Álgebra Lineal y Aplicaciones

Segunda Edición

Colección manuales uex - 112

Pedro
Sancho de Salas

112

ÁLGEBRA LINEAL Y APLICACIONES

MANUALES UEX

112

PEDRO SANCHO DE SALAS

ÁLGEBRA LINEAL Y APLICACIONES

Segunda Edición

UNIVERSIDAD DE EXTREMADURA

2024

UNIVERSIDAD DE EXTREMADURA

Edita

Universidad de Extremadura. Servicio de Publicaciones
C./ Caldereros, 2 - Planta 2ª - 10071 Cáceres (España)
Telf. 927 257 041 - Fax 927 257 046
publicac@unex.es
www.unex.es/ publicaciones
ISSN XXX
ISBN de méritos XXX

Índice general

Introducción

El presente manual se ha escrito como texto de referencia del curso de Álgebra Lineal I del Grado en Química de la Uex. Esta asignatura es obligatoria en los grados de Biología, Estadística, Física, Química, Matemáticas, etc., de la facultad de Ciencias. Tiene también muchos contenidos comunes con Matemáticas I de Ingeniería Industrial. Hemos procurado escribir un texto apropiado para todas ellas, de modo que el estudiante o profesor pueda escoger o resaltar aquellas partes que considere más convenientes. Así, un estudiante de matemáticas dará más importancia a las demostraciones, a la coherencia y desarrollo de la teoría y un estudiante de ingeniería a las aplicaciones.

El manual está divido en cuatro temas. En cada tema incluimos una lista de problemas (con sus soluciones), prácticas de cálculos con ordenador y la biografía de un matemático relevante (en inglés).

Hagamos una brevísima introducción al curso.

En el Álgebra Lineal confluyen la visión geométrica, con la analítica y la algebraica. Así, por ejemplo, el conjunto de soluciones del sistema de ecuaciones

$$\begin{array}{rcl} 2x_1+x_2-3x_3 & = & 0 \\ x_1+5x_2+x_3 & = & 0 \end{array}$$

puede verse como la intersección del plano de $\mathbb{R}^3$ de ecuación $2x_1+x_2-3x_3=0$ con el plano $2x_1+x_2-3x_3=0$; puede verse como el conjunto de puntos donde se anulan las funciones $2x_1+x_2-3x_3$ y $2x_1+x_2-3x_3$, donde x_1,x_2,x_3 son las tres funciones coordenadas de $\mathbb{R}^3$; o puede verse como un subespacio vectorial de $\mathbb{R}^3$, es decir, un subconjunto $V\subset\mathbb{R}^3$ que cumple que si $v,v'\in V$ entonces $\lambda\cdot v+\mu\cdot v'\in V$ para todo $\lambda,\mu\in\mathbb{R}$.

Por otra parte, muchos problemas de la Física, Química, Economía etc., admiten una modelización lineal para su resolución aproximada.

El concepto principal del capítulo segundo y de todo el curso es la noción de linealidad, la noción de espacio vectorial. Los espacios vectoriales son conjuntos cuyos elementos (llamados vectores) se pueden sumar entre sí y multiplicar por escalares. Las aplicaciones lineales son las aplicaciones entre espacios vectoriales que respetan la linealidad. Es fundamental en el estudio y clasificación de los espacios vectoriales y sus aplicaciones lineales la introducción y definición de base de un espacio vectorial. Con las bases, podemos expresar los vectores como n coordenadas y las aplicaciones lineales como matrices ($n \cdot m$ coordenadas). Introducimos el método de la cascada de Gauss para calcular bases y resolver sistemas de ecuaciones lineales. Si $T : \mathbb{R}^n \to \mathbb{R}^n$ es una aplicación lineal y $V \subset \mathbb{R}^n$ es un objeto de volumen v entonces $T(V)$ es un objeto de volumen $\det(T) \cdot v$, y se dice que $\det(T)$ es el determinante de T. Con el determinante calculamos cuándo n vectores son linealmente independientes, cuándo T tiene inversa y resolvemos por la regla de Cramer los sistemas de ecuaciones lineales. Damos diversas aplicaciones. Por ejemplo, balanceamos ecuaciones químicas, calculamos la presión de vapor de una mezcla de líquidos volátiles, las intensidades de corrientes en los distintas ramas de un circuito eléctrico con resistencias y baterías, etc.

Entre las aplicaciones lineales destacan las aplicaciones lineales de un espacio vectorial en sí mismo, es decir, los operadores lineales. Buscaremos bases donde las matrices asociadas a los operadores lineales sean muy sencillas. Para todo ello introduciremos los conceptos de autovectores, autovalores y polinomio característico del operador lineal. Dentro de este marco trataremos múltiples problemas como los que siguen: La resolución de ciertas ecuaciones diferenciales y ecuaciones en diferencias finitas. La ecuación de movimiento de un objeto en caída libre; la ecuación del movimiento armónico amortiguado. El pago de una hipoteca a interés fijo. La ley de desintegración radiactiva. La evolución de poblaciones. La función de onda de una partícula en un pozo. La modelización del efecto de un terremoto sobre un edificio, etc.

Por último, en el cuarto capítulo estudiamos los espacios vectoriales euclídeos. Es decir, espacios vectoriales donde hay una noción de distancia como la que tenemos en $\mathbb{R}^3$. En términos matemáticos, estudiamos los espacios vectoriales donde hay definido un producto escalar. Definimos las bases ortonormales y damos el método de Gram-Schmidt para calcularlas. En estas bases, las distancias, las proyecciones ortogonales, las aplicaciones lineales que conservan las distancias, los volúmenes, etc. se calculan de modo sencillo. Dentro de este marco calculamos las ecuaciones reducidas de las cuádricas (elipses, hiperboloides, etc.) y la solución aproximada óptima de un sistema de ecuaciones lineales. Definimos los operadores hermíticos y cómo se aplican en Mecánica Cuántica. Hacemos una introducción a la Relatividad Especial. Etc.

Éste es el desafío que tenemos por delante.

Querido lector: “Aliquando bonus dormitat Homerus". Si encuentras alguna errata puedes escribirme para que la corrija a la dirección: sancho@unex.es.

Capítulo 1

El cuerpo de los números complejos

1.1. Preliminares

Un conjunto está totalmente determinado por sus elementos, que se pueden contar, aunque puedan ser infinitos. El conjunto vacío, que es el conjunto que no tiene ningún elemento, se denota $\emptyset$.

Si X es un conjunto, $x \in X$ significa que x es un elemento de X.

Si X e Y son conjuntos, $Y \subseteq X$ significa que Y es un subconjunto de X, i.e. que todos los elementos de Y son elementos de X. Si Z es otro subconjunto de X, se define unión e intersección con Y como sigue:

$$Y \cup Z := \{x \in X \text{ tales que } x \in Y \text{ ó } x \in Z\}.$$
$$Y \cap Z := \{x \in X \text{ tales que } x \in Y \text{ y } x \in Z\}.$$

El producto directo o cartesiano de los conjuntos X_1 y X_2, se denota $X_1 \times X_2$ y es el conjunto definido como sigue

$$X_1 \times X_2 := \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\}$$

Es decir, $X_1 \times X_2$ es el conjunto formado por todas las parejas (x_1, x_2), para todo $x_1 \in X_1$ y todo $x_2 \in X_2$.

Los conjuntos más importantes son

Números naturales	$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$
Números enteros	$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
Números racionales	$\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$
Números reales	$\mathbb{R} = \left\{\begin{array}{l}\text{Números con infinitos decimales } n'n_1n_2\ldots, \\ \text{con } n \in \mathbb{Z} \text{ y } 0 \leq n_i \leq 9, \forall i.\end{array}\right\}$

Recordemos que $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

1.2. Números Complejos

1. Definiciones: El conjunto de los números **complejos**, que denotamos por $\mathbb{C}$, es el conjunto definido por

$$\mathbb{C} := \{x + y \cdot i, \text{ para todo } x, y \in \mathbb{R}\} = \{5 + 7 \cdot i, \pi + 4 \cdot i, 3 + \pi \cdot i, -2 + 7 \cdot i, \dots\}$$

Dado un número complejo $z = x + yi$, se dice que $x \in \mathbb{R}$ es la **parte real** de z e $y \in \mathbb{R}$ la **parte imaginaria**. Seguiremos las siguientes notaciones $a + 0 \cdot i = a$ ($a \in \mathbb{R}$) y $0 + b \cdot i = b \cdot i$ ($b \in \mathbb{R}$). Observemos que $\mathbb{R} \subset \mathbb{C}$, $a = a + 0 \cdot i$.

