entonces

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \tag{1.41}$$

Caso II: $\lambda_1 = \lambda_2$ ($a^2 = 4b$): La única raíz $\lambda = -a/2$ origina la solución $y_1(t) = e^{\lambda t}$. La segunda solución $y_2(t)$ podemos hallarla con el método (1.33). Obtenemos aquí

$$v(t) = \int \frac{e^{-\int a dt}}{e^{2\lambda t}} dt = \int \frac{e^{-at}}{e^{-at}} dt = \int 1 dt = t$$

donde hemos descartado las constantes de integración (ver página previa). Por lo tanto, una segunda solución linealmente independiente de $y_1(t)$ es $y_2(t) = v(t)y_1(t) = te^{\lambda t}$. Obtenemos así el par de soluciones linealmente independientes

$$y_1(t) = e^{\lambda t}, \qquad y_2(t) = te^{\lambda t} \tag{1.42}$$

y la solución general es

$$y(t) = c_1 e^{\lambda t} + c_2 t\, e^{\lambda t} \tag{1.43}$$

Raíces complejas. Supondremos a y b reales. El caso I comprende dos subcasos:

I.1. $a^2 > 4b$: Ambas raíces λ_1, λ_2 son reales. Las soluciones son de tipo exponencial (creciente o decreciente) o constante (si λ_1 o λ_2 es nulo).

I.2. $a^2 < 4b$: Las raíces son complejas conjugadas:

$$\lambda_1 = \alpha + i\omega, \quad \lambda_2 = \alpha - i\omega = \overline{\lambda}_1, \quad \text{con } \alpha = \frac{-a}{2}, \quad \omega = \frac{\sqrt{4b - a^2}}{2}$$

y α, ω **reales**. La solución general sigue siendo (1.41), pero esta es ahora una representación compleja de la misma. Aplicando la fórmula de Euler (1.23) y el punto **4.** de la página 8, podemos obtener dos soluciones reales tomando la parte real e imaginaria de $e^{\lambda_1 t}$:

$$y_1(t) = \mathrm{Re}[e^{\lambda_1 t}] = e^{\alpha t}\cos(\omega t), \;\; y_2(t) = \mathrm{Im}[e^{\lambda_1 t}] = e^{\alpha t}\,\mathrm{sen}\,(\omega t) \tag{1.44}$$

que son linealmente independientes si $\omega \neq 0$. Obtenemos así una expresión real de la solución general:

$$y(t) = c_1 e^{\alpha t}\cos(\omega t) + c_2 e^{\alpha t}\,\mathrm{sen}\,(\omega t) \tag{1.45}$$

El significado de las raíces complejas de (1.39) es pues el de soluciones oscilatorias de frecuencia $\omega/2\pi$ con una amplitud que puede crecer ($\alpha > 0$) o decrecer ($\alpha < 0$) exponencialmente, o bien permanecer constante ($\alpha = 0$, o sea λ_1 y λ_2 imaginarios). Las partes real e imaginaria de $e^{\lambda_2 t}$ no originan nuevas soluciones linealmente independientes.

El espacio generado por las soluciones (1.40) o (1.44) **es el mismo** cuando se consideran escalares complejos: Se deja como ejercicio probar que

$$c_1 e^{\alpha t}\cos(\omega t) + c_2 e^{\alpha t}\,\mathrm{sen}\,(\omega t) = c_+ e^{(\alpha + i\omega)t} + c_- e^{(\alpha - i\omega)t} \tag{1.46}$$

con

$$c_\pm = \frac{c_1 \mp i c_2}{2} \tag{1.47}$$

Amplitud y fase. La solución (1.45) suele escribirse en forma más clara como

$$y(t) = A e^{\alpha t}\cos(\omega t + \phi) \tag{1.48}$$

donde ω representa la frecuencia angular, A la amplitud a $t = 0$ y ϕ una constante de fase. Esto muestra que toda solución (1.45) corresponde a un movimiento oscilatorio de período $T = 2\pi/\omega$ y amplitud efectiva $Ae^{\alpha t}$, que decrece ($\alpha < 0$) o crece ($\alpha > 0$) exponencialmente o bien es constante ($\alpha = 0$). ϕ indica la diferencia de fase de la oscilación respecto de $\cos(\omega t)$. Como $\cos(\omega t + \phi) = \cos\omega t\cos\phi - \mathrm{sen}\,(\omega t)\,\mathrm{sen}\,(\phi)$, entonces

$$A\cos\phi = c_1, \quad -A\sin\phi = c_2$$

y por lo tanto (probar!)

$$A = \sqrt{c_1^2 + c_2^2}, \quad \tan\phi = -c_2/c_1 \tag{1.49}$$

En (1.44), $y_1(t)$ corresponde a $A = 1$, $\phi = 0$, mientras que $y_2(t)$ a $A = 1$, $\phi = -\pi/2$.

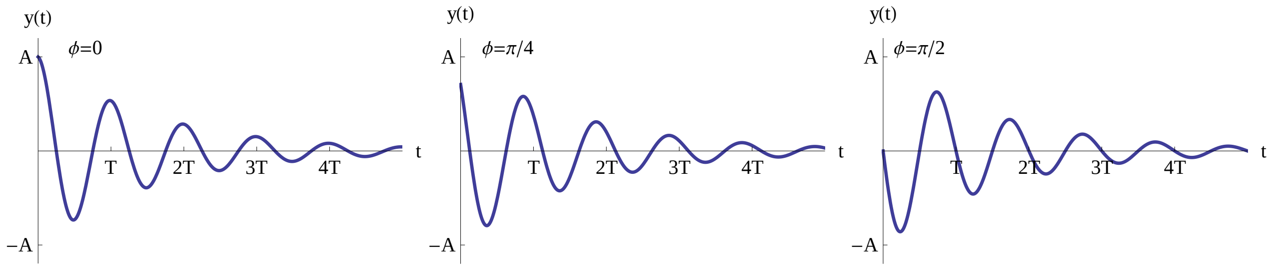


Gráfico de la solución (1.48), $y(t) = Ae^{\alpha t}\cos(\omega t + \phi)$, para $\alpha = -\omega/10$ y distintos valores de la fase ϕ. $T = 2\pi/\omega > 0$ es el período de la función $\cos(\omega t + \phi)$.

Ejemplo 1:

$$y'' - y = 0 \tag{1.50}$$

Planteando una solución de la forma $y(t) = e^{\lambda t}$, se obtiene la ecuación $\lambda^2 - 1 = 0$, cuyas raíces son $\lambda_1 = 1$, $\lambda_2 = -1$. Por lo tanto,

$$y_1(t) = e^t, \quad y_2(t) = e^{-t} \tag{1.51}$$

son dos soluciones linealmente independientes de (1.50) y la solución general es

$$y(t) = c_1 e^t + c_2 e^{-t}$$

Ejemplo 2:

$$y'' + 2y' + y = 0 \tag{1.52}$$

Planteando una solución de la forma $y(t) = e^{\lambda t}$, se obtiene la ecuación $\lambda^2 + 2\lambda + 1 = 0$, es decir $(\lambda + 1)^2 = 0$, cuya única raíz es $\lambda = -1$. Por lo tanto, utilizando (1.42),

$$y_1(t) = e^{-t}, \quad y_2(t) = te^{-t} \tag{1.53}$$

son dos soluciones linealmente independientes de (1.52) y la solución general es

$$y(t) = c_1 e^{-t} + c_2 t e^{-t}$$

Ejemplo 3:

$$y'' + 2y' + 5y = 0 \tag{1.54}$$

Planteando una solución de la forma $y(t) = e^{\lambda t}$, se obtiene la ecuación $\lambda^2 + 2\lambda + 5 = 0$, cuyas raíces son $\lambda_1 = \frac{-2+\sqrt{4-20}}{2} = -1 + 2i$, $\quad \lambda_2 = \frac{-2-\sqrt{4-20}}{2} = -1 - 2i$. Por lo tanto, utilizando (1.44),

$$y_1(t) = \mathrm{Re}[e^{(-1+2i)t}] = e^{-t}\cos(2t), \quad y_2(t) = \mathrm{Im}[e^{(-1+2i)t}] = e^{-t}\,\mathrm{sen}\,(2t) \tag{1.55}$$

son dos soluciones reales linealmente independientes de (1.54) y la solución general es

$$y(t) = c_1 e^{-t}\cos(2t) + c_2 e^{-t}\,\mathrm{sen}\,(2t)$$

Esta solución general puede expresarse también como $y(t) = c_+ e^{(-1+2i)t} + c_- e^{(-1-2i)t}$ o $y(t) = Ae^{-t}\cos(2t + \phi)$, donde $c_\pm$ y A, ϕ se relacionan con c_1, c_2 mediante (1.47) y (1.49).

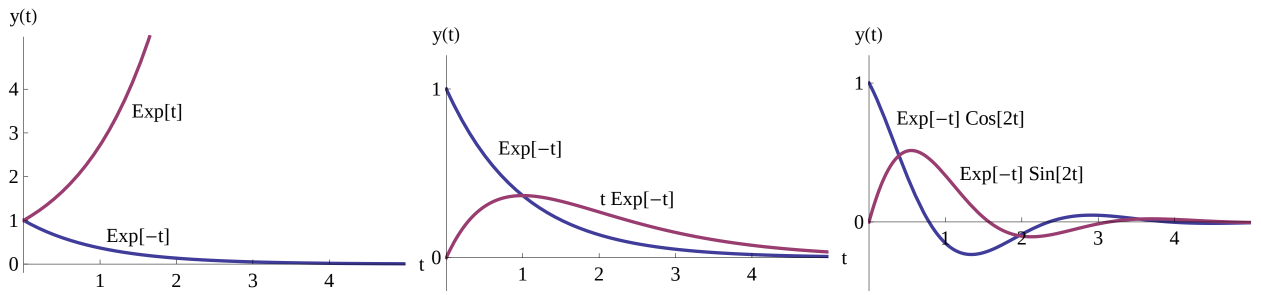


Gráfico de las soluciones (1.51)–(1.53)–(1.55) de las ecuaciones (1.50)–(1.52)–(1.54).

Problema 11: Halle las soluciones particulares de (1.50), (1.52) y (1.54) que satisfacen $\begin{cases} y(0) = 1 \\ y'(0) = 0 \end{cases}$. Grafique estas soluciones.

Problema 12: Si $W[y_1, y_2] = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$ denota el wronskiano (1.32), muestre que

a) $W[e^{\lambda_1 t}, e^{\lambda_2 t}] = (\lambda_2 - \lambda_1)e^{(\lambda_1+\lambda_2)t}$
b) $W[e^{\lambda t}, te^{\lambda t}] = e^{2\lambda t}$
c) $W[e^{\alpha t}\cos(\omega t), e^{\alpha t}\,\mathrm{sen}\,(\omega t)] = \omega e^{2\alpha t}$

Ejercicios I:

1) Hallar la solución general de las siguientes ecuaciones diferenciales. Dar una expresión real de la misma.

a) $y'' - 4y' + 4y = 0$	*b*) $y'' - 3y' - 4y = 0$	*c*) $y'' + 4y = 0$
d) $y'' + 4y' = 0$	*e*) $y'' + 4y' + 5y = 0$	*f*) $y'' = 0$
g) $y''' = y'$	*h*) $y'' + 2\alpha y' + \alpha^2 y = 0$,	*i*) $y'' - \omega^2 y = 0$, $\omega \neq 0$

2) Excepto en g), hallar las soluciones particulares de las ecuaciones anteriores que satisfacen a) $y(0) = 0$, $y'(0) = 1$, b) $y(0) = 1$, $y'(0) = 0$. Grafíquelas.
En g) considerar $y(0) = y'(0) = y''(0) = 1$.

3) Mostrar que si $y(t)$ es solución de la ecuación $y'' + by' + ay = 0$, con a, b constantes, $w(t) = y'(t)$ es también solución de la misma ecuación. Exprese las condiciones iniciales que satisface w en términos de las de y.

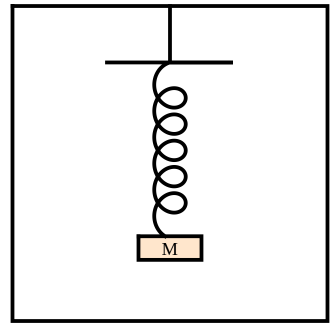
4) Muestre que si $y(t)$ es solución de $y'' + by' + ay = 0$ y satisface $y(0) = y_0$, $y'(0) = v_0$, la solución que satisface $y(t_0) = y_0$, $y'(t_0) = v_0$ es $y(t - t_0)$. Generalizar al caso de orden n.

5) a) Sabiendo que t es solución de la ecuación

$$y'' - 2y'/t + 2y/t^2 = 0$$

encontrar una segunda solución linealmente independiente de t para $t \in (0, \infty)$ y escribir la solución general.
b) Mostrar que existen en este caso infinitas soluciones que satisfacen $y(0) = 0$, $y'(0) = 1$, y ninguna que satisface $y(0) = 1$, $y'(0) = 0$. Explicar la causa de esta no unicidad y no existencia !! ¿Sucede lo mismo para $y(1) = 0$, $y'(1) = 1$ y $y(1) = 1$, $y'(1) = 0$?

1.7. Aplicaciones



Problema 13: El oscilador armónico. Determinar el movimiento de una masa $m > 0$ unida a un resorte de constante $k > 0$, de modo que la fuerza neta sobre la masa es $F = -ky$, con y la posición medida desde el punto de equilibrio.

Hemos visto que la ecuación de movimiento es (véase ecuación (1.5))

$$y'' + \omega^2 y = 0, \quad \omega = \sqrt{k/m} > 0 \tag{1.56}$$

a) Planteando una solución $y(t) = e^{\lambda t}$, mostrar que $\lambda = \pm i\omega$ y que por lo tanto, las soluciones reales linealmente independientes son

$$y_1(t) = \mathrm{Re}[e^{i\omega t}] = \cos(\omega t), \;\; y_2(t) = \mathrm{Im}[e^{i\omega t}] = \operatorname{sen}(\omega t)$$

y la solución general es

$$y(t) = c_1 \cos(\omega t) + c_2 \operatorname{sen}(\omega t)$$

b) Usando (1.48), reescribir esta solución como

$$y(t) = A\cos(\omega t + \phi)$$

Esta expresión permite ver claramente que toda solución corresponde a un movimiento oscilatorio de amplitud A y frecuencia angular ω. La frecuencia de oscilación es

$$f = \frac{\omega}{2\pi}$$

y el período

$$T = \frac{1}{f} = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{k}}$$

ya que $\omega T = 2\pi$ y por lo tanto $A\cos(\omega(t+T)+\phi) = A\cos(\omega t+\phi)$ (probar!). Observar que el período es *independiente de la amplitud A y la fase ϕ*, es decir, de las condiciones iniciales. Estas determinan A y ϕ, pero no ω.

c) Probar que si $y(0) = y_0$, $y'(0) = v_0$, entonces

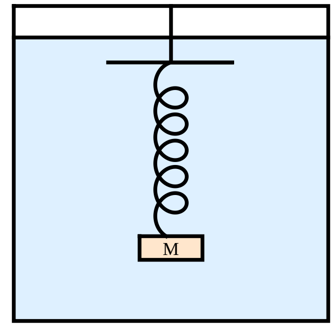
$$c_1 = y_0, \;\; c_2 = v_0/\omega$$

y por lo tanto

$$A = \sqrt{y_0^2 + v_0^2/\omega^2}, \;\; \tan\phi = -v_0/(\omega y_0)$$

c) Analizar el caso $\omega = 0$ en (1.56). Determinar la solución general.

Problema 14: El oscilador armónico amortiguado.
Determinar el movimiento de una masa $m > 0$ unida a un resorte de constante $k > 0$ en un medio viscoso (en el que la fuerza de roce es proporcional a la velocidad), de modo que la fuerza neta sobre la masa es $F = -ky - \mu y'$, con y la posición medida desde el punto de equilibrio y $\mu > 0$ un coeficiente de roce.



La ecuación de movimiento resultante puede escribirse en la forma

$$y'' + 2\gamma y' + \omega^2 y = 0, \quad \omega = \sqrt{\frac{k}{m}} > 0, \;\; \gamma = \frac{\mu}{2m} > 0$$

que es nuevamente una ecuación diferencial **lineal** homogénea de segundo orden. Tanto γ como ω tienen las mismas unidades (tiempo^{-1}).
a) Proponiendo una solución $y(t) = e^{\lambda t}$, mostrar que la ecuación característica es

$$\lambda^2 + 2\gamma\lambda + \omega^2 = 0$$

y que sus soluciones son

$$\lambda_\pm = -\gamma \pm \sqrt{\gamma^2 - \omega^2}$$

b) Mostrar que existen tres casos diferentes:
I. $\boldsymbol{\gamma > \omega}$: Movimiento sobreamortiguado. Probar que en este caso ambas raíces $\lambda_\pm$ son **reales, negativas y distintas**, por lo que la solución general es combinación lineal de exponenciales decrecientes:

$$y(t) = c_1 e^{\lambda_+ t} + c_2 e^{\lambda_- t} = c_1 e^{(-\gamma + \sqrt{\gamma^2 - \omega^2})t} + c_2 e^{(-\gamma - \sqrt{\gamma^2 - \omega^2})t}$$

El roce es lo suficientemente fuerte como para suprimir las oscilaciones. Describa cualitativamente el movimiento resultante.

II. $\boldsymbol{\gamma = \omega}$: Amortiguamiento crítico. Probar que en este caso ambas raíces son **coincidentes:** $\lambda_\pm = -\gamma < 0$, y que la solución general es

$$y(t) = c_1 e^{-\gamma t} + c_2 t e^{-\gamma t}$$

Describa en forma cualitativa el movimiento resultante.

III. $\boldsymbol{\gamma < \omega}$: Movimiento subamortiguado. Probar que en este caso las raíces son **complejas conjugadas,**

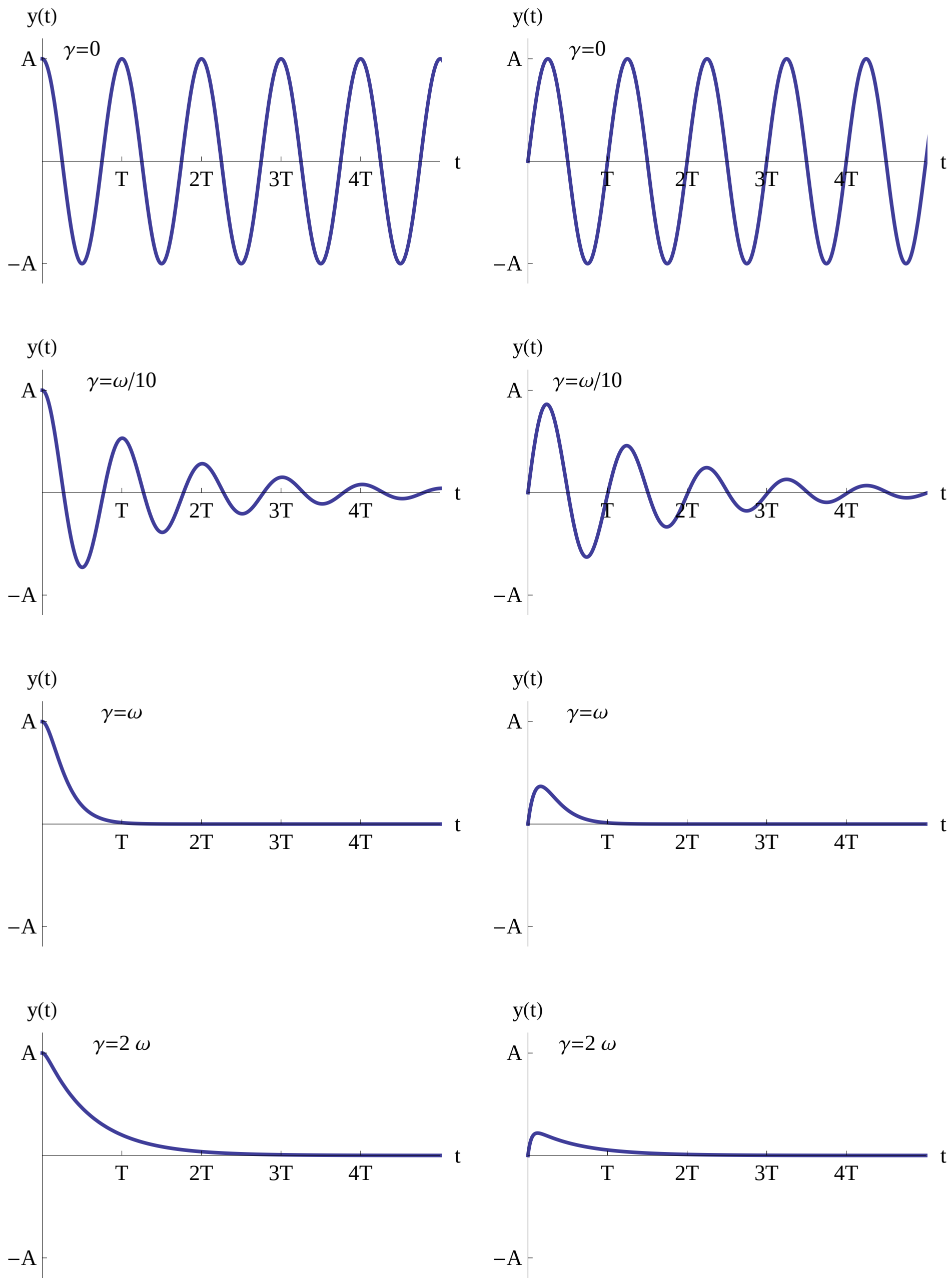
$$\lambda_\pm = -\gamma \pm i\omega_\gamma, \;\; \omega_\gamma = \sqrt{\omega^2 - \gamma^2} > 0$$

con ω_γ real, y que la solución general es entonces

$$y(t) = c_1 e^{-\gamma t}\cos(\omega_\gamma t) + c_2 e^{-\gamma t}\,\mathrm{sen}\,(\omega_\gamma t) = A e^{-\gamma t}\cos(\omega_\gamma t + \phi)$$

Describa en forma cualitativa el movimiento resultante, mostrando que corresponde a un movimiento oscilatorio con una amplitud que decrece exponencialmente al aumentar t, y una frecuencia angular efectiva $\omega_\gamma < \omega$. $\tau = 1/\gamma$ es el tiempo de decaimiento (o relajación).

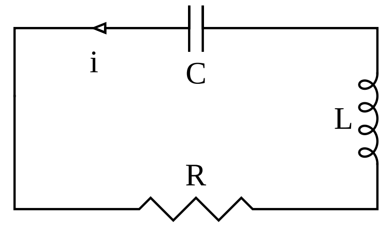
c) Hallar la solución en los casos I, II y III que satisface las condiciones iniciales:
i) Desplazamiento inicial A, velocidad inicial nula: $y(0) = A$, $y'(0) = 0$.
ii) Desplazamiento inicial nulo, velocidad inicial v_0: $y(0) = 0$, $y'(0) = v_0$.



$y(0) = A, \;\; y'(0) = 0$ $\qquad$ $y(0) = 0, \;\; y'(0) = A\omega$

Movimiento oscilatorio armónico y amortiguado: Gráficas de la solución $y(t)$. La columna izquierda corresponde a condiciones iniciales $y(0) = A$, $y'(0) = 0$ (velocidad inicial nula, partiendo de A) y la derecha a $y(0) = 0$, $y'(0) = A\omega$ (velocidad inicial fija, partiendo del punto de equilibrio). La primer fila representa el movimiento oscilatorio armónico ($\gamma = 0$), la segunda el movimiento subamortiguado ($0 < \gamma < \omega$), la tercera el caso crítico ($\gamma = \omega$) y la cuarta el movimiento sobreamortiguado ($\gamma > \omega$). En todos los casos $T = 2\pi/\omega$ es el período del movimiento oscilatorio en ausencia de roce.

Problema 15: Circuito LCR en serie.
Determinar la descarga de un capacitor de capacitancia C en un circuito como el de la figura, con resistencia R e inductancia L.



a) Si q es la carga y $dq/dt = i$ la corriente, probar que la ecuación del circuito conduce a la ecuación diferencial

$$L\frac{d^2q}{dt^2} + \frac{dq}{dt}R + \frac{q}{C} = 0$$

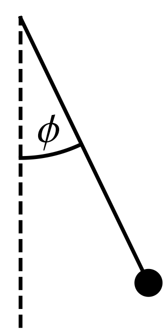
Llamamdo $y(t) = q(t)$, podemos escribir esta ecuación en la forma

$$y'' + 2\gamma y' + \omega^2 y = 0, \quad \gamma = \frac{R}{2L} > 0, \quad \omega = \frac{1}{\sqrt{LC}} > 0$$

que es exactamente igual a la del movimiento oscilatorio amortiguado: La inductancia L reemplaza a la masa m, la resistencia R al coeficiente de roce viscoso μ, la capacitancia C a la inversa de la constante de resorte k y la carga al desplazamiento respecto del punto de equilibrio. Es un ejemplo de analogía física: Sistemas muy distintos quedan descriptos por la misma ecuación diferencial, y exhiben por lo tanto el mismo comportamiento en las variables correspondientes. Podemos pues simular un circuito LCR en serie mediante una masa unida a un resorte en un líquido viscoso o viceversa!

b) Discutir los tres casos posibles de evolución, indicando para qué valores de L, C, R el sistema será sub y sobre-amortiguado, y en que caso será crítico.
c) Determinar explícitamente la solución para una carga inicial q_0 y corriente inicial $i_0 = 0$ (circuito inicialmente abierto, que se cierra en $t = 0$).

Problema 16: Péndulo simple.
Determinar el movimiento de una masa m unida a un hilo o barra de longitud L (y masa despreciable).



Puede mostrarse que la ecuación que describe el movimiento es

$$mL\phi'' = -mg \operatorname{sen}\phi$$

donde ϕ es el ángulo que forma el hilo con la vertical, ya que la fuerza en la dirección tangencial es $-mg \operatorname{sen}\phi$. Podemos reescribir esta ecuación como

$$\phi'' + \omega^2 \operatorname{sen}\phi = 0, \quad \omega^2 = g/L$$

Esta es una ecuación diferencial de segundo orden **no lineal**, por lo que en principio no podemos aplicar los métodos presentes. Sin embargo, para **pequeños ángulos** $|\phi| \ll 1$, podemos escribir, utilizando el desarrollo de Maclaurin de $\operatorname{sen}\phi$,

$$\operatorname{sen}\phi = \phi + O(\phi^3)$$

Despreciando los términos $O(\phi^3)$, obtenemos la aproximación lineal $\operatorname{sen}\phi \approx \phi$, que conduce a la **ecuación diferencial lineal**

$$\phi'' + \omega^2\phi = 0$$

Esta es exactamente igual a la del oscilador armónico. Por lo tanto, para pequeños ángulos el movimiento del péndulo es similar al de una masa unida a un resorte.
a) Mostrar, en base a la discusión anterior, que el período de oscilación para pequeños ángulos iniciales será independiente de la amplitud:

$$T = 2\pi/\omega = 2\pi\sqrt{L/g}$$

b) Determinar $\phi(t)$ para $\phi(0) = \phi_0$, $\phi'(0) = 0$.

Problema 17: Pequeñas oscilaciones en torno al equilibrio

El caso del péndulo muestra el comportamiento general de sistemas estables en torno a un punto de equilibrio. En dicho punto, la fuerza es nula. Si la fuerza es una función de la posición x, su desarrollo de Taylor en torno a dicho punto (denotado como x_0) conduce a

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + O((x - x_0)^2)$$

Como $F(x_0) = 0$, despreciando los términos $O((x - x_0)^2)$ obtenemos

$$F(x) \approx -k(x - x_0), \quad k = -f'(x_0)$$

que es la aproximación lineal a F en x_0. La segunda ley de Newton conduce entonces a la ecuación $mx'' = -k(x - x_0)$, que es válida para describir el movimiento en la vecindad del punto de equilibrio si $k \neq 0$.
a) Probar que la ecuación para el desplazamiento $y = x - x_0$ respecto del punto de equilibrio es

$$y'' + \frac{k}{m}y = 0$$

b) Mostrar que si $k > 0$ ($f'(x_0) < 0$) el sistema ejecuta oscilaciones de frecuencia angular $\omega = \sqrt{\frac{k}{m}}$ si se lo aparta de la posición de equilibrio. Este es el caso de equilibrio **estable**, en el que la fuerza es restitutiva.
c) Mostrar que si $k < 0$ ($f'(x_0) > 0$) el sistema se aleja en general en forma exponencial ($y \propto e^{\sqrt{|k|/m}\,t}$) de la posición de equilibrio si es apartado. Este es el caso de equilibrio **inestable**, en el que la fuerza tiende a alejar al sistema del punto de equilibrio.
d) Discutir el caso $k = 0$.

Ejercicios II:

1) Determinar el movimiento de una masa $m = 1kg$ sujeta a un resorte de constante $k = 1N/m$, con un coeficiente de roce viscoso μ: Halle la solución general, el período en ausencia de roce, y el valor de μ a partir del cual el sistema cesa de exhibir oscilaciones.
2) Determine la solución de la ecuación diferencial

$$y'' + 2\gamma y' - \alpha^2 y = 0$$

que corresponde a $k < 0$ en el problema 17, asumiendo $\alpha > 0$, $\gamma > 0$. Indique si el roce ($\mu = 2m\gamma$) logra evitar el alejamiento de la posición de equilibrio.

Problema 18: Ecuación de Euler de 2^o orden.
Esta ecuación es de la forma

$$y'' + b\frac{y'}{x} + a\frac{y}{x^2} = 0 \tag{1.57}$$

donde a y b son constantes y $x \in (0, \infty)$ denota la variable independiente. Esta ecuación lineal puede resolverse exactamente y surge en diversas aplicaciones.
a) Mostrar que la función

$$y(x) = x^\lambda$$

será solución de (1.57) si y sólo si λ es raíz de la ecuación

$$\lambda^2 + \lambda(b-1) + a = 0$$

b) Si dicha ecuación posee dos raíces distintas λ_1, λ_2, justifique que la solución general es

$$y(x) = c_1 x^{\lambda_1} + c_2 x^{\lambda_2} \tag{1.58}$$

c) Si en cambio $\lambda_1 = \lambda_2 = \lambda$, demuestre, utilizando el método (1.33), que la segunda solución es $y_2(x) = x^\lambda \ln x$ y que la solución general en este caso es

$$y(x) = c_1 x^\lambda + c_2 x^\lambda \ln x \tag{1.59}$$

d) Si λ es complejo, entonces $\lambda = \alpha + i\omega$ y $x^\lambda = e^{\lambda \ln x} = x^\alpha e^{i\omega \ln x}$. Aplicando la fórmula de Euler, muestre que la expresión real de la solución general es

$$y(x) = c_1 x^\alpha \cos(\omega \ln x) + c_2 x^\alpha \operatorname{sen}(\omega \ln x)$$

e) La ecuación (1.57) puede también resolverse reemplazando $x = e^t$. Muestre que esta sustitución transforma (1.57) en una ecuación lineal de 2^o orden con coeficientes constantes en $t = \ln x$:

$$\frac{d^2y}{dt^2} + (b-1)\frac{dy}{dt} + ay = 0$$

y que como $e^{\lambda t} = x^\lambda$, se vuelven a obtener las soluciones anteriores en la variable x.

f) Utilizando los resultados anteriores, halle la solución general de las ecuaciones

i) $y'' + y'/x - 4y/x^2 = 0$, *ii)* $y'' - 3y'/x + 4y/x^2 = 0$

g) Pruebe que la solución general de

$$y'' + 2y/x - l(l+1)y/x^2 = 0$$

para $l \neq -1/2$ es $y(x) = c_1 x^l + c_2 x^{-l-1}$. ¿ Que sucede si $l = -1/2$?

1.8. Ecuación lineal homogénea de orden n con coeficientes constantes

Los resultados anteriores se generalizan en forma inmediata a la ecuación

$$y^{(n)} + a_{(n-1)}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0$$

donde $a_0, \ldots, a_{n-1}$ son constantes. Proponiendo una solución $y(t) = e^{\lambda t}$, obtenemos $e^{\lambda t}[\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0] = 0$, lo que conduce a la ecuación característica

$$\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0$$

Si esta ecuación posee n raíces distintas λ_i, $i = 1, \ldots, n$, las n soluciones linealmente independientes serán $y_i(t) = e^{\lambda_i t}$ y la solución general será

$$y(t) = c_1 e^{\lambda_1 t} + \ldots + c_n e^{\lambda_n t}$$

Por el contrario, si una raíz λ tiene multiplicidad m, pueden obtenerse de ella m soluciones linealmente independientes,

$$e^{\lambda t},\;\; te^{\lambda t}, \ldots, t^{m-1}e^{\lambda t}$$

Como la suma de todas las multiplicidades de las raíces distintas es n ($\sum_i m_i = 1$) se obtienen así n soluciones linealmente independientes.

Las raíces complejas se tratan de forma similar. Si los a_i, $i = 1, \ldots, n$, son todos reales, las raíces complejas aparecen en **pares conjugados**: $\lambda = \alpha \pm i\omega$. De cada par de multiplicidad m pueden obtenerse $2m$ soluciones reales $t^k e^{\alpha t}\cos(\omega t)$, $t^k e^{\alpha t}\operatorname{sen}(\omega t)$, $k = 0, \ldots, m-1$. Importante: **Todas** las raíces, tanto reales como complejas, deben ser incluidas en la solución general.

Ejemplo 1: Encontrar la solución general de la ecuación de cuarto orden

$$y'''' - y = 0$$

Proponiendo $y(t) = e^{\lambda t}$, obtenemos la ecuación

$$\lambda^4 - 1 = 0$$

Las 4 raíces de esta ecuación son $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = i$, $\lambda_4 = -i$. Tomando las partes real e imaginaria de e^{it}, la solución general será

$$y(t) = c_1 e^t + c_2 e^{-t} + c_3 \cos t + c_4 \operatorname{sen} t$$

Los cuatro valores iniciales $y(t_0)$, $y'(t_0)$, $y''(t_0)$ y $y'''(t_0)$ determinan las cuatro constantes.

Ejemplo 2: Encontrar la solución general de la ecuación de tercer orden

$$y''' + 3y'' + 3y' + y = 0$$

Proponiendo $y(t) = e^{\lambda t}$, obtenemos la ecuación $\lambda^3 + 3\lambda^2 + 3\lambda + 1 = 0$, es decir $(\lambda + 1)^3 = 0$, cuya única raíz es $\lambda = -1$, con multiplicidad 3. La solución general es entonces

$$y(t) = c_1 e^{-t} + c_2 t e^{-t} + c_3 t^2 e^{-t}$$

Ejercicios III: Determine la solución general de las ecuaciones

a) $y'''' - 2y'' + y = 0$ *b*) $y''' = 0$

1.9. La ecuación diferencial lineal no homogénea

Consideremos ahora el caso general

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = f(t) \tag{1.60}$$

es decir,

$$L[y] = f(t) \tag{1.61}$$

donde $a_1(t), \ldots, a_n(t)$ y $f(t)$ continuas en intervalo abierto I.

1. La solución general de (1.60) está dada por la suma de la solución general $y_h(t)$ de la ecuación **homogénea** más una solución particular $y_p(t)$ de la ecuación **no homogénea**:

$$y(t) \;\; = \;\; y_h(t) + y_p(t) \tag{1.62}$$

donde

$$y_h(t) = c_1 y_1(t) + \ldots + c_n y_n(t) \tag{1.63}$$

es la solución general (1.13) y satisface $L[y_h] = 0$, mientras que $y_p(t)$ satisface $L[y_p] = f(t)$.

Demostración: Dado que $L[y_h] = 0$ y $L[y_p] = f(t)$, vemos que (1.62) es también solución de (1.60), pues al ser L lineal,

$$L[y_h + y_p] = L[y_h] + L[y_p] = 0 + f(t) = f(t)$$

Además, **toda** solución $y(t)$ de (1.60) es de la forma (1.62), pues si $L[y(t)] = f(t)$, entonces

$$L[y - y_p] = L[y] - L[y_p] = f(t) - f(t) = 0$$

lo que muestra que la diferencia $y(t) - y_p(t)$ **es una solución** $y_h(t)$ **de la ecuación homogénea**. Por lo tanto $y(t) - y_p(t) = y_h(t)$ y entonces $y(t) = y_h(t) + y_p(t)$.

Para resolver la ecuación (1.60) debemos pues resolver la ecuación homogénea y luego encontrar alguna solución particular $y_p(t)$ de (1.60) mediante algún método.

El conjunto de soluciones de (1.60) **no es un espacio vectorial** si $f(t) \neq 0$ (no contiene a la función nula y no es cerrado bajo suma de funciones o multiplicación por un escalar). Sin embargo, la linealidad de L implica el siguiente resultado:

2. Si $y_{p1}(t)$, $y_{p2}(t)$ son soluciones de (1.60) para $f(t) = f_1(t)$ y $f(t) = f_2(t)$ respectivamente, una solución particular de (1.60) para la **combinación lineal** $f(t) = c_1 f_1(t) + c_2 f_2(t)$ es la combinación lineal de soluciones particulares:

$$L[y_{p1}] = f_1(t), \;\; L[y_{p2}] = f_2(t) \;\; \Rightarrow \;\; L[c_1 y_{p1} + c_2 y_{p2}] = c_1 f_1(t) + c_2 f_2(t)$$

ya que $L[c_1 y_{p1} + c_2 y_{p2}] = c_1 L[y_{p1}] + c_2 L[y_{p2}] = c_1 f_1(t) + c_2 f_2(t)$.
En particular, $L[y_{p1} + y_{p2}] = f_1(t) + f_2(t)$ y $L[c y_{pi}(t)] = c f_i(t)$, $i = 1, 2$.
Por lo tanto si $f(t)$ es suma de términos más simples, podemos hallar una solución particular sumando las soluciones para c/uno de los términos.