Los números complejos se suman y multiplican con las siguientes reglas:

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

En particular,

$$\boxed{i^2 = (0 + 1 \cdot i) \cdot (0 + 1 \cdot i) = -1}$$

Nota: $\mathbb{C}$ es un conjunto importante en Matemáticas. El teorema de D'Alambert dice que si $p(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$ es un polinomio con coeficientes complejos de grado n (es decir, $a_i \in \mathbb{C}$ para todo i y $a_0 \neq 0$) entonces existen $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ únicos (que se denominan las raíces de $p(x)$) de modo que

$$p(x) = a_0 \cdot (x - \alpha_1) \cdot (x - \alpha_2) \cdots (x - \alpha_n).$$

Por ejemplo, $x^2 + 1$ no tiene raíces reales pero sí complejas: $x^2 + 1 = (x - i)(x + i)$; con otras palabras, $\sqrt{-1} = i \in \mathbb{C}$ y $\sqrt{-1} \notin \mathbb{R}$. Los números complejos son muy utilizados en Álgebra, Análisis, Electromagnetismo, Mecánica Cuántica, Química Cuántica, etc.

2. Ejercicio: Calcula $(1 + 2i) \cdot (1 - 3i) + 7 + 5i$.

El **conjugado** de $z = x + yi$ es el número complejo $\bar{z} := x - yi$. Algunas propiedades de la conjugación:

$$\overline{z + u} = \bar{z} + \bar{u}, \quad \overline{zu} = \bar{z}\bar{u}, \quad \bar{\bar{z}} = z.$$

El **módulo** de $z = x + yi$ es el número real $|z| := \sqrt{z \cdot \bar{z}} = \sqrt{x^2 + y^2} \geq 0$. Algunas propiedades del módulo:

$$|z| = |\bar{z}|, \quad |z| = 0 \Leftrightarrow z = 0, \quad |zu| = |z| \cdot |u|, \quad |z + u| \leq |z| + |u|.$$

Para demostrar la última propiedad, basta ver que $|z+u|^2 \le (|z|+|u|)^2$: Para ello, empecemos observando que si $w = c+di$ es un número complejo entonces $w+\bar{w} = 2c \le 2|c| \le 2\sqrt{c^2+d^2} = 2|w|$. Sea $w := z\bar{u}$. Entonces,

$$\begin{aligned}
|z+u|^2 &= (z+u)(\overline{z+u}) = (z+u)(\bar{z}+\bar{u}) = |z|^2+|u|^2+z\bar{u}+\bar{z}u \\
&= |z|^2+|u|^2+z\bar{u}+\overline{z\bar{u}} = |z|^2+|u|^2+w+\bar{w} \\
(|z|+|u|)^2 &= |z|^2+|u|^2+2|z|\cdot|u| = |z|^2+|u|^2+2|z|\cdot|\bar{u}| = |z|^2+|u|^2+2|w|.
\end{aligned}$$

Luego, $|z+u|^2 \le (|z|+|u|)^2$.

3. Ejercicio: Sean $z, z' \in \mathbb{C}$. Probad que $z\cdot z' = 0$ si y solo si $z = 0$ ó $z' = 0$.

Sea $z = x+yi$ un número complejo no nulo y $z' = x'+y'i$. Calculemos $\frac{z'}{z}$:

$$\frac{z'}{z} = \frac{z'\bar{z}}{z\bar{z}} = \frac{(x'x-y'y)+(x'y+y'x)i}{x^2+y^2} = \frac{x'x-y'y}{x^2+y^2} + \frac{x'y+y'x}{x^2+y^2}\cdot i.$$

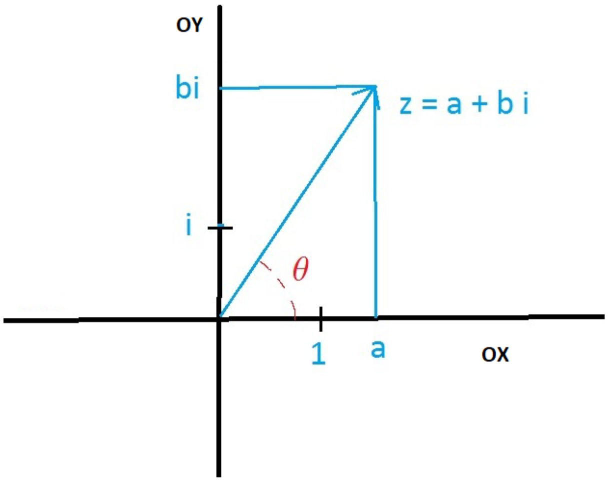
En particular, $\frac{1}{z} = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2}\, i$.

4. Ejercicio: Calcula $\frac{1+2i}{3+5i}$.

1.3. Representación gráfica de $\mathbb{C}$

Los números reales se representan como los puntos de una recta: Fijado un origen en la recta, es decir, el cero, y el punto 1 (a la derecha del origen, a una distancia de una unidad), cada número real se corresponde con un punto de la recta, de manera que los números reales llenan por completo la recta.

Para representar a los números complejos vamos a considerar un plano. El número complejo $z = a+bi$ se representa como el punto (a,b) del plano (o equivalentemente, como el vector de origen el punto $(0,0)$ y extremo el punto (a,b)). Por tanto, hay tantos números complejos como puntos del plano, y podemos identificar $\mathbb{C}$ con los puntos del plano. Los números reales $a+0i = a$ se representan por puntos $(a,0)$ del eje OX, llamado eje real. Los números imaginarios puros bi se sitúan en el eje OY, llamado eje imaginario.



Observemos que $|z| = \sqrt{a^2+b^2}$ es la longitud del vector z que dibujamos en el plano. Observemos que el **ángulo θ** entre los vectores 1 y z es el número real (comprendido entre 0 y 2π)[1] que cumple que $\cos\theta = \frac{a}{|z|}$ y $\operatorname{sen}\theta = \frac{b}{|z|}$. Por tanto,

$$\boxed{z = a + bi = |z|\cdot(\cos\theta + i\cdot\operatorname{sen}\theta)}$$

(Forma trigonométrica)

Diremos que θ es el **argumento** de z. Puede comprobarse que $\theta = \operatorname{sign}(b)\cdot\arccos\frac{a}{|z|}$. Si ρ es el módulo de un número complejo z y θ es su argumento, se suele seguir la siguiente notación

$$\boxed{z = \rho_\theta}$$

(Forma polar)

Observemos que un número complejo z tiene módulo 1 si solo si $z = \cos x + i\operatorname{sen} x$ para cierto $x \in \mathbb{R}$ (donde x es el argumento de z). Por tanto, los números complejos de módulo 1 se corresponden con los puntos de la circunferencia de radio 1.

1. Proposición: *Se cumple que*

$$\rho_\theta\cdot\rho'_{\theta'} = (\rho\cdot\rho')_{\theta+\theta'}.$$

Demostración. Recordemos las fórmulas de trigonometría: $\cos(x+x') = \cos x\cdot\cos x' - \operatorname{sen} x\cdot\operatorname{sen} x'$ y $\operatorname{sen}(x+x') = \cos x\cdot\operatorname{sen} x' + \operatorname{sen} x\cdot\cos x'$. Entonces,

$$\begin{aligned}\rho_\theta\cdot\rho'_{\theta'} &= \rho(\cos\theta + i\operatorname{sen}\theta)\cdot\rho'(\cos\theta' + i\operatorname{sen}\theta')\\ &= \rho\rho'(\cos\theta\cos\theta' - \operatorname{sen}\theta\operatorname{sen}\theta' + i(\cos\theta\operatorname{sen}\theta' + \operatorname{sen}\theta\cos\theta'))\\ &= \rho\rho'(\cos(\theta+\theta') + i\operatorname{sen}(\theta+\theta')) = (\rho\cdot\rho')_{\theta+\theta'}.\end{aligned}$$

□

[1]Seguiremos la siguiente convención: el ángulo α es el mismo que $\alpha + n\cdot 2\pi$ para todo $n \in \mathbb{Z}$.

2. Corolario: *Se cumple que*

1. $(\rho_\theta)^n = (\rho^n)_{n\cdot\theta}$.
2. $\frac{\rho_\theta}{\rho'_{\theta'}} = (\rho/\rho')_{\theta-\theta'}$.

3. Fórmula de De Moivre: *Se cumple que*

$$(\cos x + i \operatorname{sen} x)^n = \cos nx + i \cdot \operatorname{sen} nx.$$

Demostración. El número complejo $\cos x + i \operatorname{sen} x$ es de módulo 1 y argumento x, luego $(\cos x + i \operatorname{sen} x)^n$ es un número complejo de módulo 1 y argumento nx, es decir, es igual a $\cos nx + i \operatorname{sen} nx$. □

4. Ejercicio: Expresa $\cos 3x$ en función de $\cos x$ y $\operatorname{sen} x$.

Solución: Usemos la fórmula de De Moivre:

$$\cos 3x + i \operatorname{sen} 3x = (\cos x + i \operatorname{sen} x)^3 = ((\cos x)^3 - 3(\cos x)(\operatorname{sen} x)^2) + i(3(\cos x)^2 \operatorname{sen} x - (\operatorname{sen} x)^3).$$

Por tanto, $\cos 3x = (\cos x)^3 - 3(\cos x)(\operatorname{sen} x)^2$.

1.4. Forma exponencial de un número complejo

1. Definición: Sea $t \in \mathbb{R}$, definimos $e^{ti} := \cos t + i \cdot \operatorname{sen} t$. [2]

2. Ejercicio: Comprobad la **identidad de Euler** (1707-1783)

$$e^{\pi i} = -1.$$

3. Ejercicio: Probad que $e^{2\pi i} = 1$.

Si $z = \rho_\theta$ (donde $\rho = |z|$ y θ es el argumento de z) entonces

$$z = \rho \cdot (\cos\theta + i \operatorname{sen}\theta) = \rho \cdot e^{\theta i} \qquad \text{(Forma exponencial de } z\text{).}$$

Por tanto, $|z| = 1$ si y solo si $z = e^{\theta i}$ para un cierto $\theta \in [0, 2\pi)$.

[2] Hay razones que explican esta notación: si escribimos los desarrollos de Taylor infinitos en el origen $e^x = 1 + x + x^2/2! + x^3/3! + x^4/4! + \cdots$, $\operatorname{sen} x = 0 + x - x^3/3! + \cdots$ y $\cos x = 1 - x^2/2! + x^4/4! + \cdots$, entonces se puede comprobar que $e^{ti} = \cos t + i \cdot \operatorname{sen} t$.

4. Proposición: $\boxed{e^{\theta i} \cdot e^{\theta' i} = e^{(\theta+\theta')i}}$.

Demostración. En efecto,

$$e^{\theta i} \cdot e^{\theta' i} = 1_\theta \cdot 1_{\theta'} = 1_{\theta+\theta'} = e^{(\theta+\theta')i}.$$

□

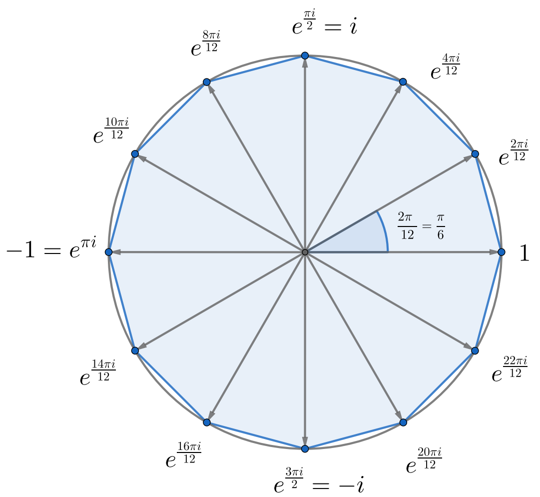
Raíces n-ésimas: Sea $z = \rho e^{i\theta}$ un número complejo. Queremos calcular los números complejos $u = \rho' e^{i\theta'}$ tales que $u^n = z$ (para cierto número natural $n \geq 2$). Es decir, queremos calcular las raíces n-ésimas de z. Tenemos

$$\rho e^{i\theta} = z = u^n = \rho'^n e^{in\theta'}.$$

Entonces, $\rho = \rho'^n$ luego $\rho' = \sqrt[n]{\rho}$; y $\theta + 2k\pi = n\theta'$ luego $\theta' = \frac{\theta}{n} + \frac{2k\pi}{n}$ (y claramente basta tomar $k = 0, \ldots, n-1$). Luego,

$$\boxed{u = \sqrt[n]{\rho}\, e^{\left(\frac{\theta+2k\pi}{n}\right)i};\ k = 0, \ldots, n-1.}$$

Todo número complejo no nulo $z = \rho e^{i\theta}$ tiene n raíces n-ésimas complejas (que forman un polígono regular de n vértices, inscrito en el círculo de radio $\sqrt[n]{\rho}$ centrado en el 0)



En particular, las raíces n-ésimas de la unidad son

$$\{e^{\frac{2k\pi}{n}i}, \quad k = 0, \ldots, n-1\}.$$

En la imagen de la izquierda hemos dibujado las raíces duodécimas de la unidad, que son los vértices del dodecágono inscrito en la circunferencia unidad.

Las raíces n-ésimas de un número complejo no nulo $\rho e^{\theta i}$ se obtienen multiplicando una de ellas por las n raíces n-ésimas de la unidad, que son las sucesivas potencias de $\varepsilon_n := e^{\frac{2\pi}{n}i}$: En efecto,

$$\sqrt[n]{\rho}\, e^{\left(\frac{\theta+2k\pi}{n}\right)i} = \sqrt[n]{\rho}\, e^{\left(\frac{\theta}{n}\right)i} \cdot e^{\frac{2k\pi}{n}i} = \sqrt[n]{\rho}\, e^{\left(\frac{\theta}{n}\right)i} \cdot (e^{\frac{2\pi}{n}i})^k.$$

5. Ejemplos: Las raíces n-ésimas de la unidad complejas, cuando $n = 2,3,4,6$ y 8, son:

$\varepsilon_2 = -1$, $\varepsilon_2^2 = 1$.

$\varepsilon_3 = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$, $\varepsilon_3^2 = -\frac{1}{2} - \frac{\sqrt{3}}{2} i$, $\varepsilon_3^3 = 1$.

$\varepsilon_4 = i$, $\varepsilon_4^2 = -1$, $\varepsilon_4^3 = -i$, $\varepsilon_4^4 = 1$.

$\varepsilon_6, \varepsilon_6^5 = \frac{1}{2} \pm \frac{\sqrt{3}}{2} i$, $\varepsilon_6^2, \varepsilon_6^4 = \frac{-1}{2} \pm \frac{\sqrt{3}}{2} i$, $\varepsilon_6^3 = -1$, $\varepsilon_6^6 = 1$.

$\varepsilon_8, \varepsilon_8^7 = \frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}$, $\varepsilon_8^2, \varepsilon_8^6 = \pm i$, $\varepsilon_8^3, \varepsilon_8^5 = \frac{-1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}$, $\varepsilon_8^4 = -1$, $\varepsilon_8^8 = 1$.

1.5. History of Complex Numbers

Many mathematicians contributed to the development of complex numbers. The rules for addition, subtraction, multiplication, and root extraction of complex numbers were developed by the Italian mathematician Rafael Bombelli. A more abstract formalism for the complex numbers was further developed by the Irish mathematician William Rowan Hamilton, who extended this abstraction to the theory of quaternions.

The earliest fleeting reference to square roots of negative numbers can perhaps be said to occur in the work of the Greek mathematician Hero of Alexandria in the 1st century AD, where in his Stereometrica he considers, apparently in error, the volume of an impossible frustum of a pyramid to arrive at the term $\sqrt{81-144} = 3i\sqrt{7}$ in his calculations, although negative quantities were not conceived of in Hellenistic mathematics and Hero merely replaced it by its positive ($\sqrt{144-81} = 3\sqrt{7}$).

The impetus to study complex numbers as a topic in itself first arose in the 16th century when algebraic solutions for the roots of cubic and quartic polynomials were discovered by Italian mathematicians (Niccolò Fontana Tartaglia, Gerolamo Cardano, etc.). It was soon realized that these formulas, even if one was only interested in real solutions, sometimes required the manipulation of square roots of negative numbers. As an example, Tartaglia's formula for a cubic equation of the form $x^3 = px + q$ gives the solution to the equation $x^3 = x$ as

$$\frac{1}{\sqrt{3}}\left(\left(\sqrt{-1}\right)^{1/3} + \left(\sqrt{-1}\right)^{-1/3}\right).$$

At first glance this looks like nonsense. However formal calculations with complex numbers show that the equation $z^3 = i$ has solutions $-i, \frac{\sqrt{3}}{2} + \frac{1}{2}i$ and $\frac{-\sqrt{3}}{2} + \frac{1}{2}i$. Substituting these in turn for $\sqrt{-1}^{1/3}$ in Tartaglia's cubic formula and simplifying, one gets

0, 1 and -1 as the solutions of $x^3 - x = 0$. Of course this particular equation can be solved at sight but it does illustrate that when general formulas are used to solve cubic equations with real roots then, as later mathematicians showed rigorously, the use of complex numbers is unavoidable. Rafael Bombelli was the first to explicitly address these seemingly paradoxical solutions of cubic equations and developed the rules for complex arithmetic trying to resolve these issues.