1.10. La ecuación lineal no homogénea de primer orden

La ecuación es

$$y' + p(t)y = f(t) \tag{1.64}$$

Vimos previamente que la solución general de la ecuación homogénea es

$$y_h(t) = cy_1(t), \quad y_1(t) = e^{-\int p(t)dt}$$

Para hallar una solución particular, utilizamos el método usualmente denominado **variación de parámetros**: Como en (1.33), consiste en proponer una solución de la forma

$$y_p(t) = v(t)y_1(t)$$

con $v(t)$ una función a determinar. Reemplazando en (1.64) se obtiene

$$v'y_1(t) + v[y_1'(t) + p(t)y_1(t)] = f(t)$$

pero como $y_1'(t) + p(t)y_1(t) = 0$ obtenemos $v'y_1(t) = f(t)$. Por lo tanto, $v' = \frac{f(t)}{y_1(t)}$ y

$$v(t) = \int \frac{f(t)}{y_1(t)}dt = \int e^{\int p(t)dt} f(t)dt$$

Una solución particular es entonces

$$y_p(t) \quad = \quad y_1(t)\int \frac{f(t)}{y_1(t)}dt = e^{-\int p(t)dt}\int e^{\int p(t)dt} f(t)dt \tag{1.65}$$

y la solución general de (1.64) es

$$\begin{aligned} y(t) &= cy_1(t) + y_1(t)\int \frac{f(t)}{y_1(t)}dt && (1.66)\\ &= ce^{-\int p(t)dt} + e^{-\int p(t)dt}\int e^{\int p(t)dt} f(t)dt && (1.67) \end{aligned}$$

Problema 19: Probar que la solución particular que se anula en $t = t_0$ puede escribirse como

$$y_p(t) = \int_{t_0}^{t} g(t,t')f(t')dt', \;\; \text{con} \;\; g(t,t') = y_1(t)/y_1(t') = e^{-\int_{t'}^{t} p(s)ds}$$

y que $g(t,t')$ es la solución de la ecuación homogénea que satisface $y(t') = 1$.

Ejemplo: Utilizando (1.67), vemos que una solución particular de $y' + yt = t$ es

$$y_p(t) = e^{-t^2/t}\int e^{t^2/2}tdt = e^{-t^2/2}e^{t^2/2} = 1$$

y que la solución general es entonces $y(t) = ce^{-t^2/2} + 1$.

1.11. La ecuación lineal no homogénea de segundo orden

Consideremos ahora

$$y'' + p(t)y' + q(t)y = f(t) \tag{1.68}$$

Si se conocen dos soluciones linealmente independientes $y_1(t)$, $y_2(t)$ de la ecuación homogénea, es siempre posible encontrar una solución particular utilizando nuevamente el método de **variación de parámetros**. El resultado es

$$y_p(t) = -y_1(t)\int \frac{y_2(t)f(t)}{W(t)}dt + y_2(t)\int \frac{y_1(t)f(t)}{W(t)}dt \tag{1.69}$$

donde

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix} = y_1(t)y_2'(t) - y_1'(t)y_2(t) \tag{1.70}$$

es el wronskiano de las soluciones (hemos visto ya que $W(t) \neq 0 \ \forall \ t \in I$).
La solución general de (1.68) es entonces

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + y_p(t)$$

Demostración: Proponemos una solución particular de la forma

$$y_p(t) = v_1(t)y_1(t) + v_2(t)y_2(t)$$

donde $y_1(t)$ e $y_2(t)$ son dos soluciones linealmente independientes de la solución homogénea y $v_1(t)$, $v_2(t)$ funciones a determinar. Imponiendo la condición $v_1'(t)y_1(t) + v_2'(t)y_2(t) = 0$, obtenemos $y_p' = v_1y_1' + v_2y_2'$ y entonces $y_p'' = v_1'y_1' + v_2'y_2' + v_1y_1'' + v_2y_2''$. Reemplazando en (1.68) se obtiene entonces

$$v_1'y_1' + v_2'y_2' + v_1[y_1'' + p(t)y_1' + q(t)y_1] + v_2[y_2'' + p(t)y_2' + q(t)y_2] = f(t)$$

pero como y_1 e y_2 son solución de la ecuación homogénea, los dos últimos términos del primer miembro se anulan. Por lo tanto, la condición impuesta y la ecuación anterior conducen al sistema

$$\left\{ \begin{array}{rcl} v_1'(t)y_1(t) + v_2'(t)y_2(t) & = & 0 \\ v_1'(t)y_1'(t) + v_2'(t)y_2'(t) & = & f(t) \end{array} \right. \tag{1.71}$$

Despejando $v_1'(t)$ y $v_2'(t)$ y luego integrando, obtenemos

$$v_1(t) = -\int \frac{y_2(t)}{W(t)}f(t)dt\,, \quad v_2(t) = \int \frac{y_1(t)}{W(t)}f(t)dt \tag{1.72}$$

donde $W(t)$ es el wronskiano (1.70). Esto conduce a la solución (1.69).
Problema 20: Probar (1.72) a partir de (1.71).

Problema 21: Probar a partir de (1.69) que la solución particular que satisface $y_p(t_0) = y_p'(t_0) = 0$ es

$$y_p(t) = \int_{t_0}^{t} g(t,t')b(t')dt', \quad g(t,t') = \frac{-y_1(t)y_2(t') + y_2(t)y_1(t')}{W(t')}$$

siendo $g(t,t')$ la solución de la ecuación homogénea que satisface $y(t') = 0$, $y'(t') = 1$.

Ejercicios III. Halle la solución general de las siguientes ecuaciones diferenciales. Utilice resultados de ejercicios previos.

a) $y'' + y'/t - 4y'/t^2 = t$ *b)* $y'' - y = e^{-t}$

1.12. La ecuación lineal no homogénea de segundo orden con coeficientes constantes

Consideremos ahora el caso en que $p(t) = a$ y $q(t) = b$ son constantes:

$$y'' + ay' + by = f(t) \tag{1.73}$$

es decir,

$$L[y] = f(t), \quad L[y] = y'' + ay' + by$$

Método general:
Para obtener una solución particular en el caso de una $f(t)$ general, podemos aplicar el método de variación de parámetros, cuyo resultado final es la ecuación (1.69).

Problema 22: Probar que si las soluciones linealmente independientes son $y_1(t) = e^{\lambda_1 t}$ y $y_2(t) = e^{\lambda_2 t}$, con $\lambda_1 \neq \lambda_2$, la ecuación (1.69) implica la solución particular

$$y_p(t) = \frac{1}{\lambda_1 - \lambda_2}[e^{\lambda_1 t}\int e^{-\lambda_1 t} f(t)dt - e^{\lambda_2 t}\int e^{-\lambda_2 t} f(t)dt] \tag{1.74}$$

Problema 23: Probar que si las soluciones linealmente independientes son $y_1(t) = e^{\lambda t}$ y $y_2(t) = te^{\lambda t}$, entonces

$$y_p(t) = e^{\lambda t}[t\int e^{-\lambda t} f(t)dt - \int te^{-\lambda t} f(t)dt] \tag{1.75}$$

Problema 24: a) Probar que si las soluciones linealmente independientes son $y_1(t) = e^{\alpha t}\cos(\omega t)$ y $y_2(t) = e^{\alpha t}\sin(\omega t)$, $\omega \neq 0$, entonces

$$y_p(t) = \frac{e^{\alpha t}}{\omega}[-\cos(\omega t)\int e^{-\alpha t}\sin(\omega t) f(t)dt + \sin(\omega t)\int e^{-\alpha t}\cos(\omega t) f(t)dt] \tag{1.76}$$

Problema 25: Muestre que la solución particular que satisface $y_p(t_0) = y_p'(t_0) = 0$ puede escribirse en todos los casos como

$$y_p(t) = \int_{t_0}^{t} g(t - t')b(t')dt',$$

con $g(t)$ la solución de la ecuación **homogénea** que satisface $y(0) = 0$, $y'(0) = 1$.

Ejemplo : Resolver

$$y'' + 4y' + 4y = f(t)$$

Proponiendo $y(t) = e^{\lambda t}$ para la solución de la ecuación homogénea se obtiene $\lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0$, o sea $\lambda = -2$. La solución general de la ecuación homogénea es entonces

$$y_h(t) = c_1 e^{-2t} + c_2 t e^{-2t}$$

Tomando $y_1(t) = e^{-2t}$, $y_2(t) = te^{-2t}$, tenemos $W(y_1, y_2) = e^{-4t}$ y aplicando (1.69) o directamente la expresión final (1.75), obtenemos la solución particular

$$y_p(t) = e^{-2t}[t \int e^{2t} f(t) dt - \int t e^{2t} f(t) dt]$$

Por ejemplo, si $f(t) = e^{-2t}$, se obtiene

$$y_p(t) = e^{-2t}[t \int dt - \int t dt] = e^{-2t}[t^2 - t^2/2] = \frac{t^2}{2} e^{-2t}$$

La solución general de la ecuación

$$y'' + 4y' + 4y = e^{-2t}$$

es entonces

$$y(t) = c_1 e^{-2t} + c_2 t e^{-2t} + \frac{t^2}{2} e^{-2t}$$

Si las condiciones iniciales son $y(0) = y'(0) = 0$, se obtiene el sistema $\begin{cases} c_1 + 0 = 0 \\ -2c_1 + c_2 = 0 \end{cases}$ cuya única solución es $c_1 = c_2 = 0$. La única solución particular de la ecuación anterior que satisface esta condición inicial es por lo tanto $y(t) = \frac{t^2}{2} e^{-2t}$.

Ejercicios IV. Halle la solución general de las siguientes ecuaciones diferenciales:

a) $y'' + 2y' + y = e^{-t}$ *b*) $y'' + 2y' + y = te^{-t}$

c) $y'' + y = A\cos(t)$ *d*) $y'' - y = Ae^{-t}$

e) $y'' + 2y' + y = \frac{e^{-t}}{1+t^2}$ *f*) $y'' - y = e^{-t}\cos t$

1.13. Método de coeficientes indeterminados para ecuaciones con coeficientes constantes

En el caso de coeficientes constantes, este método facilita el cálculo de una solución particular de $L[y] = f(t)$ cuando $f(t)$ es exponencial, seno, coseno, polinomio o producto de estas funciones. Consiste en proponer una solución particular $y_p(t)$ del mismo tipo que $f(t)$, si es que tal propuesta no contiene sumandos que sean soluciones de la ecuación homogénea. Damos a continuación una tabla de propuestas básicas de solución particular $y_p(t)$ de $L[y] = f(t)$, válidas para el caso general de orden n con coeficientes constantes.

$f(t)$	$y_p(t)$	Condición
$Ae^{\alpha t}$	$Be^{\alpha t}$	$L[e^{\alpha t}] \neq 0$
$A_0 + A_1 t + \ldots + A_m t^m$	$B_0 + B_1 t + \ldots + B_m t^m$	$L[1] \neq 0$
$e^{\alpha t}(A_0 + A_1 t + \ldots + A_m t^m)$	$e^{\alpha t}(B_0 + B_1 t + \ldots + B_m t^m)$	$L[e^{\alpha t}] \neq 0$
$A \cos \omega t$ o $A \operatorname{sen} \omega t$	$B \cos \omega t + C \operatorname{sen} \omega t$	$L[e^{i\omega t}] \neq 0$
$Ae^{i\omega t}$	$Be^{i\omega t}$	$L[e^{i\omega t}] \neq 0$
$Ae^{\alpha t} \cos \omega t$ o $Ae^{\alpha t} \operatorname{sen} \omega t$	$e^{\alpha t}[B \cos \omega t + C \operatorname{sen} \omega t]$	$L[e^{(\alpha + i\omega t)}] \neq 0$
$Ae^{(\alpha+i\omega)t}$	$Be^{(\alpha+i\omega)t}$	$L[e^{(\alpha+i\omega)t}] \neq 0$

donde A, B, C, etc. son **constantes** a determinar. Si no se cumple la condición indicada, debe multiplicarse la propuesta anterior por t o en general, t^k, con k el menor número natural tal que ningún sumando de $t^k y_p(t)$ sea solución de $L[y] = 0$.

Por la linealidad de L (punto **2.** de la página 23), si $f(t)$ es suma de varios términos del tipo anterior, $y_p(t)$ será la suma de las soluciones particulares para cada término.

El método puede justificarse tanto a partir de la solución general (1.69) (ver página previa) como de la observación que L aplicado a cada una de estas funciones da una función del mismo tipo. Por ejemplo, $L[e^{\alpha t}] = P(\alpha)e^{\alpha t}$, donde $P(\alpha)$ es el polinomio característico (Probar!).

Ejemplo 1:

$$y'' + ay' + by = Ae^{\alpha t} \tag{1.77}$$

Proponiendo $y_p(t) = Be^{\alpha t}$ y reemplazando, se obtiene

$$Be^{\alpha t}[\alpha^2 + a\alpha + b] = Ae^{\alpha t}$$

y por lo tanto,

$$B = \frac{A}{\alpha^2 + a\alpha + b}$$

si $P(\alpha) = \alpha^2 + a\alpha + b \neq 0$, es decir, si α no es raíz del polinomio característico ($L[e^{\alpha t}] \neq 0$). La solución general de (1.77) es entonces

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + \frac{A}{\alpha^2 + a\alpha + b} e^{\alpha t} \qquad (L[e^{\alpha t}] \neq 0)$$

donde y_1 e y_2 son dos soluciones linealmente independientes de la ecuación homogénea.
Importante: Las condiciones iniciales $y(0) = y_0$, $y'(0) = v_0$, se deben ajustar con **la solución completa**, que contiene a la solución particular.

Problema 26: Mostrar que una solución particular de $y'' + y = Ae^{-t}$ es $y_p(t) = \frac{1}{2}Ae^{-t}$.

Si en cambio α es raíz de la ecuación característica ($L[e^{\alpha t}] = 0$), se propone

$$y_p(t) = Bte^{\alpha t}$$

Obtenemos $y_p' = Be^{\alpha t}(1 + \alpha t)$, $y_p'' = Be^{\alpha t}(2\alpha + \alpha^2 t)$ y entonces, reemplazando en (1.77),

$$Be^{\alpha t}[t(\alpha^2 + a\alpha + b) + a + 2\alpha] = Ae^{\alpha t}$$

Como $\alpha^2 + a\alpha + b = 0$, obtenemos

$$B = \frac{A}{a + 2\alpha}$$

si $a + 2\alpha \neq 0$. La solución general de (1.77) en este caso es entonces

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + \frac{A}{a + 2\alpha}\, te^{\alpha t} \qquad (L[e^{\alpha t}] = 0, \;\; L[te^{\alpha t}] \neq 0)$$

Si también $a + 2\alpha = 0$, α será justamente raíz doble de la ecuación característica, en el que $te^{\alpha t}$ es también solución (Probar!). En este caso se debe proponer $y_p = Bt^2 e^{\alpha t}$, o utilizar el método general (1.69).

Problema 27: Hallar la solución general de (1.77) si α es raíz doble ($a^2 = 4b$, $\alpha = -a/2$).

Problema 28: Mostrar que una solución particular de $y'' - y = Ae^{-t}$ es $y_p(t) = -\frac{1}{2}Ate^{-t}$.

Ejemplo 2:

$$y'' + ay' + by = A_0 + A_1 t + A_2 t^2\,, \quad A_2 \neq 0 \tag{1.78}$$

Proponiendo

$$y_p(t) = B_0 + B_1 t + B_2 t^2$$

se obtiene $y_p' = B_1 + 2B_2 t$, $y_p'' = 2B_2$ y reemplazando en (1.78),

$$(2B_2 + aB_1 + bB_0) + (2aB_2 + bB_1)t + bB_2 t^2 = A_0 + A_1 t + A_2 t^2$$

Igualando los coeficientes de igual grado, se llega al sistema

$$\left\{\begin{array}{rcl} bB_0 + aB_1 + 2B_2 & = & A_0 \\ bB_1 + 2aB_2 & = & A_1 \\ bB_2 & = & A_2 \end{array}\right.$$

que es compatible con solución única si $b \neq 0$. Puede resolverse directamente comenzando con la última ecuación.

Importante: El polinomio propuesto debe ser **completo**, aun si $A_0 = 0$ o $A_1 = 0$.

Problema 29: Mostrar que una solución particular de $y''+2y'+y = t^2$ es $y_p(t) = 6-4t+t^2$.

Si en cambio $b = 0$, caso en el que $L[1] = 0$ (probar!) se debe proponer una solución particular

$$y_p(t) = t(B_0 + B_1 t + B_2 t^2)$$

Problema 30: Mostrar que una solución particular de $y''+y' = t^2$ es $y_p(t) = t(2-t+t^2/3)$.

Ejemplo 3:

$$y'' + ay' + by = A\cos\omega t \tag{1.79}$$

Tenemos dos formas de resolverlo: 1) Se propone

$$y_p(t) = B\cos\omega t + C\,\mathrm{sen}\,\omega t$$

Se obtiene $y_p' = \omega(C\cos\omega t - B\,\mathrm{sen}\,\omega t)$, $y_p''(t) = -\omega^2 y_p$ y reemplazando en la ecuación,

$$((b-\omega^2)B + aC\omega)\cos\omega t + ((b-\omega^2)C - aB\omega)\,\mathrm{sen}\,\omega t = B\cos\omega t$$

que conduce al sistema

$$\left\{\begin{array}{l} (b-\omega^2)B + a\omega C = A \\ -a\omega B + (b-\omega^2)C = 0 \end{array}\right.$$

Este posee solución única salvo que $a = 0$ y $\omega^2 = b > 0$ (asumiendo ω real), caso en el que $\cos\omega t$ es solución.
Problema 31: Probar que una solución particular de $y'' + 2y' + y = A\cos t$
es $y_p(t) = \frac{1}{2}A\,\mathrm{sen}\,t$.

2) La segunda forma, asumiendo a, b, A y ω reales, es resolver la ecuación

$$y'' + ay' + by = Ae^{i\omega t}$$

ya que $e^{i\omega t} = \cos\omega t + i\,\mathrm{sen}\,\omega t$. Se propone así la solución particular $Be^{i\omega t}$. Se obtiene, reemplazando en la ecuación,

$$B[-\omega^2 + ai\omega + b]e^{i\omega t} = Ae^{i\omega t}$$

de donde, asumiendo $\omega^2 \neq b$ si $a = 0$,

$$B = \frac{A}{b - \omega^2 + ai\omega}$$

Dado que $\cos\omega t = \mathrm{Re}[e^{i\omega t}]$, la solución particular para $f(t) = A\cos\omega t$ es entonces

$$y_p(t) = \mathrm{Re}[Be^{i\omega t}]$$

Las ventajas de esta segunda forma, que es la utilizada en Ingeniería, son varias:
i) Es necesario determinar un sólo coeficiente: B.
ii) Escribiendo B en forma polar $B = |B|e^{i\phi}$, se obtiene $Be^{i\omega t} = |B|e^{i(\omega t+\phi)}$ y entonces,

$$y_p(t) = \mathrm{Re}[|B|e^{i(\omega t+\phi)}] = |B|\cos(\omega t + \phi)$$

Esta expresión nos da explícitamente tanto la **amplitud** $|B|$ como la **diferencia de fase** ϕ de la solución con respecto a la entrada $A\cos\omega t$, que son justamente los datos buscados de la solución en las aplicaciones prácticas de este tipo de problemas.
iii) La parte imaginaria de la solución particular compleja es una solución particular de

$$y'' + ay' + by = A\,\mathrm{sen}\,\omega t$$

por lo que se resuelven ambos problemas ($L[y] = A\cos\omega t$ y $L[y] = A\,\mathrm{sen}\,\omega t$) simultáneamente con un sólo coeficiente B.
Problema 32: Probar que si $L[y] = f(t)$, con $f(t)$ compleja, y L es real y lineal, entonces

$$L[\mathrm{Re}(y)] = \mathrm{Re}[f(t)], \quad L[\mathrm{Im}(y)] = \mathrm{Im}[f(t)]$$

Problema 33: Mostrar que una solución particular de $y'' + 2y' + y = Ae^{it}$ es $y_p(t) = \frac{1}{2i}Ae^{it} = -\frac{i}{2}Ae^{it}$, y que por lo tanto,
a) $y_p(t) = \frac{1}{2}A \operatorname{sen} t$ es solución particular de $y'' + 2y' + y = A\cos t$
b) $y_p(t) = -\frac{1}{2}A\cos t$ es solución particular de $y'' + 2y' + y = A \operatorname{sen} t$.
Determinar la diferencia de fase de la solución particular con respecto a la entrada.

Problema 34: Dada la ecuación

$$y'' + \omega^2 y = A\cos(\omega t),$$

muestre que es necesario proponer una solución particular de la forma $y_p(t) = t[B\cos\omega t + E \operatorname{sen}\omega t]$ o $Be^{i\omega t}$, y que el resultado es

$$y_p(t) = \frac{1}{2\omega}A \operatorname{sen}\omega t$$

Ejercicios V.
1) Halle la solución general de las siguientes ecuaciones diferenciales.

a) $y'' + 2y' + y = e^{-2t}$	*b*) $y'' + 2y' + y = 4e^{-2t} + t + t^2$
c) $y'' + 4y = A\cos(\omega t), \;\; \omega^2 \neq 4$	*d*) $y'' + 4y = A\cos(2t)$
e) $y'' + 4y' + 5y = Ae^{-2t} + Bte^{-t}$	*f*) $y'' + 4y' + 5y = 4 + 12t - 5t^2$
g) $y'' - 4y = e^{-2t} + 1$	*h*) $y'' + y = e^{-t}\cos t$
i) $y'' = e^{-t} + 1$	*j*) $y''' + y' = e^{-t}$
k) $y'''' - y = e^{2t}$	*i*) $y'' + 2y' + y = e^{-t}$

2) Determine la solución de las ecuaciones a), b), c) y d) que satisface i) $y(0) = y'(0) = 0$ ii) $y(0) = 1$, $y'(0) = 0$. iii) Determine la solución de la ecuación j) que satisface $y(0) = y'(0) = y''(0) = 0$.

3) Si $y(t)$ es solución de $y'' + ay' + by = f(t)$, con a, b constantes, determine la ecuación diferencial que satisface i) $y'(t)$, ii) $y(t - t_0)$, iii) $Ay(t) + C$.

4) a) Muestre que en una ecuación diferencial lineal de segundo orden, si $f(t) \to \alpha f(t)$, la solución particular obtenida por el método general o por el de coeficientes indeterminados, se multiplica por α.
b) ¿ Es válida dicha propiedad para la solución particular de la misma ecuación que satisface condiciones iniciales fijas no nulas $y(0) = y_0$, $y'(0) = v_0$? ¿Qué sucede si $y_0 = v_0 = 0$?

5) Determine la solución general de las siguientes ecuaciones.

a) $y'' + y'/t - y''/t^2 = t^n, \;\; t > 0$ *b*) $y'' + 2y' + y = e^{-t}t^n$ *c*) $y'' + y = \tan(t)$

1.14. Aplicaciones

Problema 34. El oscilador armónico en presencia de una fuerza constante Supongamos que se aplica una fuerza constante F, tal como la fuerza de gravedad $F = -mg$, sobre una masa unida a un resorte. Muestre que la ecuación que describe la posición y (medida a partir de la posición de equilibrio del resorte para $F = 0$) es

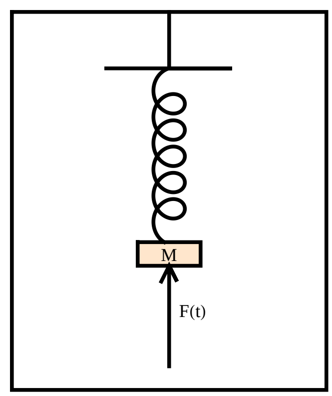
$$y'' + \omega^2 y = f\,, \qquad \omega^2 = k/m > 0, \quad f = F/m > 0$$

a) Muestre que una solución particular de esta ecuación es la constante

$$y_p(t) = y_0, \;\; y_0 = f/\omega^2 = F/k$$

y que y_0 representa la nueva posición de equilibrio en presencia de la fuerza constante.
b) Escriba la solución general de la ecuación, y muestre que el único efecto de la fuerza constante fue desplazar el punto de equilibrio, es decir, el centro de las oscilaciones.
c) Muestre que la ecuación diferencial que satisface $\tilde{y} = y - y_0$ (posición medida a partir del nuevo punto de equilibrio) es idéntica a la ecuación que satisface y para $F = 0$.

Problema 35.
El oscilador armónico forzado. Resonancia.



Considerar la aplicación de una fuerza externa $F(t)$ a una masa unida a un resorte.
a) Mostrar que la ecuación que describe el movimiento es

$$y'' + \omega^2 y = f(t), \quad f(t) = F(t)/m$$

b) Dar una expresión de la solución general para una fuerza continua arbitraria.
c) Considerar ahora una fuerza externa de la forma $F(t) \propto \cos(\omega_{ex} t)$, tal que la ecuación de movimiento es

$$y'' + \omega^2 y = F\cos(\omega_{ex} t)$$

Muestre que si $\omega_{ex} \neq \omega$, la solución general es

$$y(t) = c_1 \cos(\omega t) + c_2 \,\mathrm{sen}\,(\omega t) + \frac{F}{\omega^2 - \omega_{ex}^2}\cos(\omega_{ex} t), \quad (\omega_{ex} \neq \omega)$$

y la solución que satisface $y(0) = 0$, $y'(0) = 0$ es

$$y(t) = \frac{F}{\omega^2 - \omega_{ex}^2}[\cos(\omega_{ex} t) - \cos(\omega t)]\,, \quad (\omega_{ex} \neq \omega) \tag{1.80}$$

Grafique esta solución y discuta su comportamiento (tipo "batido") al acercarse ω_{ex} a ω. Muestre que puede escribirse como

$$y(t) = \frac{2F}{\omega^2 - \omega_{ex}^2}\,\mathrm{sen}\,(\omega_1 t)\,\mathrm{sen}\,(\omega_2 t)$$

con $\omega_1 = (\omega + \omega_{ex})/2$, $\omega_2 = (\omega - \omega_{ex})/2$.

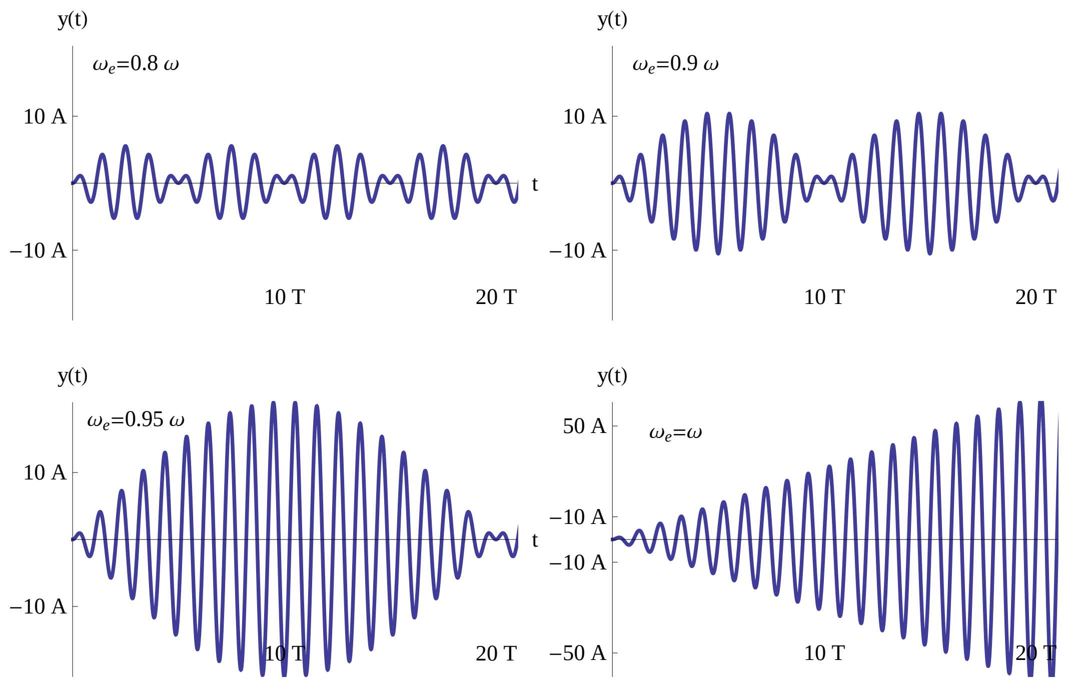
d) Mostrar que si $\omega_{ex} = \omega$ (**resonancia**), la solución general es

$$y(t) = c_1 \cos(\omega t) + c_2 \operatorname{sen}(\omega t) + \frac{A}{2\omega} \operatorname{sen}(\omega t)\,, \qquad (\omega_{ex} = \omega) \tag{1.81}$$

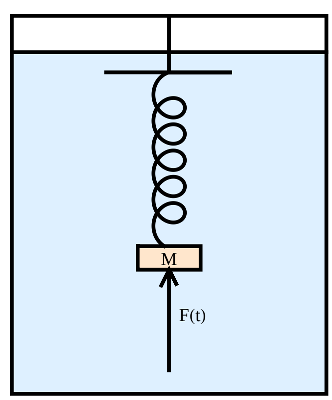
y la solución que satisface $y(0) = y'(0) = 0$ es

$$y(t) = \frac{A}{2\omega} t \operatorname{sen}(\omega t), \qquad (\omega_{ex} = \omega) \tag{1.82}$$

Grafique e interprete esta solución, y discuta el fenómeno de resonancia.
e) Muestre que el límite de la solución (1.80) para $\omega_{ex} \to \omega$ (y t fijo) es la solución (1.81).



Gráficos de la solución (1.80) para frecuencias externas ω_e próximas a la frecuencia propia ω, e intensidad F fija . Nótese el cambio de escala en el caso resonante $\omega_e = \omega$.
$T = 2\pi/\omega$ es el período y $A = F/\omega^2$.



Problema 36.
Oscilador forzado amortiguado.

Examinemos ahora el sistema anterior en presencia de una fuerza de roce viscosa $F_r = -\mu y'$.
a) Mostrar que la ecuación que describe el movimiento en presencia de una fuerza externa $F(t)$ es

$$y'' + 2\gamma y' + \omega^2 y = f(t)$$

donde $f(t) = F(t)/m$ y $\gamma = \mu/(2m) > 0$.

b) Mostrar que para

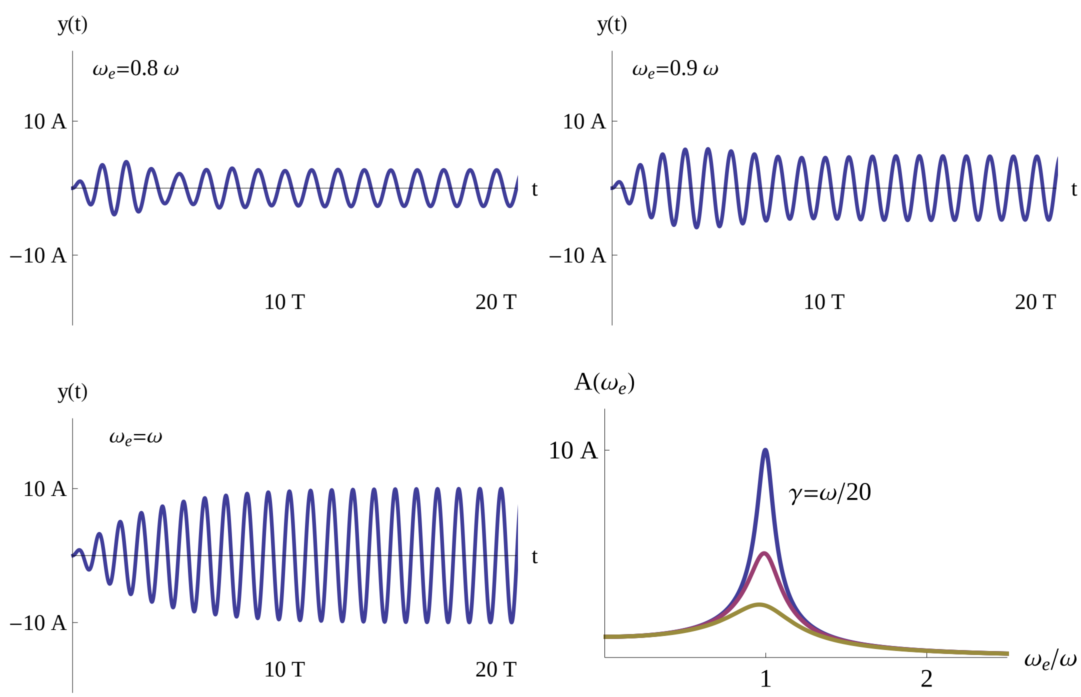
$$f(t) = F\cos(\omega_{ex}t)$$

una solución particular es

$$y_p(t) = F\frac{(\omega^2 - \omega_{ex}^2)\cos(\omega_{ex}t) + 2\gamma\omega_{ex}\,\mathrm{sen}\,(\omega_{ex}t)}{(\omega^2 - \omega_{ex}^2)^2 + 4\gamma^2\omega_{ex}^2} \tag{1.83}$$

$$= \frac{F}{\sqrt{(\omega^2 - \omega_{ex}^2)^2 + 4\gamma^2\omega_{ex}^2}}\cos(\omega_{ex}t + \phi) \tag{1.84}$$

Dado que la solución general de la ecuación homogénea disminuye ahora exponencialmente al aumentar t, para tiempos grandes (mayores que el tiempo de relajación) **sólo subsiste la solución particular.** Para $\omega = \omega_{ex}$, la amplitud de la oscilación resultante crece inicialmente pero se estabiliza rápidamente en el valor determinado por la solución particular.
Es también muy importante destacar que el sistema termina oscilando con la frecuencia **externa** y no la frecuencia propia.
Haga un gráfico de la amplitud de la oscilación resultante y muestre que disminuye al aumentar el coeficiente de roce α, y es máxima para ω_{ex} cercano a ω.



Gráficos de la solución en presencia de roce viscoso con $\gamma = \omega/20$ para $y(0) = y'(0) = 0$. La última figura (1.80) muestra la amplitud de la solución particular en función de la frecuencia externa para $\gamma = \omega/20$, $\omega/10$ y $\omega/5$. Es máxima para $\omega_{ex} \approx \omega$.

Problema 37. Discuta el fénomeno de resonancia y el efecto de la resistencia ("roce") para el circuito LCR en serie con una fuente que suministra una fem $V(t) = V\cos(\omega_{ex}t)$.

La ecuación no homogénea de orden n.
El método (1.33) para obtener una solución particular se puede extender al caso general

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = f(t) \tag{1.85}$$

Si $y_1(t)$, ..., $y_n(t)$ son n soluciones linealmente independientes de la ecuación homogénea, se puede proponer una solución particular de la forma

$$y_p(t) = v_1(t)y_1(t) + \ldots + v_n(t)y_n(t)$$

con $v_1(t), \ldots, v_n(t)$ funciones a determinar. Imponiendo las $n-1$ condiciones $v_1'y_1^{(k)} + \ldots + v_n'y_n^{(k)} = 0 \ \forall \ t \in I$, $k = 0, \ldots, n-2$, y reemplazando en (1.85), se obtiene el sistema

$$\left\{ \begin{array}{rcrcl} v_1'y_1+ & \ldots & +v_n'y_n & = & 0 \\ & & & \vdots & \\ v_1'y_1^{(n-2)}+ & \ldots & +v_n'y_n^{(n-2)} & = & 0 \\ v_1'y_1^{(n-1)}+ & \ldots & +v_n'y_n^{(n-1)} & = & f(t) \end{array} \right.$$

cuya solución es (probar!)

$$v_i(t) = \int (-1)^{i+n} \frac{W_i(t)}{W(t)} f(t)dt, \quad i = 1, \ldots, n$$

donde $W(t) = \begin{vmatrix} y_1 & \ldots & y_n \\ & \ddots & \\ y_1^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix}$, es el Wronskiano y $W_i(t)$ el determinante de la matriz obtenida suprimiendo la fila n y la columna i de la matriz que define a $W(t)$.

2. Sistemas de ecuaciones diferenciales ordinarias

Un sistema de ecuaciones diferenciales ordinarias de primer orden es un sistema con n funciones incógnitas $y_1(t), y_2(t), \ldots, y_n(t)$, que deben satisfacer las n ecuaciones

$$\left\{ \begin{array}{c} y_1' = f_1(t, y_1, \ldots, y_n) \\ \vdots \\ y_n' = f_n(t, y_1, \ldots, y_n) \end{array} \right. \tag{2.1}$$

Las n ecuaciones diferenciales están en general *acopladas*, ya que f_i depende no sólo de y_i sino también de las funciones restantes.

Podemos escribir el sistema (2.1) en forma vectorial como

$$\boldsymbol{Y}' = \boldsymbol{F}(t, \boldsymbol{Y}), \quad \boldsymbol{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{F}(t, \boldsymbol{Y}) = \begin{pmatrix} f_1(t, \boldsymbol{Y}) \\ \vdots \\ f_n(t, \boldsymbol{Y}) \end{pmatrix} \tag{2.2}$$

Si ninguna de las funciones f_i depende explícitamente de la variable independiente t, el sistema se denomina **autónomo**: Es de la forma $\boldsymbol{Y}' = \boldsymbol{F}(\boldsymbol{Y})$.

Toda ecuación diferencial ordinaria de orden superior puede expresarse como un sistema de ecuaciones ordinarias de primer orden.

Una ecuación diferencial ordinaria de orden n,

$$y^{(n)} = f(t, y, y', \ldots, y^{(n-1)}) \tag{2.3}$$

puede escribirse como un sistema de n ecuaciones diferenciales de **primer orden:** Introduciendo las funciones

$$y_1(t) = y(t), \quad y_2(t) = y'(t), \quad \ldots, \quad y_n(t) = y^{(n-1)}(t) \tag{2.4}$$

vemos que la ecuación (2.3) es equivalente al sistema de ecuaciones de primer orden

$$\left\{ \begin{array}{lcl} y_1' & = & y_2 \\ y_2' & = & y_3 \\ & \vdots & \\ y_{n-1}' & = & y_n \\ y_n' & = & f(t, y_1, y_2, \ldots, y_n) \end{array} \right. \tag{2.5}$$

ya que $y_1'(t) = y'(t) = y_2(t)$, $y_2'(t) = y''(t) = y_3(t)$, ..., $y_n'(t) = y^{(n)}(t)$.

Por ejemplo, la ecuación diferencial de segundo orden

$$y'' = f(t, y, y')$$

es equivalente al sistema de dos ecuaciones de primer orden

$$\left\{\begin{array}{l} y_1' = y_2 \\ y_2' = f(t, y_1, y_2) \end{array}\right.$$

donde $y_1 = y$ y $y_2 = y'$. Si $y(t)$ representa una posición, $y_2(t) = y'(t)$ es la velocidad.
Problema 1: Escribir la ecuación $y'' + y = t$ como un sistema de primer orden.