The term "imaginary" for these quantities was coined by René Descartes in 1637, although he was at pains to stress their imaginary nature.

A further source of confusion was that the equation $\sqrt{-1}^2 = \sqrt{-1}\,\sqrt{-1} = -1$ seemed to be capriciously inconsistent with the algebraic identity $\sqrt{a}\,\sqrt{b} = \sqrt{ab}$, which is valid for non-negative real numbers a and b, and which was also used in complex number calculations with one of a, b positive and the other negative. The incorrect use of this identity (and the related identity $\frac{1}{\sqrt{a}} = \sqrt{\frac{1}{a}}$ in the case when both a and b are negative even bedeviled Euler. This difficulty eventually led to the convention of using the special symbol i in place of $\sqrt{-1}$ to guard against this mistake. Even so, Euler considered it natural to introduce students to complex numbers much earlier than we do today. In his elementary algebra text book, Elements of Algebra, he introduces these numbers almost at once and then uses them in a natural way throughout.

In the 18th century complex numbers gained wider use, as it was noticed that formal manipulation of complex expressions could be used to simplify calculations involving trigonometric functions. For instance, in 1730 Abraham de Moivre noted that the complicated identities relating trigonometric functions of an integer multiple of an angle to powers of trigonometric functions of that angle could be simply re-expressed by the following well-known formula which bears his name, de Moivre's formula:

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

In 1748 Leonhard Euler went further and obtained Euler's formula of complex analysis:

$$\cos\theta + i\sin\theta = e^{i\theta}$$

by formally manipulating complex power series and observed that this formula could be used to reduce any trigonometric identity to much simpler exponential identities.

The idea of a complex number as a point in the complex plane (above) was first described by Caspar Wessel in 1799, although it had been anticipated as early as 1685 in Wallis's De Algebra tractatus.

Wessel's memoir appeared in the Proceedings of the Copenhagen Academy but went largely unnoticed. In 1806 Jean-Robert Argand independently issued a pamphlet on complex numbers and provided a rigorous proof of the fundamental theorem of algebra. Carl Friedrich Gauss had earlier published an essentially topological proof of the theorem in 1797 but expressed his doubts at the time about "the true metaphysics of the square root of -1". It was not until 1831 that he overcame these doubts and published his treatise on complex numbers as points in the plane, largely establishing modern notation and terminology. In the beginning of the 19th century, other mathematicians discovered independently the geometrical representation of the complex numbers: Buée, Mourey, Warren, Français and his brother, Bellavitis.

The English mathematician G.H. Hardy remarked that Gauss was the first mathematician to use complex numbers in "a really confident and scientific way" although mathematicians such as Niels Henrik Abel and Carl Gustav Jacob Jacobi were necessarily using them routinely before Gauss published his 1831 treatise.

Augustin Louis Cauchy and Bernhard Riemann together brought the fundamental ideas of complex analysis to a high state of completion, commencing around 1825 in Cauchy's case.

The common terms used in the theory are chiefly due to the founders. Argand called $\cos\phi + i\sin\phi$ the direction factor, and $r = \sqrt{a^2+b^2}$ the modulus; Cauchy (1828) called $\cos\phi + i\sin\phi$ the reduced form (l'expression réduite) and apparently introduced the term argument; Gauss used i for $\sqrt{-1}$, introduced the term complex number for $a+bi$, and called a^2+b^2 the norm. The expression direction coefficient, often used for $\cos\phi + i\sin\phi$, is due to Hankel (1867), and absolute value, for modulus, is due to Weierstrass.

Later classical writers on the general theory include Richard Dedekind, Otto Hölder, Felix Klein, Henri Poincaré, Hermann Schwarz, Karl Weierstrass and many others. (From Wikipedia: Complex Numbers. https://en.wikipedia.org/wiki/Complex_number)

1.6. Cálculos con ordenador

Vamos a indicar algunos comandos convenientes en Mathematica para realizar cálculos con números complejos. En Wolfram Alpha (https://www.wolframalpha.com/) podemos escribir estos comandos, más aún, permite múltiples modos de escribir nuestros cálculos, incluso a menudo corrige nuestros errores. Escribe:

1. $(1+2*I)/(1+3*I)$.

2. $(1+2*I)\hat{}(-5)$.

3. $Conjugate[1+I]*Conjugate[1-I]$.

4. $Abs[1+I]$ (que es el módulo o valor absoluto del número complejo $1+i$). En Wolfram Alpha, $|1+i|$, $abs(1+i)$.

5. $Arg[1+I]$ (que es el argumento del número complejo $1+i$).

6. $ArcTan[Sqrt[3]]$, $Cos[Pi/2]$, $Sin[Pi/3]$, $Cot[Pi/3]$, $N[Tan[1]]$.

7. $3*Exp[Pi*I]$. En Wolfram Alpha, $3e\hat{}(pi*i)$.

8. $3*Exp[x*I]$.

9. $Re[E\hat{}(2*Pi*I/3)]$, $Im[E\hat{}(2*Pi*I/3)]$ (parte real e imaginaria de $e^{2\pi i/3}$). En Wolfram Alpha, $1\hat{}(1/3)$.

1.7. Problemas

1. Realiza las siguientes operaciones con números complejos y escribe el resultado en forma binomial $a+bi$, siendo a y b números reales:

$$\frac{1+i}{1-i}, \qquad \frac{1+\sqrt{3}i}{(\sqrt{3}+i)^2}, \qquad (1+i)^2(1-i\sqrt{3})^2,$$

$$\frac{(2-i)^2}{(-3i)^3}, \qquad (3+2i)(2-i)+\frac{2-3i}{4-i}, \qquad \frac{1}{i+\frac{1}{i+\frac{1}{1+i}}}.$$

2. Calcula el valor del número real a para que z sea un número real, donde

$$z=\frac{3-2ai}{4-3i}.$$

3. Calcula para qué valores reales de b y c se cumple que $(2+bi)(c+3i)=-1+7i$.

4. Dado el número complejo $z=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$, prueba que

   a) $1+z+z^2=0$;

   b) $\frac{1}{z}=z^2$.

5. Calcula el módulo y el argumento de los números complejos: i, $1+i$, $1-i$, $-1-i$, $-1+i$.

6. Uno de los vértices de un hexágono regular inscrito en una circunferencia con centro en el origen de coordenadas tiene coordenadas $(1,1)$. Halla las coordenadas de los otros cinco vértices.

7. Escribe en forma polar los siguientes números complejos: $-1+i$, $3i$, -5, $2\sqrt{3}-2i$.

8. Calcula la forma binomial de los siguientes números complejos $e^{i7\pi}$, $2e^{i\pi/4}$, $6e^{i\pi/6}$, $e^{i\pi/4}-e^{-i\pi/4}$, $\frac{1-e^{i\pi/2}}{1+e^{i\pi/2}}$.

9. Realiza las siguientes operaciones y expresa el resultado en forma polar y en forma binomial $1_{\frac{5\pi}{6}}\cdot 5_{\frac{\pi}{6}}$, $(2_{\frac{\pi}{6}})^3$, $(2_{\frac{5\pi}{12}})^3$, $(2_{\frac{3\pi}{2}})^3$.

10. Determina en qué cuadrante del plano complejo[3] está (el afijo correspondiente a) cada uno de los siguientes números complejos:

$$(\mathrm{e}^{\frac{3\pi}{4}i})^7, \qquad (\mathrm{e}^{\frac{9\pi}{7}i})^{11}, \qquad (1_{\frac{2\pi}{3}})^{10}, \qquad (2_{\frac{5\pi}{6}})^8.$$

11. Escribe en forma binomial i^{37}, i^{126}, $z=i^{2009}-i^{2010}$.

12. Halla las raíces cuadradas de -4, -2, $3i$, $-4i$, $1+i$.

13. Calcula las raíces cuadradas de $a+bi$.

14. Resuelve las siguientes ecuaciones en $\mathbb{C}$:

$$z^2-\sqrt{2}z+1=0, \qquad z^2-2z+2=0.$$

15. Dados $z_1=\frac{\sqrt{2}}{2}+\frac{\sqrt{6}}{2}i$, $z_2=1+i$ calcula $z=\frac{z_1^{100}}{z_2^{104}}$, y calcula las raíces cuartas de z.

[3]Dividimos el plano el plano en cuatro cuadrantes: forman el primer cuadrante los números complejos de argumento θ tal que $0\le\theta<\frac{\pi}{2}$, el segundo cuadrante los números complejos de argumento θ tal que $\frac{\pi}{2}\le\theta<\pi$, etc.