De la misma manera, un sistema de m ecuaciones de orden n,

$$\boldsymbol{Y}^{(n)} = \boldsymbol{F}(t, \boldsymbol{Y}, \boldsymbol{Y}', \ldots, \boldsymbol{Y}^{(n-1)})$$

donde $\boldsymbol{Y} = (y_1, \ldots, y_m)$, $\boldsymbol{F} = (f_1, \ldots, f_m)$, puede escribirse, introduciendo las funciones

$$\boldsymbol{Z}_1 = \boldsymbol{Y}, \quad \boldsymbol{Z}_2 = \boldsymbol{Y}', \quad \ldots, \quad \boldsymbol{Z}_n = \boldsymbol{Y}^{(n-1)}$$

como un sistema de $\boldsymbol{n \times m}$ ecuaciones de primer orden:

$$\left\{\begin{array}{lcl} \boldsymbol{Z}_1' & = & \boldsymbol{Z}_2 \\ \boldsymbol{Z}_2' & = & \boldsymbol{Z}_3 \\ & \vdots & \\ \boldsymbol{Z}_{n-1}' & = & \boldsymbol{Z}_n \\ \boldsymbol{Z}_n' & = & \boldsymbol{F}(t, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_n) \end{array}\right. \tag{2.6}$$

para $\boldsymbol{Z}_1(t), \ldots, \boldsymbol{Z}_n(t)$ ($n \times m$ funciones incógnitas en total).

Por ejemplo, la segunda ley de Newton para el movimiento en tres dimensiones de una partícula de masa $m > 0$,

$$m\frac{d^2\boldsymbol{r}}{dt^2} = \boldsymbol{F}(t, \boldsymbol{r}, \frac{d\boldsymbol{r}}{dt})$$

donde $\boldsymbol{F}(t, \boldsymbol{r}, d\boldsymbol{r}/dt)$ es la fuerza neta que actúa sobre la misma, es un sistema de tres ecuaciones diferenciales de 2^o orden para las coordenadas x, y, z de $\boldsymbol{r} = (x, y, z)$. Introduciendo la velocidad $\boldsymbol{v} = d\boldsymbol{r}/dt$, resulta equivalente al sistema de primer orden

$$\left\{\begin{array}{rcl} \frac{d\boldsymbol{r}}{dt} & = & \boldsymbol{v} \\ m\frac{d\boldsymbol{v}}{dt} & = & \boldsymbol{F}(t, \boldsymbol{r}, \boldsymbol{v}) \end{array}\right.$$

que es un sistema de seis ecuaciones diferenciales de primer orden para las seis funciones $x(t), y(t), z(t), v_x(t), v_y(t), v_z(t)$.

Por lo tanto, en lo sucesivo consideraremos sólo sistemas de primer orden.

Teorema de Existencia y Unicidad de la Solución

Consideremos el sistema de n ecuaciones de primer orden con **condiciones iniciales:**

$$\left\{\begin{array}{lcl} \boldsymbol{Y}' & = & \boldsymbol{F}(t, \boldsymbol{Y}) \\ \boldsymbol{Y}(t_0) & = & \boldsymbol{Y}_0 \end{array}\right. \tag{2.7}$$

donde $\boldsymbol{Y}_0 = (y_{10}, \ldots, y_{n0})^t$. Si las n funciones f_i de $\boldsymbol{F}$ son continuas en una región $R = \{(t, \boldsymbol{Y}) \mid |t - t_0| \leq a,\ |\boldsymbol{Y} - \boldsymbol{Y}_0| \leq b,\ a > 0,\ b > 0\} \subset \mathbb{R}^{n+1}$, y si las derivadas parciales $\partial f_i/\partial y_j$ están acotadas en dicha región $\forall\ i, j$, **existe una única solución $\boldsymbol{Y}(t)$** en un cierto intervalo I centrado en t_0, que satisface la condición inicial $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$.

2.1. Sistemas lineales de primer orden

Un sistema de n ecuaciones diferenciales de primer orden es lineal si las n funciones $f_i(t, \boldsymbol{Y})$ son funciones lineales de y_1, y_2, ..., y_n (aunque no necesariamente de t). Un sistema lineal de primer orden puede escribirse en la forma

$$\left\{ \begin{array}{rcl} y_1' &=& a_{11}(t)y_1 + \ldots a_{1n}(t)y_n + f_1(t) \\ &\vdots& \\ y_n' &=& a_{n1}(t)y_1 + \ldots + a_{nn}(t)y_n + f_n(t) \end{array} \right. \tag{2.8}$$

con $a_i(t)$, $f_i(t)$, $i = 1, \ldots, n$, definidas en un cierto intervalo I. El sistema es **homogéneo** si $f_i(t) = 0 \ \forall \ t \in I$ para $i = 1, \ldots, n$.
Podemos escribir (2.8) en forma **matricial** como

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \ldots & a_{nn}(t) \\ & \ddots & \\ a_{n1}(t) & \ldots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} \tag{2.9}$$

es decir,

$$\boldsymbol{Y}' = \boldsymbol{A}(t)\boldsymbol{Y} + \boldsymbol{F}(t) \tag{2.10}$$

con $\boldsymbol{Y}(t)$, $\boldsymbol{F}(t)$ vectores columna de $1 \times n$ y $\boldsymbol{A}(t)$ una matriz de $n \times n$.

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{A}(t) = \begin{pmatrix} a_{11}(t) & \ldots & a_{nn}(t) \\ & \ddots & \\ a_{n1}(t) & \ldots & a_{nn}(t) \end{pmatrix}, \quad \boldsymbol{F}(t) = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \tag{2.11}$$

En este caso el teorema de existencia y unicidad asegura, si todos los elementos $a_{ij}(t)$ y $f_i(t)$ son funciones **continuas** en un intervalo I, que el problema de condiciones iniciales

$$\left\{ \begin{array}{rcl} \boldsymbol{Y}' &=& \boldsymbol{A}(t)\boldsymbol{Y} + \boldsymbol{F}(t) \\ \boldsymbol{Y}(t_0) &=& \boldsymbol{Y}_0 \end{array} \right. \tag{2.12}$$

tendrá solución **única** $\forall \ \boldsymbol{Y}_0 \in \mathbb{R}^n$, $\forall \, t_0 \in I$.
Problema 2: Mostrar que la unicidad implica que si $\boldsymbol{Y}_1(t)$, $\boldsymbol{Y}_2(t)$ son dos soluciones de (2.10) y cumplen $\boldsymbol{Y}_1(t_0) \neq \boldsymbol{Y}_2(t_0)$, entonces $\boldsymbol{Y}_1(t) \neq \boldsymbol{Y}_2(t) \ \forall \ t \in I$.

Ejemplo: El sistema

$$\left\{ \begin{array}{rcl} x' &=& x - y \\ y' &=& -x + y + 2t \end{array} \right. \tag{2.13}$$

es un sistema de dos ecuaciones diferenciales lineales de primer orden para las funciones incógnitas $x(t)$, $y(t)$. Podemos escribirlo en forma matricial como

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 2t \end{pmatrix} \tag{2.14}$$

o sea, $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{F}(t)$, con

$$\boldsymbol{Y} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \boldsymbol{F} = \begin{pmatrix} 0 \\ 2t \end{pmatrix}$$

2.2. Sistema lineal homogéneo de primer orden

Consideremos ahora el sistema **homogéneo**

$$\boldsymbol{Y}' = \boldsymbol{A}(t)\boldsymbol{Y} \tag{2.15}$$

Asumiremos que todos elementos $a_{ij}(t)$ de $\boldsymbol{A}(t)$ son funciones continuas para $t \in I$.

Teorema 1. El conjunto S de soluciones de un sistema de n ecuaciones diferenciales lineales de primer orden homogéneo es un espacio vectorial de **dimensión $\boldsymbol{n}$**.
Que sea un espacio vectorial implica que si $\boldsymbol{Y}_1(t)$ e $\boldsymbol{Y}_2(t)$ son dos soluciones de (2.15), **toda** combinación lineal

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + c_2\boldsymbol{Y}_2(t)$$

con c_1, c_2 constantes, es también solución de (2.15) (propiedad de **superposición**). Y que sea de dimensión n, implica que existen n soluciones linealmente independientes
$\boldsymbol{Y}_1(t)$, …, $\boldsymbol{Y}_n(t)$ tales que **toda** solución de (2.15) puede escribirse como

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + \ldots + c_n\boldsymbol{Y}_n(t) \tag{2.16}$$

El conjunto $\{\boldsymbol{Y}_1(t), \ldots, \boldsymbol{Y}_n(t)\}$ es una base de S y se denomina sistema **fundamental** de soluciones. La expresión (2.16) es la **solución general** de (2.15).

Demostración: La solución nula $\boldsymbol{Y}(t) = \boldsymbol{0}\ \forall\ t \in I$ (solución trivial) es solución de (2.15), pues $\boldsymbol{Y}'(t) = \boldsymbol{0} = \boldsymbol{A}\boldsymbol{0} = \boldsymbol{0}$. Y si $\boldsymbol{Y}_1(t)$ e $\boldsymbol{Y}_2(t)$ son soluciones, $c_1\boldsymbol{Y}_1(t) + c_2\boldsymbol{Y}_2(t)$ es también solución pues

$$(c_1\boldsymbol{Y}_1 + c_2\boldsymbol{Y}_2)' = c_1\boldsymbol{Y}_1' + c_2\boldsymbol{Y}_2' = c_1\boldsymbol{A}(t)\boldsymbol{Y}_1 + c_2\boldsymbol{A}(t)\boldsymbol{Y}_2 = \boldsymbol{A}(t)(c_1\boldsymbol{Y}_1 + c_2\boldsymbol{Y}_2)$$

S es pues cerrado bajo la operaciones de suma de funciones y producto por escalar, siendo entonces un subespacio del espacio de funciones vectoriales $\boldsymbol{F} : I \Rightarrow \mathbb{R}^n$ derivables.
Podemos también escribir (2.15) como

$$L[\boldsymbol{Y}] = \boldsymbol{0}\,, \qquad L[\boldsymbol{Y}] = \boldsymbol{Y}' - \boldsymbol{A}(t)\boldsymbol{Y}$$

con L un operador **lineal:** $L[c_1\boldsymbol{Y}_1 + c_2\boldsymbol{Y}_2] = c_1L[\boldsymbol{Y}_1] + c_2L[\boldsymbol{Y}_2]$ (Probar!).
El conjunto de soluciones $S = \{\boldsymbol{Y}(t) \mid L[\boldsymbol{Y}] = \boldsymbol{0}\}$ es entonces el núcleo de L y es por lo tanto un espacio vectorial.

Demostremos ahora que la dimensión de S es n. Consideremos n condiciones iniciales $\boldsymbol{Y}_1^0$, …, $\boldsymbol{Y}_n^0 \in \mathbb{R}^n$, **linealmente independientes**, y sean $\boldsymbol{Y}_1(t), \ldots, \boldsymbol{Y}_n(t)$ las soluciones de (2.15) que satisfacen las condiciones iniciales

$$\boldsymbol{Y}_1(t_0) = \boldsymbol{Y}_1^0, \ldots, \boldsymbol{Y}_n(t_0) = \boldsymbol{Y}_n^0 \tag{2.17}$$

Estas soluciones **existen y son únicas** por el teorema anterior. Además, son **linealmente independientes**, ya que si existen c_1, …, c_n tales que

$$c_1\boldsymbol{Y}_1(t) + \ldots + c_n\boldsymbol{Y}_n(t) = \boldsymbol{0} \tag{2.18}$$

$\forall\ t \in I$, para $t = t_0$ (2.18) implica $c_1\boldsymbol{Y}_1^0 + \ldots + c_n\boldsymbol{Y}_n^0 = \boldsymbol{0}$, por lo que $c_1 = \ldots = c_n = 0$, por ser $\{\boldsymbol{Y}_1^0, \ldots, \boldsymbol{Y}_n^0\}$ un conjunto linealmente independiente de vectores de $\mathbb{R}^n$. La dimensión de S es entonces no menor que n.

Además, como vectores de $\mathbb{R}^n$, $\boldsymbol{Y}_1(t)$, ... $\boldsymbol{Y}_n(t)$ permanecen **linealmente independientes** $\forall\ t \in I$, ya que si la suma (2.18) se anula para algún $t \in I$, debe coincidir, por unicidad, con la solución trivial $\boldsymbol{Y}(t) = \mathbf{0}\ \forall\ t \in I$, incluyendo $t = t_0$, y por lo tanto $c_1 = \ldots = c_n = 0$.

Consideremos ahora una solución cualquiera $\boldsymbol{Y}(t)$ de (2.15), que satisface $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$. Como el conjunto de vectores iniciales $\{\boldsymbol{Y}_1^0, \ldots, \boldsymbol{Y}_n^0\}$ es una base de $\mathbb{R}^n$ (pues son n vectores linealmente independientes), el valor inicial $\boldsymbol{Y}_0$ puede escribirse como

$$\boldsymbol{Y}_0 = c_1\boldsymbol{Y}_1^0 + \ldots c_n\boldsymbol{Y}_n^0$$

Por lo tanto, $\boldsymbol{Y}(t)$ debe coincidir, por el teorema de unicidad, con la combinación lineal

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + \ldots c_n\boldsymbol{Y}_n(t) \tag{2.19}$$

ya que (2.19) satisface $\boldsymbol{Y}(t_0) = c_1\boldsymbol{Y}_1^0 + \ldots + c_n\boldsymbol{Y}_n^0 = \boldsymbol{Y}_0$ y es solución del sistema por ser combinación lineal de soluciones. Esto implica que la dimensión de S es n y no mayor.

Si $\{\boldsymbol{Y}_1(t), \ldots, \boldsymbol{Y}_n(t)\}$ son n soluciones **linealmente independientes** de (2.15) ($\boldsymbol{Y}_i' = \boldsymbol{A}(t)\boldsymbol{Y}_i$, $i = 1, \ldots, n$), la **matriz fundamental** de soluciones se define como

$$\boldsymbol{M}(t) = (\boldsymbol{Y}_1(t), \ldots, \boldsymbol{Y}_n(t)) \tag{2.20}$$

tal que la columna i de $\boldsymbol{M}(t)$ es la solución $\boldsymbol{Y}_i(t)$. Es de $n \times n$ y satisface

$$\boldsymbol{M}' = \boldsymbol{A}(t)\boldsymbol{M} \tag{2.21}$$

Como los vectores $\{\boldsymbol{Y}_1(t), \ldots, \boldsymbol{Y}_n(t)\}$ son linealmente independientes $\forall\ t \in I$, $\boldsymbol{M}(t)$ es **no singular**: $\det[\boldsymbol{M}(t)] \neq 0\ \forall\ t \in I$. La solución **general** (2.16) puede así escribirse como

$$\boldsymbol{Y}(t) = \boldsymbol{M}(t)\boldsymbol{C}\,, \qquad \boldsymbol{C} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Ejemplo: Consideremos el sistema

$$\begin{cases} x' &= \quad x - y \\ y' &= -x + y \end{cases} \tag{2.22}$$

es decir, $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$, con $\boldsymbol{Y} = \binom{x}{y}$, $\boldsymbol{A} = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$, Es fácil ver que

$$\boldsymbol{Y}_1(t) = \begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix}, \quad \boldsymbol{Y}_2(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

son dos soluciones L.I. de (2.22) (probar!). La solución general de (2.22) es entonces

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + c_2\boldsymbol{Y}_2(t) = c_1\begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

es decir, $x(t) = c_1e^{2t} + c_2$, $y(t) = -c_1e^{2t} + c_2$, y la matriz fundamental asociada es

$$\boldsymbol{M}(t) = \begin{pmatrix} e^{2t} & 1 \\ -e^{2t} & 1 \end{pmatrix}$$

2.3. Sistema lineal homogéneo con coeficientes constantes

Consideremos ahora el caso, de gran importancia práctica, en el que todos los coeficientes de la matriz A son **constantes**, es decir, independientes de t ($a_{ij}(t) = a_{ij}\ \forall\ i, j$):

$$\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} \tag{2.23}$$

o en forma explícita,

$$\begin{cases} y_1' & = & a_{11}y_1 + \ldots + a_{1n}y_n \\ & \vdots & \\ y_n' & = & a_{n1}y_1 + \ldots + a_{nn}y_n \end{cases}$$

Al no depender explícitamente del tiempo, el sistema es **autónomo**. Podemos entonces tomar $I = \mathbb{R}$ y el teorema general asegura la existencia de n soluciones linealmente independientes de dicho sistema válidas $\forall\ t \in \mathbb{R}$.

Estos sistemas pueden además resolverse en forma **exacta**.

> El sistema lineal homogéneo (2.23) admite una solución no nula de la forma
>
> $$\boldsymbol{Y}(t) = e^{\lambda t}\boldsymbol{V}, \quad \boldsymbol{V} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \tag{2.24}$$
>
> con λ y $\boldsymbol{V}$ independientes de t, si y sólo si λ es **autovalor** de $\boldsymbol{A}$ y $\boldsymbol{V}$ un **autovector asociado**.

Demostración: Proponiendo una solución de la forma (2.24), tenemos $\mathbf{Y}' = \lambda e^{\lambda t}\boldsymbol{V}$ y reemplazando en (2.23) obtenemos

$$\lambda e^{\lambda t}\boldsymbol{V} = \boldsymbol{A}e^{\lambda t}\boldsymbol{V}$$

lo que conduce a

$$\boldsymbol{A}\boldsymbol{V} = \lambda\boldsymbol{V}$$

Si exigimos $\boldsymbol{V} \neq 0$ (solución no trivial) esta igualdad implica que λ debe ser **autovalor** de $\mathbf{A}$ ($\text{Det}[\mathbf{A} - \lambda I] = 0$) y $\boldsymbol{V}$ un **autovector asociado**.

Si buscamos ahora la **solución general,** tenemos dos casos:

a) La matriz $\boldsymbol{A}$ es **diagonalizable**. En este caso existen n autovectores linealmente independientes $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n$, asociados a n autovalores $\lambda_1, \ldots, \lambda_n$ (no necesariamente distintos):

$$\boldsymbol{A}\boldsymbol{V}_i = \lambda_i \boldsymbol{V}_i, \quad i = 1, \ldots, n$$

Por lo tanto, un sistema fundamental de soluciones de (2.23) es

$$\boldsymbol{Y}_1(t) = e^{\lambda_1 t}\boldsymbol{V}_1, \ldots, \boldsymbol{Y}_n(t) = e^{\lambda_n t}\boldsymbol{V}_n$$

dado que son n soluciones **linealmente independientes**.
Toda solución de (2.23) es entonces de la forma

$$\boldsymbol{Y}(t) = c_1 e^{\lambda_1 t}\boldsymbol{V}_1 + \ldots + c_n e^{\lambda_n t}\boldsymbol{V}_n \tag{2.25}$$

Esta expresión es la **solución general** del sistema.

Recordemos que este es, por ejemplo, el caso de:
i) Matrices $\boldsymbol{A}$ de $n \times n$ que poseen $\boldsymbol{n}$ **autovalores distintos**.
ii) Matrices **reales simétricas** ($\boldsymbol{A}^t = \boldsymbol{A}$) o **complejas hermíticas** ($\bar{\boldsymbol{A}}^t = \boldsymbol{A}$), aun con autovalores repetidos. Estas matrices tienen además todos los autovalores reales.
iii) Matrices reales antisimétricas ($\boldsymbol{A}^t = -\boldsymbol{A}$) u ortonormales ($\boldsymbol{A}^{-1} = \boldsymbol{A}^t$), o en general toda matriz que satisfaga $\bar{\boldsymbol{A}}^t\boldsymbol{A} = \boldsymbol{A}\bar{\boldsymbol{A}}^t$, aun con autovalores repetidos. Los autovalores de estas matrices son en general complejos.

Ejemplo: Consideremos nuevamente el sistema (2.22),

$$\left\{ \begin{array}{lcl} x' & = & x - y \\ y' & = & -x + y \end{array} \right. \tag{2.26}$$

que corresponde a

$$\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

Esta matriz es real y simétrica. Por lo tanto, sabemos que será **diagonalizable**. Se deja como ejercicio probar que los autovalores y autovectores asociados son

$$\lambda_1 = 2, \quad \boldsymbol{V}_1 = \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \ \alpha \neq 0, \qquad \lambda_2 = 0, \quad \boldsymbol{V}_2 = \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \ \beta \neq 0$$

Por lo tanto, $\boldsymbol{Y}_1(t) = e^{2t}\binom{1}{-1}$, $\boldsymbol{Y}_2(t) = e^{0t}\binom{1}{1} = \binom{1}{1}$ (constante) forman un sistema fundamental de soluciones linealmente independientes y la solución general es

$$\boldsymbol{Y}(t) = c_1 e^{2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

es decir, $x(t) = c_1 e^{2t} + c_2$, $y(t) = -c_1 e^{2t} + c_2$.
Problema 7: Mostrar que en el caso diagonalizable, una matriz fundamental es

$$\boldsymbol{M}(t) = (e^{\lambda_1 t}\boldsymbol{V}_1, \ldots, e^{\lambda_n t}\boldsymbol{V}_n)$$

y su determinante

$$\det \boldsymbol{M}(t) = e^{(\lambda_1 + \ldots + \lambda_n)t}\, \mathrm{Det}\, \boldsymbol{M}(0)$$

verificándose que $\det \boldsymbol{M}(t) \neq 0 \ \forall \ t$ si los vectores $\{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n\}$ son L.I.

b) La matriz A es **no diagonalizable**. En este caso $\boldsymbol{A}$ posee $k < n$ autovectores linealmente independientes $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_k$, asociados a k autovalores $\lambda_1, \ldots, \lambda_k$ (no necesariam. distintos). El método previo proporciona aquí sólo $k < n$ soluciones linealmente independientes

$$\boldsymbol{Y}_i(t) = e^{\lambda_i t}\boldsymbol{V}_i, \quad i = 1, \ldots, k$$

Dejaremos para el final el tratamiento completo de este caso. Mencionamos aquí que las restantes soluciones linealmente independientes son de la forma $e^{\lambda t}\boldsymbol{V}(t)$, con λ autovalor de $\boldsymbol{A}$ de multiplicidad algebraica $m > 1$ y $\boldsymbol{V}(t)$ un polinomio en t de grado menor que m.

Tanto en el caso diagonalizable como no diagonalizable, debemos tener en cuenta la posibilidad de que existan autovalores complejos!

Autovalores complejos.
Los autovalores λ pueden ser complejos, aún si la matriz $\boldsymbol{A}$ es real! Si

$$\lambda = \alpha + i\omega$$

con $\omega \neq 0$ y α, ω reales, debemos aplicar, como vimos previamente, la fórmula de Euler

$$e^{(\alpha+i\omega)t} = e^{\alpha t}[\cos(\omega t) + i\sin(\omega t)]$$

Si $\boldsymbol{A}$ es **real**, los autovalores complejos aparecerán en **pares conjugados** con autovectores **conjugados:**

$$\boldsymbol{A}\boldsymbol{V} = \lambda\boldsymbol{V}, \quad \boldsymbol{A}\overline{\boldsymbol{V}} = \overline{\lambda}\,\overline{\boldsymbol{V}}$$

donde $\overline{\lambda} = \alpha - i\omega$, $\boldsymbol{V} = \boldsymbol{U} + i\boldsymbol{W}$, $\overline{\boldsymbol{V}} = \boldsymbol{U} - i\boldsymbol{W}$ y $\boldsymbol{V}$, $\boldsymbol{W}$ reales. Si $\omega \neq 0$,

$$\boldsymbol{Y}(t) = e^{\lambda t}\boldsymbol{V}, \quad \overline{\boldsymbol{Y}}(t) = e^{\overline{\lambda} t}\overline{\boldsymbol{V}}$$

formarán un par de soluciones complejas linealmente independientes (pues $\lambda \neq \overline{\lambda}$).

Para obtener un par de soluciones **reales** linealmente independientes, basta con tomar las partes **real** e **imaginaria** de una de las dos, por ejemplo

$$\boldsymbol{Y}_1(t) = \mathrm{Re}[\boldsymbol{Y}(t)] = \frac{\boldsymbol{Y}(t) + \overline{\boldsymbol{Y}}(t)}{2}, \quad \boldsymbol{Y}_2(t) = \mathrm{Im}[\boldsymbol{Y}(t)] = \frac{\boldsymbol{Y}(t) - \overline{\boldsymbol{Y}}(t)}{2i}$$

Un par de autovalores conjugados origina pues un par de soluciones linealmente independientes, que pueden elegirse siempre reales si la matriz $\boldsymbol{A}$ es real!
El significado de autovalores complejos es el de soluciones oscilatorias, con una amplitud que crece ($\alpha > 0$) o decrece ($\alpha < 0$) exponencialmente o permanece constante ($\alpha = 0$).

Problema 8: Probar que si $\boldsymbol{A}$ es real y $\boldsymbol{Y}(t)$ es una solución compleja del sistema homogéneo $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$, las partes real $\mathrm{Re}[\boldsymbol{Y}(t)]$ e imaginaria $\mathrm{Im}[\boldsymbol{Y}(t)]$ son también soluciones de dicho sistema.
Problema 9: Probar que si $\boldsymbol{Y}(t)$ y $\overline{\boldsymbol{Y}}(t)$ son soluciones linealmente independientes, entonces $\mathrm{Re}[\boldsymbol{Y}(t)]$ y $\mathrm{Im}[\boldsymbol{Y}(t)]$ son también linealmente independientes.

Ejemplo: Consideremos el sistema,

$$\left\{\begin{array}{lcl} x' & = & x-y \\ y' & = & x+y \end{array}\right. \tag{2.27}$$

que corresponde a

$$\boldsymbol{A}=\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Esta matriz es real antisimétrica, y por lo tanto diagonalizable aunque con autovalores complejos. Los autovalores y autovectores asociados son

$$\lambda_1=1+i,\quad \boldsymbol{V}_1=\alpha\begin{pmatrix} 1 \\ -i \end{pmatrix},\ \alpha\neq 0,\qquad \lambda_2=1-i,\quad \boldsymbol{V}_2=\beta\begin{pmatrix} 1 \\ i \end{pmatrix},\ \ \beta\neq 0$$

verificándose que $\lambda_2=\overline{\lambda}_1$, $\overline{\boldsymbol{V}}_2=\boldsymbol{V}_1$. Por lo tanto,

$$\boldsymbol{Y}(t)=e^{(1+i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix},\ \ \overline{\boldsymbol{Y}}(t)=e^{(1-i)t}\begin{pmatrix} 1 \\ i \end{pmatrix}$$

son un par de soluciones complejas linealmente independientes, y

$$\boldsymbol{Y}_1(t)=\mathrm{Re}[e^{(1+i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix}]=\begin{pmatrix} e^t\cos t \\ e^t\sin t \end{pmatrix},\quad \boldsymbol{Y}_2(t)=\mathrm{Im}[e^{(1+i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix}]=\begin{pmatrix} e^t\sin t \\ -e^t\cos t \end{pmatrix}$$

un par de soluciones **reales** linealmente independientes del mismo sistema.
La solución general puede entonces expresarse en forma real como

$$\boldsymbol{Y}(t)=c_1\boldsymbol{Y}_1(t)+c_2\boldsymbol{Y}_2(t)=c_1\begin{pmatrix} e^t\cos t \\ e^t\sin t \end{pmatrix}+c_2\begin{pmatrix} e^t\sin t \\ -e^t\cos t \end{pmatrix}$$

es decir, $x(t)=e^t(c_1\cos t+c_2\sin t)$, $y(t)=e^t(c_1\sin t-c_2\cos t)$.

Problema 10: Mostrar que la solución general puede expresarse también como

$$\boldsymbol{Y}(t)=c_+e^{(1+i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix}+c_-e^{(1-i)t}\begin{pmatrix} 1 \\ i \end{pmatrix}$$

donde $c_\pm=(c_1\mp ic_2)/2$.
Problema 11: Mostrar que la solución general puede expresarse también como

$$\boldsymbol{Y}(t)=Ae^t\begin{pmatrix} \cos(t+\phi) \\ \sin(t+\phi) \end{pmatrix}=Ae^t\mathrm{Re}[e^{i(t+\phi)}\begin{pmatrix} 1 \\ -i \end{pmatrix}]$$

donde $A^2=c_1^2+c_2^2$, $\tan\phi=-c_2/c_1$. Interpretar esta expresión.

Problema 12: Probar que si $\boldsymbol{Y}(t)=e^{(\alpha+i\omega)t}(\boldsymbol{U}+i\boldsymbol{W})$, con $\boldsymbol{U}$ y $\boldsymbol{W}$ reales,

$$\mathrm{Re}[\boldsymbol{Y}(t)]=e^{\alpha t}[\boldsymbol{U}\cos\omega t-\boldsymbol{W}\,\mathrm{sen}\,\omega t],\quad \mathrm{Im}[\boldsymbol{Y}(t)]=e^{\alpha t}[\boldsymbol{U}\,\mathrm{sen}\,\omega t+\boldsymbol{W}\cos\omega t]$$

2.4. Representación diagonal y desacoplamiento

El caso diagonalizable puede también tratarse pasando a una representacíón **diagonal** del sistema. En este caso la matriz $\boldsymbol{A}$ es semejante a una matriz diagonal $\boldsymbol{D}$, de modo que existe una matriz $\boldsymbol{S}$ de $n \times n$ no singular tal que

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}, \qquad \boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad \boldsymbol{S} = (\mathbf{V}_1, \dots, \boldsymbol{V}_n)$$

donde las columnas de $\boldsymbol{S}$ son n autovectores linealmente independientes de $\boldsymbol{A}$. Multiplicando el sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$ a izquierda por la inversa $\boldsymbol{S}^{-1}$, se obtiene

$$\boldsymbol{S}^{-1}\boldsymbol{Y}' = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{Y} = (\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S})(\boldsymbol{S}^{-1}\boldsymbol{Y})$$

o sea,

$$\boldsymbol{Z}' = \boldsymbol{D}\boldsymbol{Z}, \quad \text{con} \quad \boldsymbol{Z} = \boldsymbol{S}^{-1}\boldsymbol{Y} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

Por ser $\boldsymbol{D}$ diagonal, las n funciones $z_1(t), \dots, z_n(t)$ satisfacen un sistema de n ecuaciones diferenciales de primer orden **desacopladas:**

$$\boldsymbol{Z}' = \boldsymbol{D}\ \ \boldsymbol{Z} \Leftrightarrow \begin{cases} z_1' = \lambda_1 z_1 \\ \quad\vdots \\ z_n' = \lambda_n z_n \end{cases}$$

donde la derivada de cada z_i depende sólo de z_i, siendo independiente de las demás. La solución de estas ecuaciones es inmediata:

$$z_1(t) = c_1 e^{\lambda_1 t}, \ \dots, z_n(t) = c_n e^{\lambda_n t}$$

Podemos entonces escribir la solución general para $\boldsymbol{Z}$ como

$$\boldsymbol{Z}(t) = \begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix} = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix}$$

Y dado que $\boldsymbol{Z} = \boldsymbol{S}^{-1}\boldsymbol{Y} \Rightarrow \boldsymbol{Y} = \boldsymbol{S}\boldsymbol{Z}$ y la solución general para $\boldsymbol{Y}$ es

$$\begin{aligned} \boldsymbol{Y}(t) = \boldsymbol{S}\boldsymbol{Z}(t) &= z_1(t)\boldsymbol{V}_1 + \dots + z_n(t)\boldsymbol{V}_n && (2.28) \\ &= c_1 e^{\lambda_1 t}\boldsymbol{V}_1 + \dots + c_n e^{\lambda_n t}\boldsymbol{V}_n && (2.29) \end{aligned}$$

que coincide con el resultado previo (2.25). Si recordamos el formalismo de cambio de base, vemos de (2.28) que $\boldsymbol{Z}(t)$ es el vector de **coordenadas** de la solución $\boldsymbol{Y}(t)$ **en la base de autovectores** $\{\boldsymbol{V}_1, \dots, \boldsymbol{V}_n\}$ de $\boldsymbol{A}$. La evolución desacoplada se obtiene pues representando $\boldsymbol{Y}(t)$ en esta base, en la que $\boldsymbol{A}$ pasa a ser **diagonal.**

Ejemplo: Consideremos nuevamente el sistema (2.26),

$$\left\{ \begin{array}{l} x' = x - y \\ y' = -x + y \end{array} \right.$$

En este caso $\lambda_1 = 2$, $\lambda_2 = 0$ y $\boldsymbol{S} = (\boldsymbol{V}_1, \boldsymbol{V}_2) = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$. Entonces $\boldsymbol{S}^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)/2$ y el vector de coordenadas de $\boldsymbol{Y} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ en la base de autovectores de $\boldsymbol{A}$ es

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x - y \\ x + y \end{pmatrix}$$

o sea, $z_1 = (x - y)/$, $z_2 = (x + y)/2$. Estas variables satisfacen el sistema **desacoplado**

$$\left\{ \begin{array}{l} z_1' = 2z_1 \\ z_2' = 0 \end{array} \right.$$

ya que $z_1' = \frac{x'-y'}{2} = x - y = 2z_1$, $z_2' = \frac{x'+y'}{2} = 0$. La solución de este sistema es

$$z_1(t) = c_1 e^{2t}, \;\; z_2(t) = c_2$$

y entonces, se obtiene el resultado previo

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = z_1(t)\boldsymbol{V}_1 + z_2(t)\boldsymbol{V}_2 = c_1 e^{2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Ejercicios VI.

1) Determinar la solución general de los siguientes sistemas de primer orden homogéneos. Dar una expresión real de la solución y escribir la matriz fundamental. Identificar también las variables en las que el sistema queda desacoplado.

$$a) \left\{ \begin{array}{l} x' = -2x + y \\ y' = x - 2y \end{array} \right. \qquad b) \left\{ \begin{array}{l} x' = -2x + y \\ y' = -x - 2y \end{array} \right.$$

$$c) \left\{ \begin{array}{l} x' = x - 2y + z \\ y' = -2x - 2z \\ z' = x - 2y + z \end{array} \right. \qquad d) \left\{ \begin{array}{l} x' = y \\ y' = -y + z \\ z' = -y - z \end{array} \right.$$

2) Hallar la solución de 1 a) y de 1 b) que satisface $x(0) = 1$, $y(0) = 0$, y la solución de 1 c) y de 1 d) que satisface $x(0) = 1$, $y(0) = 0 = z(0) = 1$.
Indique también si alguno de los sistemas del ej. 1) tiene soluciones constantes no nulas.

3) Determine la solución general del sistema de segundo orden

$$\left\{ \begin{array}{l} x'' = -y \\ y' = -2x' - y \end{array} \right.$$

planteándolo como un sistema de primer orden.

4) a) Indique para qué valores reales de α puede garantizarse que el sistema

$$\left\{ \begin{array}{l} x' = -x - \alpha y \\ y' = -\alpha x - y \end{array} \right.$$

tiene un punto de **equilibrio estable** en $x = y = 0$, tal que si se lo aparta inicialmente del mismo, retorna a el.
b) Repita el análisis si se cambia la segunda ecuación por $y' = \alpha x - y$.

5) Si $\boldsymbol{A}$ es una matriz arbitraria de $n \times n$ (independiente de t), halle la solución de $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$ que satisface $\boldsymbol{Y}(0) = \boldsymbol{V}$, con $\boldsymbol{V}$ autovector de $\boldsymbol{A}$.

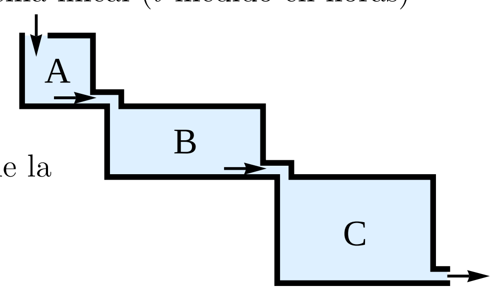
6) Muestre que si $\boldsymbol{Y}(t)$ es solución de $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$, con $\boldsymbol{A}$ independiente de t y $\boldsymbol{Y}(0) = \boldsymbol{Y}_0$,
a) $\boldsymbol{Y}(t - t_0)$ es la única solución de dicha equación que satisface $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$.
b) $\boldsymbol{Y}'(t)$ es también solución de dicha ecuación y satisface $\boldsymbol{Y}'(0) = \boldsymbol{A}\boldsymbol{Y}_0$.

7) Muestre que los autovalores de la matriz obtenida al representar la ecuación de segundo orden $y'' + ay' + by = 0$ como un sistema de primer orden, son las raíces de la ecuación característica $\lambda^2 + a\lambda + b = 0$.

2.5. Aplicaciones

1. Evolución de la concentración de un soluto.
Considere los tanques A, B, C de volúmenes 10, 20 y 30 litros, inicialmente llenos con agua salada. Si entra agua pura (sin sal) a razón de 10 lts por hora en el tanque A, saliendo agua por el orificio inferior de C tal que el volumen de agua en cada uno de los tanques permanece constante, muestre que la cantidad de sal x, y, z, en A, B y C (asumiendo distribución uniforme en c/tanque) satisface el sistema lineal (t medido en horas)

$$\begin{cases} dx/dt = -x \\ dy/dt = x - y/2 \\ dz/dt = y/2 - z/3 \end{cases}$$



Encuentre la solución general de este sistema y halle la solución para concentración inicial uniforme c.
Grafique la concentración en cada tanque.

2. Partícula en un campo magnético. Resolver el problema de una partícula de carga q y masa m en un campo magnético $\boldsymbol{B}$. La fuerza que se ejerce sobre la partícula es $\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}$ (fuerza de Lorentz).
a) Mostrar que la ecuación de movimiento que determina su velocidad,

$$m d\boldsymbol{v}/dt = q\boldsymbol{v} \times \boldsymbol{B}$$

constituye un sistema de 3 ecuaciones lineales homogéneas de primer orden. Escribirlo explícitamente.
b) Probar que si el campo está dirigido en la dirección z, $\boldsymbol{B} = (0, 0, B)$, entonces el sistema anterior se reduce a

$$\begin{cases} v'_x = \omega v_y \\ v'_y = -\omega v_x \\ v'_z = 0 \end{cases}$$

donde $\omega = qB/m$ es la denominada **frecuencia del ciclotrón**.
c) Mostrar que la solución general del sistema anterior es

$$\begin{pmatrix} v_x(t) \\ v_y(t) \\ v_z(t) \end{pmatrix} = c_+ e^{i\omega t} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} + c_- e^{-i\omega t} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \quad A \begin{pmatrix} \cos(\omega t + \phi) \\ -\sin(\omega t + \phi) \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.30}$$

Interpretar esta solución, mostrándo que corresponde a velocidad constante en la dirección z, y movimiento circular uniforme en el plano x, y, de frecuencia angular ω.

d) Resuelva el sistema para un campo general $\boldsymbol{B} = (B_x, B_y, B_z)$. Muestre que los autovalores son (y deben ser!) $\lambda = 0$ y $\lambda = \pm i\omega$, con $\omega = q|\boldsymbol{B}|/m$ y $|\boldsymbol{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2}$.

3. Importante: Sistemas lineales de segundo orden.
Consideremos el sistema lineal de segundo orden homogéneo asociado a una matriz $\boldsymbol{A}$ de $n \times n$,

$$\boldsymbol{Y}'' = \boldsymbol{A}\boldsymbol{Y} \tag{2.31}$$

donde hemos supuesto que $\boldsymbol{Y}''$ no depende explícitamente de $\boldsymbol{Y}'$. Este caso surge, por ejemplo, al aplicar la 2^a ley de Newton a sistemas descriptos por fuerzas que dependen linealmente de la posición, tales como masas unidas por resortes y en general sistemas cercanos a un punto de equilibrio. El sistema (2.31) es equivalente al sistema de $2n$ ecuaciones de primer orden

$$\begin{cases} \boldsymbol{Y}' = \boldsymbol{V} \\ \boldsymbol{V}' = \boldsymbol{A}\boldsymbol{Y} \end{cases}$$

donde $\boldsymbol{V} = \boldsymbol{Y}'$. Es posible, no obstante, resolver (2.31) en forma directa.
a) Para $\boldsymbol{A}$ independiente de t, proponiendo una solución $\boldsymbol{Y}(t) = e^{\alpha t}\boldsymbol{V}$ y reemplazando directamente en (2.31), muestre que se obtiene la ecuación

$$\boldsymbol{A}\boldsymbol{V} = \alpha^2 \boldsymbol{V}$$

y que esto implica que el sistema (2.31) posee soluciones de la forma

$$\boldsymbol{Y}^+(t) = e^{\sqrt{\lambda}\,t}\boldsymbol{V}, \quad \boldsymbol{Y}^-(t) = e^{-\sqrt{\lambda}t}\boldsymbol{V}$$

donde λ es un **autovalor** de $\boldsymbol{A}$ y $\boldsymbol{V}$ un **autovector** asociado.
b) Si la matriz $\boldsymbol{A}$ es **diagonalizable**, tal que $\boldsymbol{A}\boldsymbol{V}_i = \lambda_i \boldsymbol{V}_i$, $i = 1, \ldots, n$, con $\{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n\}$ linealmente independientes, muestre, utilizando la representación diagonal de $\boldsymbol{A}$, que un conjunto completo de soluciones de (2.31) es

$$\boldsymbol{Y}_i^+(t) = e^{\sqrt{\lambda_i}t}\,\boldsymbol{V}_i, \;\; \boldsymbol{Y}_i^-(t) = e^{-\sqrt{\lambda_i}t}\,\boldsymbol{V}_i$$

si $\lambda_i \neq 0$, y

$$\boldsymbol{Y}_i^+(t) = \boldsymbol{V}_i, \;\; \boldsymbol{Y}_i^-(t) = t\boldsymbol{V}_i$$

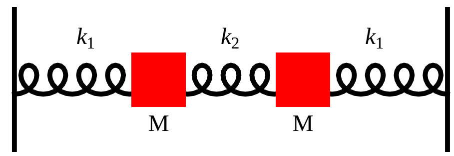
si $\lambda_i = 0$, para $i = 1, \ldots, n$. Escriba la solución general.
Pruebe también que estas soluciones proveen $2n$ soluciones linealmente independientes del sistema asociado de primer orden.
c) Muestre que si $\boldsymbol{A}$ es real y $\lambda_i < 0$, un par de soluciones reales independientes asociado a λ_i es

$$\boldsymbol{Y}_i^c(t) = \cos(\omega_i t)\,\boldsymbol{V}_i, \;\; \boldsymbol{Y}_i^s = \operatorname{sen}(\omega_i t)\,\boldsymbol{V}_i\,, \quad \omega_i = \sqrt{-\lambda_i} > 0$$

Interprete estas soluciones.

4. Utilizando el método anterior, resuelva el problema de dos masas iguales (m) unidas cada una a una pared por un resorte de constate k_1 y unidas entre sí por un resorte de constante k_2.
a) Muestre que si y_1, y_2 denotan las posiciones de las masas a partir de los respectivos puntos de equilibrio, las ecuaciones de movimiento son

$$\begin{cases} my_1'' = -k_1 y_1 + k_2(y_2 - y_1) \\ my_2'' = k_2(y_1 - y_2) - k_1 y_2 \end{cases}$$



y constituyen un sistema de dos ecuaciones lineales homogéneas de segundo orden de la forma $\boldsymbol{Y}'' = \boldsymbol{A}\boldsymbol{Y}$. Identificar la matriz $\boldsymbol{A}$.
b) Pruebe que los autovalores de $\boldsymbol{A}$ son

$$\lambda_1 = -(k_1 + 2k_2)/m, \quad \lambda_2 = -k_1/m$$

ambos negativas, y que por lo tanto, las frecuencias de oscilación del sistema (asumiendo $k_1 > 0$, $k_2 > 0$, $m > 0$) son

$$\omega_1 = \sqrt{\frac{k_1 + 2k_2}{m}}, \qquad \omega_2 = \sqrt{\frac{k_1}{m}}$$

c) Muestre que la solución general del sistema es

$$\begin{aligned} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} &= (c_1^+ e^{i\omega_1 t} + c_1^- e^{-i\omega_1 t}) \begin{pmatrix} 1 \\ -1 \end{pmatrix} + (c_2^+ e^{i\omega_2 t} + c_2^- e^{-i\omega_2 t}) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= A_1 \cos(\omega_1 t + \phi_1) \begin{pmatrix} 1 \\ -1 \end{pmatrix} + A_2 \cos(\omega_2 t + \phi_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned} \tag{2.32}$$

Interprete esta solución y discuta los **modos normales** de vibración.
d) Determine y grafique la solución que satisface las condiciones iniciales
i) $y_1(0) = A$, $y_2(0) = -A$, $y_1'(0) = y_2'(0) = 0$
ii) $y_1(0) = A$, $y_2(0) = A$, $y_1'(0) = y_2'(0) = 0$.

e) Resuelva el caso especial $k_1 = 0$. Discuta e interprete el resultado.

5. Resuelva el problema anterior pero considerando que la segunda masa no está unida a la pared. Determine las frecuencias y los modos normales de vibración.

6. El movimiento de un masa unida a un resorte en un plano está descripto por la ecuación

$$m\frac{d^2\boldsymbol{r}}{dt^2} = -k\boldsymbol{r}$$

con $\boldsymbol{r} = (x, y)$, que es nuevamente un sistema lineal homogéneo de segundo orden. Determine su solución e identifique las trayectorias posibles.

7. Importante: Trayectorias en el plano de fase

Considere el sistema lineal homogéneo de primer orden,

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} \qquad a, b, c, d \in \mathbb{R} \tag{2.33}$$

Las soluciones reales $x(t)$, $y(t)$ del sistema pueden visualizarse graficando las "trayectorias" $(x(t), y(t))$ para $t \in \mathbb{R}$, en el plano x, y, denominado plano (o espacio) de fase.

a) Muestre que las trayectorias correspondientes a soluciones distintas no pueden cruzarse para tiempos t finitos.

b) Pruebe que la ecuación que define las trayectorias es $\frac{dy}{dx} = \frac{cx+dy}{ax+by}$.

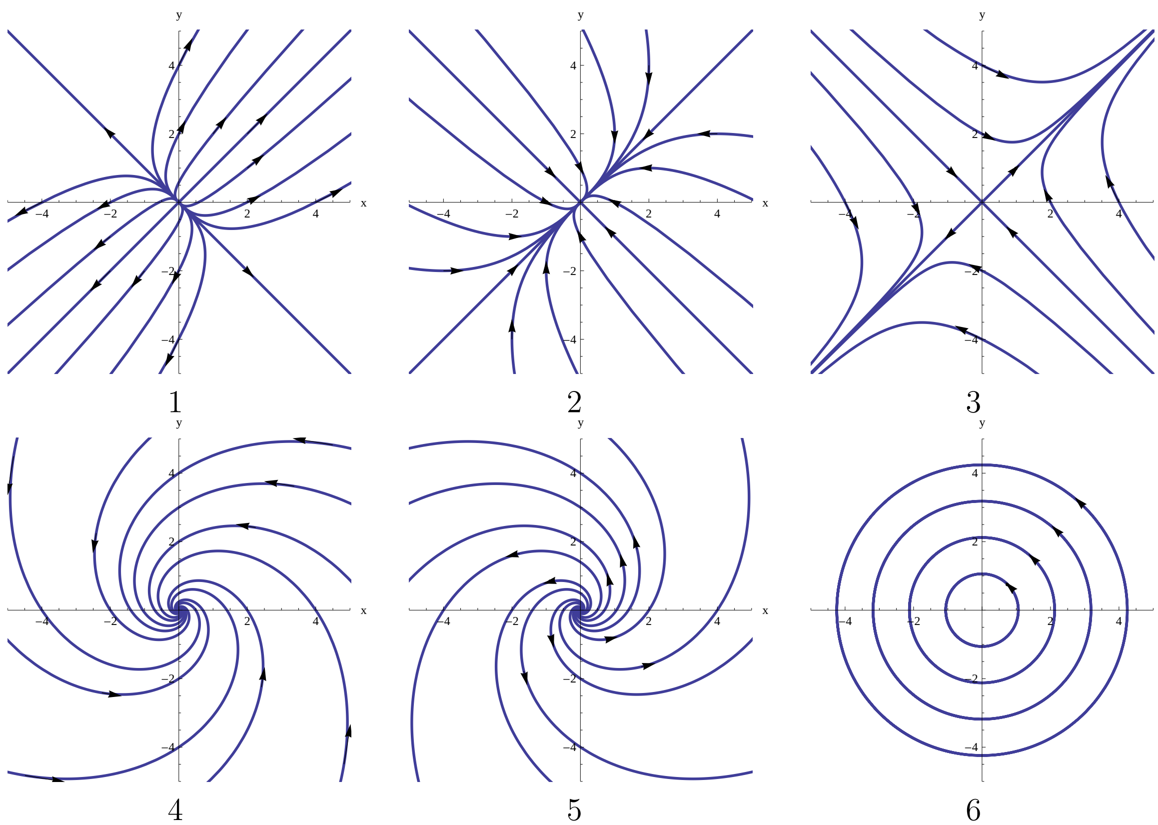
c) Muestre que si $\binom{x(0)}{y(0)}$ coincide con un autovector de la matriz asociado a un autovalor real $\lambda \neq 0$, la trayectoria es recta. Indique en qué casos se alejará, y en que casos se acercará al origen.

d) Considere ahora el caso $a = d = \alpha$, $b = c = \beta$. Muestre que los autovalores de la matriz son $\lambda_\pm = \alpha \pm \beta$ y grafique e interprete las trayectorias para:

i) $\alpha = 1$, $\beta = 1/2$, ii) $\alpha = -1$, $\beta = 1/2$, iii) $\alpha = -1$, $\beta = 2$. Identifique los casos en que el origen $(x, y) = (0, 0)$ es un punto de equilibrio estable y aquellos en que es inestable.

e) Considere ahora el caso $a = d = \alpha$, $c = -b = \omega$, en el que los autovalores de la matriz son $\lambda_\pm = \alpha \pm i\omega$. Grafique e interprete las trayectorias para:

i) $\alpha = -1$, $\omega = 1$, ii) $\alpha = 1$, $\omega = 1$, iii) $\alpha = 0$, $\omega = 1$. Muestre que en este último caso las trayectorias son círculos centrados en el origen.



Gráficos de las soluciones del sistema (2.33) para $a = d$, $|\alpha| = |\beta|$. Las flechas indican el sentido del movimiento. Los gráficos corresponden a: 1) Ambos autovalores reales y positivos, 2) Ambos reales y negativos, 3) Uno positivo y uno negativo, 4) Complejos con parte real negativa, 5) Complejos con parte real positiva, 6) Imaginarios.

2.6. El caso no diagonalizable

Tratemos ahora en detalle este caso. Existe al menos un autovalor repetido λ de la matriz $\boldsymbol{A}$ cuya multiplicidad algebraica m (multiplicidad como raíz del polinomio característico) es mayor que su multiplicidad geométrica l (dimensión del espacio propio correspondiente).

Los l autovectores linealmente independientes asociados proveen $l < m$ soluciones linealmente independientes del sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$. Para obtener las soluciones restantes, tenemos el siguiente teorema general:

Si λ es un autovalor de $\boldsymbol{A}$ con multiplicidad algebraica m, existen m vectores linealmente independientes $\boldsymbol{V}_i$, $i = 1\ldots,m$, que verifican

$$(\boldsymbol{A} - \lambda I)^{k_i}\boldsymbol{V}_i = \boldsymbol{0}, \quad \text{con} \quad (\boldsymbol{A} - \lambda I)^{k_i-1}\boldsymbol{V}_i \neq \boldsymbol{0} \tag{2.34}$$

para algún entero positivo no nulo $k_i \leq m$, siendo I la matriz identidad. Estos m vectores originan m soluciones linealmente independientes del sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$, de la forma

$$\boldsymbol{Y}_i(t) = e^{\lambda t}[\boldsymbol{V}_i + t(\boldsymbol{A} - \lambda I)\boldsymbol{V}_i + \ldots + \frac{t^{k_i-1}}{(k_i-1)!}(\boldsymbol{A} - \lambda I)^{k_i-1}\boldsymbol{V}_i] \tag{2.35}$$

Como la suma de las multiplicidades algebraicas de todos los autovalores (distintos) es n, se obtienen así n soluciones linealmente independientes del sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$.

Si $k_i = 1$, $\boldsymbol{V}_i$ es un autovector de $\boldsymbol{A}$ con autovalor λ, ya que satisface $(\boldsymbol{A} - \lambda I)\boldsymbol{V}_i = \boldsymbol{0}$, con $(\boldsymbol{A} - \lambda I)^0\boldsymbol{V}_i = \boldsymbol{V}_i \neq \boldsymbol{0}$, y la solución (2.35) se reduce a $\boldsymbol{Y}(t) = e^{\lambda t}\boldsymbol{V}_i$.

Si en cambio $k_i > 1$ el vector $\boldsymbol{V}_i$ se denomina **autovector generalizado** asociado a λ.

Problema 13: Probar que si se cumple (2.34), entonces (2.35) es solución de $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$.

Problema 14: Probar que si se cumple (2.34), $(\boldsymbol{A} - \lambda I)^{k_i-1}\boldsymbol{V}_i$ es autovector de $\boldsymbol{A}$.

Por lo tanto, para encontrar m soluciones linealmente independientes asociadas a un autovalor λ de multiplicidad algebraica m y multiplicidad geométrica $l < m$, un método es el siguiente: Se encuentran primero l autovectores linealmente independientes $\boldsymbol{V}_i$ $((\boldsymbol{A} - \lambda I)\boldsymbol{V}_i = \boldsymbol{0}$, $\boldsymbol{V}_i \neq \boldsymbol{0})$, que generan las l soluciones $\boldsymbol{Y}_i(t) = e^{\lambda t}\boldsymbol{V}_i$, $i = 1,\ldots,l$.

Para hallar las $m - l$ restantes, encontramos luego todos los vectores $\boldsymbol{V}$ que satisfacen

$$(\boldsymbol{A} - \lambda I)^2\boldsymbol{V} = \boldsymbol{0} \quad \text{con} \quad (\boldsymbol{A} - \lambda I)\boldsymbol{V} \neq \boldsymbol{0}$$

y que formen, junto con los l autovectores previos, un conjunto **linealmente independiente**. Cada uno de estos nuevos vectores (que a lo sumo serán l) genera una solución adicional de la forma

$$\boldsymbol{Y}(t) = e^{\lambda t}[\boldsymbol{V} + t(\boldsymbol{A} - \lambda I)\boldsymbol{V}]$$

Estas nuevas soluciones formarán, junto con las anteriores, un conjunto ampliado de soluciones linealmente independientes.

Si aún no se tienen m soluciones linealmente independientes, se determinan todos los vectores $\boldsymbol{V}$ linealmente independientes que satisfacen

$$(\boldsymbol{A}-\lambda I)^3\boldsymbol{V}=\boldsymbol{0} \ \text{ con } \ (\boldsymbol{A}-\lambda I)^2\boldsymbol{V}\neq\boldsymbol{0}$$

y que formen, junto con todos los anteriores, un conjunto linealmente independiente. Cada uno de estos nuevos vectores (a lo sumo habrá l) aportará una solución adicional

$$\boldsymbol{Y}(t)=e^{\lambda t}[\boldsymbol{V}+t(\boldsymbol{A}-\lambda I)\boldsymbol{V}+\frac{t^2}{2!}(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}]$$

generándose así un nuevo conjunto ampliado de soluciones linalmente independientes.

Este proceso se continúa hasta obtener m soluciones linealmente independientes.

Por ejemplo, si λ es un autovalor de multiplicidad algebraica $m=2$ y multiplicidad geométrica $l=1$, dos soluciones linealmente independientes son

$$\begin{aligned}\boldsymbol{Y}_1(t) &= e^{\lambda t}\boldsymbol{V}_1 \qquad (2.36)\\ \boldsymbol{Y}_2(t) &= e^{\lambda t}[\boldsymbol{V}_2+t(\boldsymbol{A}-\lambda I)\boldsymbol{V}_2] \qquad (2.37)\end{aligned}$$

con $\boldsymbol{V}_1$ autovector asociado a λ y $\boldsymbol{V}_2$ un vector que satisface

$$(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}_2=\boldsymbol{0}, \quad (\boldsymbol{A}-\lambda I)\boldsymbol{V}_2\neq\boldsymbol{0}$$

Si determinamos primero $\boldsymbol{V}_2$, podemos directamente elegir

$$\boldsymbol{V}_1=(\boldsymbol{A}-\lambda I)\boldsymbol{V}_2 \qquad (2.38)$$

ya que cumple $\boldsymbol{V}_1\neq\boldsymbol{0}$ y $(\boldsymbol{A}-\lambda I)\boldsymbol{V}_1=(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}_2=\boldsymbol{0}$.
Y si determinamos primero el autovector $\boldsymbol{V}_1$, podemos elegir $\boldsymbol{V}_2$ como cualquier vector que satisface (2.38), ya que en tal caso $(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}_2=(\boldsymbol{A}-\lambda I)\boldsymbol{V}_1=\boldsymbol{0}$.

Problema 15: Mostrar que proponiendo una solución de la forma

$$\boldsymbol{Y}(t)=e^{\lambda t}(\boldsymbol{V}+t\boldsymbol{W})$$

con $\boldsymbol{V}$, $\boldsymbol{W}$ constantes, reemplazando en $\boldsymbol{Y}'=\boldsymbol{A}\boldsymbol{Y}$ se obtiene el sistema

$$\left\{\begin{aligned}\boldsymbol{A}\boldsymbol{W} &= \lambda\boldsymbol{W},\\ (\boldsymbol{A}-\lambda I)\boldsymbol{V} &= \boldsymbol{W}\end{aligned}\right.$$

que indica precisamente que $\boldsymbol{W}$ debe ser autovector de $\boldsymbol{A}$ con autovalor λ y $\boldsymbol{V}$ un vector que satisface $(\boldsymbol{A}-\lambda I)\boldsymbol{V}=\boldsymbol{W}$ (y por lo tanto $(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}=\boldsymbol{0}$).

Y en general, si λ es un autovalor de multiplicidad algebraica $m > 1$ y multiplicidad geométrica $l = 1$, las m soluciones linealmente independientes son

$$\begin{aligned}
\boldsymbol{Y}_1(t) &= e^{\lambda t}\boldsymbol{V}_1 && (2.39)\\
\boldsymbol{Y}_2(t) &= e^{\lambda t}[\boldsymbol{V}_2 + t(\boldsymbol{A}-\lambda I)\boldsymbol{V}_2] && (2.40)\\
\boldsymbol{Y}_3(t) &= e^{\lambda t}[\boldsymbol{V}_3 + t(\boldsymbol{A}-\lambda I)\boldsymbol{V}_3 + \frac{t^2}{2!}(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}_3] && (2.41)\\
&\vdots\\
\boldsymbol{Y}_m(t) &= e^{\lambda t}[\boldsymbol{V}_m + t(\boldsymbol{A}-\lambda I)\boldsymbol{V}_m + \ldots + \frac{t^{m-1}}{(m-1)!}(\boldsymbol{A}-\lambda I)^{m-1}\boldsymbol{V}_m] && (2.42)
\end{aligned}$$

donde $\boldsymbol{V}_k$, $k = 1, \ldots, m$, satisface

$$(\boldsymbol{A}-\lambda I)^k\boldsymbol{V}_k = \boldsymbol{0}, \quad (\boldsymbol{A}-\lambda I)^{k-1}\boldsymbol{V}_k \neq \boldsymbol{0} \tag{2.43}$$

Si determinamos primero $\boldsymbol{V}_m$, tal que $(\boldsymbol{A}-\lambda I)^m\boldsymbol{V}_m = \boldsymbol{0}$, con $(\boldsymbol{A}-\lambda I)^{m-1}\boldsymbol{V}_m \neq \boldsymbol{0}$, podemos en este caso obtener **todos** los vectores restantes directamente como

$$\boldsymbol{V}_k = (\boldsymbol{A}-\lambda I)^{m-k}\boldsymbol{V}_m\,, \quad k = 1, \ldots, m-1 \tag{2.44}$$

ya que $(\boldsymbol{A}-\lambda I)^k\boldsymbol{V}_k = (\boldsymbol{A}-\lambda I)^m\boldsymbol{V}_m = \boldsymbol{0}$ y $(\boldsymbol{A}-\lambda I)^{k-1}\boldsymbol{V}_k = (\boldsymbol{A}-\lambda I)^{m-1}\boldsymbol{V}_m \neq \boldsymbol{0}$.

Ejemplo. Consideremos el sistema

$$\left\{\begin{array}{l} x' = x + by \\ y' = y \end{array}\right. \tag{2.45}$$

que puede escribirse en forma matricial como

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$$

La ecuación característica es $\det(\mathbf{A}-\lambda I) = (\lambda-1)^2 = 0$, que conduce al único autovalor $\lambda = 1$, con multiplicidad algebraica 2.

Si $b \neq 0$, la matriz $\boldsymbol{A}$ **no es diagonalizable**, pues el rango de

$$\mathbf{A} - \lambda I = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$$

es 1 y por lo tanto la nulidad es 1. Esto indica que existirá un sólo autovector linealmente independiente, $\boldsymbol{V}_1 \propto \binom{1}{0}$. Eligiendo $\boldsymbol{V}_1 = \binom{1}{0}$, se obtiene la solución

$$\boldsymbol{Y}_1(t) = e^t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Buscamos ahora un vector $\boldsymbol{V}_2$ linealmente independiente de $\boldsymbol{V}_1$ tal que

$$(\boldsymbol{A}-\lambda I)^2\boldsymbol{V}_2 = \boldsymbol{0}, \quad (\boldsymbol{A}-\lambda I)\boldsymbol{V}_2 \neq \boldsymbol{0}.$$

Como en este caso $(\boldsymbol{A}-\lambda I)^2 = \binom{0\,0}{0\,0}$, basta con encontrar un vector $\boldsymbol{V}_2$ linealmente independiente de $\boldsymbol{V}_1$, por ejemplo $\boldsymbol{V}_2 = \binom{0}{1}$, que satisface, dado que $b \neq 0$,

$$(\boldsymbol{A}-\lambda I)\boldsymbol{V}_2 = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \neq \boldsymbol{0}$$

Obtenemos así la solución linealmente independiente de $\boldsymbol{Y}_1(t)$,

$$\boldsymbol{Y}_2(t) = e^t\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} b \\ 0 \end{pmatrix}\right] = e^{at}\begin{pmatrix} bt \\ 1 \end{pmatrix}$$

La solución general es entonces

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + c_2\boldsymbol{Y}_2(t) = c_1e^t\begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2e^t\begin{pmatrix} bt \\ 1 \end{pmatrix}$$

es decir, $x(t) = e^t(c_1 + btc_2)$, $\quad y(t) = c_2e^t$.

Una matriz fundamental de soluciones es entonces

$$\boldsymbol{M}(t) = \begin{pmatrix} e^t & bte^t \\ 0 & e^t \end{pmatrix}$$

que satisface $\boldsymbol{M}(0) = I$, $\mathrm{Det}[\boldsymbol{M}(t)] = e^{2t} \neq 0 \ \forall\ t$.

En este ejemplo, esta solución puede obtenerse directamente resolviendo primero la ecuación $y' = y$, que está desacoplada, y luego la ecuación $x' = x + y$, reemplazando en y la solución previa. No es posible aquí encontrar combinaciones lineales de x e y en las que el sistema resulte totalmente desacoplado.

Observación: Si $b \neq 0$, la matriz $\boldsymbol{A} = \binom{a\ b}{\varepsilon\ a}$ es diagonalizable $\forall\ \varepsilon \neq 0$, ya que tendrá dos autovalores distintos $\lambda_\pm = a \pm \sqrt{b\varepsilon}$.

El caso no diagonalizable ocurre únicamente cuando $\varepsilon = 0$. Los autovalores serán ambos reales si $b\varepsilon \geq 0$, y complejos si $b\varepsilon < 0$. Por lo tanto, el caso no diagonalizable corresponde al punto "crítico" $\varepsilon = 0$, en el que se produce la transición de autovalores reales ($b\varepsilon > 0$) a autovalores complejos ($b\varepsilon < 0$), es decir, de soluciones $\boldsymbol{Y}(t)$ de tipo exponencial a soluciones de tipo oscilatorio.

Ejercicios VII

1) Mostrar que la solución general del sistema

$$\begin{cases} x' = -x - y \\ y' = x - 3y \end{cases}$$

es

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{-2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-2t} \left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right]$$

es decir, $x(t) = c_1 e^{-2t} + c_2 e^{-2t}(1+2t)$, $y(t) = c_1 e^{-2t} + c_2 e^{-2t}(-1+2t)$.

2) Mostrar que la solución general del sistema

$$\begin{cases} x' = y + z \\ y' = y - z \\ z' = y - z \end{cases}$$

es

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \left[\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right] + c_3 \left[\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + \frac{t^2}{2} \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \right]$$

o sea, $x(t) = c_1 + 2c_2 t + 2c_3 t^2$, $y(t) = c_2 + c_3(1+2t)$, $z(t) = c_2 + c_3(-1+2t)$.

3) Mostrar que la posición y y velocidad v de un móvil en ausencia de fuerzas satisfacen un sistema de ecuaciones diferenciales de primer orden homogéneo asociado a una matriz no diagonalizable. Determine su solución.

4) Hallar la solución del sistema

$$\begin{cases} x' = 3x - 4y \\ y' = x - y \end{cases}$$

que satisface $x(0) = 1$, $y(0) = 1$.

5) Mostrar que si la ecuación de segundo orden con coeficientes constantes $y'' + ay' + by = 0$ tiene raíces iguales ($a^2 = 4b$) el sistema lineal de primer orden asociado

$$\begin{cases} z_1' = z_2 \\ z_2' = -bz_1 - az_2 \end{cases}$$

donde $z_1 = y$, $z_2 = y'$, corresponde necesariamente a una matriz no diagonalizable.

6) Encontrar la solución general del sistema

$$\begin{cases} x' = 3x - 4z \\ y' = x + y \\ z' = x - z \end{cases}$$

2.7. Matriz fundamental

Tanto para el **caso (a)** ($\boldsymbol{A}$ diagonalizable) como para el **caso (b)** ($\boldsymbol{A}$ no diagonalizable), la matriz

$$\exp[\boldsymbol{A}t] = \sum_{k=0}^{\infty} \frac{\boldsymbol{A}^k t^k}{k!}$$

es una matriz fundamental del sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$ con $\boldsymbol{A}$ independiente de t, y es la matriz fundamental que satisface $\boldsymbol{M}(0) = I$, con I la matriz identidad.

Esto puede demostrarse directamente a partir de la serie, notando que

$$\exp[\boldsymbol{A}t]' = \sum_{k=1}^{\infty} \frac{k\boldsymbol{A}^k t^{k-1}}{k!} = \boldsymbol{A}\sum_{k=1}^{\infty} \frac{\boldsymbol{A}^{k-1} t^{k-1}}{(k-1)!} = \boldsymbol{A}\sum_{k=0}^{\infty} \frac{\boldsymbol{A}^k t^k}{k!} = \boldsymbol{A}\exp[\boldsymbol{A}t]$$

donde hemos tenido en cuenta que $\boldsymbol{A}^0 t^0/0! = I$ (matriz identidad).

Por lo tanto, cada columna $\boldsymbol{Y}_i(t)$ de $\exp[\boldsymbol{A}t]$ satisface también $\boldsymbol{Y}_i'(t) = \boldsymbol{A}\boldsymbol{Y}_i(t)$, siendo entonces solución del sistema homogéneo. Además, si $t = 0$, el primer término de la serie es el único término no nulo y por lo tanto

$$\exp[\boldsymbol{A}(0)] = I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

lo que implica que las n columnas $\boldsymbol{Y}_i(t)$ son linealmente independientes, pues los n vectores $\boldsymbol{Y}_i(0)$ son linealmente independientes, formando la base canónica de $\mathbb{R}^n$: $\boldsymbol{Y}_i(0) = \boldsymbol{e}_i$.

Esto demuestra que $\exp[\boldsymbol{A}t]$ es una matriz fundamental del sistema en todos los casos, y es la única que satisface $\boldsymbol{M}(0) = I$. Resumiendo,

$$\exp[\boldsymbol{A}t] = (\boldsymbol{Y}_1(t), \dots, \boldsymbol{Y}_n(t))$$

donde la columna $\boldsymbol{Y}_i(t)$, $i = 1, \dots, n$, es la solución del sistema que satisface $\boldsymbol{Y}_i(0) = \boldsymbol{e}_i$.

La solución general del sistema $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$ puede entonces expresarse como

$$\boldsymbol{Y}(t) = c_1\boldsymbol{Y}_1(t) + \dots + c_n\boldsymbol{Y}_n(t) = \exp[\boldsymbol{A}t]\boldsymbol{C}$$

donde $\boldsymbol{C} = \boldsymbol{Y}(0)$ es directamente el valor inicial (en $t = 0$). Esta expresión de la solución es análoga a la solución general del caso elemental de primer orden $y' = ay$, con a constante, que es $y(t) = e^{at}c$, con $c = y(0)$.

En el caso diagonalizable, $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$, con $\boldsymbol{D}$ diagonal y $\boldsymbol{S} = (\boldsymbol{V}_1, \dots, \boldsymbol{V}_n)$ una matriz de autovectores linealmente independientes. Por lo tanto, $\boldsymbol{A}^k = \boldsymbol{S}\boldsymbol{D}^k\boldsymbol{S}^{-1}$ y entonces

$$\exp[\boldsymbol{A}t] = \exp[\boldsymbol{S}(\boldsymbol{D}t)\boldsymbol{S}^{-1}] = \boldsymbol{S}\exp[\boldsymbol{D}t]\boldsymbol{S}^{-1}$$

Y como $\boldsymbol{D}$ es diagonal,

$$\boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \Rightarrow \exp[\boldsymbol{D}t] = \sum_{k=0}^{\infty} \frac{\boldsymbol{D}^k t^k}{k!} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}$$

por lo que $\exp[\boldsymbol{A}t]$ puede evaluarse directamente. En el caso no diagonalizable, la evaluación se realiza mediante la denominada forma canónica de Jordan de la misma, que es la representación en la base de autovectores generalizados.

Es importante destacar que al tomar la exponencial como matriz fundamental, resulta muy fácil evaluar la matriz inversa $\boldsymbol{M}^{-1}(t)$:

$$(\exp[\boldsymbol{A}t])^{-1} = \exp[-\boldsymbol{A}t]$$

ya que $\exp[\boldsymbol{A}t]\exp[-\boldsymbol{A}t] = \exp[\boldsymbol{A}(t-t)] = \exp[\boldsymbol{0}] = I$.

Observación: En general $\exp[\boldsymbol{A}]\exp[\boldsymbol{B}] \neq \exp[\boldsymbol{A}]\exp[\boldsymbol{B}]$ para matrices. La igualdad vale cuando $\boldsymbol{AB} = \boldsymbol{BA}$, como en el caso previo ($\boldsymbol{B} = -\boldsymbol{A}$), pero no en el caso general.

Problema 16: Considere el sistema (2.22), en el que

$$\boldsymbol{A} = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right), \quad \boldsymbol{D} = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right), \quad \boldsymbol{S} = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right)$$

Pruebe que

$$\exp[\boldsymbol{A}t] = \boldsymbol{S}\exp[\boldsymbol{D}t]\boldsymbol{S}^{-1} = \boldsymbol{S}\left(\begin{smallmatrix} e^{2t} & 0 \\ 0 & 1 \end{smallmatrix}\right)\boldsymbol{S}^{-1} = \tfrac{1}{2}\begin{pmatrix} 1+e^{2t} & 1-e^{2t} \\ 1-e^{2t} & 1+e^{2t} \end{pmatrix}$$

Por lo tanto, las columnas $\boldsymbol{Y}_1(t) = \frac{1}{2}\begin{pmatrix} 1+e^{2t} \\ 1-e^{2t} \end{pmatrix}$, $\boldsymbol{Y}_2(t) = \frac{1}{2}\begin{pmatrix} 1-e^{2t} \\ 1+e^{2t} \end{pmatrix}$, son soluciones linealmente independientes del sistema, que satisfacen $\boldsymbol{Y}_1(0) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $\boldsymbol{Y}_2(0) = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$.

Problema 17. Consideremos ahora el caso no diagonalizable (2.45), en el que

$$\boldsymbol{A} = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} = I + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$$

Si bien en general $\exp[\boldsymbol{B}+\boldsymbol{C}] \neq \exp[\boldsymbol{B}]\exp[\boldsymbol{C}]$, si $\boldsymbol{B} = I$ tenemos $I\boldsymbol{C} = \boldsymbol{C}I \ \forall\ \boldsymbol{C}$ y por lo tanto $\exp[(I+\boldsymbol{B})t] = \exp[It]\exp[\boldsymbol{B}t] = e^t\exp[\boldsymbol{B}t]$.
Utilizando el resultado anterior, probar que en este caso

$$\exp[\boldsymbol{A}t] = \exp[(I + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix})t] = e^t\exp[\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}t] = e^t\begin{pmatrix} 1 & bt \\ 0 & 1 \end{pmatrix}$$

ya que $\left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right)^2 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y entonces $\left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right)^k = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \ \forall\ k \geq 2$.
Por lo tanto, las columnas $\boldsymbol{Y}_1(t) = e^t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\boldsymbol{Y}_2(t) = e^t\begin{pmatrix} bt \\ 1 \end{pmatrix}$ son soluciones linealmente independientes del sistema $\boldsymbol{Y}' = \boldsymbol{AY}$ que satisfacen $\boldsymbol{Y}_1(0) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, $\boldsymbol{Y}_2(0) = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$.

Problema 18. Si $\boldsymbol{A}$ es de 2×2 y no diagonalizable, puede mostrarse que $(\boldsymbol{A}-\lambda I)^2 = \boldsymbol{0}$. A partir de este resultado, muestre que si $\boldsymbol{A}$ es no diagonalizable,

$$\exp[\boldsymbol{A}t] = e^{\lambda t}[I + (\boldsymbol{A}-\lambda I)t]$$

donde λ es el único autovalor de $\boldsymbol{A}$.

2.8. Sistema lineal no homogéneo general de primer orden

Consideremos ahora el sistema no homogéneo

$$\boldsymbol{Y}' = \boldsymbol{A}(t)\boldsymbol{Y} + \boldsymbol{F}(t) \tag{2.46}$$

que podemos escribir como $L[\boldsymbol{Y}] = \boldsymbol{F}(t)$, con $L[\boldsymbol{Y}] = \boldsymbol{Y}' - \boldsymbol{A}(t)\boldsymbol{Y}$.

Al igual que en el caso de la ecuación lineal de orden n, es válido el siguiente teorema:

Teorema 2. La solución general de (2.46) está dada por la suma de la solución general $\boldsymbol{Y}_h(t)$ del sistema homogéneo más una solución particular $\boldsymbol{Y}_p(t)$ del sistema no homogéneo:

$$\boldsymbol{Y}(t) = \boldsymbol{Y}_h(t) + \boldsymbol{Y}_p(t) \tag{2.47}$$

donde $\boldsymbol{Y}_h(t) = c_1\boldsymbol{Y}_1(t) + \ldots c_n\boldsymbol{Y}_n(t)$ es la solución general (2.16) (satisface $\boldsymbol{Y}_h' = \boldsymbol{A}(t)\boldsymbol{Y}_h$) mientras que $\boldsymbol{Y}_p(t)$ satisface $\boldsymbol{Y}_p' = \boldsymbol{A}(t)\boldsymbol{Y}_p' + \boldsymbol{F}(t)$.

Demostración: Es similar a la realizada para la ecuación no homogénea de orden n. Si $\boldsymbol{Y}_p(t)$ es una solución de (2.46), $\boldsymbol{Y}_h(t) + \boldsymbol{Y}_p(t)$ es también solución pues

$$(\boldsymbol{Y}_h + \boldsymbol{Y}_p)' = \boldsymbol{Y}_h' + \boldsymbol{Y}_p' = \boldsymbol{A}(t)\boldsymbol{Y}_h + \boldsymbol{A}(t)\boldsymbol{Y}_p + \boldsymbol{F}(t) = \boldsymbol{A}(t)(\boldsymbol{Y}_h + \boldsymbol{Y}_p) + \boldsymbol{F}(t)$$

Y si $\boldsymbol{Y}(t)$ es otra solución de (2.46), entonces

$$(\boldsymbol{Y} - \boldsymbol{Y}_p)' = \boldsymbol{Y}' - \boldsymbol{Y}_p' = \boldsymbol{A}(t)\boldsymbol{Y} + \boldsymbol{F}(t) - [\boldsymbol{A}(t)\boldsymbol{Y}_p + \boldsymbol{F}(t)] = \boldsymbol{A}(t)(\boldsymbol{Y} - \boldsymbol{Y}_p)$$

por lo que $\boldsymbol{Y} - \boldsymbol{Y}_p$ es solución del sistema homogéneo y entonces $\boldsymbol{Y}(t) = \boldsymbol{Y}_h(t) + \boldsymbol{Y}_p(t)$.

Si se conoce la solución general del sistema homogéneo, es siempre posible obtener $\boldsymbol{Y}_p(t)$ por integración!