16. Calcula todas las raíces del polinomio $P(z) = z^6 - 2z^3 + 2$.

17. Determina el valor de $\cos\frac{\pi}{12}$ y $\operatorname{sen}\frac{\pi}{12}$ a partir del cociente $1_{\frac{\pi}{3}} : 1_{\frac{\pi}{4}}$.

18. Dado el número complejo $z = (1 - \sqrt{3}i)^{-1}$, calcula:

    a) El módulo y el argumento de z.

    b) Los valores del número entero n para los que se cumple que z^n es un número real.

19. Calcula $\int e^x \cdot \cos^3 x\, dx$.

1.8. Solución de los problemas

P1. Tenemos $\frac{1+i}{1-i} = \frac{(1+i)^2}{(1-i)(1+i)} = \frac{2i}{2} = i$.

$\frac{1+\sqrt{3}i}{(\sqrt{3}+i)^2} = \frac{1+\sqrt{3}i}{(2+2\sqrt{3}i)} = \frac{1}{2}$.

$(1+i)^2(1-i\sqrt{3}) = 2i(1-i\sqrt{3}) = 2\sqrt{3}+2i$.

$\frac{(2-i)^2}{(-3i)^3} = \frac{3-4i}{27i} = \frac{-4}{27} + \frac{-1}{9}i$.

$(3+2i)(2-i) + \frac{2-3i}{4-i} = 8+i+\frac{(2-3i)(4+i)}{17} = 8+i+\frac{11-10i}{17} = \frac{147}{17}+\frac{7}{17}i$.

$\frac{1}{i+\frac{1}{i+\frac{1}{1+i}}} = \frac{1}{i+\frac{1}{i+\frac{1-i}{2}}} = \frac{1}{i+\frac{2}{i+1}} = \frac{1}{i+\frac{2(1-i)}{i+1}} = \frac{1+i}{1-i} = \frac{(1+i)^2}{2} = i$.

P2. Como $z = \frac{(3-2ai)(4+3i)}{25} = \frac{(12+6a)+(9-8a)i}{25}$, entonces $a = \frac{9}{8}$.

P3. Como $-1+7i = (2+bi)(c+3i) = (2c-3b)+(6+bc)i$, entonces $-1 = 2c-3b$ y $7 = 6+bc$. Resolviendo sale $c = 1$ y $b = 1$.

P4. Como $z^2 = (\frac{1}{4}-\frac{3}{4}) - \frac{\sqrt{3}}{2}i = \frac{-1}{2} - \frac{\sqrt{3}}{2}i = -1-z$, obtenemos a). Como $\frac{1}{z} = \frac{\bar{z}}{1} = \frac{-1}{2} - \frac{\sqrt{3}}{2}i = z^2$ obtenemos b).

P5. Tenemos $|i| = 1$, $|1+i| = \sqrt{2}$, $|1-i| = \sqrt{2}$, $|-1-i| = \sqrt{2}$ y $|-1+i| = \sqrt{2}$ y $\arg(i) = \pi/2$, $\arg(1+i) = \pi/4$, $\arg(1-i) = 7\pi/4$, $\arg(-1-i) = 5\pi/4$, $\arg(-1+i) = 3\pi/4$.

P6. Tenemos que $e^{\frac{2\pi}{6}i} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$. Consideremos los números complejos

$$\begin{array}{l} 1+i, \quad e^{\frac{2\pi}{6}}(1+i) = (\frac{1}{2} + \frac{\sqrt{3}}{2}i)(1+i) = \frac{1-\sqrt{3}}{2} + \frac{1+\sqrt{3}}{2}i, \\ e^{\frac{4\pi}{6}}(1+i) = (-\frac{1}{2} + \frac{\sqrt{3}}{2}i)(1+i)) = \frac{-1-\sqrt{3}}{2} + \frac{-1+\sqrt{3}}{2}i \\ -1(1+i) = -1-i, \quad e^{\frac{8\pi}{6}}(1+i) = (-\frac{1}{2} - \frac{\sqrt{3}}{2}i)(1+i)) = \frac{-1+\sqrt{3}}{2} + \frac{-1-\sqrt{3}}{2}i, \\ e^{\frac{10\pi}{6}}(1+i) = (\frac{1}{2} - \frac{\sqrt{3}}{2}i)(1+i)) = \frac{1+\sqrt{3}}{2} + \frac{1-\sqrt{3}}{2}i. \end{array}$$

Las coordenadas de los vértices son

$$\begin{array}{l} (1,1), \ (\frac{1+\sqrt{3}}{2}, \frac{1-\sqrt{3}}{2}), \ (\frac{-1-\sqrt{3}}{2}, \frac{-1-\sqrt{3}}{2}), \\ (-1,-1), \ (\frac{-1+\sqrt{3}}{2}, \frac{-1-\sqrt{3}}{2}), \ (\frac{1+\sqrt{3}}{2}, \frac{1-\sqrt{3}}{2}). \end{array}$$

P7. $|-1+i| = \sqrt{2}$ y $\arg(-1+i) = \frac{3\pi}{4}$ (porque $\arctan(-1) = -\pi/4$), luego $-1+i = \sqrt{2}_{\frac{3\pi}{4}}$; $|3i| = 3$ y $\arg(3i) = \frac{\pi}{2}$, luego $3i = 3_{\frac{\pi}{2}}$; $-5 = 5_{\pi}$; $|2\sqrt{3}-2i| = 4$ y $\arg(2\sqrt{3}-2i) = \frac{11\pi}{6}$ (porque $\arctan(\frac{-1}{\sqrt{3}}) = -\pi/6$), luego $2\sqrt{3}-2i = 4_{\frac{11\pi}{6}}$.

P8. Observemos que $e^{i7\pi} = e^{i\pi} = -1$, $2e^{\frac{i\pi}{4}} = \frac{2}{\sqrt{2}} + \frac{2}{\sqrt{2}}i = \sqrt{2} + \sqrt{2}i$, $6e^{\frac{i\pi}{6}} = 6(\frac{\sqrt{3}}{2} + \frac{1}{2}i) = 3\sqrt{3}+3i$, $e^{\frac{i\pi}{4}} - e^{\frac{-i\pi}{4}} = 2\cos\frac{\pi}{4} = \sqrt{2}$, $\frac{1-e^{\frac{i\pi}{2}}}{1+e^{\frac{i\pi}{2}}} = \frac{1-i}{1+i} = \frac{(1-i)^2}{2} = -i$.

P9. Tenemos que $1_{\frac{5\pi}{6}} \cdot 5_{\frac{\pi}{6}} = 5_{\pi} = -5$, $(2_{\frac{\pi}{6}})^3 = 8_{\frac{\pi}{2}} = 8i$, $(2_{\frac{5\pi}{12}})^3 = 8_{\frac{15\pi}{12}} = 8_{\frac{5\pi}{4}} = -4\sqrt{2}-4\sqrt{2}i$, $(2_{\frac{3\pi}{2}})^3 = 8_{\frac{9\pi}{2}} = 8_{\frac{\pi}{2}} = 8i$.

P10. Observemos que $(e^{\frac{3\pi}{4}})^7 = e^{\frac{21\pi}{4}} = 1_{\frac{5\pi}{4}}$ que está en el tercer cuadrante. $(e^{\frac{2\pi}{7}})^{11} = e^{\frac{22\pi}{7}} = 1_{\frac{8\pi}{7}}$ que está en el tercer cuadrante. $(1_{\frac{2\pi}{3}})^{10} = 1_{\frac{20\pi}{3}} = 1_{\frac{2\pi}{3}}$ que está en el segundo cuadrante. $(2_{\frac{5\pi}{6}})^8 = (2^8)_{\frac{40\pi}{6}} = (2^8)_{\frac{4\pi}{6}}$ que está en el segundo cuadrante.

P11. Tenemos que $i^{37} = i^{9\cdot 4+1} = i$, $i^{126} = i^{31\cdot 4+2} = i^2 = -1$, $z = i^{2009} - i^{2010} = i^1 - i^2 = i+1$.

P12. Tenemos que $\sqrt{-4} = \pm 2\cdot i$, $\sqrt{-2} = \pm\sqrt{2}i$, $\sqrt{3i} = \sqrt{3}\cdot\sqrt{1_{\frac{\pi}{2}}} = \pm\sqrt{3}\cdot 1_{\frac{\pi}{4}} = \pm(\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{\sqrt{2}}i)$, $\sqrt{-4i} = \pm 2\sqrt{1_{\frac{3\pi}{2}}} = \pm 21_{\frac{3\pi}{4}} = \pm(-\sqrt{2}+\sqrt{2}i)$, $\sqrt{1+i} = \sqrt{(\sqrt{2})_{\frac{\pi}{4}}} = \pm\sqrt[4]{2}1_{\frac{\pi}{8}} = \pm\sqrt[4]{2}(\cos\frac{\pi}{8} + \operatorname{sen}\frac{\pi}{8})$.