Teorema 3. Si $\boldsymbol{M}(t)$ es la matriz fundamental (2.20) del sistema homogéneo, una solución particular de (2.46) es

$$\boldsymbol{Y}_p(t) = \boldsymbol{M}(t)\int \boldsymbol{M}^{-1}(t)\boldsymbol{F}(t)dt \tag{2.48}$$

Demostración: Proponiendo una solución de la forma

$$\boldsymbol{Y}_p(t) = \boldsymbol{M}(t)\boldsymbol{V}(t), \quad \boldsymbol{V}(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix}$$

obtenemos, dado que $\boldsymbol{M}' = \boldsymbol{A}(t)\boldsymbol{M}$ (ecuación (2.21)),

$$\boldsymbol{Y}_p' - \boldsymbol{A}(t)\boldsymbol{Y}_p = \boldsymbol{M}'\boldsymbol{V} + \boldsymbol{M}\boldsymbol{V}' - \boldsymbol{A}(t)\boldsymbol{M}\boldsymbol{V} = \boldsymbol{M}\boldsymbol{V}' = \boldsymbol{F}(t)$$

Como $\boldsymbol{M}(t)$ es no singular, $\boldsymbol{V}' = \boldsymbol{M}^{-1}(t)\boldsymbol{F}(t)$ y entonces

$$\boldsymbol{V}(t) = \int \boldsymbol{M}^{-1}(t)\boldsymbol{F}(t)\,dt$$

que conduce a la solución (2.48). La solución general (2.47) puede pues escribirse como

$$\boldsymbol{Y}(t) = \boldsymbol{M}(t)\boldsymbol{C} + \boldsymbol{M}(t)\int \boldsymbol{M}^{-1}(t)\boldsymbol{F}(t)dt \tag{2.49}$$

Nótese la semejanza con la solución (1.66) de la ecuación lineal de primer orden.
Problema 3: Escribir (2.49) para el caso $n = 1$ y verificar que se reduce a (1.66).

Problema 4: Mostrar que la solución particular de (2.46) que satisface $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$ es

$$\boldsymbol{Y}(t) = \boldsymbol{M}(t)\left[\boldsymbol{M}^{-1}(t_0)\boldsymbol{Y}_0 + \int_{t_0}^{t} \boldsymbol{M}^{-1}(t')\boldsymbol{F}(t')dt'\right]$$

4. Superposición: Si $\boldsymbol{Y}_{p1}(t)$ e $\boldsymbol{Y}_{p2}(t)$ son soluciones de (2.46) para $\boldsymbol{F}_1(t)$ y $\boldsymbol{F}_2(t)$, una solución particular para la combinación lineal $\boldsymbol{F}(t) = c_1\boldsymbol{F}_1(t) + c_2\boldsymbol{F}_2(t)$ es la combinación lineal de soluciones $\boldsymbol{Y}_p(t) = c_1\boldsymbol{Y}_{p1}(t) + c_2\boldsymbol{Y}_{p2}(t)$:

$$\left\{\begin{array}{l} \boldsymbol{Y}'_{p1} = \boldsymbol{A}(t)\boldsymbol{Y}_{p1} + \boldsymbol{F}_1(t) \\ \boldsymbol{Y}'_{p2} = \boldsymbol{A}(t)\boldsymbol{Y}_{p2} + \boldsymbol{F}_2(t) \end{array}\right. \;\Rightarrow\; (c_1\boldsymbol{Y}_{p1} + c_2\boldsymbol{Y}_{p2})' = \boldsymbol{A}(t)(c_1\boldsymbol{Y}_{p1} + c_2\boldsymbol{Y}_{p2}) + c_1\boldsymbol{F}_1(t) + c_2\boldsymbol{F}_2(t)$$

Este resultado, similar al de la ecuación no homogénea de orden n, es consecuencia inmediata de la linealidad del sistema y su demostración se deja como ejercicio.
También se ve directamente de la expresión (2.48) (probar!).
Implica nuevamente que si podemos descomponer $\boldsymbol{F}(t)$ en términos simples, podemos hallar una solución particular encontrando soluciones para cada uno de los términos.
Y si $\boldsymbol{F}(t) \to \alpha\boldsymbol{F}(t) \Rightarrow \boldsymbol{Y}_p(t) \to \alpha\boldsymbol{Y}_p(t)$

Ejemplo: Hallar la solución general del sistema lineal de primer orden

$$\left\{\begin{array}{rcl} x' &=& x - y + f(t) \\ y' &=& -x + y + g(t) \end{array}\right. \tag{2.50}$$

En forma matricial, corresponde al sistema

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$$

es decir, $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{F}(t)$, con $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $\boldsymbol{F}(t) = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$.
Una matriz fundamental del sistema homogéneo es (ver ejemplo previo) $\boldsymbol{M}(t) = \begin{pmatrix} e^{2t} & 1 \\ -e^{2t} & 1 \end{pmatrix}$.
Por lo tanto,

$$\boldsymbol{M}^{-1}(t) = \frac{1}{2}\begin{pmatrix} e^{-2t} & -e^{-2t} \\ 1 & 1 \end{pmatrix}$$

Aplicando (2.48), obtenemos

$$\boldsymbol{Y}_p(t) = \boldsymbol{M}(t)\int \frac{1}{2}\begin{pmatrix} e^{-2t} & -e^{-2t} \\ 1 & 1 \end{pmatrix}\begin{pmatrix} f(t) \\ g(t) \end{pmatrix} dt = \frac{1}{2}\boldsymbol{M}(t)\begin{pmatrix} \int e^{-2t}(f(t)-g(t))dt \\ \int (f(t)+g(t))dt \end{pmatrix} \tag{2.51}$$

Definiendo las funciones

$$v(t) = \frac{1}{2}e^{2t}\int e^{-2t}(f(t)-g(t))dt, \quad w(t) = \frac{1}{2}\int (f(t)+g(t))dt$$

la solución general es entonces

$$\boldsymbol{Y}(t) \;=\; c_1\begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} v(t)+w(t) \\ -v(t)+w(t) \end{pmatrix} \tag{2.52}$$

Problema 5: Mostrar que a solución general de (2.50) para $f(t) = 0$, $g(t) = 1$ es

$$\begin{aligned} x(t) &= c_1 e^{2t} + c_2 + 1/4 + t/2 \\ y(t) &= -c_1 e^{2t} + c_2 - 1/4 + t/2 \end{aligned}$$

Determinar también la única solución que cumple $x(0) = 0$, $y(0) = 0$.

Ecuaciones lineales de orden n como sistema lineal de primer orden:
Las ecuaciones diferenciales **lineales** de orden n pueden escribirse como un sistema **lineal** de n ecuaciones de primer orden n (probar!, en base a lo ya visto para ecuaciones generales de orden n). Todos los resultados de la primer parte para ecuaciones lineales de orden n pueden entonces también obtenerse a partir de los resultados generales para sistemas lineales de primer orden.

Problema 6:
a) Mostrar que la ecuación lineal de segundo orden no homogénea

$$y'' + p(t)y' + q(t)y = f(t) \tag{2.53}$$

puede expresarse, definiendo $v = y'$, como el sistema lineal de primer orden no homogéneo

$$\begin{pmatrix} y' \\ v' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -q(t) & -p(t) \end{pmatrix}\begin{pmatrix} y \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix} \tag{2.54}$$

b) Justificar que el teorema 2.1 asegura que existen dos soluciones linealmente independientes $y_1(t)$, $y_2(t)$, de (2.53).
c) Mostrar que el determinante de la matriz fundamental $\boldsymbol{M}(t)$ de soluciones del sistema homogéneo asociado a (2.54), es el Wronskiano: $\det \boldsymbol{M}(t) = W(t) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$.
d) Mostrar que la solución particular (2.48) para el sistema (2.54) implica la solución particular

$$y_p(t) = -y_1(t)\int \frac{y_2(t)}{W(t)}f(t)dt + y_2(t)\int \frac{y_1(t)}{W(t)}f(t)dt$$

de (2.53).

2.9. Sistema lineal no homogéneo con coeficientes constantes. Método general y representación diagonal.

Consideremos ahora el caso no homogéneo de primer orden con $\boldsymbol{A}$ independiente de t,

$$\boldsymbol{Y}' = \boldsymbol{AY} + \boldsymbol{F}(t) \tag{2.55}$$

Si conocemos una matriz fundamental $\boldsymbol{M}(t)$, podemos aplicar directamente el resultado general (2.48) para determinar la solución particular $\boldsymbol{Y}_p(t)$:

$$\boldsymbol{Y}_p(t) = \boldsymbol{M}(t) \int \boldsymbol{M}^{-1}(t)\boldsymbol{F}(t)dt \tag{2.56}$$

Si utilizamos $\boldsymbol{M}(t) = \exp[\boldsymbol{A}t]$, $\boldsymbol{M}^{-1}(t) = \exp[-\boldsymbol{A}t]$ y entonces

$$\boldsymbol{Y}_p(t) = \exp[\boldsymbol{A}t] \int \exp[-\boldsymbol{A}t]\boldsymbol{F}(t)dt \tag{2.57}$$

La solución general de (2.55) puede así escribirse como

$$\boldsymbol{Y}(t) = \exp[\boldsymbol{A}t]\,\boldsymbol{C} + \exp[\boldsymbol{A}t] \int \exp[-\boldsymbol{A}t]\boldsymbol{F}(t)dt \tag{2.58}$$

que es completamente análoga! a la solución general $y(t) = e^{at}c + e^{at}\int e^{-at}f(t)dt$ de la ecuación de primer orden $y' = ay + f(t)$: Se reemplaza el número a por la matriz $\boldsymbol{A}$.

Representación diagonal del sistema no homogéneo. Si $\boldsymbol{A}$ es diagonalizable, tal que $\boldsymbol{S}^{-1}\boldsymbol{AS} = \boldsymbol{D}$, con $\boldsymbol{S} = (\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n)$ la matriz de autovectores y $\boldsymbol{D}$ la matriz diagonal de autovalores, podemos entender mejor la solución (2.58) pasando a la representación diagonal de (2.55). Multiplicando a izquierda por $\boldsymbol{S}^{-1}$ (véase sección 2.5), se obtiene

$$\boldsymbol{Z}' = \boldsymbol{DZ} + \boldsymbol{G}(t) \tag{2.59}$$

donde $\boldsymbol{Z} = \boldsymbol{S}^{-1}\boldsymbol{Y}$ es el vector de coordenadas de $\boldsymbol{Y}$ en la base de autovectores y

$$\boldsymbol{G}(t) = \boldsymbol{S}^{-1}\boldsymbol{F}(t) = \begin{pmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{pmatrix}$$

el vector de coordenadas de $\boldsymbol{F}(t)$ en la misma base de autovectores, tal que

$$\boldsymbol{F}(t) = g_1(t)\boldsymbol{V}_1 + \ldots + g_n(t)\boldsymbol{V}_n$$

Como $\boldsymbol{D}$ es diagonal, (2.59) implica n ecuaciones no homogéneas de primer orden **desacopladas**:

$$z_i' = \lambda_i z_i + g_i(t), \quad i = 1, \ldots, n$$

La solución particular de esta ecuación es

$$z_{ip}(t) = e^{\lambda_i t} \int e^{-\lambda_i t} g_i(t)dt \tag{2.60}$$

y por lo tanto, la solución particular del sistema es

$$\boldsymbol{Y}_p = \boldsymbol{SZ}_p(t) = z_{1p}(t)\boldsymbol{V}_1 + \ldots + z_{np}(t)\boldsymbol{V}_n \tag{2.61}$$

Problema 19: Mostrar que en el caso diagonalizable, (2.57) conduce a la expresión (2.61) con los coeficientes (2.60). Utilizar $\exp[\boldsymbol{A}t] = \boldsymbol{S}\exp[\boldsymbol{D}t]\boldsymbol{S}^{-1}$.

Ejemplo: Resolver el sistema

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x \\ y \end{pmatrix} + \boldsymbol{F}(t)\,, \quad \text{con } \boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \boldsymbol{F}(t) = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$$

Este sistema ya fue resuelto en 2.8 mediante la aplicación de la ecuación (2.56). Para aplicar el método (2.61), en la base de autovectores de $\boldsymbol{A}$ formada por $\boldsymbol{V}_1 = \binom{1}{-1}$, $\boldsymbol{V}_2 = \binom{1}{1}$, con $\lambda_1 = 2$, $\lambda_2 = 0$, tenemos $\binom{g_1(t)}{g_2(t)} = \boldsymbol{S}^{-1}\binom{f(t)}{g(t}) = \frac{1}{2}\binom{f(t)-g(t)}{f(t)+g(t)}$ y por lo tanto

$$\boldsymbol{F}(t) = g_1(t)\boldsymbol{V}_1 + g_2(t)\boldsymbol{V}_2 = \frac{f(t)-g(t)}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{f(t)+g(t)}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

Las coordenadas $z_1(t)$, $z_2(t)$ de la solución en la base de autovectores satisfacen entonces las ecuaciones desacopladas $z_1' = \lambda_1 z_1 + g_1(t)$, $z_2' = \lambda_2 z_2 + g_2(t)$, o sea,

$$\begin{cases} z_1' = 2z_1 + \frac{f(t)-g(t)}{2}, \\ z_2' = \frac{f(t)+g(t)}{2} \end{cases}$$

cuyas soluciones generales son

$$z_1(t) = c_1 e^{2t} + e^{2t}\int e^{-2t}[\tfrac{f(1)-g(t)}{2}]dt = c_1 e^{2t} + v(t) \tag{2.62}$$

$$z_2(t) = c_2 + \int \tfrac{f(t)+g(t)}{2}dt = c_2 + w(t) \tag{2.63}$$

donde $v(t)$ y $w(t)$ son las funciones definidas en (2.28). La solución general es entonces

$$\boldsymbol{Y}(t) = z_1(t)\begin{pmatrix} 1 \\ -1 \end{pmatrix} + z_2(t)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = c_1\begin{pmatrix} e^{2t} \\ -e^{2t} \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} v(t)+w(t) \\ -v(t)+w(t) \end{pmatrix}$$

que coincide con la expresión (2.52).

2.10. Método de coeficientes indeterminados

Al igual que en el caso de la ecuación de orden n, si $\boldsymbol{F}(t)$ es exponencial, seno, coseno, polinomio o producto de estas funciones y la matriz $\boldsymbol{A}$ es independiente de t, es posible proponer una solución particular del mismo tipo que $\boldsymbol{F}(t)$. La razón es que nuevamente, L aplicado a tal función ($L[\boldsymbol{Y}] = \boldsymbol{Y}' - \boldsymbol{A}\boldsymbol{Y}$) será una función del mismo tipo, si se cumplen ciertas condiciones. No obstante, la propuesta involucra ahora vectores de coeficientes a ser determinados. Damos a continuación una tabla de propuestas, similar a la de 1.13.

$\boldsymbol{F}(t)$	$\boldsymbol{Y}_p(t)$	Condición
$\boldsymbol{F}e^{\alpha t}$	$\boldsymbol{V}e^{\alpha t}$	α no es autovalor de $\boldsymbol{A}$
$\boldsymbol{F}_0 + \boldsymbol{F}_1 t + \ldots + \boldsymbol{F}_m t^m$	$\boldsymbol{V}_0 + \boldsymbol{V}_1 t + \ldots + \boldsymbol{V}_m t^m$	$\boldsymbol{A}$ no singular
$e^{\alpha t}(\boldsymbol{F}_0 + \boldsymbol{F}_1 t + \ldots + \boldsymbol{F}_m t^m)$	$e^{\alpha t}(\boldsymbol{V}_0 + \boldsymbol{V}_1 t + \ldots + \boldsymbol{V}_m t^m)$	α no es autovalor de $\boldsymbol{A}$
$\boldsymbol{F}\cos\omega t$ o $\boldsymbol{F}\,\mathrm{sen}\,\omega t$	$\boldsymbol{V}_1\cos\omega t + \boldsymbol{V}_2\,\mathrm{sen}\,\omega t$	$i\omega$ no es autovalor de $\boldsymbol{A}$
$\boldsymbol{F}e^{i\omega t}$	$\boldsymbol{V}e^{i\omega t}$	$i\omega$ no es autovalor de $\boldsymbol{A}$
$\boldsymbol{F}e^{(\alpha+i\omega)t}$	$\boldsymbol{V}e^{(\alpha+i\omega)t}$	$\alpha + i\omega$ no es autovalor de $\boldsymbol{A}$

donde $\boldsymbol{V}$, $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, etc. son vectores con coeficientes a determinar e independientes de t. Si no se cumple la condición se debe multiplicar la propuesta por un polinomio en t (con coeficientes vectores). Nuevamente es más efectivo trabajar en forma compleja cuando $\boldsymbol{F}(t)$ es seno o coseno. La propuesta debe ser completa, aún cuando $\boldsymbol{F}$ tenga componentes no nulas sólo en ciertas filas.

Caso Exponencial. Resolver

$$\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + e^{\alpha t}\boldsymbol{F}, \tag{2.64}$$

con $\boldsymbol{F}$ un vector independiente de t y α un número real o complejo, que no sea autovalor de $\boldsymbol{A}$. Proponemos entonces una solución particular de la forma

$$\boldsymbol{Y}_p(t) = e^{\alpha t}\boldsymbol{V}$$

con $\boldsymbol{V}$ un vector independiente de t. Tenemos $\boldsymbol{Y}_p' = \alpha e^{\alpha t}\boldsymbol{V}$ y reemplazando en (2.64), se obtiene

$$\alpha e^{\alpha t}\boldsymbol{V} = \boldsymbol{A}e^{\alpha t}\boldsymbol{V} + \boldsymbol{F}e^{\alpha t}$$

Por lo tanto, debemos resolver el sistema $\alpha\boldsymbol{V} = \boldsymbol{A}\boldsymbol{V} + \boldsymbol{F}$, es decir,

$$(\alpha I - \boldsymbol{A})\boldsymbol{V} = \boldsymbol{F}$$

que es un sistema lineal no homogéneo para el vector $\boldsymbol{V}$. El sistema tendrá **solución única** $\forall\ \boldsymbol{F}$ sii $\alpha I - \boldsymbol{A}$ es **no singular**, es decir, si α **no es autovalor** de $\boldsymbol{A}$ ($\mathrm{Det}\,(\boldsymbol{A} - \alpha I) \neq 0$), lo que justifica la condición requerida. En tal caso,

$$\boldsymbol{V} = (\alpha I - \boldsymbol{A})^{-1}\boldsymbol{F}$$

Si en cambio α coincide con un autovalor de $\boldsymbol{A}$, el sistema anterior puede no tener solución, por ser la matriz $\boldsymbol{A} - \alpha I$ singular. En tal caso se puede proponer una solución de la forma

$$\boldsymbol{Y}_p(t) = e^{\alpha t}(\boldsymbol{V}_1 + t\boldsymbol{V}_2)$$

con $\boldsymbol{V}_1$ y $\boldsymbol{V}_2$ vectores constantes a determinar. Reemplazando en (2.55) se obtiene el sistema

$$(\alpha I - \boldsymbol{A})\boldsymbol{V}_2 = \boldsymbol{0}, \quad (\alpha I - \boldsymbol{A})\boldsymbol{V}_1 = \boldsymbol{F} - \boldsymbol{V}_2,$$

que muestra que $\boldsymbol{V}_2$ debe ser un autovector de $\boldsymbol{A}$ con autovalor α, tal que la segunda ecuación sea compatible. Si no existe tal vector, se debe agregar un tercer término $\boldsymbol{V}_3 t^2$ en $\boldsymbol{Y}_p$, aunque en estos casos resulta ya más conveniente utilizar el método general (2.57) o (2.60)-(2.61).

Ejemplo. Consideremos

$$\begin{cases} x' = x - y + e^{\alpha t} \\ y' = -x + y \end{cases}$$

que corresponde a $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $\boldsymbol{F} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Proponemos una solución particular

de la forma
$$\boldsymbol{Y}_p(t) = e^{\alpha t}\boldsymbol{V}, \quad \text{con} \quad \boldsymbol{V} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$
Reemplazando, obtenemos el sistema $(\alpha I - \boldsymbol{A})\boldsymbol{V} = \boldsymbol{F}$, es decir,
$$\begin{pmatrix} \alpha - 1 & 1 \\ 1 & \alpha - 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
Los autovalores de $\boldsymbol{A}$ son 2 y 0. Si $\alpha \neq 2$ o $\alpha \neq 0$ la matriz anterior es no singular y tiene la solución única
$$\boldsymbol{V} = (\alpha I - \boldsymbol{A})^{-1}\boldsymbol{F}, \quad (\alpha \neq 2, \;\; \alpha \neq 0)$$
o sea,
$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{\alpha(\alpha - 2)} \begin{pmatrix} \alpha - 1 \\ -1 \end{pmatrix}$$
La solución particular es entonces
$$\boldsymbol{Y}_p(t) = \frac{e^{\alpha t}}{\alpha(\alpha - 2)} \begin{pmatrix} \alpha - 1 \\ 1 \end{pmatrix} \quad (\alpha \neq 2, \;\; \alpha \neq 0)$$
que coincide con el resultado obtenido de (2.63)–(2.62) (probar!).

Caso Polinomial. Resolver
$$\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{F}_0 + \boldsymbol{F}_1 t + \boldsymbol{F}_2 t^2, \tag{2.65}$$
con $\boldsymbol{F}_0$, $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ vectores independientes de t y $\boldsymbol{A}$ no singular.
Proponemos una solución particular de la forma
$$\boldsymbol{Y}_p(t) = \boldsymbol{V}_0 + \boldsymbol{V}_1 t + \boldsymbol{V}_2 t^2$$
con $\boldsymbol{V}_i$ vectores independientes de t. Tenemos $\boldsymbol{Y}_p' = \boldsymbol{V}_1 + 2t\boldsymbol{V}_2$ y reemplazando en (2.64), se obtiene
$$\boldsymbol{V}_1 + 2t\boldsymbol{V}_2 = \boldsymbol{A}(\boldsymbol{V}_0 + t\boldsymbol{V}_1 + t^2\boldsymbol{V}_2) + \boldsymbol{F}_0 + \boldsymbol{F}_1 t + \boldsymbol{F}_2 t^2$$
Igualando los términos de igual grado en t de los dos miembros, obtenemos el sistema
$$\begin{cases} \boldsymbol{A}\boldsymbol{V}_0 = \boldsymbol{V}_1 - \boldsymbol{F}_0 \\ \boldsymbol{A}\boldsymbol{V}_1 = 2\boldsymbol{V}_2 - \boldsymbol{F}_1 \\ \boldsymbol{A}\boldsymbol{V}_2 = -\boldsymbol{F}_2 \end{cases}$$
Si $\boldsymbol{A}$ es **no singular**, comenzando con la última ecuación se obtiene el resultado
$$\boldsymbol{V}_2 = -\boldsymbol{A}^{-1}\boldsymbol{F}_2, \;\; \boldsymbol{V}_1 = \boldsymbol{A}^{-1}(2\boldsymbol{V}_2 - \boldsymbol{F}_1), \;\; \boldsymbol{V}_0 = \boldsymbol{A}^{-1}(\boldsymbol{V}_1 - \boldsymbol{F}_0)$$

Si en cambio $\boldsymbol{A}$ es singular, se debe proponer un polinomio de grado 3 o más alto.

Ejemplo. Consideremos $\left\{\begin{array}{l} x' = -x + y + t \\ y' = -x - y + 1 \end{array}\right.$
que corresponde a $\boldsymbol{A} = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$, $\boldsymbol{F} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
Proponemos una solución particular

$$\boldsymbol{Y}_p(t) = \boldsymbol{V}_0 + t\boldsymbol{V}_1 = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + t \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$$

Reemplazando, obtenemos

$$\boldsymbol{V}_1 = \boldsymbol{A}(\boldsymbol{V}_0 + t\boldsymbol{V}_1) + \binom{0}{1} + t\binom{1}{0}$$

que conduce al sistema

$$\left\{\begin{array}{l} \boldsymbol{A}\boldsymbol{V}_0 = \boldsymbol{V}_1 - \binom{0}{1} \\ \boldsymbol{A}\boldsymbol{V}_1 = -\binom{1}{0} \end{array}\right.$$

Como $\boldsymbol{A}$ es no singular, $\boldsymbol{A}^{-1} = \frac{1}{2}\left(\begin{smallmatrix} -1 & -1 \\ 1 & -1 \end{smallmatrix}\right)$ y

$$\boldsymbol{V}_1 = \boldsymbol{A}^{-1} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}, \quad \boldsymbol{V}_0 = \boldsymbol{A}^{-1}(\boldsymbol{V}_1 - \begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}$$

La solución general del sistema es entonces (probar!)

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{-t} \begin{pmatrix} \cos t \\ -\operatorname{sen} t \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} \operatorname{sen} t \\ \cos t \end{pmatrix} + \begin{pmatrix} (1+t)/2 \\ 1 - t/2 \end{pmatrix}$$

Caso Periódico. Consideremos ahora una fuente de frecuencia angular ω, tal que

$$\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{F}e^{i\omega t}$$

con $\boldsymbol{A}$ y $\boldsymbol{F}$ independientes de t. Supongamos que $i\omega$ no es autovalor de $\boldsymbol{A}$. Proponiendo

$$\boldsymbol{Y}_p(t) = \boldsymbol{V}e^{i\omega t}$$

se obtiene $i\omega\boldsymbol{V}e^{\omega t} = \boldsymbol{A}\boldsymbol{V}e^{\omega t} + \boldsymbol{F}e^{\omega t}$ y por lo tanto

$$(i\omega I - \boldsymbol{A})\boldsymbol{V} = \boldsymbol{F}$$

Dado que $i\omega I - \boldsymbol{A}$ es no singular, podemos determinar $\boldsymbol{V}$ como

$$\boldsymbol{V} = (i\omega I - \boldsymbol{A})^{-1}\boldsymbol{F}$$

Si $\boldsymbol{A}$ y $\boldsymbol{F}$ son reales, podemos así obtener una solución particular real de

$$\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{F}\cos\omega t$$

como

$$\boldsymbol{Y}_p^r(t) = \operatorname{Re}[(i\omega I - \boldsymbol{A})^{-1}\boldsymbol{F}e^{i\omega t}]$$

Las componentes de esta solución oscilan con la frecuencia de la fuente $\boldsymbol{F}(t)$ con una amplitud que depende de la frecuencia, y pueden presentar un desfasaje respecto de $\boldsymbol{F}(t)$.

Si en cambio $i\omega$ es autovalor (resonancia!) entonces debemos proponer una solución $\boldsymbol{Y}_p(t) = e^{i\omega t}(\boldsymbol{V}_1 + t\boldsymbol{V}_2)$.

Ejemplo. Resolver

$$\left\{ \begin{array}{l} x' = y + e^{i\omega t} \\ y' = -x \end{array} \right. \qquad \omega^2 \neq 1$$

que corresponde a $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $\boldsymbol{F}(t) = \boldsymbol{F}e^{i\omega t}$, con $\boldsymbol{F} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$. Proponiendo una solución particular $\boldsymbol{Y}_p(t) = \boldsymbol{V}e^{i\omega t}$ obtenemos, siguiendo el método anterior (probar!)

$$\boldsymbol{V} = (i\omega I - \boldsymbol{A})^{-1}\boldsymbol{F} = \frac{1}{\omega^2 - 1}\left(\begin{array}{c} -i\omega \\ 1 \end{array} \right)$$

Por lo tanto,

$$\boldsymbol{Y}_p(t) = \frac{1}{\omega^2 - 1}\left(\begin{array}{c} -i\omega \\ 1 \end{array} \right) e^{i\omega t}$$

es una solución particular del sistema anterior, y la parte real

$$\boldsymbol{Y}_p^r(t) = \mathrm{Re}[\boldsymbol{Y}_p(t)] = \frac{1}{\omega^2 - 1}\left(\begin{array}{c} \omega \sin \omega t \\ \cos \omega t \end{array} \right)$$

una solución particular del sistema

$$\left\{ \begin{array}{l} x' = y + \cos \omega t \\ y' = -x \end{array} \right.$$

Problema 20: a) Determinar la solución general de este sistema para $\omega \neq 1$.
b) Determine una solución particular en el caso resonante $\omega = 1$.
Nota: Este sistema puede también resolverse como una ecuación de 2^o orden (probar!).

Ejercicios VIII

1) Muestre que las soluciones generales de los siguientes sistemas,

$$a) \left\{ \begin{array}{l} x' = -x + 2y + 1 \\ y' = -x + y + 2 \end{array} \right. \qquad b) \left\{ \begin{array}{l} x' = 6x + y - 6t \\ y' = 4x + 3y - 3 - 4t \end{array} \right.$$

$$c) \left\{ \begin{array}{l} x' = -x + 2y + e^{-t} \\ y' = -x + y + 2 \end{array} \right. \qquad d) \left\{ \begin{array}{l} x' = 5x + 3y - 9e^{-t} \\ y' = -x + y - e^{-t} \end{array} \right.$$

son:
a) $x(t) = c_1(\cos t + \sin t) + c_2(\sin t - \cos t) + 3$, $y(t) = c_1 \cos t + c_2 \sin t + 1$
b) $x(t) = c_1 e^{2t} + c_2 e^{7t} + t$, $y(t) = -4c_1 e^{2t} + c_2 e^{7t} + 1$
c) $x(t) = c_1(\cos t + \sin t) + c_2(\sin t - \cos t) + 4 - e^{-t}$, $y(t) = c_1 \cos t + c_2 \sin t + 2 - e^{-t}/2$.
d) $x(t) = c_1 e^{2t} + 3c_2 e^{4t} + e^{-t}$, $y(t) = -c_1 e^{2t} - c_2 e^{4t} + e^{-t}$.

2) Halle la solución de los sistemas anteriores que satisface $x(0) = y(0) = 0$.

3) Determine la solución general de los sistemas

$$a)\ \begin{cases} x' = 2y + A\,\mathrm{sen}\,\omega t \\ y' = -2x \end{cases} \qquad b)\ \begin{cases} x' = -2x + z \\ y' = -x + z \\ z' = x - 2z + e^{-2t} \end{cases} \qquad c)\ \begin{cases} x' = -x + y \\ y' = -x - y + t \\ z' = x - z + e^{-2t} \end{cases}$$

Dé una expresión real en todos los casos. En a) considere los casos $\omega \neq \pm 2$ y $\omega = 2$.

4) Halle la solución general de los sistemas

$$a)\ \begin{cases} x' = 2y + f(t) \\ y' = -2x \end{cases} \qquad b)\ \begin{cases} x' = -2x + y \\ y' = x - 2y + f(t) \end{cases}$$

2.11. Aplicaciones

5) Resolver el problema de una partícula de carga q en un campo magnético $\boldsymbol{B}$ constante y campo eléctrico $\boldsymbol{E}$. La fuerza es $q(\boldsymbol{v} \times \boldsymbol{B} + \boldsymbol{E})$ y la ecuación

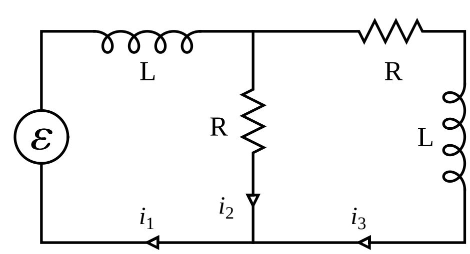
$$m d\boldsymbol{v}/dt = q(\boldsymbol{v} \times \boldsymbol{B} + \boldsymbol{E})$$

que es un sistema de 1^{er} orden no homogéneo. Considerar que $\boldsymbol{B}$ está en la dirección z y
a) $\boldsymbol{E}$ es constante y en la dirección z.
b) $\boldsymbol{E}$ es constante y en la dirección x.
c) $\boldsymbol{E}(t) = \boldsymbol{E} \cos \omega_e t$, con $\boldsymbol{E}$ en la dirección x. ¿Existe resonancia para algún valor de ω_e?
Grafique e interprete las soluciones.

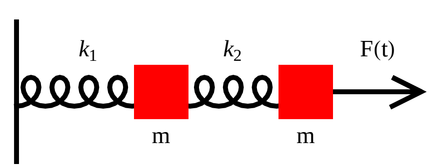
6) Determinar las corrientes $i_1(t)$, $i_2(t)$ $i_3(t)$ en el circuito de la figura. Plantear el sistema de ecuaciones lineales de primer orden correspondiente y resolverlo para a) un fem constante ε_0 y b) una fem alterna $\varepsilon = V_0 \cos \omega t$, para condición inicial $i_1(0) = i_2(0) = i_3(0) = 0$. Grafique las soluciones. Muestre que el sistema es

$$\begin{cases} \varepsilon = L di_1/dt + R(i_1 - i_3) \\ 0 = L di_3/dt + 2R\, i_3 - R\, i_1 \end{cases}$$

con $i_1 = i_2 + i_3$.



7) Resolver el problema de dos masas acopladas sujetas a una fuerza externa $F(t)$ aplicada sobre la segunda masa. Considerar los casos:
a) $F(t) = F$ (constante)
b) $F(t) = F \cos \omega_{ex} t$.
Determine en b) las frecuencias resonantes.
El sistema homogéneo ya fue estudiado en la sección 2.6.



Temario por clase:

Clase 1: Funciones periódicas. Funciones ortogonales. Series de Fourier.

Clase 2: Ecuaciones diferenciales en derivadas parciales. Ecuación de difusión. Ecuación de ondas. Ecuación de Laplace. Resolución por separación de variables y serie de Fourier.

Clase 3: Ejercicios y Aplicaciones. Discusión.

Bibliografía:

1. R. Churchill, Series de Fourier y Problemas de Contorno, McGraw-Hill
2. S. Salvioli, Ecuaciones Diferenciales (CEILP).
3. C.H. Edwards, D.E. Penney, Ecuaciones Diferenciales Elementales y Problemas con Condiciones en la Frontera, Prentice-Hall Hispanoamericana.
4. D.G. Zill, Ecuaciones Diferenciales con aplicaciones, Editorial Iberoamérica.
5. Ecuaciones Diferenciales en Física, C.M. Naón, M. Santangelo, R. Rossignoli, EDULP (2014).

Facultad de Ingeniería
UNLP

Matemática C

X Series de Fourier
y
Ecuaciones Diferenciales en Derivadas Parciales

2022

I. Series de Fourier

1. Introducción

Las series de Fourier cumplen un rol fundamental en el análisis de funciones periódicas. Permiten escribir estas funciones como una serie de funciones sinusoidales (senos y cosenos), y así obtener una descripción precisa de las mismas por medio de los coeficientes de la serie. Constituyen, por lo tanto, una herramienta fundamental no solo en Matemática, sino también en diversas áreas de la Ingeniería, Física, Astronomía y otras Ciencias, tales como análisis y generación de señales, telecomunicaciones, ingeniería eléctrica, acústica, óptica, procesamiento de imágenes, compresión de datos, mecánica cuántica, etc. Por ejemplo, permite comprender en forma inmediata el origen de la diferencia entre un "La" de frecuencia 440 Hz emtido por un piano y un "La" de la misma frecuencia emitido por otro instrumento, a través los distintos coeficientes de la serie correspondiente (distribución de armónicos).

Como veremos en la sección siguiente, las series de Fourier juegan también un rol esencial en la resolución de ecuaciones diferenciales en derivadas parciales, de importancia en Ingeniería, Física y Matemática, y surgieron de hecho en relación con estas ecuaciones. El nombre proviene del matemático y físico francés Jean-Baptiste Joseph Fourier (1768-1830), quien desarrolló la teoría con el objeto de resolver la ecuación del calor, una ecuación diferencial en derivadas parciales. Otros matemáticos y físico-matemáticos que contribuyeron (antes y después) al desarrollo de la teoría incluyen a Leonhard Euler (1707-1783), D'Alembert, Daniel Bernoulli, Gauss, Dirichlet y Riemann. Y la idea de descomponer oscilaciones complejas en oscilaciones simples es aun más antigua, existiendo antecedentes incluso en el sigo 3 a.C.

Por otro lado, el desarrollo de la teoría general continúa hoy en diversas formas. El análisis de Fourier dió origen a teorías generales de análisis de señales y series temporales. Generalizaciones recientes incluyen las denominadas transformadas wavelet, transformadas de Fourier fraccionarias, etc. Y el desarrollo reciente de la computación cuántica fue impulsado precisamente por el descubrimiento de un algoritmo cuántico eficiente para implementar la transformada de Fourier en sistemas cuánticos.

2. Funciones periódicas

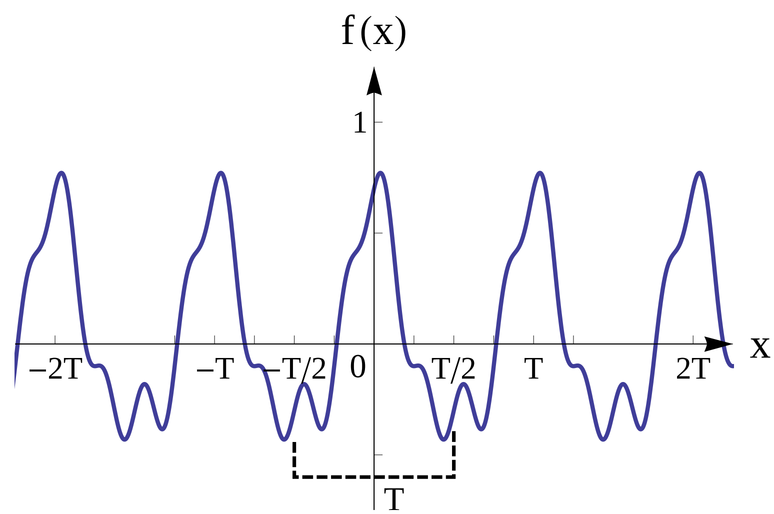
Una función $f : \mathbb{R} \to \mathbb{R}$ es periódica con período $T > 0$ si satisface

$$f(x+T) = f(x) \quad \forall x \in \mathbb{R}$$

En tal caso también cumple $f(x+2T) = f(x+T+T) = f(x+T) = f(x)$, y en general

$$f(x+nT) = f(x) \;\; \forall \, x \in \mathbb{R}\,, \quad n = 0, \pm 1, \pm 2, \ldots$$

Basta entonces conocer los valores de $f(x)$ en cualquier intervalo de longitud T (por ejemplo, $(-T/2, T/2]$) para determinar su valor $f(x)$ en cualquier $x \in \mathbb{R}$.



Ejemplos bien conocidos de funciones periódicas con período T son

$$f_1(x) = \cos(2\pi x/T) \quad (1)$$

$$f_2(x) = \operatorname{sen}(2\pi x/T) \quad (2)$$

ya que $\cos(\frac{2\pi}{T}(x+T)) = \cos(\frac{2\pi}{T}x + 2\pi) = \cos(\frac{2\pi}{T}x) \; \forall \; x$ (igualmente para $\operatorname{sen}(\frac{2\pi}{T}x)$). También

$$f(x) = A\cos(2\pi x/T - \phi) \quad (3)$$

es periódica con período T para cualquier valor de A (amplitud) y ϕ (fase). Y una función constante $f(x) = c \; \forall \; x$ es un caso trivial de función periódica. Esto incluye en particular a la función nula $0(x) = 0 \; \forall \; x$.

Si $f_1(x)$ y $f_2(x)$ son dos funciones periódicas con el mismo período T, toda combinación lineal

$$f(x) = c_1 f_1(x) + c_2 f_2(x) \quad (4)$$

es también una función periódica con el mismo período T, como es muy fácil probar. El conjunto de funciones periódicas con período T es entonces un **subespacio** (de dimensión infinita) del espacio vectorial de funciones generales $f : \mathbb{R} \to \mathbb{R}$.

Por ejemplo, la función (3) es una combinación lineal de las funciones (1) y (2):

$$A\cos(2\pi x/T - \phi) = (A\cos\phi)\,\cos(2\pi x/T) + (A\operatorname{sen}\phi)\operatorname{sen}(2\pi x/T)$$

Toda función periódica con período T/n, con $n \geq 1$ natural, es también periódica con período T: $f(x+T) = f(x + n(T/n)) = f(x)\ \forall\ x$. Por lo tanto, las funciones

$$\cos(2n\pi x/T)\,, \qquad \mathrm{sen}(2n\pi x/T)\,, \qquad n = 1, 2, 3, \ldots \tag{5}$$

que pueden escribirse como $\cos(2\pi x/T_n)$, $\mathrm{sen}(2\pi x/T_n)$, con $T_n = T/n$, son también periódicas con período T, ya que su período es $T_n = T/n$.

Cuando x representa un tiempo t, suelen escribirse las funciones (5) como

$$\cos(n\omega t)\,, \qquad \mathrm{sen}(n\omega t)\,, \qquad \omega = 2\pi/T, \quad n = 1, 2, 3, \ldots \tag{6}$$

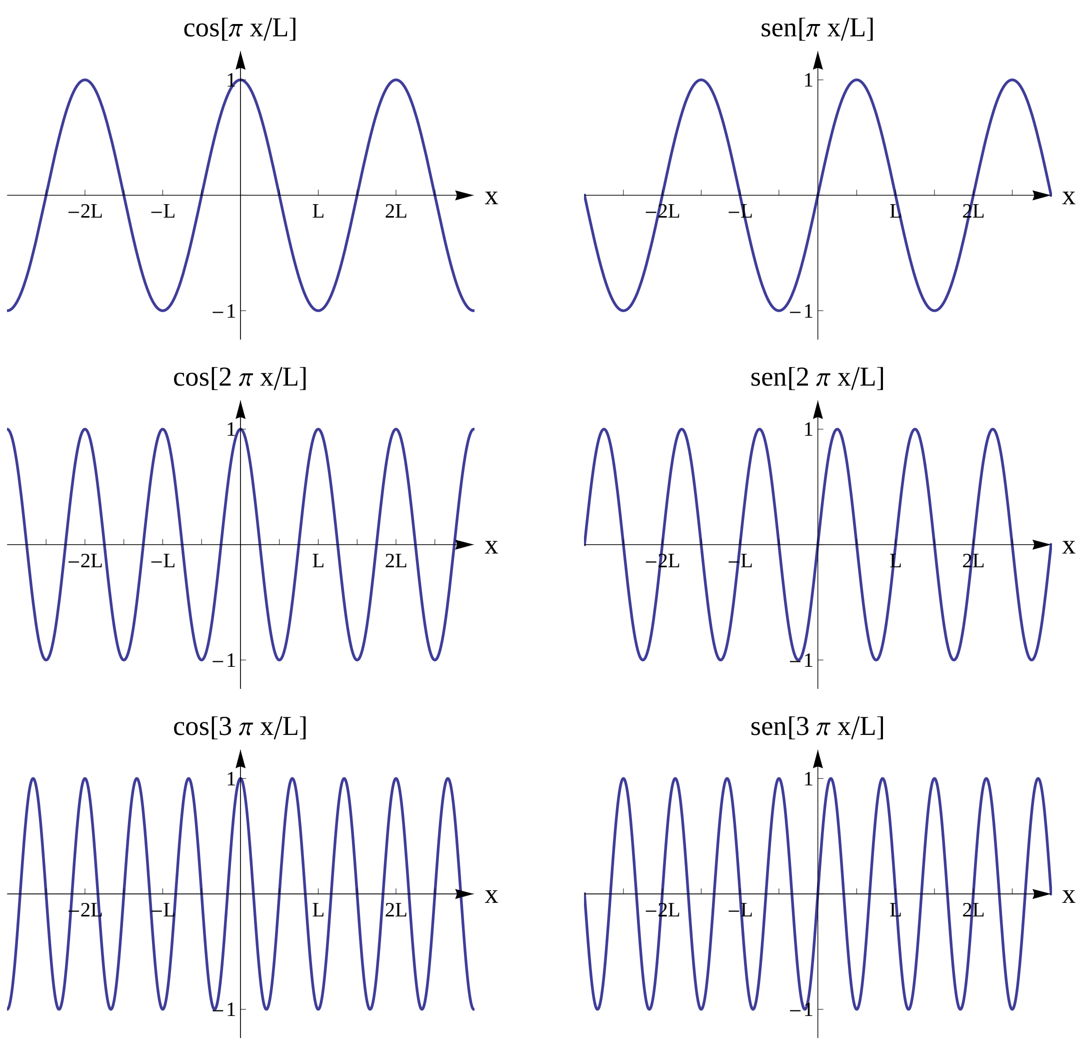
donde $\omega = 2\pi f$ es la frecuencia angular y $f = 1/T$ la frecuencia fundamental.

Y en el contexto de series de Fourier, suele escribirse $T = 2L$ y las funciones (5) como

$$\cos(n\pi x/L)\,, \qquad \mathrm{sen}(n\pi x/L)\,, \qquad L = T/2\,, \quad n = 1, 2, 3, \ldots \tag{7}$$

Se muestran abajo los gráficos de estas funciones para $n \leq 3$. Notar que:

- $\cos(n\pi x/L)$ es una función **par**: satisface $\boldsymbol{f(-x) = f(x)}\ \forall\ x$
- $\mathrm{sen}(n\pi x/L)$ es una función **impar**: satisface $\boldsymbol{f(-x) = -f(x)}\ \forall\ x$.



3. Conjunto ortogonal de funciones

Para definir ortogonalidad entre funciones, debemos definir primero un producto interno (análogo al producto escalar entre vectores de $\mathbb{R}^n$). Dadas dos funciones periódicas reales f y g de período $T = 2L > 0$, continuas, definimos aquí su producto interno como

$$(f, g) = \frac{1}{L}\int_{-L}^{L} f(x)g(x)dx \tag{8}$$

que satisface $(f, g) = (g, f)$. Si $g = f$, se tiene

$$(f, f) = \frac{1}{L}\int_{-L}^{L}[f(x)]^2 dx \geq 0$$

La **norma** de una función f se define entonces como $||f|| = \sqrt{(f, f)}$. Satisface: $||f|| \geq 0$, con $||f|| = 0$ para f continua sólo si f es la función nula. Además, se verifica:
1) $||\alpha f|| = |\alpha|\, ||f||$ ($\alpha \in \mathbb{R}$) y 2) $|(f, g)| \leq ||f||\, ||g||$ (desigualdad de Cauchy-Schwarz).
La propiedad (2) implica (probar!) $|||f_1|| - ||f_2||| \leq ||f_1 + f_2|| \leq ||f_1|| + ||f_2||$.

El producto (8) puede extenderse a funciones seccionalmente continuas (es decir, que tengan a lo sumo un número finito de discontinuidades finitas en un período) o integrables.

Con la definición (8), dos funciones f, g son **ortogonales** si y solo si

$$(f, g) = \frac{1}{L}\int_{-L}^{L} f(x)g(x)dx = 0 \tag{9}$$

Una propiedad notable del conjunto de funciones $\{1, \cos(n\pi x/L), \mathrm{sen}(n\pi x/L),\ n = 1, 2, \ldots\}$, que incluye la funcón constante $1 = \cos(0\pi x/L)$, es que resultan **ortogonales** con este producto interno, con **norma** 1 para $n \geq 1$: Utilizando las fórmulas

$$\begin{aligned}
\cos(a)\cos(b) &= \frac{1}{2}[\cos(a-b) + \cos(a+b)] && (10)\\
\mathrm{sen}(a)\,\mathrm{sen}(b) &= \frac{1}{2}[\cos(a-b) - \cos(a+b)] && (11)\\
\mathrm{sen}(a)\cos(b) &= \frac{1}{2}[\mathrm{sen}(a-b) + \mathrm{sen}(a+b)] && (12)
\end{aligned}$$

es muy fácil probar que para $n, m = 0, 1, 2, \ldots$, se cumple

$$\begin{aligned}
\frac{1}{L}\int_{-L}^{L}\cos(\frac{n\pi x}{L})\cos(\frac{m\pi x}{L})dx &= \begin{cases} 0 & m \neq n \\ 1 & m = n \geq 1 \\ 2 & m = n = 0 \end{cases} && (13)\\
\frac{1}{L}\int_{-L}^{L}\mathrm{sen}(\frac{n\pi x}{L})\,\mathrm{sen}(\frac{m\pi x}{L})dx &= \begin{cases} 0 & m \neq n \\ 1 & m = n \geq 1 \end{cases} && (14)\\
\frac{1}{L}\int_{-L}^{L}\mathrm{sen}(\frac{n\pi x}{L})\cos(\frac{m\pi x}{L})dx &= 0 \ \ \forall\, m, n && (15)
\end{aligned}$$

El resultado (15) es además obvio por **paridad**: el coseno es una función par, el seno impar, y su producto es entonces impar. Y la integral de toda función impar en un intervalo $[-L, L]$ es 0 (justificar!).

Por lo tanto, dada una función periódica $f(x)$ definida por la combinación lineal

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{N}[a_n \cos(\frac{n\pi x}{L}) + b_n \operatorname{sen}(\frac{n\pi x}{L})] \tag{16}$$

donde $N \geq 1$ y hemos escrito el coeficiente correspondiente a la función constante $1 = \cos(0\pi x/L)$ como $a_0/2$, se obtiene, usando las fórmulas (13)–(15) (probar!)

$$\begin{aligned} \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{m\pi x}{L})dx &= a_m, \quad m = 0, 1, 2, \ldots, N & (17)\\ \frac{1}{L}\int_{-L}^{L} f(x)\operatorname{sen}(\frac{m\pi x}{L})dx &= b_m, \quad m = 1, 2, \ldots, N & (18) \end{aligned}$$

Por lo tanto, si se conoce $f(x)$ en un período completo y se sabe que es de la forma (16), los coeficientes a_n, b_n pueden obtenerse como

$$\begin{aligned} a_n &= \frac{1}{L}\int_{-L}^{L} f(x)\cos(n\pi x/L)dx \qquad n = 0, 1, 2, \ldots, N & (19)\\ b_n &= \frac{1}{L}\int_{-L}^{L} f(x)\operatorname{sen}(n\pi x/L)dx \qquad n = 1, 2, \ldots, N & (20) \end{aligned}$$

Tal como ocurre con los vectores de $\mathbb{R}^n$ cuando se los escribe en la base canónica, $\boldsymbol{v} = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$, con $x_i = \boldsymbol{v} \cdot \boldsymbol{e_i}$, las coordenadas a_n, b_n de la función f en este conjunto ortogonal de funciones pueden determinarse mediante productos escalares: $a_n = (f, \cos(n\pi x/L))$, $b_n = (f, \operatorname{sen}(n\pi x/L))$.

Y así como $x_i \boldsymbol{e}_i$ es la **proyección ortogonal** de $\boldsymbol{v}$ sobre $\boldsymbol{e}_i$, $a_n \cos(n\pi x/L)$ es la proyección ortogonal de f sobre $\cos(n\pi x/L)$, y $b_n \operatorname{sen}(n\pi x/L)$ aquella sobre $\operatorname{sen}(n\pi x/L)$.

Para $n = 0$, (19) implica $a_0 = \frac{1}{L}\int_{-L}^{L} f(x)dx$, por lo que el primer término en (16),

$$\frac{a_0}{2} = \frac{1}{2L}\int_{-L}^{L} f(x)dx \tag{21}$$

representa el **valor medio** de f en el intervalo $[-L, L]$, es decir, en un período.

Otra consecuencia de la ortogonalidad es que usando nuevamente (13)–(15), o directamente (16) y (19)–(20), el cuadrado de la norma de (16) puede expresarse como (probar!)

$$||f||^2 = (f, f) \;=\; \frac{1}{L}\int_{-L}^{L} f(x)f(x)dx = \frac{a_0^2}{2} + \sum_{n=1}^{N}(a_n^2 + b_n^2) \tag{22}$$

Esencialmente, esta es la extensión de la expresión $|\boldsymbol{v}|^2 = \boldsymbol{v}\cdot\boldsymbol{v} = \sum_{i=1}^{n} x_i^2$ para el cuadrado de la longitud de un vector $\boldsymbol{v} \in \mathbb{R}^n$, al presente caso.

Observemos también que si $f(x)$ es **par** ($f(-x) = f(x)$), (20) implica $b_n = 0 \ \forall\ n$, mientras que si $f(x)$ es **impar** ($f(-x) = -f(x)$), (19) implica $a_n = 0 \ \forall\ n$.

Surge ahora naturalmente la pregunta ¿Es posible representar cualquier función periódica f de período $2L$ por una suma de la forma (16) con $N \to \infty$?

4. Desarrollo en serie de Fourier

El desarrollo en serie de Fourier de una función periódica f consiste esencialmente en tomar el límite $N \to \infty$ en la suma (16).

Sea $f : \mathbb{R} \to \mathbb{R}$ una función **periódica** de período $T = 2L$, continua y de variación acotada (por ejemplo, posee a lo sumo un número finito de máximos y mínimos aislados en un período). Entonces se la puede representar mediante la serie de Fourier

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}[a_n \cos(n\pi x/L) + b_n \sin(n\pi x/L)] \tag{23}$$

donde los coeficientes a_n, b_n están dados por

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(n\pi x/L)dx \qquad n = 0, 1, 2, \ldots, \tag{24}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\,\mathrm{sen}(n\pi x/L)dx \qquad n = 1, 2, \ldots, \tag{25}$$

El primer término de la serie, $a_0/2$, es el **valor medio** de f en un período:

$$\frac{a_0}{2} = \frac{1}{2L}\int_{-L}^{L} f(x)dx = \langle f\rangle \tag{26}$$

Además, vale la igualdad (Teorema de Parseval)

$$(f,f) = \frac{1}{L}\int_{-L}^{L}[f(x)]^2dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) \tag{27}$$

La convergencia de la serie de Fourier (23) puede extenderse a otros casos:
1) En las condiciones anteriores, si f tiene un número finito de discontinuidades finitas en un período, existiendo los límites laterales de f en los puntos de discontinuidad (f seccionalmente continua), la serie de Fourier (23) con los coeficientes (24)–(25),

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) + b_n \sin(n\pi x/L) \tag{28}$$

converge a $f(x)$ en los puntos x donde f es continua, y al punto medio

$$\frac{f(x_i^+) + f(x_i^-)}{2} \tag{29}$$

en los puntos x_i en los que f es discontinua, donde $f(x_i^\pm) = \lim\limits_{x \to x_i^\pm} f(x)$ son los límites laterales.
2) Si f no cumple las condiciones anteriores pero $(f,f) = \frac{1}{L}\int_{-L}^{L}[f(x)]^2dx$ existe, entonces $\lim\limits_{N\to\infty}\int_{-L}^{L}[f(x) - S_N(x)]^2dx = 0$ (convergencia en media), donde $S_N(x)$ es la suma parcial (16) de la serie (28). Esto no garantiza la convergencia puntual ($S(x) = f(x)$) en todo punto.

Observaciones importantes:

1. Si $f(x)$ es **par**, $f(-x) = f(x)\ \forall\ x$, $b_n = 0\ \forall\ n$ y la serie de Fourier se reduce a

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) \tag{30}$$

con

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(n\pi x/L)dx = \frac{2}{L}\int_{0}^{L} f(x)\cos(n\pi x/L)dx \tag{31}$$

2. Si $f(x)$ es **impar**, $f(-x) = -f(x)\ \forall\ x$, $a_n = 0\ \forall\ n$ y la serie de Fourier se reduce a

$$S(x) = \sum_{n=1}^{\infty} b_n \operatorname{sen}(n\pi x/L) \tag{32}$$

con

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\operatorname{sen}(n\pi x/L)dx = \frac{2}{L}\int_{0}^{L} f(x)\operatorname{sen}(n\pi x/L)dx \tag{33}$$

3. Si f no es periódica o está definida en $[-L, L]$, la serie de Fourier (28) con los coeficientes (24)–(25) convergerá fuera de $(-L, L)$ **a la extensión periódica de la función**.

4. Si $f(-L) \neq f(L)$, la extensión periódica de f será **discontinua** en $x = \pm L$ (y en $\pm 3L, \pm 5L, \ldots$). En estos puntos la serie de Fourier (28) convergerá al punto medio:

$$S(\pm L) = \frac{f(L) + f(-L)}{2} \tag{34}$$

5. Cuando f es periódica, las integrales en (24)–(25) pueden tomarse en cualquier intervalo de longitud $2L$: $\int_{-L}^{L} f(x)\cos(n\pi x/L)dx = \int_{-L+c}^{L+c} f(x)\cos(n\pi x/L)dx$, $\int_{-L}^{L} f(x)\operatorname{sen}(n\pi x/L)dx = \int_{-L+c}^{L+c} f(x)\operatorname{sen}(n\pi x/L)dx$, $\forall\ c \in \mathbb{R}$.

6. Los términos con $n = 1$ se denominan componentes armónicos fundamentales o primer armónico. Los restantes son los componentes armónicos de orden superior.

7. Si a_n, b_n y a'_n, b'_n son los coeficientes de Fourier de dos funciones f y g que satisfacen las condiciones de convergencia, el resultado (27) se extiende a

$$(f, g) = \frac{1}{L}\int_{-L}^{L} f(x)g(x)dx = \frac{a_0 a'_0}{2} + \sum_{n=1}^{\infty}(a_n a'_n + b_n b'_n) \tag{35}$$

8. Si se desea aproximar $f(x)$ por una suma finita de la forma

$$S_N(x) = c_0/2 + \sum_{n=1}^{N}[c_n \cos(n\pi x/L) + d_n \operatorname{sen}(n\pi x/L)] \tag{36}$$

la **mejor aproximación** en el sentido de minimizar $||f - S_N||^2 = \frac{1}{L}\int_{-L}^{L}[f(x) - S_N(x)]^2 dx$, se obtiene para los **coeficientes de Fourier:** $c_n = a_n$, $b_n = d_n\ \forall\ n$.

Ejemplo 1: $f(x) = x^2/L^2$, $|x| \leq L$.

Como $f(L) = f(-L) = 1$, la extensión periódica de esta función es **continua** (ver figura). El desarrollo en serie de Fourier de f convergerá a esta extensión, dada para $x \in \mathbb{R}$ por $f(x) = f(x - m_x L)$, con m_x un entero tal que $x - m_x L \in (-L, L]$.

Como $f(x)$ es par, $b_n = 0\ \forall\ n$ y (ver (30)–(31))

$$a_0 = \frac{2}{L}\int_0^L \frac{x^2}{L^2}dx = \frac{2}{3}, \qquad a_n = \frac{2}{L}\int_0^L \frac{x^2}{L^2}\cos(n\pi x/L)dx = \frac{4(-1)^n}{n^2\pi^2}\,, \quad n \geq 1$$

donde para $n \geq 1$ se ha integrado por partes dos veces. Se obtiene

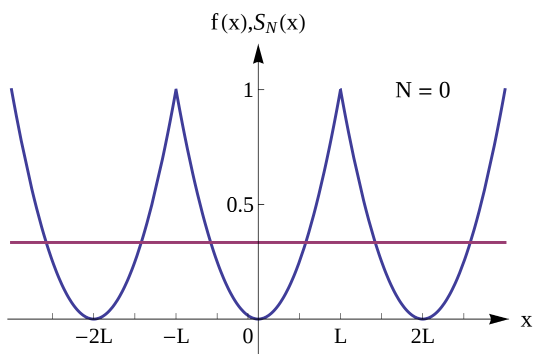
$$f(x) \;=\; \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) = \frac{1}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^2\pi^2}\cos\left(\frac{n\pi}{L}x\right) \tag{37}$$

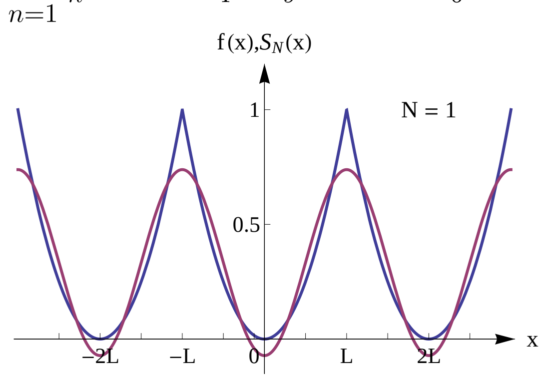
Se muestran en las figuras las primeras sumas parciales $S_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(\frac{n\pi}{L}x)$ junto con la extensión periódica de $f(x)$ y los coeficientes a_n. La serie converge absolutamente y la convergencia es rápida, ya que $a_n \propto n^{-2}$ tiende rápidamente a 0 para $n \to \infty$.

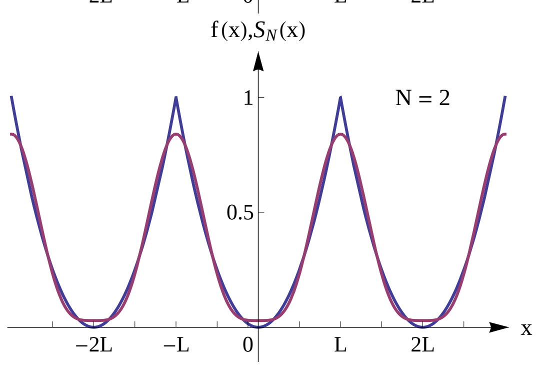
Este desarrollo permite también obtener en forma analítica la suma de series numéricas notables. Por ejemplo, si $x = 0$, $f(0) = 0$ y (37) implica $0 = \frac{1}{3} + \frac{4}{\pi^2}\sum_{n=1}^{\infty}(-1)^n \frac{1}{n^2}$, o sea

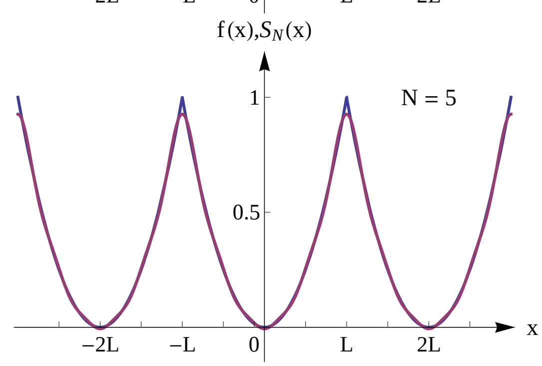
$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n^2} = 1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \ldots = \frac{\pi^2}{12}$$

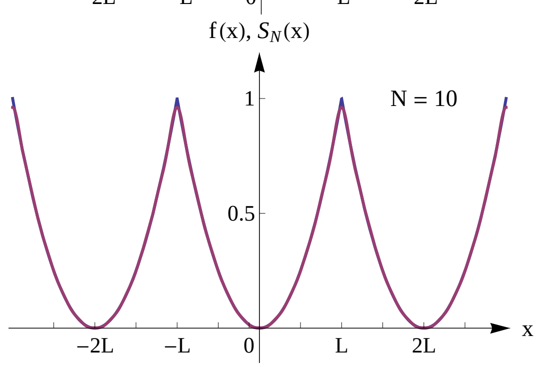
Y si $x = L$, $f(L) = 1$ y (37) implica (probar!) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \ldots = \frac{\pi^2}{6}$

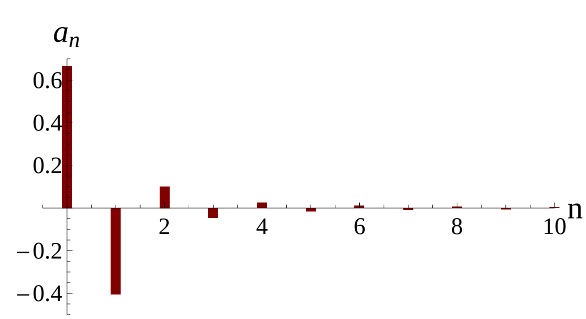












Ejemplo 2: $f(x) = x/L$, $|x| \leq L$.

En este caso $f(-L) = -1 \neq f(L) = 1$, por lo que la extensión periódica de f es **discontinua** en $x = \pm L, \pm 3L, \ldots$, con $f(L_-) = 1$, $f(L_+) = -1$ (ver figura).

Como $f(x)$ es impar, $a_n = 0 \; \forall \; n$ y (ver (32)–(33))

$$b_n = \frac{2}{L} \int_0^L \frac{x}{L} \operatorname{sen} \frac{n\pi x}{L} dx = \frac{2(-1)^{n+1}}{n\pi}\,, \; n \geq 1$$

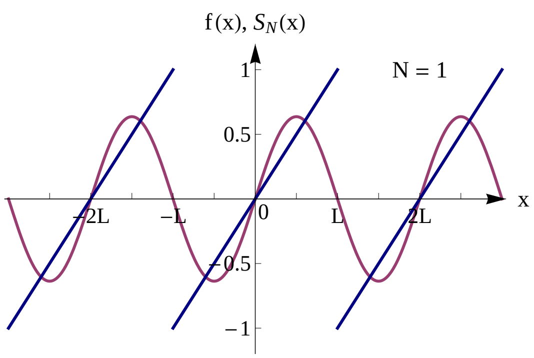
donde se ha integrado por partes. La serie de Fourier es entonces

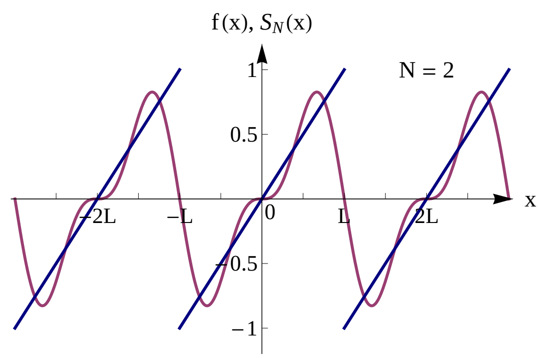
$$S(x) \;=\; \sum_{n=1}^{\infty} b_n \operatorname{sen}(n\pi x/L) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n\pi} \operatorname{sen}\left(\frac{n\pi}{L} x\right) \tag{38}$$

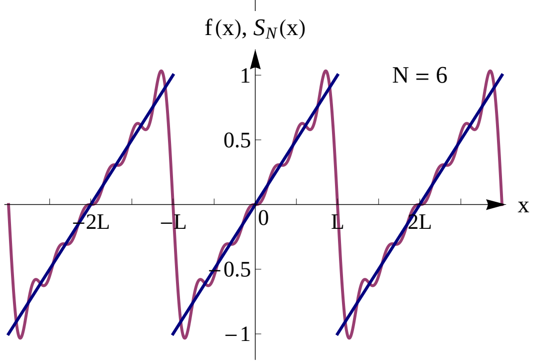
y converge condicionalmente. Esta serie converge a x/L para $x \in (-L, L)$ y al **punto medio** $\frac{f(L)+f(-L)}{2} = 0$ en $x = \pm L$, como es fácil verificar.

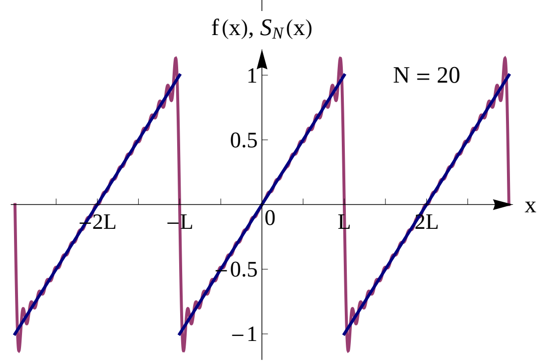
Se muestran abajo las sumas parciales $S_N(x) = \sum_{n=1}^{N} b_n \operatorname{sen}(n\pi x/L)$ junto con la extensión periódica de $f(x)$ y los coeficientes b_n. La convergencia a f es ahora más lenta, especialmente en la vecindad de la discontinuidad, observándose el **fenómeno de Gibbs**.

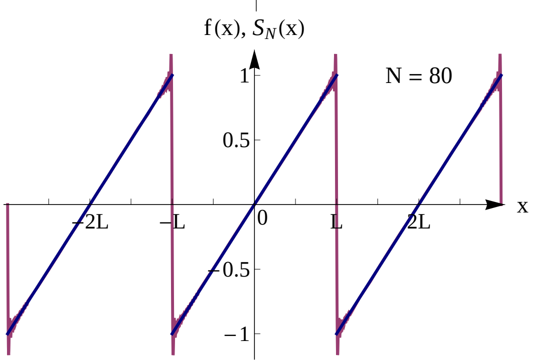
Dado que $\frac{x}{L} = \frac{L}{2}(\frac{x^2}{L^2})'$, la serie (38) puede obtenerse derivando el desarrollo (37) y multiplicándolo por $L/2$. Esto es válido pues la extensión periódica de x^2/L^2 no presenta discontinuidades. En cambio, la derivada término a término del desarrollo (38) no es una serie convergente, debido a las discontinuidades de la extensión periódica de x/L (que implican convergencia no uniforme de (38); tal derivada converge, no obstante, como *distribución* a $1 - 2\sum_{m=-\infty}^{\infty} \delta(x - L + 2mL)$, tema que no profundizaremos en este curso).

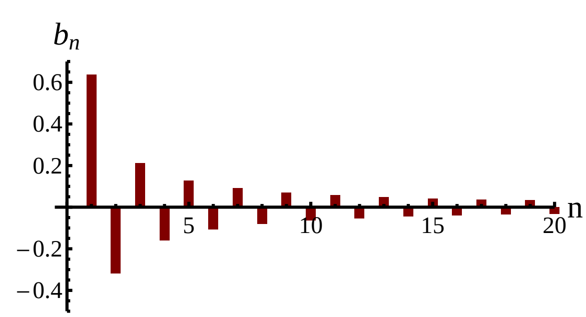


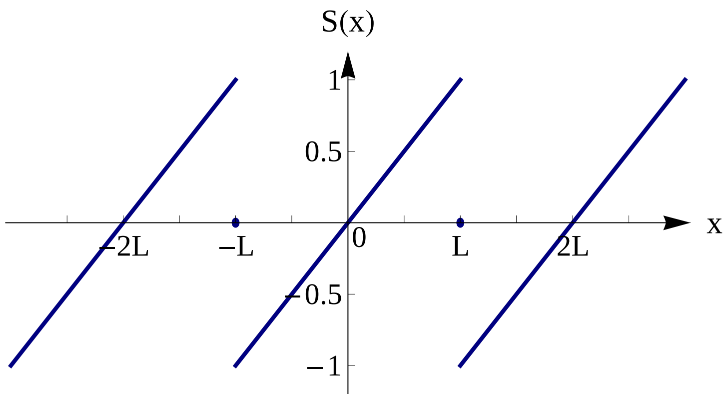


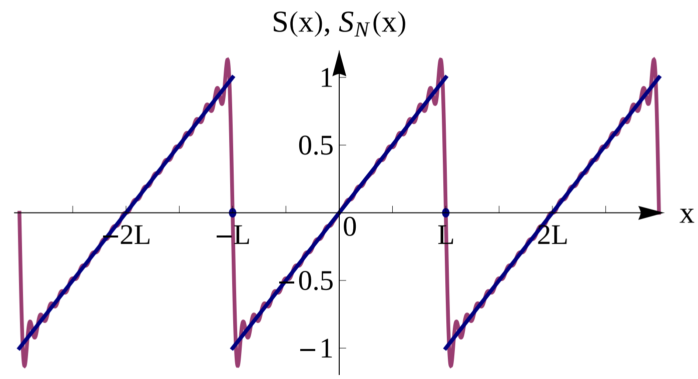












Fenómeno de Gibbs. En estas figuras se muestra a la izquierda la suma exacta $S(x)$ de la serie de Fourier (38) del ejemplo anterior, que es discontinua, convergiendo en las discontinuidades $x = \pm L, \pm 3L, \ldots$ al punto medio 0, y a la derecha $S(x)$ junto con la suma parcial $S_N(x)$ para $N = 20$. $S_N(x)$ es continua para cualquier N finito y exhibe un máximo y un mínimo en los bordes de cada discontinuidad. El exceso en estos bordes sobre el valor exacto es del orden del 9 % del salto y **no disminuye** al aumentar N. No obstante, el ancho de estos máximos o mínimos (el intervalo donde ocurre el exceso) disminuye al aumentar N, tendiendo a 0 para $N \to \infty$.

Este exceso de las sumas parciales finitas de Fourier ocurre siempre en los bordes de toda discontinuidad finita y se denomina fenómeno de Gibbs. Es consecuencia de la convergencia no uniforme. El valor exacto del porcentaje de exceso es $\gamma = \frac{1}{2}\int_0^\pi \frac{\operatorname{sen}\theta}{\theta} d\theta = 0{,}089489\ldots$.

Ejemplo 3: $f(x) = 4 + 2\cos(3\pi x/L) - 6\operatorname{sen}(2\pi x/L)$.

En este caso se trata de una función periódica que es combinación lineal de las funciones ortogonales $\cos(n\pi x/L)$ y $\operatorname{sen}(n\pi x/L)$. Por lo tanto la expresión que la define ya es su desarrollo en serie de Fourier. Usando la ortogonalidad se verifica que

$$\begin{aligned}
a_0 &= \frac{1}{L}\int_{-L}^{L} f(x)dx = 8, \qquad a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos(\frac{n\pi}{L}x)dx = \left\{\begin{array}{ll} 2 & n = 3 \\ 0 & n \neq 3,\ n \geq 1 \end{array}\right. \\
b_n &= \frac{1}{L}\int_{-L}^{L} f(x)\operatorname{sen}(\frac{n\pi}{L}x)dx = \left\{\begin{array}{ll} -6 & n = 2 \\ 0 & n \neq 2 \end{array}\right.
\end{aligned} \tag{39}$$

por lo que la serie se reduce a una suma finita.

Ejemplo 4: $f(x) = \begin{cases} 1 & |x| \leq d \\ 0 & d < |x| \leq L \end{cases}$

Esta función (pulso rectangular, ver figura) posee discontinuidades finitas en $x = \pm d$. Como es par, $b_n = 0 \ \forall \ n$ y

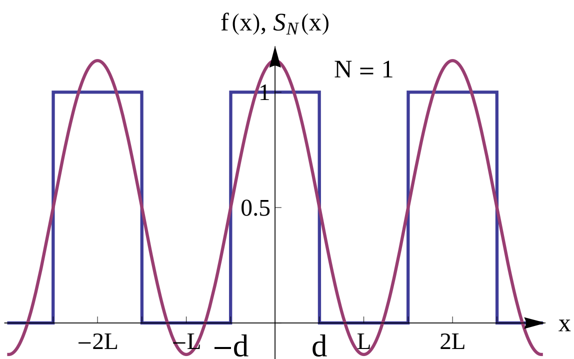
$$a_0 = \frac{2}{L}\int_0^d dx = \frac{2d}{L}, \quad a_n = \frac{2}{L}\int_0^d \cos(n\pi x/L)dx = \frac{2}{n\pi}\,\mathrm{sen}(n\pi d/L)\,, \quad n \geq 1 \qquad (40)$$

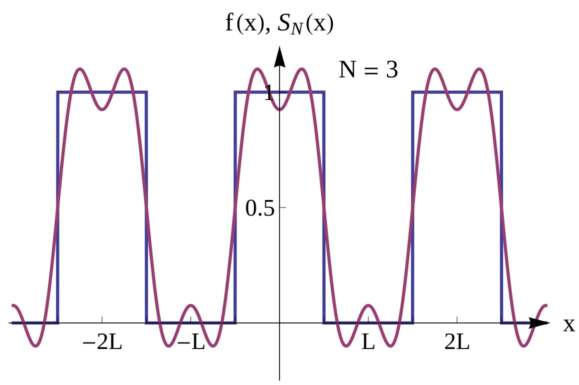
Por lo tanto, la serie de Fourier es

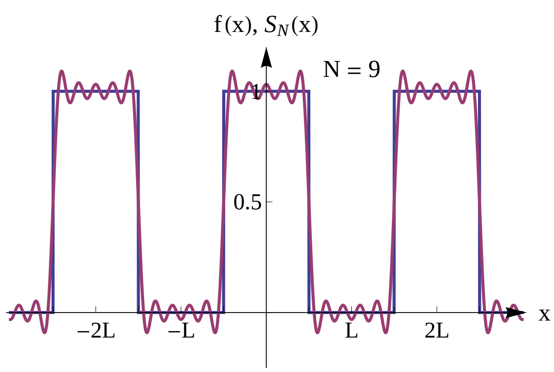
$$S(x) = \frac{d}{L} + \sum_{n=1}^{\infty} \frac{2}{n\pi}\,\mathrm{sen}(n\pi d/L)\cos(n\pi x/L)$$

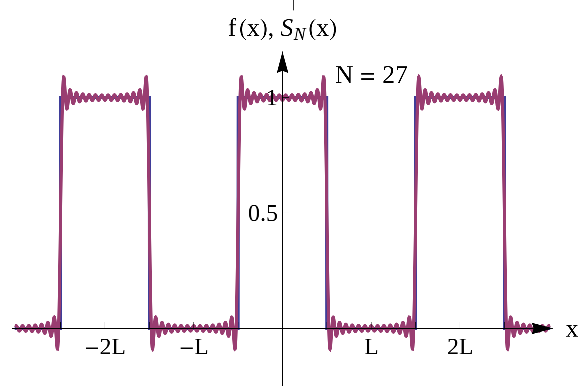
y converge condicionalmente. En $[-L, L]$, converge a $f(x)$ para $|x| \neq d$ y al punto medio $1/2$ para $x = \pm d$. Nuevamente, se observa en las figuras, obtenidas para $d = L/2$, la convergencia lenta en los bordes de la discontinuidad y el fenómeno de Gibbs.