P13. El número complejo $\sqrt{a^2+b^2}$ es de argumento cero y tiene el mismo módulo que

$a+bi$. Entonces, el argumento de $\sqrt{a^2+b^2}+a+bi$ es igual a $\frac{\arg(a+bi)}{2}$. Luego,

$$\sqrt{a+bi} = \frac{\sqrt{a^2+b^2}+a+bi}{|\sqrt{a^2+b^2}+a+bi|}\cdot\sqrt[4]{a^2+b^2} = \frac{\sqrt{a^2+b^2}+a+bi}{\sqrt{2(a^2+b^2+a\sqrt{a^2+b^2})}}\cdot\sqrt[4]{a^2+b^2}$$
$$= \frac{\sqrt{a^2+b^2}+a+bi}{\sqrt{2}\cdot\sqrt{\sqrt{a^2+b^2}+a}} = \frac{\sqrt{\sqrt{a^2+b^2}+a}}{\sqrt{2}} + \frac{bi}{\sqrt{2}\sqrt{\sqrt{a^2+b^2}+a}}.$$

P14. Las soluciones de la primera son $\frac{\sqrt{2}\pm\sqrt{-2}}{2} = \frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}$ y las de la segunda son $\frac{2\pm\sqrt{-4}}{2} = 1\pm i$.

P15. Como $|z_1| = \sqrt{2}$ y $\arg(z_1) = \frac{\pi}{3}$, $z_1 = \sqrt{2}_{\frac{\pi}{3}}$. Como $|z_2| = \sqrt{2}$ y $\arg(z_2) = \frac{\pi}{4}$, $z_2 = \sqrt{2}_{\frac{\pi}{4}}$. Entonces, $z = (\frac{1}{4})_{\frac{100\pi}{3}-\frac{104\pi}{4}} = (\frac{1}{4})_{\frac{4\pi}{3}} = \frac{-1}{8} - \frac{\sqrt{3}i}{8}$: las raíces cuartas son $(\frac{1}{\sqrt{2}})_{\frac{\pi}{3}} = \frac{z_1}{2}$, multiplicada por ± 1 y $\pm i$.

P16. Consideremos el cambio de variable $x = z^3$. Entonces, $P(z) = x^2-2x+2$ y las raíces de este polinomio son $x = 1\pm i = \{\sqrt{2}_{\frac{\pi}{4}}, \sqrt{2}_{\frac{3\pi}{4}}\}$, luego las raíces del polinomio $P(z)$ son $z = \{\sqrt[6]{2}_{\frac{\pi}{12}}, \sqrt[6]{2}_{\frac{3\pi}{12}}\}\cdot\{1, 1_{\frac{2\pi}{3}}, 1_{\frac{4\pi}{3}}\}$, es decir,

$$\{\frac{1+\sqrt{3}}{2\sqrt[3]{2}} + \frac{i(\sqrt{3}-1)}{2\sqrt[3]{2}}, \frac{1+\sqrt{3}}{2\sqrt[3]{2}} - \frac{i(\sqrt{3}-1)}{2\sqrt[3]{2}}\}\cdot\{1, \frac{-1}{2}+\frac{i\sqrt{3}}{2}, \frac{-1}{2}-\frac{i\sqrt{3}}{2}\}.$$

P17. Observemos que $1_{\pi/3} : 1_{\pi/4} = 1_{\pi/12}$. Por otra parte, $1_{\pi/3} : 1_{\pi/4} = \frac{\frac{1}{2}+\frac{\sqrt{3}}{2}}{\frac{1}{\sqrt{2}}+\frac{i}{\sqrt{2}}} = \frac{1+\sqrt{3}}{2\sqrt{2}} - \frac{1-\sqrt{3}}{2\sqrt{2}}i$. Por lo tanto, $\cos(\pi/12) = \frac{1+\sqrt{3}}{2\sqrt{2}}$ y $\operatorname{sen}(\pi/12) = \frac{-1+\sqrt{3}}{2\sqrt{2}}$.

P18. $|1-\sqrt{3}i| = \sqrt{1+3} = 2$, $\arg(1-\sqrt{3}i) = \arctan(-\sqrt{3}) = \frac{-\pi}{3}$. Por tanto, $|z| = \frac{1}{2}$ y $\arg(z) = \frac{\pi}{3}$. Por último, z^n es real cuando $n\cdot\arg(z) = \arg(z^n) \in \mathbb{N}\cdot\pi$, es decir, cuando n sea múltiplo de 3.

P19. Observemos que $\cos x = \frac{e^{ix}+e^{-ix}}{2}$. Luego, $\cos^3 x = (\frac{e^{ix}+e^{-ix}}{2})^3 = \frac{e^{3ix}+3e^{ix}+3e^{-ix}+e^{-3ix}}{8}$ y

$$\int e^x\cdot\cos^3 x\cdot dx = \int \frac{e^{(1+3i)x}+3e^{(1+i)x}+3e^{(1-i)x}+e^{(1-3i)x}}{8}\cdot dx$$
$$= \frac{1}{8}(\frac{e^{(1+3i)x}}{1+3i} + \frac{e^{(1-3i)x}}{1-3i} + \frac{3e^{(1+i)x}}{1+i} + \frac{3e^{(1-i)x}}{1-i})$$
$$= \frac{e^x}{8}\cdot(\frac{\cos(3x)+3\operatorname{sen}(3x)}{5} + 3\cos x + 3\operatorname{sen} x).$$

Capítulo 2

Espacios vectoriales

2.1. Introducción

En este capítulo estudiamos los sistemas de ecuaciones lineales, las aplicaciones lineales, las matrices y las subvariedades afines de $\mathbb{R}^n$.

El método algorítmico principal en el que nos apoyaremos es el método de la cascada de Gauss. Otra herramienta que estudiaremos y usaremos para responder a los distintos problemas y preguntas que vayan apareciendo es el determinante de una matriz cuadrada.

Es fundamental observar que constantemente sumaremos "objetos" y consideraremos combinaciones lineales de ellos. Los conjuntos en los que sus elementos se suman y admiten combinaciones lineales son los espacios vectoriales. Las aplicaciones entre espacios vectoriales que respetan su estructura son las aplicaciones lineales. El marco conceptual en el que todos los problemas, conceptos y algoritmos antes mencionados se entienden y se explican bien es el de los espacios vectoriales.

2.2. Espacio vectorial

Sea $\mathbb{R}$ el cuerpo de los números reales. " Un $\mathbb{R}$-espacio vectorial es un conjunto en el que podemos sumar sus elementos, podemos multiplicar sus elementos por números reales, y estas operaciones cumplen propiedades muy naturales ". Del mismo modo podemos definir los $\mathbb{Q}$-espacios vectoriales y los $\mathbb{C}$-espacios vectoriales.

1. Definición : Un $\mathbb{R}$-*espacio vectorial* es un conjunto, E, dotado de dos operaciones, una llamada suma $E \times E \to E$ y se escribe $(e, e') \mapsto e + e'$, y otra llamada producto por escalares $\mathbb{R} \times E \to E$ y se escribe $(\lambda, e) \mapsto \lambda \cdot e$, que cumplen:

1. $(E,+)$ es un grupo abeliano, es decir,

   a) $e+(e'+e'')=(e+e')+e''$, para todo $e,e',e''\in E$.

   b) Existe un elemento en E que denotamos por 0 y decimos que es el vector cero que cumple que $0+e=e+0=e$, para todo $e\in E$.

   c) Para cada $e\in E$ existe otro elemento que denotamos $-e$ tal que $e+(-e)=0$.

   d) $e+e'=e'+e$, para todo $e,e'\in E$.

2. $\lambda\cdot(e+v)=\lambda\cdot e+\lambda\cdot v$, $\forall\lambda\in\mathbb{R}$, $e,v\in E$

3. $(\lambda+\mu)\cdot e=\lambda\cdot e+\mu\cdot e$, $\forall\lambda,\mu\in\mathbb{R}$, $e\in E$

4. $(\lambda\cdot\mu)\cdot e=\lambda\cdot(\mu\cdot e)$, $\forall\lambda,\mu\in\mathbb{R}$, $e\in E$

5. $1\cdot e=e$, $\forall e\in E$.