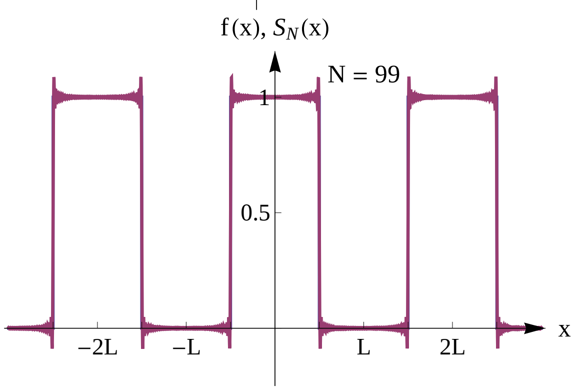
Una observación importante es que $|a_n|/a_0$ es pequeño recien para $n \gg L/d$. Así, el número de términos necesarios para representar el pulso correctamente aumenta al disminuir d en relación a L. O sea, cuanto más corto es el pulso en relación al período. mayor es el número de frecuencias necesarias para representarlo. De ahi el nombre de “ultra wideband” para la tecnología de comunicación basada en pulsos cortos. Esta conclusión es general: Mayor concentración en tiempo $\Rightarrow$ mayor dispersión en frecuencia, y viceversa.

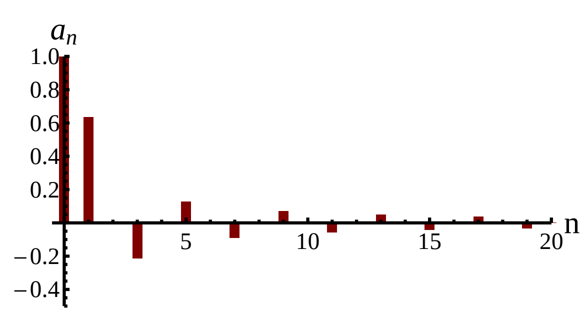




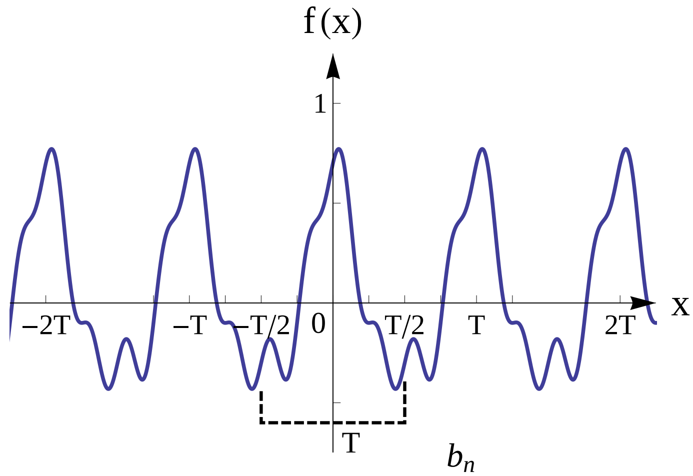


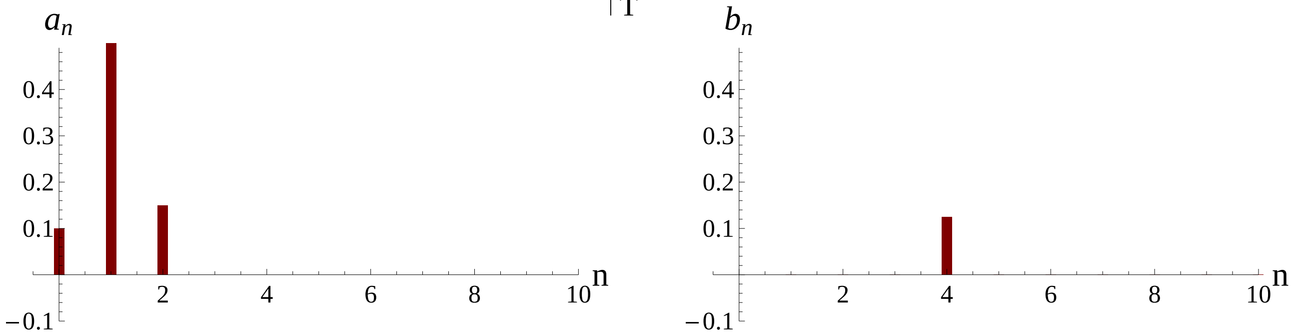






Ejemplo 5.





Análisis de Fourier ("radiografía") de la señal periódica de la figura 1. El cálculo de los coeficientes de Fourier muestra que no es más que la superposición de 3 armónicos y un término constante. Se deja como ejercicio escribir explícitamente la función a partir de estos gráficos.

Problemas 4.

1) a) Verificar los resultados de los ejemplos 1, 2, 3 y 4. Indicar a qué valor convergen las series de Fourier en $x = 0$, L y $3L/2$ (usar $d = L/2$ en 4).

b) Mostrar que la igualdad de Parseval (27) en los desarrollos (37) y (38) implica las identidades $\sum_{n=1}^{\infty} 1/n^4 = \pi^4/90$ y $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$.

2) Hallar el desarrollo en serie de Fourier de $f(x) = \left\{ \begin{array}{rl} 1 & 0 \leq x \leq L \\ -1 & -L < x < 0 \end{array} \right.$ para $L = 1$. Indicar a qué valor converge el desarrollo en $x = 0$, 1, 2 y $3/2$. Graficar f y las primeras sumas parciales ($n \leq 5$) en el intervalo $[-2, 2]$.

3) Hallar el desarrollo en serie de Fourier de $f(x) = |x|$ en el intervalo $[-1, 1]$. Indicar a qué valor converge la serie en $x = 0$, 1, 2 y $x = 3/2$. Graficar f y las primeras sumas parciales.

4) Hallar el desarrollo en serie de Fourier de $\mathrm{sen}^2(x)$ (tomar $L = \pi$ y recordar que $\mathrm{sen}^2(x) = (1 - \cos(2x))/2$).

5) Mostrar que los coeficientes de Fourier a_n, b_n son funciones lineales de la función f. ¿Cual es el desarrollo en serie de Fourier de $3x/L - 2x^2/L^2$? (usar los ejemplos 1 y 2).

6) Probar que la suma finita (36) que minimiza $||f - S_N||^2$ corresponde a $c_n = a_n$ y $d_n = b_n$, con a_n, b_n los coeficientes (24)–(25). ¿Cual es el valor mínimo de $||f - S_N||^2$?

5. Desarrollos de medio rango

Si $f(x)$ está definida en el intervalo $[0, L]$, es posible desarrollarla en serie de Fourier completándola en forma par ($f(-x) = f(x)$) o impar ($f(-x) = -f(x)$) para $x \in [-L, 0)$.

Si se la completa en forma **par**, se obtiene el desarrollo de medio rango en **cosenos**:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x/L) \tag{41}$$

con (ver (30)–(31))

$$a_n = \frac{2}{L}\int_0^L f(x)\cos(n\pi x/L)dx \tag{42}$$

Esta opción se utiliza normalmente cuando f tiene derivada nula en $x = 0$ y $x = L$.

Si se la completa en forma **impar**, se obtiene el desarrollo de medio rango en **senos**:

$$f(x) = \sum_{n=1}^{\infty} b_n \operatorname{sen}(n\pi x/L) \tag{43}$$

con (ver (32)–(33))

$$b_n = \frac{2}{L}\int_0^L f(x)\operatorname{sen}(n\pi x/L)dx \tag{44}$$

Esta opción se utiliza normalmente cuando $f(x)$ se anula en $x = 0$ y $x = L$. Notar que si $f(0) \neq 0$, la función extendida tendrá una discontinuidad en $x = 0$. Y si $f(L) \neq 0$, su extensión periódica tendrá una discontinuidad en $x = \pm L, \pm 3L, \ldots$.

Estos desarrollos serán utilizados al resolver ecuaciones diferenciales en derivadas parciales.

Problemas 5.

1) Hallar el desarrollo en serie de Fourier de medio rango en i) cosenos y ii) senos, en el intervalo $[0, L]$, de las siguientes funciones:
a) $f(x) = 1$
b) $f(x) = x/L$
c) $f(x) = x^2/L^2$.
d) Indicar a qué valor convergen los desarrollos anteriores en $x = 0$ y $x = L$, junto con los límites laterales en dichos puntos.

6. Forma compleja del desarrollo

Recordando que

$$e^{\pm in\pi x/L} = \cos(n\pi x/L) \pm i \operatorname{sen}(n\pi x/L) \tag{45}$$

es posible reescribir la serie de Fourier (28) en la forma compacta

$$S(x) = \sum_{n=-\infty}^{\infty} C_n e^{in\pi x/L} \tag{46}$$

donde $\sum\limits_{n=-\infty}^{\infty} = \lim\limits_{N\to\infty} \sum\limits_{n=-N}^{N}$ y

$$C_n = \frac{1}{2L}\int_{-L}^{L} e^{-in\pi x/L} f(x)dx\,, \quad n = 0, \pm 1, \pm 2, \dots \tag{47}$$

o sea,

$$C_0 = a_0/2 = \langle f \rangle, \qquad C_{\pm n} = (a_n \mp i b_n)/2 \quad (n \geq 1)$$

El resultado (46)–(47) vale además para cualquier función periódica compleja $f : \mathbb{R} \to \mathbb{C}$ que satisfaga las condiciones de convergencia. Si la función f es real, $C_{-n} = C_n^*$ (conjugado).

El desarrollo (46) puede obtenerse directamente notando que las funciones (45) forman también un **conjunto ortogonal** con el producto interno definido para funciones complejas $f : \mathbb{R} \to \mathbb{C}$ de período $2L$ como

$$(f, g) = \frac{1}{L}\int_{-L}^{L} f^*(x)g(x)dx$$

Este producto satisface $(f, g) = (g, f)^*$, con $(f, f) = \frac{1}{L}\int_{-L}^{L} |f(x)|^2 dx \geq 0$ para cualquier f compleja, donde $|f(x)|^2 = f^*(x)f(x)$ es el cuadrado del módulo de $f(x)$. Con este producto,

$$(e^{in\pi x/L}, e^{im\pi x/L}) = \frac{1}{L}\int_{-L}^{L} e^{-in\pi x/L} e^{im\pi x/L} dx = \begin{cases} 0 & n \neq m \\ 2 & n = m \end{cases} \tag{48}$$

por lo que $\frac{1}{2L}\int_{-L}^{L} e^{-in\pi x/L} S(x)dx = C_n$.
En la forma compleja, las igualdades (27) y (35) toman la forma compacta

$$\frac{1}{2L}\int_{-L}^{L} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |C_n|^2 \tag{49}$$

$$\frac{1}{2L}\int_{-L}^{L} f^*(x)g(x)dx = \sum_{n=-\infty}^{\infty} C_n^* C'_n \tag{50}$$

donde C'_n son los coeficientes de Fourier complejos de la función g.

7. Transformada discreta de Fourier (opcional)

Dada una función $f : [0,T] \to \mathbb{C}$, pueden considerarse sus valores $f_j = f(t_j)$ en n tiempos equiespaciados $0 \le t_1 < t_2 < \ldots < t_n \le T$. Resulta entonces válido el desarrollo

$$f_j = \sum_{k=1}^{n} e^{i2\pi jk/n} F_k \,, \quad j = 1, \ldots, n \tag{51}$$

donde los n coeficientes F_k están dados por

$$F_k = \frac{1}{n} \sum_{j=1}^{n} e^{-i2\pi kj/n} f_j \,, \quad k = 1, \ldots, n \tag{52}$$

Estos coeficientes constituyen la *transformada de Fourier discreta* de los f_j. La demostración de (51)–(52) se basa en la igualdad

$$\frac{1}{n} \sum_{k=1}^{n} e^{i2\pi k(j-j')/n} = \begin{cases} 0 & j \neq j' \\ 1 & j = j' \end{cases} \tag{53}$$

válida para $j, j' \in \{1, \ldots, n\}$, cuya demostración se deja como ejercicio.

Notemos que para k y m enteros arbitrarios, (52) implica $F_{k+mn} = F_k$, ya que $e^{-i2\pi(k+mn)j/n} = e^{-i2\pi kj/n}$. En particular, $F_0 = F_n = \frac{1}{n}\sum_{j=1}^{n} f_j$ es nuevamente el valor medio de los f_j. Resulta también válida la igualdad

$$\frac{1}{n} \sum_{j=1}^{n} |f_j|^2 = \sum_{k=1}^{n} |F_k|^2 \tag{54}$$

Como ejemplo, si $f_j = c$ (constante) $\forall\ j = 1, \ldots, n \Rightarrow F_k = 0$ para $k = 1, \ldots, n-1$ mientras que $F_n = c$. Y si $f_j = 0$ para $j = 1, \ldots, n-1$ y $f_n = c \Rightarrow F_k = c/n\ \forall\ k$.

De esta forma, los coeficientes (52) posibilitan un análisis de Fourier de la función $f(t)$ restringida a estos n tiempos. Y para n suficientemente grande y f de variación acotada, el esquema discreto puede proporcionar una buena aproximación al desarrollo en serie de Fourier de f. Su ventaja es que requiere solo el conocimiento de f en un número finito de puntos (y no su forma funcional exacta). Además el cálculo de los n coeficientes F_k puede realizarse en forma eficiente (requiere $O(n)$ pasos).

II. Ecuaciones Diferenciales en Derivadas Parciales

1. Introducción

Las ecuaciones diferenciales en derivadas parciales juegan un rol fundamental en Ingeniería y Física. Permiten determinar funciones de varias variables a partir de ciertas condiciones iniciales y/o de contorno. Como ejemplos típicos podemos mencionar:

a) **La ecuación de difusión,** que en una dimensión es de la forma

$$\frac{\partial U(x,t)}{\partial t} - \alpha \frac{\partial^2 U(x,t)}{\partial x^2} = 0 \tag{1}$$

Esta ecuación, de segundo orden en x y primer orden en t, permite determinar, por ejemplo, la temperatura $U(x,t)$ de una barra conductora de calor de longitud L en función de la posición x y el tiempo t, si se conocen la distribución inicial de temperatura $U(x,0)$ y las temperaturas en los bordes $U(0,t)$ y $U(L,t)$ (se asume que la barra está térmicamente aislada salvo en los extremos). En este caso α es una constante positiva que depende de la conductividad térmica, el calor especiífico y la densidad del material.

b) **La ecuación de ondas**, que en una dimensión es de la forma

$$\frac{\partial^2 U(x,t)}{\partial t^2} - v^2 \frac{\partial^2 U(x,t)}{\partial x^2} = 0 \tag{2}$$

Esta ecuación, de segundo orden en ambas variables x, t, permite determinar, por ejemplo, la elongación $U(x,t)$ de una cuerda tensa de longitud L sujeta en ambos extremos, en función de la posición x y el tiempo t, a partir de su elongación inicial $U(x,0)$ y su velocidad inicial $\frac{\partial U(x,t)}{\partial t}|_{t=0}$. Aquí v representa la velocidad de propagación de la onda.

c) **La ecuación de Laplace**, que en dos dimensiones es de la forma

$$\frac{\partial^2 U(x,y)}{\partial x^2} + \frac{\partial^2 U(x,y)}{\partial y^2} = 0 \tag{3}$$

El objetivo es determinar una función $U(x,y)$ que satisfaga esta ecuación en una cierta región R del plano $\mathbb{R}^2$, conociendo sus valores en el borde de R. Esta ecuación de segundo orden en x e y, permite determinar. por ejemplo, la temperatura *estacionaria* $U(x,y)$ (es decir, de equilibrio y por lo tanto independiente del tiempo t) de una placa conductora de calor en una región R sin fuentes de calor, si se conocen sus valores en el borde de R.

La ecuación (3) tiene también varias otras aplicaciones: descripción de membranas elásticas en equilibrio, de flujos estacionarios irrotacionales en dos dimensiones, etc. Su versión tridimensional,

$$\frac{\partial^2 U(x,y,z)}{\partial x^2} + \frac{\partial^2 U(x,y,z)}{\partial y^2} + \frac{\partial^2 U(x,y,z)}{\partial z^2} = 0 \tag{4}$$

es la ecuación que satisface el potencial electrostático $U(x,y,z)$ en una región R del espacio $\mathbb{R}^3$ sin cargas eléctricas. Si el potencial no depende de z, (4) se reduce a la ecuación bidimensional (3).

Las tres ecuaciones anteriores son ejemplos de **ecuaciones diferenciales lineales homogéneas de segundo orden en derivadas parciales**. En este curso nos concentraremos en este tipo de ecuaciones.

2. Definición general y propiedades fundamentales

Una ecuación diferencial lineal de segundo orden en derivadas parciales para una función de n variables $U(x_1,\ldots,x_n)$ ($U:\mathbb{R}^n\to\mathbb{R}$) es de la forma

$$\sum_{i,j}\alpha_{ij}\frac{\partial^2 U}{\partial x_i\partial x_j}+\sum_i\beta_i\frac{\partial U}{\partial x_i}+\gamma U=f(x_1,\ldots,x_n) \tag{5}$$

donde los coeficientes α_{ij}, β_i y γ pueden ser constantes o también funciones continuas de las variables $x_1,\ldots,x_n$. El caso **homogéneo** corresponde a $f(x_1,\ldots,x_n)=0$.

Para $n=1$, la ecuación (5) se reduce a la ya conocida ecuación diferencial lineal *ordinaria* de segundo orden $\alpha\frac{d^2U}{dx^2}+\beta\frac{dU}{dx}+\gamma U=f(x)$, donde la función incógnita $U(x)$ depende de una sola variable $x=x_1$. Por ser lineal, el caso general $n\geq 2$ posee ciertas propiedades similares a las del caso $n=1$.

Propiedades del caso homogéneo.

$$\sum_{i,j}\alpha_{ij}\frac{\partial^2 U}{\partial x_i\partial x_j}+\sum_i\beta_i\frac{\partial U}{\partial x_i}+\gamma U=0 \tag{6}$$

1) La función nula $U(x_1,\ldots,x_n)=0\ \forall\ x_1,\ldots,x_n$ es siempre una solución de (6) (solución trivial).

2) Es válida la **superposición** de soluciones:
Si $U_1(x_1,\ldots,x_n)$ y $U_2(x_1,\ldots,x_n)$ son soluciones de (6), entonces toda combinación lineal

$$U(x_1,\ldots,x_n)=c_1\,U_1(x_1,\ldots,x_n)+c_2\,U_2(x_1,\ldots,x_n) \tag{7}$$

con $c_1,c_2\in\mathbb{R}$, es también solución de (6). La demostración es obvia a partir de la propiedad de linealidad de las derivadas parciales.

Estas dos propiedades implican que el conjunto de soluciones de (6) es un *subespacio* S del espacio vectorial V de funciones $F:\mathbb{R}^n\to\mathbb{R}$. No obstante, a diferencia del caso $n=1$, para $n\geq 2$ la dimensión de S es en general *infinita*, existiendo infinitas soluciones linealmente independientes si no se imponen condiciones iniciales o de contorno.

Problemas 2.

1. Verificar que para cualquier valor (constante) de A, λ y ϕ, ϕ_1, ϕ_2,
a) $U(x,t)=Ae^{-\alpha\lambda^2 t}\operatorname{sen}(\lambda x+\phi)$ es solución de la ecuación de difusión (1).
b) $U(x,t)=A\cos(\lambda vt+\phi_1)\operatorname{sen}(\lambda x+\phi_2)$ es solución de la ecuación de ondas (2).
c) $U(x,y)=Ae^{-\lambda x}\operatorname{sen}(\lambda x+\phi)$ es solución de la ecuación de Laplace (3).

2. a) Probar que $U(x,t)=f(x-vt)+g(x+vt)$, con f y g funciones reales arbitrarias con derivada segunda definida, es solución de (2). Interpretar $f(x-vt)$ y $g(x+vt)$ como ondas que se propagan hacia la derecha e izquierda sin deformación con velocidad v.
b) Escribir $U(x,t)=A\cos(\lambda vt+\phi_1)\operatorname{sen}(\lambda x+\phi_2)$ como una suma $f(x-vt)+g(x-vt)$.

3. Mostrar que la solución general del caso no homogéneo (5) es de la forma $U(x_1,\ldots,x_n)=U_h(x_1,\ldots,x_n)+U_p(x_1,\ldots,x_n)$, donde $U_h(x_1,\ldots,x_n)$ es la solución general del caso homogéneo (6) y $U_p(x_1,\ldots,x_n)$ una solución particular de (5).

3. Método de separación de variables y solución por serie de Fourier

Emplearemos aquí el método de separación de variables y las series de Fourier para resolver las ecuaciones diferenciales homogéneas (1), (2) y (3) con condiciones de iniciales y/o de contorno en regiones rectangulares.

3.1. Ecuación de difusión en una dimensión

Comenzaremos por resolver la ecuación de difusión (1) en un intervalo finito $0 \leq x \leq L$. El objetivo es determinar la función $U(x,t)$ que satisface la ecuación

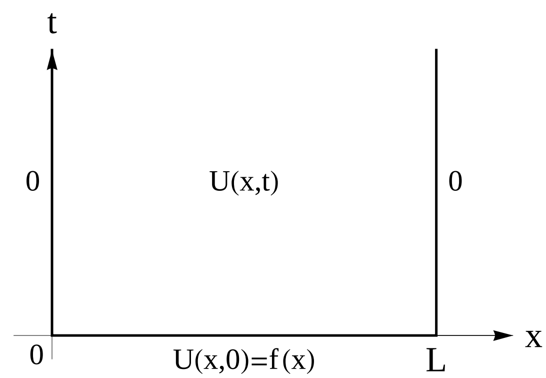
$$\frac{\partial U(x,t)}{\partial t} - \alpha \frac{\partial^2 U(x,t)}{\partial x^2} = 0, \qquad \begin{array}{c} 0 < x < L \\ t > 0 \end{array} \tag{8}$$

con $\alpha > 0$, junto con las *condiciones de contorno*

$$U(0,t) = 0, \quad U(L,t) = 0\,, \quad t > 0 \tag{9}$$

y la *condición inicial*

$$U(x,0) = f(x)\,, \quad 0 < x < L \tag{10}$$



La función $U(x,t)$ puede representar, por ejemplo, la temperatura en función de la posición y el tiempo de una barra metálica de longitud L cuyos bordes están a cero grados centigrados (por ejemplo en contacto con hielo) y está lateralmente aislada, teniendo inicialmente ($t = 0$) una distribución de temperatura dada por $f(x)$.

Si la barra está inicialmente a temperatura mayor que 0, se enfriará al aumentar t, tendiendo su temperatura a 0 para tiempos grandes. La ecuación (9) describe con todo detalle este proceso de enfriamiento. Notemos que al ser (8) de primer orden en el tiempo, se requiere conocer solamente el valor inicial $U(x,0)$ de la función $U(x,t)$.

Soluciones producto

El método de separación de variables consiste en proponer, primero, una *solución producto*

$$U(x,t) = X(x)T(t) \tag{11}$$

donde $X(x)$ depende solo de x y $T(t)$ solo de t. Reemplazando (11) en (8) se obtiene

$$X(x)T'(t) - \alpha X''(x)T(t) = 0$$

Dividiendo esta ecuación por $X(x)T(t)$, se llega a $\frac{T'(t)}{T(t)} - \alpha\frac{X''(x)}{X(x)} = 0$, o sea,

$$\frac{X''(x)}{X(x)} = \frac{1}{\alpha}\frac{T'(t)}{T(t)}$$

Como el primer cociente depende solo de x y el segundo solo de t, siendo x y t variables *independientes*, la única posibilidad es que ambos cocientes sean *constantes*:

$$\frac{X''(x)}{X(x)} = \frac{1}{\alpha}\frac{T'(t)}{T(t)} = -k^2 \tag{12}$$

La constante k^2 se denomina **constante de separación**. Se obtienen así dos ecuaciones diferenciales **ordinarias**, una de segundo orden para X y una de primer orden para T:

$$X''(x) + k^2X(x) = 0 \tag{13}$$
$$T'(t) + \alpha k^2T(t) = 0 \tag{14}$$

que ya sabemos resolver. Si $k = 0$, las soluciones son

$$X(x) = A + Bx, \quad T(t) = C \tag{15}$$

lo que conduce a una solución **estacionaria** (independiente de t) de la forma

$$X(x)T(t) = A + Bx\,. \tag{16}$$

Si $k \neq 0$, obtenemos en cambio las soluciones

$$X(x) = A\cos kx + B\,\mathrm{sen}\,kx \tag{17}$$
$$T(t) = Ce^{-\alpha k^2 t} \tag{18}$$

que conducen a soluciones no estacionarias de la forma

$$X(x)T(t) = (A\cos kx + B\,\mathrm{sen}\,kx)e^{-\alpha k^2 t}\,. \tag{19}$$

Las condiciones de contorno (9) implican ahora

$$U(0,t) = X(0)T(t) = 0, \quad U(L,t) = X(L)T(t) = 0 \tag{20}$$

$\forall\ t > 0$. Por lo tanto, si $T(t) \neq 0$, debe cumplirse $X(0) = 0$, $X(L) = 0$.

Para $k = 0$, esto implica $A = B = 0$ en (16). El valor $k = 0$ conduce en este caso solo a la solución trivial. En cambio, para $k \neq 0$, (20) implica $A = 0$ en (19) y entonces

$$B\,\mathrm{sen}(kL) = 0 \tag{21}$$

Si $B \neq 0$ esto conduce a $kL = n\pi$, con n entero:

$$k = n\pi/L\,, \quad n = 1, 2, 3, \ldots \tag{22}$$

Hemos excluido valores enteros negativos de n ya que si bien satisfacen (21), no generan nuevas soluciones linealmente independientes ($\mathrm{sen}(-kx) = -\,\mathrm{sen}(kx)$). Los valores posibles de k quedan entonces determinados por *la condición de contorno.*

Reemplazando (22) en (19) y recordando que $A = 0$, se obtienen las soluciones

$$X_n(x)T_n(t) \;=\; B_n \sin(n\pi x/L)\, e^{-\alpha(n\pi/L)^2 t}\,, \quad n = 1, 2, 3, \ldots\,, \tag{23}$$

con B_n constantes arbitrarias. Estas soluciones son *linealmente independientes* si $B_n \neq 0$. Para $t = 0$,

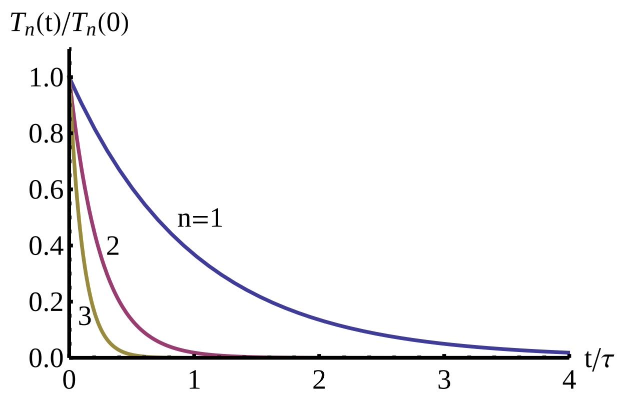
$$X_n(x)T_n(0) = B_n \sin(n\pi x/L) \tag{24}$$

por lo que estas soluciones producto corresponden a una distribución inicial $f(x)$ proporcional a $\sin(n\pi x/L)$.

Su valor absoluto $|X_n(x)T_n(t)|$ decrece **exponencialmente** al aumentar t:

$$T_n(t) = e^{-\alpha(n\pi/L)^2 t} = e^{-n^2 t/\tau}\,, \qquad \tau = L^2/(\pi^2\alpha) \tag{25}$$

donde τ representa un tiempo característico, tal que $|T_n(t)| \ll 1$ si $t \gg \tau/n^2$. A mayor n, más rápidamemte disminuirá el módulo de esta solución, como se ve en la figura.



Solución general

Es evidente que con una sola solución producto no es posible satisfacer una condición inicial general $U(x, 0) = f(x)$.

No obstante, por la propiedad de **superposición**, podemos proponer ahora una solución más general que sea una combinación lineal de las soluciones producto (23), es decir, una solución de la forma

$$U(x, t) = \sum_{n=1}^{\infty} B_n e^{-\alpha(n\pi/L)^2 t}\, \mathrm{sen}(n\pi x/L) \tag{26}$$

que también satisface (8) (asumiendo convergencia uniforme de la serie) junto con la condición de contorno (9), ya que cada término de la serie los satisface. Los coeficientes B_n pueden ser ahora determinados por la condición inicial (10):

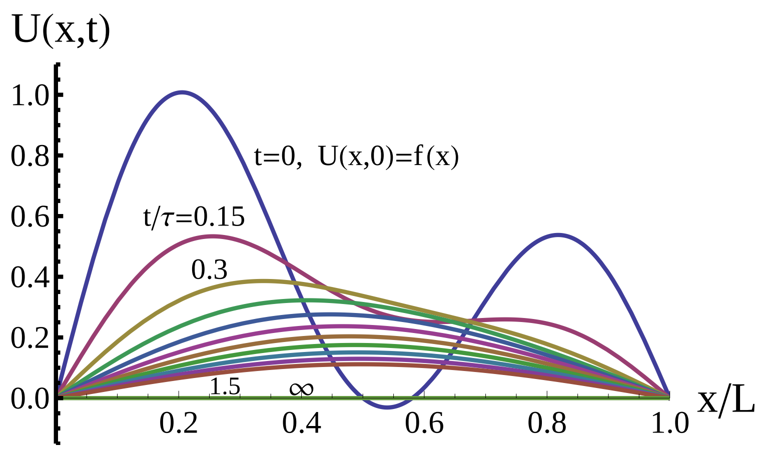
$$U(x, 0) = \sum_{n=1}^{\infty} B_n\, \mathrm{sen}(n\pi x/L) = f(x)\,, \quad 0 < x < L \tag{27}$$

que constituye *el desarrollo en serie de Fourier de medio rango en senos de $f(x)$ en el intervalo* $[0, L]$ (ecs. (43)-(44) del cap. anterior). Los coeficientes B_n están pues dados por

$$B_n \;=\; \frac{2}{L}\int_0^L f(x)\,\mathrm{sen}(n\pi x/L)dx\,, \quad n = 1, 2, 3, \ldots \tag{28}$$

La solución final **y única** del problema (8)–(9)–(10) es pues la serie (26) con los coeficientes B_n determinados por (28). Si el desarrollo (27) converge, también lo hará la serie (26) para $t > 0$ debido al factor exponencial.

Podemos entender la solución general (26)–(28) como la "suma" de las soluciones producto (23) para cada término $B_n \sin(n\pi x/L)$ del desarrollo de Fourier de medio rango de la condición inicial $f(x)$. Para $t > \tau$, el término más importante en la serie (27) será obviamente el primero ($n = 1$). Esto implica que las oscilaciones locales de temperatura desaparecen rápidamente (ver figura).



Ejemplo 1. Si $L = 1$, $\alpha = 1$ y $U(x, 0) = f(x) = 2\,\mathrm{sen}(\pi x) + \mathrm{sen}(3\pi x)$,

$$B_n = 2\int_0^1 f(x)\,\mathrm{sen}(n\pi x)dx = \begin{cases} 2 & n = 1 \\ 1 & n = 3 \\ 0 & n \neq 1,\ n \neq 3 \end{cases} \tag{29}$$

Por lo tanto,

$$U(x, t) = 2e^{-\pi^2 t}\,\mathrm{sen}(\pi x) + e^{-9\pi^2 t}\,\mathrm{sen}(3\pi x)$$

Para $t > \tau/9$, con $\tau = 1/\pi^2$, el primer término de la suma es dominante.

Ejemplo 2.. Si ahora $f(x) = cx/L$, con $\alpha > 0$ y $L > 0$ arbitrarios,

$$B_n = \frac{2}{L}\int_0^L c(x/L)\,\mathrm{sen}(n\pi x/L)dx = \frac{2c}{n\pi}(-1)^{n+1}\,, \quad n = 1, 2, 3, \ldots,$$

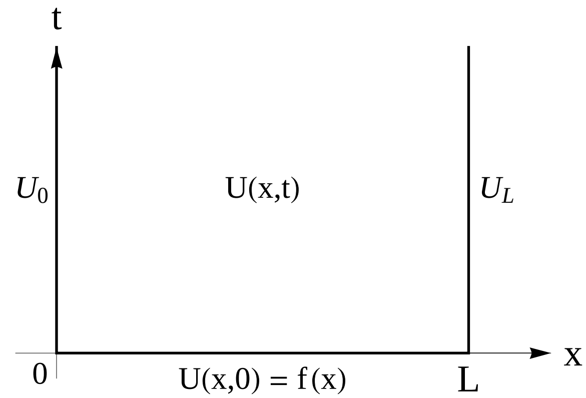
y por lo tanto

$$U(x, t) = \frac{2c}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(n\pi x/L)e^{-\alpha n^2\pi^2 t/L^2}$$

Para $t > 0$, la convergencia de la serie es absoluta. Para $t > \tau/4$, con $\tau = L^2/(\pi^2\alpha)$, esta serie es dominada por el primer término. Graficar para $c = 1$, $\alpha = 1$, $L = 1$.

Ejemplo 3. Consideremos ahora que los bordes de la barra están a temperaturas fijas generales U_0 y U_L. Esto corresponde a las condiciones de contorno

$$U(0,t) = U_0, \quad U(L,t) = U_L\,, \qquad t > 0 \tag{30}$$



Podemos utilizar la solución producto (16) para $k = 0$ (solución estacionaria),

$$X(x)T(t) = A + Bx$$

para ajustar la condición de contorno (30). Tenemos $X(0)T(t) = A = U_0$, $X(L)T(t) = A + BL = U_L$, por lo que $B = (U_L - U_0)/L$. La solución general será entonces la suma de esta solución estacionaria y la solución previa (26):

$$U(x,t) = U_0 + \frac{U_L - U_0}{L}x + \sum_{n=1}^{\infty} B_n e^{-\alpha n^2\pi^2 t/L}\,\mathrm{sen}(n\pi x/L) \tag{31}$$

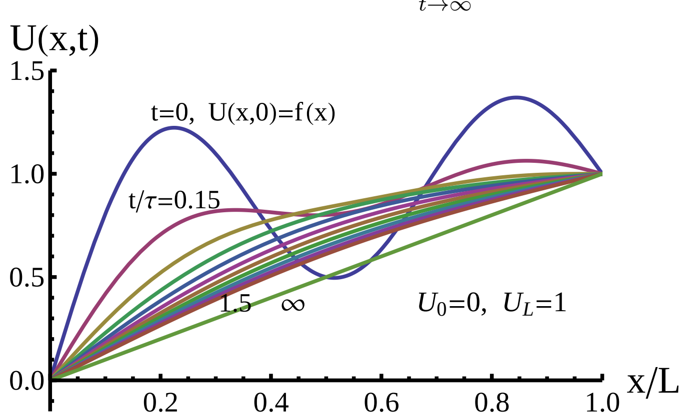
la cual satisface la ecuación (8) junto con la nueva condición de contorno (30). Ahora

$$U(x,0) = U_0 + \frac{U_L - U_0}{L}x + \sum_{n=1}^{\infty} B_n\,\mathrm{sen}(n\pi x/L) = f(x)$$

por lo que los coeficientes B_n se obtienen como

$$B_n = \frac{2}{L}\int_0^L \left(f(x) - U_0 - \frac{U_L - U_0}{L}x\right)\sin(n\pi x/L)dx \tag{32}$$

Quedan determinados por la **diferencia** entre la temperatura inicial $f(x)$ y la solución estacionaria. La principal distinción con el caso previo es que para $t \to \infty$, $U(x,t)$ no tiende a 0 sino a la solución estacionaria no nula: $\lim\limits_{t\to\infty} U(x,t) = U_0 + \frac{U_L - U_0}{L}x$



Problemas 3.1.

1. Determinar la solución $U(x,t)$ de (8) para las condiciones de contorno $U(0,t) = U(L,t) = 0$ y la condición inicial $U(x,0) = f(x)$, si:
a) $f(x) = 2\,\mathrm{sen}(\pi x/L)$
b) $f(x) = 3\,\mathrm{sen}(\pi x/L) - 2\,\mathrm{sen}(2\pi x/L) + \mathrm{sen}(3\pi x/L)$
c) $f(x) = x(L-x)/L^2$.
d) $f(x) = x/L$.
Graficar (mediante PC o cel.) $U(x,t)$ para distintos tiempos en el caso $\alpha = 1$, $L = 1$.

2. Utilizando (31)–(32), determinar la solución $U(x,t)$ para las condiciones de contorno $U(0,t) = 0$, $U(L,t) = 1$ y la condición inicial $U(x,0) = f(x)$, si:
a) $f(x) = x/L$
b) $f(x) = x/L + 3\,\mathrm{sen}(\pi x/L) - 2\,\mathrm{sen}(2\pi x/L) + \mathrm{sen}(3\pi x/L)$
c) $f(x) = x/L + x(L-x)/L^2$.
Graficar $U(x,t)$ para distintos tiempos en el caso $\alpha = 1$, $L = 1$.

3. Determinar la solución $U(x,t)$ de (8) para las condiciones de contorno

$$\left.\frac{\partial U}{\partial x}\right|_{x=0} = \left.\frac{\partial U}{\partial x}\right|_{x=L} = 0\,, \quad t > 0$$

que corresponden a una barra térmicamente aislada, y la condición inicial $U(x,0) = f(x)$. Motrar que en este caso $X(x) = A\cos(kx)$, con $k = n\pi/L$ y $n = 0, 1, 2, \ldots$, y por lo tanto

$$U(x,t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(n\pi x/L) e^{-\alpha n^2\pi^2 t/L^2}$$

Dar una expresión para A_0 y A_n en términos de $f(x)$ y determinar el límite $\lim_{t\to\infty} U(x,t)$.

4. (Opcional) Caso no homogéneo. En presencia de fuentes de calor $\propto f(x,t)$ a lo largo de la barra, la temperatura $U(x,t)$ satisface la ecuación de difusión no homogénea

$$\frac{\partial U}{\partial t} - \alpha\frac{\partial^2 U}{\partial x^2} = f(x,t)$$

Asumiendo $U(0,t) = U(L,t) = 0 \ \forall\ t > 0$, puede proponerse una solución particular

$$U_p(x,t) = \sum_{n=1}^{\infty} B_n(t)\,\mathrm{sen}(n\pi x/L)$$

Mostrar que los coeficientes de Fourier $B_n(t)$ satisfacen ahora la ecuación no homogénea

$$B_n'(t) + \alpha(n\pi/L)^2 B_n(t) = F_n(t)$$

con $F_n(t) = \frac{2}{L}\int_0^L f(x,t)\sin(\frac{n\pi}{L}x)dx$. Determinar $B_n(t)$ y la solución $U(x,t)$ si $U(x,0) = 0$.