Los elementos de un $\mathbb{R}$-espacio vectorial E se denominan vectores y los elementos de $\mathbb{R}$ se denominan escalares.

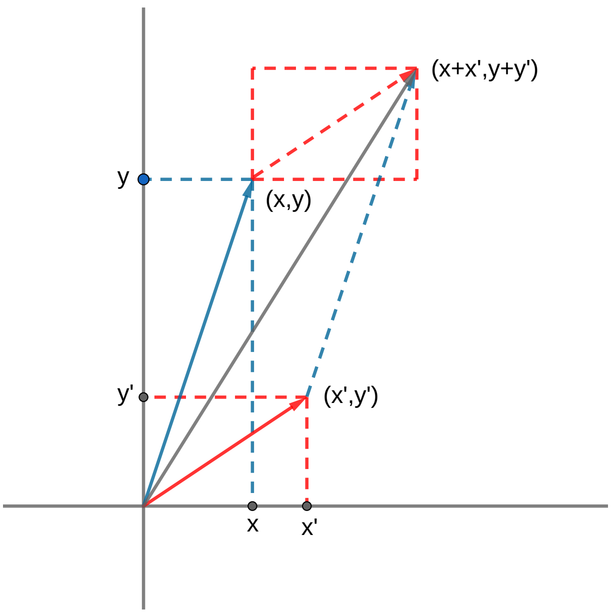
2. Ejemplos: 1. $\mathbb{R}^2 := \{(x,y),$ para todo $x,y\in\mathbb{R}\}$, con la suma de vectores

$$(x,y)+(x',y'):=(x+x',y+y')$$

y el producto de escalares por vectores

$$\lambda\cdot(x,y):=(\lambda\cdot x,\lambda\cdot y)$$

es un $\mathbb{R}$-espacio vectorial.



Interpretación geométrica de la suma de dos vectores de $\mathbb{R}^2$: La suma de dos vectores es la diagonal del paralelogramo formado por los dos vectores y sus trasladados.

2. $\mathbb{R}^3 := \{(x,y,z)$, variando $x,y,z \in \mathbb{R}\}$, que es el espacio en el que pensamos que vivimos, es un ejemplo de $\mathbb{R}$-espacio vectorial, con la suma de vectores

$$(x,y,z)+(x',y',z') := (x+x', y+y', z+z')$$

y el producto de escalares por vectores

$$\lambda \cdot (x,y,z) := (\lambda \cdot x, \lambda \cdot y, \lambda \cdot z)$$

es un $\mathbb{R}$-espacio vectorial.

3. Sea $\mathbb{R}[x] = \{a_0 + a_1 \cdot x + \cdots + a_n \cdot x^n$, variando $a_0, a_1, \ldots, a_n \in \mathbb{R}$ y variando $n \in \mathbb{N}\}$ el conjunto de los polinomios en una variable con coeficientes reales. Dados dos polinomios $p(x) = a_0 + a_1 \cdot x + \cdots + a_n \cdot x^n$, $q(x) = b_0 + b_1 \cdot x + \cdots + b_m \cdot x^m$ (supongamos $n \geq m$) se define $p(x)+q(x)$ como sigue

$$p(x)+q(x) := (a_0+b_0)+(a_1+b_1)\cdot x + \cdots + (a_m+b_m)\cdot x^m + a_{m+1}\cdot x^{m+1} + \cdots + a_n \cdot x^n$$

y se define $\lambda \cdot p(x)$ (con $\lambda \in \mathbb{R}$) como sigue

$$\lambda \cdot p(x) := (\lambda \cdot a_0) + (\lambda \cdot a_1)\cdot x + \cdots + (\lambda \cdot a_n)\cdot x^n.$$

Tenemos que $\mathbb{R}[x]$ es un $\mathbb{R}$-espacio vectorial.

4. Sea $\mathrm{Aplic}(\mathbb{R},\mathbb{R}) := \{$aplicaciones $f : \mathbb{R} \to \mathbb{R}\} = \{e^x, \cos x, \operatorname{sen} x, x^2+1, etc.\}$. Dadas $f, g \in \mathrm{Aplic}(\mathbb{R},\mathbb{R})$ se define $f+g$ como la siguiente aplicación

$$(f+g)(x) := f(x)+g(x)$$

y dado un escalar $\lambda \in \mathbb{R}$ se define $\lambda \cdot f$ como la siguiente aplicación

$$(\lambda \cdot f)(x) := \lambda \cdot f(x).$$

$\mathrm{Aplic}(\mathbb{R},\mathbb{R})$ es un $\mathbb{R}$-espacio vectorial.

En los espacios vectoriales se cumplen las siguientes propiedades:

1. $0 \cdot e = (0+0)\cdot e = 0 \cdot e + 0 \cdot e$ y por tanto $0 \cdot e = 0$.

2. Si $e + e' = 0$, sumando $-e$, obtenemos que $e' = -e$.

3. Como $0 = (1-1)\cdot e = e + (-1)\cdot e$, tenemos que $(-1)\cdot e = -e$.

2.3. Subespacio vect. generado por unos vectores

Sea E un $\mathbb{R}$-espacio vectorial. Decimos que un subconjunto $F \subset E$ no vacío es un *subespacio vectorial* de E, si

$$\begin{aligned} f + f' \in F, &\quad \text{para todo } f, f' \in F. \\ \lambda \cdot f \in F, &\quad \text{para todo } f \in F \text{ y todo } \lambda \in \mathbb{R}. \end{aligned}$$

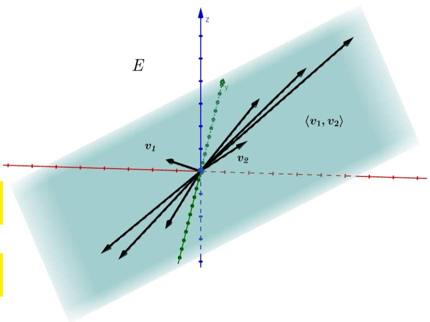
El subespacio vectorial, F, con la suma y producto por escalares (que tenemos en E) es un espacio vectorial.

1. Ejemplo: $F := \{(x,y) \in \mathbb{R}^2 : x + y = 0\} = \{(1,-1),(3,-3),(-\pi,\pi),\dots\}$ es un subespacio vectorial de $\mathbb{R}^2$. En efecto, si tomo dos vectores cualesquiera de F, $(a,-a)$ y $(b,-b)$, tenemos que $(a,-a)+(b,-b) = (a+b,-(a+b)) \in F$ y $\lambda \cdot (a,-a) = (\lambda a, -\lambda a) \in F$.

2. Definición: Sean $\{v_1,\dots,v_n\}$ unos cuantos vectores de E. Llamamos subespacio vectorial de E generado por $\{v_1,\dots,v_n\}$ y lo denotamos $\langle v_1,\dots,v_n\rangle$ a

$$\langle v_1,\dots,v_n\rangle := \{\lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n \in E,\ \forall\, \lambda_1,\dots,\lambda_n \in \mathbb{R}\}.$$

Se dice que $\lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n$ es una combinación lineal de los vectores $v_1,\dots,v_n$.



Evidentemente, el subespacio vectorial generado por unos cuantos vectores $v_1,\dots,v_n$ no depende del orden en el que se escriban estos vectores.

3. Ejercicio: ¿$(1,2,3) \in \langle (1,1,1),(0,1,1)\rangle$?

4. Proposición: *Sea E un $\mathbb{R}$-espacio vectorial y $e_1,\dots,e_n,v_1,\dots,v_m \in E$. Si $v_1,\dots,v_m \in \langle e_1,\dots,e_n\rangle$, entonces*

$$\langle v_1,\dots,v_m\rangle \subseteq \langle e_1,\dots,e_n\rangle.$$

Luego, $\langle v_1,\dots,v_m\rangle = \langle e_1,\dots,e_n\rangle$ si y solo si $v_1,\dots,v_m \in \langle e_1,\dots,e_n\rangle$ y $e_1,\dots,e_n \in \langle v_1,\dots,v_m\rangle$.

Demostración. Tenemos[1] que $v_i = \sum_{j=1}^n \lambda_{ji} \cdot e_j$, para cada i, para ciertos escalares $\lambda_{ji} \in \mathbb{R}$. Dado $v = \sum_{i=1}^m \mu_i \cdot v_i \in \langle v_1,\dots,v_m\rangle$, tendremos que

$$v = \sum_{i=1}^m \mu_i v_i = \sum_{i=1}^m \mu_i (\sum_{j=1}^n \lambda_{ji} e_j) = \sum_{i=1}^m (\sum_{j=1}^n \mu_i \lambda_{ji} e_j) = \sum_{j=1}^n (\sum_{i=1}^m \mu_i \lambda_{ji}) e_j \in \langle e_1,\dots,e_n\rangle.$$

□

[1] $\sum_{j=1}^n \lambda_{ji} \cdot e_j$ es un modo abreviado de escribir $\lambda_{1i} \cdot e_1 + \lambda_{2i} \cdot e_2 + \lambda_{3i} \cdot e_3 + \dots + \lambda_{ni} \cdot e_n$.