3.2. Ecuación de ondas en una dimensión

Veamos ahora la ecuación de ondas (2) en un intervalo finito $0 \leq x \leq L$. El objetivo es determinar la función $U(x,t)$ que satisface la ecuación

$$\frac{\partial^2 U(x,t)}{\partial t^2} - v^2 \frac{\partial^2 U(x,t)}{\partial x^2} = 0, \qquad \begin{array}{l} 0 < x < L \\ t > 0 \end{array} \tag{33}$$

junto con las *condiciones de contorno*

$$U(0,t) = 0, \quad U(L,t) = 0\,, \quad t > 0 \tag{34}$$

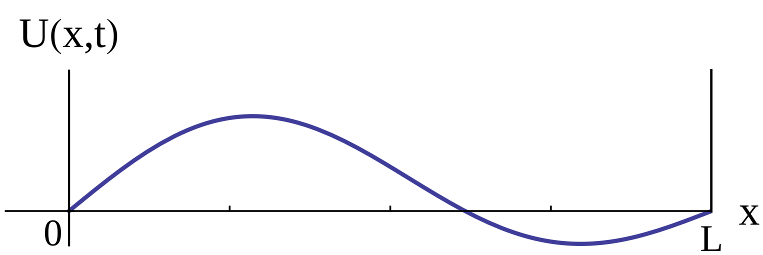
y las *condiciones iniciales*

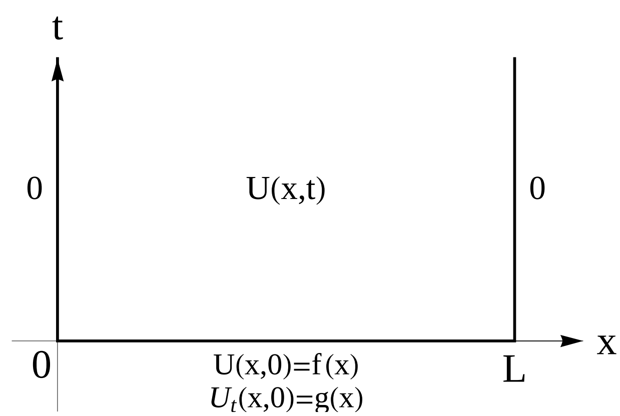
$$U(x,0) = f(x)\,, \quad 0 < x < L \tag{35}$$

$$U_t(x,0) = \left.\frac{\partial U(x,t)}{\partial t}\right|_{t=0} = g(x)\,, \quad 0 < x < L \tag{36}$$

Dado que la ecuación es de segundo orden en el tiempo, es necesario ahora especificar tanto su valor inicial $U(x,0) = f(x)$ como su derivada temporal inicial (velocidad inicial) $U_t(x,0) = g(x)$.

La función $U(x,t)$ puede representar, por ejemplo, la elongación de una cuerda tensa de longitud L fija en los bordes, siendo $U(x,0) = f(x)$ la elongación inicial de la cuerda, $U_t(x,0) = g(x)$ su velocidad inicial y v la velocidad de propagación de ondas transversales en la cuerda ($v = \sqrt{T/\mu}$, con T la tensión y μ la densidad lineal de masa). La ecuación (33) determina las oscilaciones de la cuerda a partir de estas condiciones iniciales.





Soluciones producto

Aplicando nuevamente el método de separación de variables, se propone primero una solución *producto*

$$U(x,t) = X(x)T(t) \tag{37}$$

Reemplazando (37) en (33) se obtiene $X(x)T''(t) - v^2 X''(x)T(t) = 0$, o sea,

$$\frac{X''(x)}{X(x)} = \frac{1}{v^2}\frac{T''(t)}{T(t)}$$

Y como el primer cociente depende solo de x y el segundo solo de t, debe cumplirse

$$\frac{X''(x)}{X(x)} = \frac{1}{v^2}\frac{T''(t)}{T(t)} = -k^2 \tag{38}$$

Esto conduce a las ecuaciones diferenciales ordinarias

$$X''(x) + k^2 X(x) = 0 \tag{39}$$
$$T''(t) + v^2 k^2 T(t) = 0 \tag{40}$$

Si $k = 0$, las soluciones generales son de la forma

$$X(x) = A + Bx, \quad T(t) = C + Dt \tag{41}$$

mientras que si $k \neq 0$,

$$X(x) = A\cos(kx) + B\,\mathrm{sen}(kx) \tag{42}$$
$$T(t) = C\cos(kvt) + D\,\mathrm{sen}(kvt) \tag{43}$$

Las condiciones de contorno (34) implican nuevamente $X(0) = 0,$, $X(L) = 0$. Al igual que en la ecuación de diofusión, para $k = 0$ esto implica $A = B = 0$, mientras que para $k \neq 0$, esto conduce a $A = 0$ y a la misma ecuación previa (21), $B\,\mathrm{sen}(kL) = 0$. Como se buscan soluciones no nulas e independientes, esto implica nuevamente

$$k = n\pi/L\,, \quad n = 1, 2, 3, \ldots$$

Se obtienen así las soluciones producto

$$X_n(x)T_n(t) = \mathrm{sen}(n\pi x/L)\left[C_n \cos(n\pi vt/L) + D_n\,\mathrm{sen}(n\pi vt/L)\right], \quad n = 1, 2, 3, \ldots \tag{44}$$

donde C_n y D_n son constantes arbitrarias.

A diferencia de las soluciones de la ecuación de difusión, estas soluciones producto tienen ahora un comportamiento *oscilatorio armónico* con t y representan los *modos normales de vibración de la cuerda*. Corresponden a **ondas estacionarias**. A mayor n, mayor es la frecuencia f_n de vibración y por lo tanto menor el período τ_n:

$$f_n = \frac{n\,v}{2L}, \quad \tau_n = \frac{1}{f_n} = \frac{2L}{n\,v}\,, \quad n = 1, 2, 3\ldots$$

Las soluciones (44) pueden escribirse como la superposición de ondas viajeras iguales propagándose en sentidos opuestos:

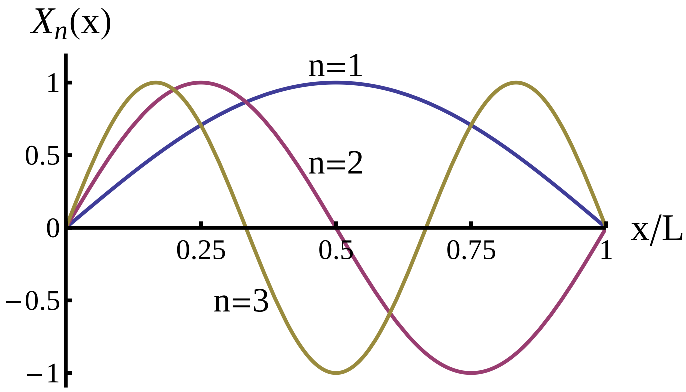
$$\mathrm{sen}(\tfrac{n\pi}{L}x)\cos(\tfrac{n\pi}{L}vt) = \tfrac{1}{2}\left[\mathrm{sen}[\tfrac{n\pi}{L}(x - vt)] + \mathrm{sen}[\tfrac{n\pi}{L}(x + vt)]\right] \tag{45}$$
$$\mathrm{sen}(\tfrac{n\pi}{L}x)\,\mathrm{sen}(\tfrac{n\pi}{L}vt) = \tfrac{1}{2}\left[\cos[\tfrac{n\pi}{L}(x - vt)] - \cos[\tfrac{n\pi}{L}(x + vt)]\right] \tag{46}$$

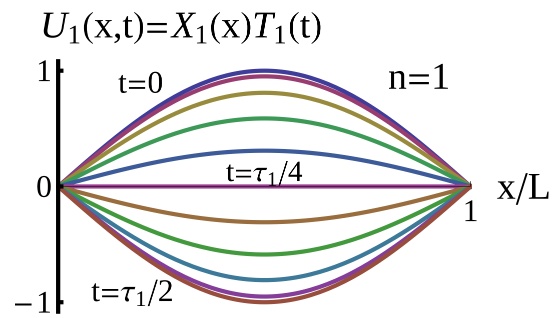
Obviamente, estas soluciones producto responden a las condiciones iniciales

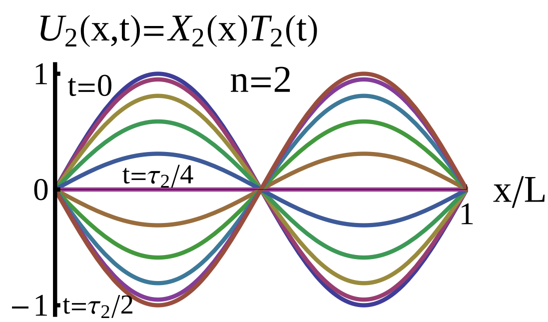
$$U(x, 0) = C_n\,\mathrm{sen}(\pi x/L) \tag{47}$$
$$U_t(x, 0) = D_n \tfrac{n\pi v}{L}\,\mathrm{sen}(n\pi x/L) \tag{48}$$

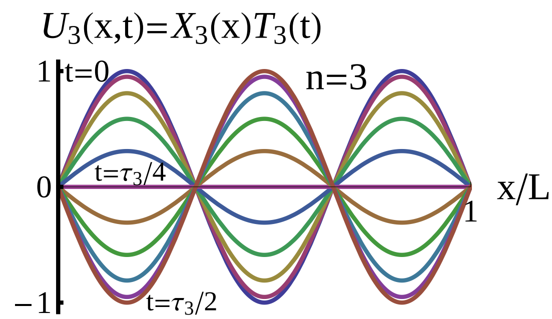
es decir, tanto la elongación inicial $f(x)$ como la velocidad inicial $g(x)$ son proporcionales a $\sin(n\pi x/L)$.

Se muestran en las figuras los primeros modos normales y sus oscilaciones.









Solución general

Para resolver el problema original con condiciones iniciales generales, debemos plantear nuevamente la superposición de soluciones producto:

$$U(x,t)=\sum_{n=1}^{\infty}\mathrm{sen}(n\pi x/L)[C_n\cos(n\pi vt/L)+D_n\,\mathrm{sen}(n\pi vt/L)] \tag{49}$$

En este caso,

$$\begin{aligned} U(x,0) &= \sum_{n=1}^{\infty}C_n\,\mathrm{sen}(n\pi x/L)=f(x) && (50)\\ U_t(x,0) &= \sum_{n=1}^{\infty}D_n\frac{n\pi v}{L}\,\mathrm{sen}(n\pi x/L)=g(x) && (51)\end{aligned}$$

Estas expresiones representan el desarrollo en serie de Fourier de medio rango de $f(x)$ y $g(x)$ respectivamente. Los coeficientes C_n y D_n están entonces determinados por

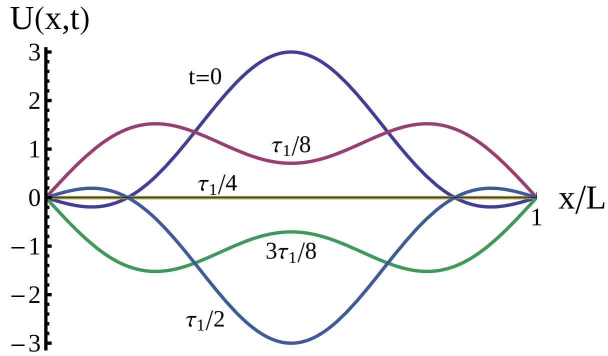
$$\begin{aligned} C_n &= \frac{2}{L}\int_0^L f(x)\,\mathrm{sen}(n\pi x/L)dx\,,\quad n=1,2,3,\ldots && (52)\\ D_n &= \frac{2}{n\pi v}\int_0^L g(x)\,\mathrm{sen}(n\pi x/L)dx\,,\quad n=1,2,3,\ldots && (53)\end{aligned}$$

La solución final **y única** del problema (33)–(36) es pues la serie (26) con los coeficientes C_n, D_n determinados por (52)–(53). Si las series (50)–(51) convergen absolutamente, también lo hará la serie (49), como es fácil probar.

Ejemplo 1. Si $U(x,0)=f(x)=2\,\text{sen}(\pi x/L)-\text{sen}(3\pi x/L)$ y $U_t(x,0)=g(x)=0$,

$$C_n=\frac{2}{L}\int_0^L f(x)\,\text{sen}(n\pi x/L)dx=\begin{cases}2 & n=1\\ -1 & n=3\\ 0 & n\neq 1,\ n\neq 3\end{cases}\quad,\quad D_n=0\ \forall\, n \tag{54}$$

Por lo tanto la solución es $U(x,t)=2\,\text{sen}(\pi x/L)\cos(\pi vt/L)-\text{sen}(3\pi x/L)\cos(3\pi vt/L)$.



Ejemplo 2. Si ahora $U(x,0)=0$ y $U_t(x,0)=Ax(L-x)/L^2$, se obtiene $C_n=0\ \forall\, n$ y

$$D_n=\frac{2}{n\pi vL^2}\int_0^L Ax(L-x)\,\text{sen}(n\pi x/L)dx=\begin{cases}0 & n \text{ par}\\ \frac{8AL}{n^4\pi^4 v} & n \text{ impar}\end{cases}$$

Por lo tanto, escribiendo $n=2k+1$ con $k=0,1,\ldots$,

$$U(x,t)=\frac{8AL}{\pi^4\,v}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^4}\text{sen}[(2k+1)\pi x/L]\,\text{sen}[(2k+1)\pi vt/L]$$

Basta aquí sumar unos pocos términos para obtener una buena estimación de $U(x,t)$.

Problemas 3.2.

1. Determinar la solución $U(x,t)$ de la ecuación de ondas para las condiciones de contorno $U(0,t)=U(L,t)=0$ y las condiciones iniciales $U(x,0)=f(x)$, $U_t(x,0)=0$, si:
a) $f(x)=4\,\text{sen}(3\pi x/L)$ b) $f(x)=2\,\text{sen}(\pi x/L)-\text{sen}(2\pi x/L)+\text{sen}(5\pi x/L)$
c) $f(x)=x(L-x)/L^2+\sin(\pi x/L)$.
Graficar $U(x,t)$ para distintos tiempos (o realizar una animación con PC) y $v=1$, $L=1$.

2. Determinar la solución $U(x,t)$ de la ecuación de ondas para las condiciones de contorno $U(0,t)=U(L,t)=0$ y las condiciones iniciales $U(x,0)=0$, $U_t(x,0)=g(x)$, si:
a) $g(x)=4\,\text{sen}(3\pi x/L)$ b) $g(x)=2\,\text{sen}(\pi x/L)-\text{sen}(2\pi x/L)+\text{sen}(5\pi x/L)$
c) $g(x)=x(L-x)/L^2+\sin(\pi x/L)$.
Graficar $U(x,t)$ para distintos tiempos t y $v=1$, $L=1$. Indicar la relación que existe entre estas soluciones y aquellas obtenidas en el problema **1.**

3. Determinar la solución $U(x,t)$ de la ecuación de ondas para la condición de contorno $\frac{\partial U}{\partial x}\big|_{x=0}=\frac{\partial U}{\partial x}\big|_{x=L}=0\ \forall\, t>0$, que corresponde a una cuerda tensa con extremos libres en la dirección vertical, con las condiciones iniciales $U(x,0)=f(x)$, $U_t(x,0)=g(x)$. Identificar los modos normales, sus frecuencias y períodos, y mostrar que en este caso

$$U(x,t)=\frac{1}{2}(A_0+B_0t)+\sum_{n=1}^{\infty}\cos(n\pi x/L)[A_n\cos(n\pi vt/L)+B_n\,\text{sen}(n\pi vt/L)]$$

3.3. Ecuación de Laplace en dos dimensiones

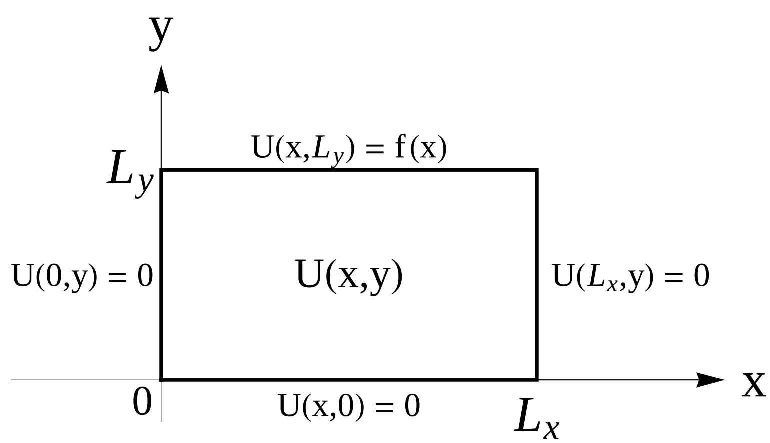
Comenzaremos por resolver la ecuación de Laplace (3) en una región rectangular con condiciones de contorno sencillas. El objetivo es determinar la función $U(x,y)$ que satisface la ecuación de Laplace en el interior de un rectángulo R,

$$\frac{\partial^2 U(x,y)}{\partial x^2} + \frac{\partial^2 U(x,y)}{\partial y^2} = 0, \quad 0 < x < L_x, \;\; 0 < y < L_y \tag{55}$$

junto con las *condiciones de contorno* (ver figura)

$$U(0,y) \;=\; 0, \quad U(L_x,y) = 0\,, \quad 0 < y < L_y \tag{56}$$

$$U(x,0) \;=\; 0, \quad U(x,L_y) = f(x)\,, \quad 0 < x < L_x \tag{57}$$



$U(x,y)$ debe ser nula en los bordes laterales y en el borde inferior, mientras que en el borde superior debe ser igual a una cierta función $f(x)$. La función $U(x,y)$ puede representar, por ejemplo, la temperatura **estacionaria** de una placa metálica rectangular, cuyos bordes laterales e inferior están a temperatura 0, mientras que su borde superior tiene una temperatura $f(x)$ dependiente de la posición. El objetivo es determinar la temperatura en el interior de la placa con estos datos.

La razón por la que esta temperatura satisface (55) es que en el caso general no estacionario, la temperatura queda determinada por la ecuación de difusión bidimensional

$$\frac{\partial U}{\partial t} - \alpha \left[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right] = 0 \tag{58}$$

En el caso estacionario, $U(x,y,t)$ **no depende del tiempo** t y por lo tanto $\partial U/\partial t = 0$. La ecuación (58) implica entonces (55).

Las funciones $U(x,y)$ que satisfacen la ecuación de Laplace (55) (y poseen derivadas segundas continuas en el interior de R) se denominan funciones **armónicas**. Puede demostrarse que estas funciones no poseen ni máximos ni mínimos en el interior del rectángulo, teniendo sus valores extremos en el borde del mismo. Puede demostrarse además que su valor en un punto interior (x,y) es el promedio de los valores que toma en cualquier círculo contenido en R que rodea dicho punto, por lo que son las funciones más "planas" posibles compatibles con las condiciones de contorno. Y dados los valores de $U(x,y)$ en el contorno, la solución de (55) es **única**.

Soluciones producto

Nuevamente comenzaremos proponiendo una solución elemental producto

$$U(x,y) = X(x)Y(y) \tag{59}$$

donde $X(x)$ depende solo de x e $Y(y)$ solo de y. Reemplazando en (55) se obtiene

$$X''(x)Y(y) + X(x)Y''(x) = 0$$

o sea, $X''(x)/X(x) = -Y''(y)/Y(y)$. Como el primer cociente depende solo de x y el segundo solo de y, siendo x e y variables independientes, esto implica

$$\frac{X''(x)}{X(x)} = -\frac{Y''(y)}{Y(y)} = -k^2 \tag{60}$$

es decir,

$$\begin{aligned} X''(x) + k^2 X(x) &= 0 & (61)\\ Y''(y) - k^2 Y(y) &= 0 & (62) \end{aligned}$$

Si $k = 0$, las soluciones generales de estas ecuaciones son

$$X(x) = A + Bx, \quad Y(y) = C + Dy \tag{63}$$

mientras que si $k \neq 0$,

$$\begin{aligned} X(x) &= A\cos kx + B \operatorname{sen} kx & (64)\\ Y(y) &= Ce^{ky} + De^{-ky} & (65) \end{aligned}$$

Se obtienen así las soluciones producto

$$X(x)Y(y) = \begin{cases} (A + Bx)(C + Dy) & (k = 0) \\ (A\cos kx + B \operatorname{sen} kx)(Ce^{ky} + De^{-ky}) & (k \neq 0) \end{cases} \tag{66}$$

Las condiciones de contorno laterales (56) implican

$$X(0) = 0, \quad X(L_x) = 0 \tag{67}$$

Nuevamente, para $k = 0$ esto conduce a $A = B = 0$, mientras que para $k \neq 0$, $A = 0$ y $B \operatorname{sen}(kL_x) = 0$ en (64), o sea,

$$k = n\pi/L_x\,, \quad n = 1, 2, 3, \ldots \tag{68}$$

Adenás, la condición de contorno inferior $U(x, 0) = X(x)Y(0) = 0$ implica $Y(0) = 0$ si $X(x) \neq 0$. Esto conduce a $C + D = 0$ en (65), o sea, $D = -C$ y entonces

$$Y(y) = C(e^{ky} - e^{-ky}) = 2C \operatorname{senh} ky$$

con k dado por (68). Se obtienen así las soluciones producto

$$X_n(x)Y_n(y) \;=\; B_n \sin(n\pi x/L_x) \operatorname{senh}(n\pi y/L_x)\,, \quad n = 1, 2, 3, \ldots \tag{69}$$

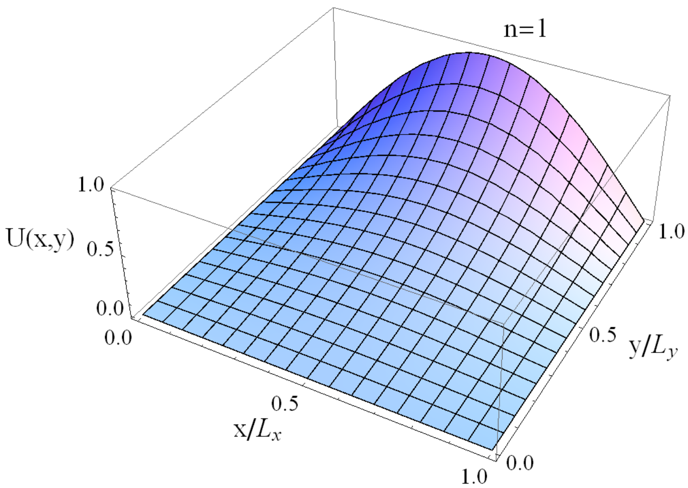
con B_n una constante arbitraria. Estas soluciones son *linealmente independientes*.

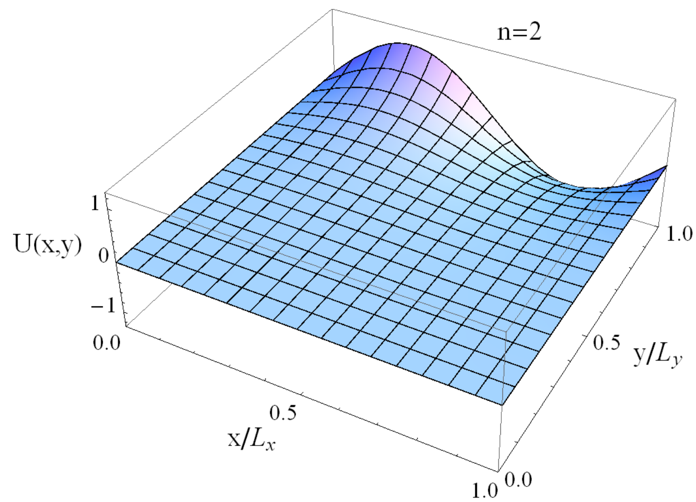
En el borde superior $y = L_y$, estas soluciones satisfacen

$$X_n(x)Y_n(L_y) = B_n \sin(n\pi x/L_x)\,\mathrm{senh}(n\pi L_y/L_x) \tag{70}$$

por lo que resuelven el problema planteado solo cuando $f(x) = F_n\,\mathrm{sen}(n\pi x/L_x)$, en cuyo caso $B_n = F_n/\,\mathrm{senh}(n\pi L_y/L_x)$.

Se muestra en los gráficos el comportamiento de las soluciones (70) para $n = 1, 2$ y $F_n = 1$. A mayor n, más rápidamente decrece $|Y_n(y)|$ al alejarse del borde $y = L_y$.





Solución general

Para $f(x)$ arbitraria, podemos ahora plantear la serie de soluciones producto,

$$u(x, y) = \sum_{n=1}^{\infty} B_n\,\mathrm{sen}(n\pi x/L_x)\,\mathrm{senh}(n\pi y/L_x) \tag{71}$$

que satisface la ecuación de Laplace (55) (asumiendo convergencia uniforme) junto con las condiciones de contorno laterales e inferior. Y en el borde superior,

$$u(x, L_y) = \sum_{n=1}^{\infty} B_n\,\mathrm{sen}(n\pi x/L_x)\,\mathrm{senh}(n\pi L_y/L_x) = f(x) \tag{72}$$

que corresponde al desarrollo en serie de Fourier de medio rango $f(x) = \sum_{n=1}^{\infty} F_n\,\mathrm{sen}(n\pi x/L_x)$, con $F_n = \frac{2}{L_x}\int_0^L f(x)\,\mathrm{sen}(n\pi x/L_x)dx$. Por lo tanto, $B_n\,\mathrm{senh}(n\pi L_y/L_x) = F_n$, o sea,

$$B_n = \frac{1}{\mathrm{senh}(n\pi L_y/L_x)} F_n \tag{73}$$

$$= \frac{1}{\mathrm{senh}(n\pi L_y/L_x)} \frac{2}{L_x} \int_0^{L_x} f(x)\,\mathrm{sen}(n\pi x/L_x)dx \tag{74}$$

La solución final **y única** es pues la serie (71) con los coeficientes (74). Se la puede ver como la suma de las soluciones producto (70) para cada término $F_n\,\mathrm{sen}(n\pi x/L_x)$ del desarrollo de Fourier de $f(x)$. Dado que $\mathrm{senh}(n\pi y/L_x) \le \mathrm{senh}(n\pi L_y/L_x)$ para $0 \le y \le L_y$, la serie (71) convergerá absolutamente si la serie de Fourier de medio rango (72) que representa a $f(x)$ converge absolutamente.

Ejemplo 1. Si $f(x) = 2\,\mathrm{sen}(\pi x/L_x) - \mathrm{sen}(3\pi x/L_x)$,

$$F_n = \frac{2}{L_x}\int_0^L f(x)\,\mathrm{sen}(n\pi x/L_x)dx = \begin{cases} 2 & n=1 \\ -1 & n=3 \\ 0 & n\neq 1,\ n\neq 3 \end{cases} \tag{75}$$

Por lo tanto la solución es

$$U(x,y) = 2\,\mathrm{sen}(\pi x/L_x)\frac{\mathrm{senh}(\pi y/L_x)}{\mathrm{senh}(\pi L_y/L_x)} - \mathrm{sen}(3\pi x/L_x)\frac{\mathrm{senh}(3\pi y/L_x)}{\mathrm{senh}(3\pi L_y/L_x)}$$

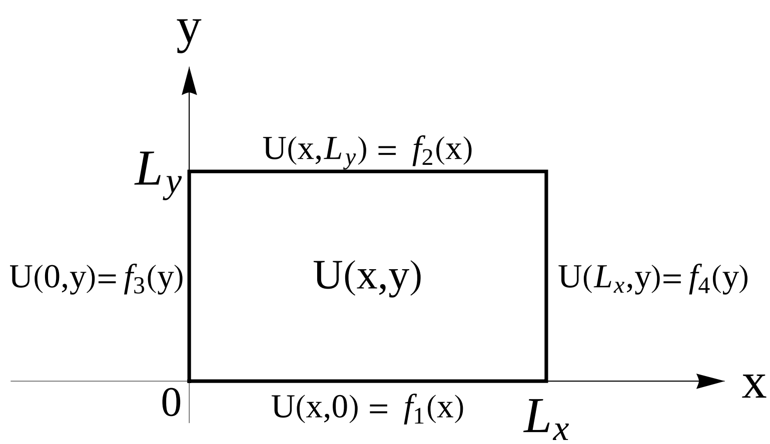
Condiciones de contorno generales

Consideremos ahora la ecuación de Laplace (55) en el rectángulo,

$$\frac{\partial^2 U(x,y)}{\partial x^2} + \frac{\partial^2 U(x,y)}{\partial y^2} = 0, \quad 0 < x < L_x,\ 0 < y < L_y \tag{76}$$

con las condiciones de contorno generales (ver figura)

$$U(x,0) = f_1(x), \quad U(x,L_y) = f_2(x)\,, \quad 0 \le x \le L_x \tag{77}$$
$$U(0,y) = f_3(y), \quad U(L_x,y) = f_4(y)\,, \quad 0 \le y \le L_y \tag{78}$$



En el caso trivial en que la temperatura es **constante** en el borde, $f_i(x) = C$ para $i = 1,2,3,4$, la solución de (76)–(78) es obviamente $U(x,y) = C$, es decir, la temperatura es **constante** e igual al valor en el contorno. Por otro lado, una función de la forma

$$U_0(x,y) = C + Ax + By + Dxy \tag{79}$$

es también una solución obvia de (76), que varía **linealmente** en los bordes:

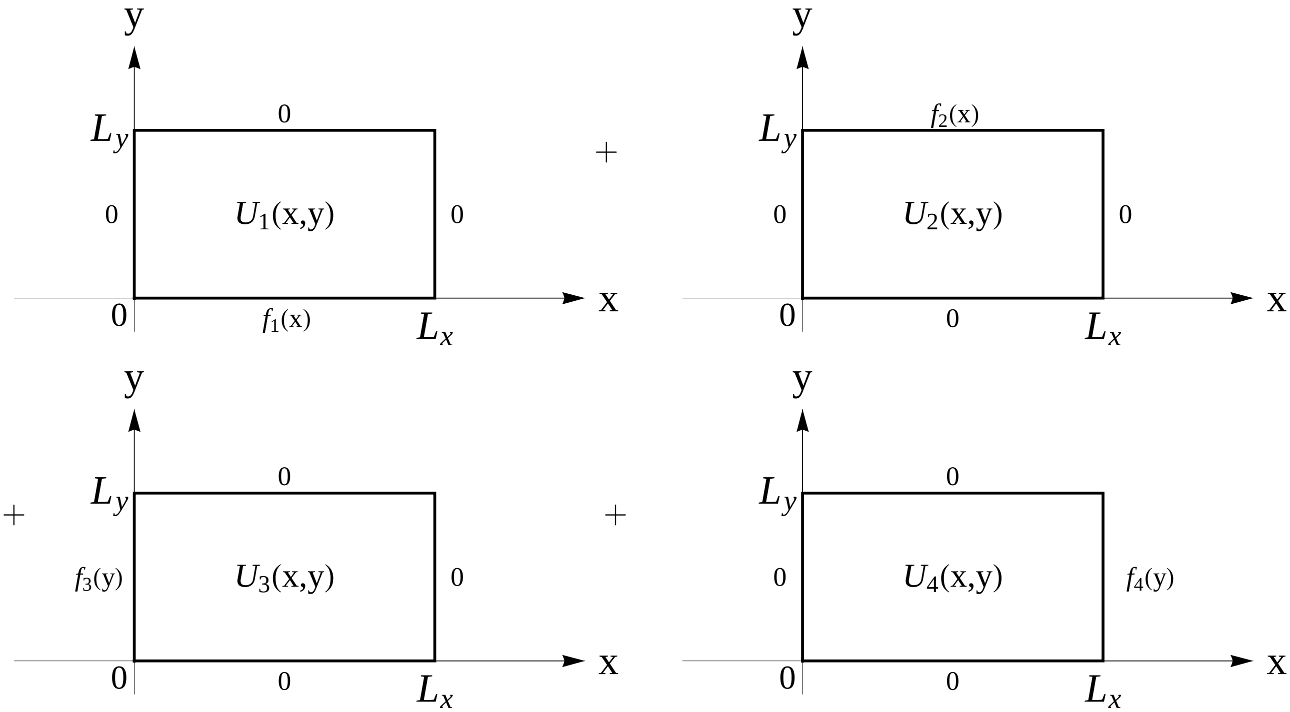
$$U_0(x,0) = C + Ax\,, \quad U_0(x,L_y) = C + BL_y + (A + DL_y)x \tag{80}$$
$$U_0(0,y) = C + By\,, \quad U_0(L_x,y) = C + AL_x + (B + DL_x)y \tag{81}$$

Es pues la solución de (76) cuando las funciones f_i son todas lineales y cumplen la condición de continuidad $f_1(0) = f_3(0)$, $f_1(L_x) = f_4(0)$, $f_2(0) = f_3(L_y)$, $f_2(L_x) = f_4(L_y)$.

En el caso general de funciones f_i arbitrarias, la solución de (76)–(78) puede obtenerse, usando la propiedad de superposición, como la suma de cuatro soluciones:

$$U(x,y) = U_1(x,y) + U_2(x,y) + U_3(x,y) + U_4(x,y) \tag{82}$$

donde $U_i(x,y)$ es la solución de (76) con $f_j(x) = 0$ para $j \neq i$ (ver figuras):



Problema 1. Probar que las soluciones anteriores $U_i(x,y)$ son de la forma

$$U_1(x,y) = \sum_{n=1}^{\infty} B_{n1} \operatorname{sen}(n\pi x/L_x) \operatorname{senh}[n\pi(L_y - y)/L_x] \tag{83}$$

$$U_2(x,y) = \sum_{n=1}^{\infty} B_{n2} \operatorname{sen}(n\pi x/L_x) \operatorname{senh}(n\pi y/L_x) \tag{84}$$

$$U_3(x,y) = \sum_{n=1}^{\infty} B_{n3} \operatorname{senh}[n\pi(L_x - x)/L_y] \operatorname{sen}(n\pi y/L_y) \tag{85}$$

$$U_4(x,y) = \sum_{n=1}^{\infty} B_{n4} \operatorname{senh}(n\pi x/L_y) \operatorname{sen}(n\pi y/L_y) \tag{86}$$

y dar una expresión para los coeficientes B_{ni} en términos de las funciones f_i.

Es conveniente suponer aquí que las funciones f_i se anulan en los extremos (vértices del rectángulo). Si esto no ocurre pero las funciones f_i cumplen la condición de continuidad en los vértices, puede usarse la solución elemental (79) para reproducir los valores de $U(x,y)$ en los 4 vértices y proponer una solución general $U(x,y) = U_0(x,y) + \sum_{i=1}^{4} U_i(x,y)$, donde las $U_i(x,y)$ serán de la forma (83)–(86) para $i = 1,2,3,4$. Así, los coeficientes B_{in} dependerán de la diferencia $f_i - U_0$ en el borde del rectángulo, que se anulará en los vértices.

Problemas 3.3.

1. Hallar la solución de la ecuación de Laplace en el rectángulo $0 < x < L_x$, $0 < y < L_y$ con las condiciones de contorno $U(0, y) = U(L_x, y) = U(x, 0) = 0$ (cero en los bordes laterales e inferior) y $U(x, L_y) = f(x)$ en el borde superior, si
a) $f(x) = \sin(\pi x/L_x) - \sin(3\pi x/L_x)$
b) $f(x) = x(L_x - x)/L^2$
c) Si ahora $U(x, y) = c$ (constante) en los bordes laterales e inferior, y $U(x, L_y) = f(x)+c$ en el borde superior, ¿como serán las soluciones? (no es necesario realizar nuevas cuentas).

2. En el rectángulo anterior, si ahora $U(0, y) = U(L_x, y) = U(x, L_y) = 0$ (cero en los bordes laterales y superior) pero $U(x, 0) = f(x)$ en el borde inferior, ¿como serán las soluciones para las mismas funciones $f(x)$ anteriores? (sugerencia: realizar un reemplazo adecuado en las soluciones obtenidas en 1. a) y 1. b), justificando el procedimiento).

3. En el rectángulo anterior, si ahora $U_x(0, y) = U_x(L_x, y) = 0 \ \forall\ y \in [0, L_y]$ (placa **térmicamente aislada** en los bordes laterales, siendo $U_x(x, y) = \frac{\partial U(x,y)}{\partial x}$), con $U(x, 0) = 0$ y $U(x, L_y) = f(x)$, mostrar que la solución es de la forma

$$U(x, y) = \frac{1}{2}A_0 y + \sum_{n=1}^{\infty} A_n \cos(n\pi x/L_x)\,\mathrm{senh}(n\pi y/L_x)$$

y determinar los coeficientes A_0, A_n.

4. Placa semi-infinita. Determinar la solución de la ecuación de Laplace en la región $0 < x < L$, $y > 0$, con las condiciones de contorno $U(0, y) = U(L_x, y) = 0$, $U(x, 0) = f(x)$, y $U(x, y) \to 0$ para $y \to \infty$. Mostrar que es de la forma

$$U(x, y) = \sum_{n=1}^{\infty} B_n\,\mathrm{sen}(n\pi x/L_x) e^{-n\pi y/L_x}$$

y determinar los coeficientes B_n.

5. (Opcional) Ecuación de Laplace en el disco $r < a$, $0 \le \theta \le 2\pi$. En coordenadas polares r, θ ($x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$), la ecuación de Laplace toma la forma

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} = 0$$

a) Mostrar que esta ecuación admite las soluciones producto

$$R(r)\Theta(\theta) = \begin{cases} (A + B\ln r)(C + D\theta) & k = 0 \\ (Ar^k + Br^{-k})(C\cos(k\theta) + D\,\mathrm{sen}(k\theta)) & k \neq 0 \end{cases}$$

b) Mostrar que en el disco, debe ser $B = D = 0$ si $k = 0$ y $B = 0$ si $k \neq 0$, con $k = 0, 1, 2, \ldots$, tal que la solución general tiene la forma

$$U(r, \theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} r^n[a_n \cos(n\theta) + b_n\,\mathrm{sen}(n\theta)]$$

Dar una expresión para los a_n, b_n si la condición de contorno es $U(a, \theta) = f(\theta)$.
c) Usando b), mostrar que $U(0, \theta)$ es el valor medio de $f(\theta)$.