5. Corolario: *Sean* $e_1,\dots,e_n,v_1,\dots,v_m,w \in E$. *Si* $\langle e_1,\dots,e_n\rangle = \langle v_1,\dots,v_m\rangle$ *entonces se cumple que* $\langle e_1,\dots,e_n,w\rangle = \langle v_1,\dots,v_m,w\rangle$.

6. Corolario: *Sea* E *un espacio vectorial y* $e,e_1,\dots,e_n \in E$. *Entonces,* $e \in \langle e_1,\dots,e_n\rangle$ *si y solo si* $\langle e_1,\dots,e_n,e\rangle = \langle e_1,\dots,e_n\rangle$.

7. Corolario: *Sea* E *un* $\mathbb{R}$*-espacio vectorial* $e_1,\dots,e_n \in E$. *Se cumple que*

$$\langle e_1,e_2,e_3,\dots,e_n\rangle = \langle e_1,e_2+\lambda\cdot e_1,e_3,\dots,e_n\rangle$$
$$\langle e_1,e_2,e_3,\dots,e_n\rangle = \langle \lambda_1 e_1,e_2,e_3,\dots,e_n\rangle, \quad (\lambda_1 \neq 0).$$

2.4. Bases

1. Definiciones: Se dice que los vectores $e_1,\dots,e_n$ de un espacio vectorial E son un sistema generador del espacio vectorial E si $\langle e_1,\dots,e_n\rangle = E$, es decir, si todo vector de E es combinación lineal de $e_1,\dots,e_n$. Se dice que son linealmente independientes si

$$\lambda_1\cdot e_1+\cdots+\lambda_n\cdot e_n \neq 0$$

siempre que algún $\lambda_i \neq 0$. Se dice que son una base de E si generan E y son linealmente independientes.

2. Ejercicio: Probad que los vectores "escalonados" de $\mathbb{R}^4$

$$\begin{array}{c}(1,2,3,4)\\(0,3,4,5)\\(0,0,6,8)\end{array}$$

son linealmente independientes.

3. Ejercicio: Probad que los vectores "escalonados" de $\mathbb{R}^4$

$$\begin{array}{c}(2,2,2,2)\\(0,0,4,5)\\(0,0,0,1)\end{array}$$

son linealmente independientes.

4. Ejercicio: Probad que los vectores $B = \{(1,0,\dots,0),(0,1,\dots,0),\dots,(0,0.\dots,1)\}$ forman una base de $\mathbb{R}^n$ (denominada "base estándar" de $\mathbb{R}^n$).

Los espacios vectoriales que consideremos estarán generados por un número finito de vectores (por sencillez), salvo que se diga lo contrario.

5. Teorema de la base: *Todo espacio vectorial $E \neq 0$ contiene alguna base. Todas las bases de E tienen el mismo número de vectores, tal número se dice que es la dimensión de E y se denota* $\dim_{\mathbb{R}} E$.

Demostración. Escribamos $E = \langle e_1, \dots, e_n \rangle$. Veamos que unos cuantos de estos e_i forman una base de E. Vamos a seguir los siguientes pasos:

Paso 1: Sea $\tilde{e}_1 := e_1$.

Paso 2: Sea $\tilde{e}_2 := \begin{cases} 0 \text{ si } e_2 \in \langle e_1 \rangle \\ e_2 \text{ si } e_2 \notin \langle e_1 \rangle \end{cases}$

Paso 3: Sea $\tilde{e}_3 := \begin{cases} 0 \text{ si } e_3 \in \langle e_1, e_2 \rangle \\ e_3 \text{ si } e_3 \notin \langle e_1, e_2 \rangle \end{cases}$

Etc. Veamos que $B = \{\tilde{e}_i : \tilde{e}_i \neq 0\}$ es una base de E: $\langle \tilde{e}_1 \rangle = \langle e_1 \rangle$, entonces $\langle \tilde{e}_1, \tilde{e}_2 \rangle = \langle e_1, \tilde{e}_2 \rangle = \langle e_1, e_2 \rangle$, luego $\langle \tilde{e}_1, \tilde{e}_2, \tilde{e}_3 \rangle = \langle e_1, e_2, \tilde{e}_3 \rangle = \langle e_1, e_2, e_3 \rangle$. Hasta que obtenemos que $\langle \tilde{e}_1, \dots, \tilde{e}_n \rangle = \langle e_1, \dots, , e_n \rangle$. Por tanto, B es un sistema generador de E. Veamos que son linealmente independientes: si $\sum_{\tilde{e}_i \neq 0} \lambda_i \cdot \tilde{e}_i = 0$, sea i máximo tal que $\lambda_i \neq 0$, entonces $\tilde{e}_i \in \langle \tilde{e}_1, \dots, \tilde{e}_{i-1} \rangle = \langle e_1, \dots, e_{i-1} \rangle$, lo que es contradictorio. Luego, todos los $\lambda_i = 0$.

Veamos que todas las bases tienen el mismo número de vectores.

Sean $\{e_1, \dots, e_n\}$ linealmente independientes y $\{v_1, \dots, v_m\}$ un sistema generador de E.

1. Veamos que $\{e_1, v_2, \dots, v_m\}$ es un sistema generador de E, reordenando si es necesario los vectores $\{v_1, \dots, v_m\}$: $e_1 = \lambda_1 v_1 + \cdots + \lambda_m v_m$, reordenando los v_i, podemos suponer que $\lambda_1 \neq 0$. Entonces, $v_1 \in \langle e_1, v_2, \dots, v_m \rangle$. Por tanto, $\langle e_1, v_2, \dots, v_m \rangle = \langle e_1, v_1, v_2, \dots, v_m \rangle = E$.

2. Veamos que $\{e_1, e_2, v_3, \dots, v_m\}$ es un sistema generador de E reordenando si es necesario los vectores $\{v_2, \dots, v_m\}$: $e_2 = \lambda_1 e_1 + \lambda_2 v_2 \cdots + \lambda_m v_m$. No pueden ser $\lambda_2 = \cdots = \lambda_n = 0$, porque e_1, e_2 son l.i. Reordenando los v_i, podemos suponer que $\lambda_2 \neq 0$. Entonces, $v_2 \in \langle e_1, e_2, v_3, \dots, v_m \rangle$. Por tanto, $\langle e_1, e_2, v_3, \dots, v_m \rangle = \langle e_1, e_2, v_2, \dots, v_m \rangle = E$.

3. Así sucesivamente $\{e_1, \dots, e_i, v_{i+1}, \dots, v_m\}$ son un sistema generador de E.

Ahora sean $\{e_1, \dots, e_n\}$ y $\{v_1, \dots, v_m\}$ dos bases de E. Si $n > m$, entonces tomando $i = m$, tenemos que $\{e_1, \dots, e_m\}$ es un sistema generador de E, luego e_{m+1} es combinación lineal de ellos, contradicción. Luego, $n \leq m$. Cambiando el papel de los e_i por el de los v_j, tendremos igualmente que $m \leq n$, luego $n = m$.

□

Si $e_1, \dots, e_i \in E$ son linealmente independientes y consideramos un sistema generador $v_1, \dots, v_m$ de E, entonces $\{e_1, \dots, e_i, v_1, \dots, v_m\}$ es también un sistema generador de E. En la demostración del teorema de la base hemos obtenido una base de E que va a estar formada por los vectores $e_1,\dots,e_i$ y algunos de los vectores de $v_1,\dots,v_m$. Luego si $E' \underset{\neq}{\subset} E$ es un subespacio vectorial, toda base de E' la podemos ampliar a una base de E, y por tanto $\dim_{\mathbb{R}} E' < \dim_{\mathbb{R}} E$.

2.5. Método de la cascada de Gauss

2.5.1. Cálculo de bases

1. Ejercicio: Sea $F = \langle (1,1,1),(1,2,3),(3,5,7)\rangle \subseteq \mathbb{R}^3$ ¿Son $(1,1,1),(1,2,3)$ y $(3,5,7)$ linealmente independientes? Calculad una base de F y ampliad esta base a una base de $\mathbb{R}^3$.

Resolución: Usando el corolario 2.3.7 vamos a obtener un conjunto de vectores "escalonados" como los de los ejercicios 2.4.2 y 2.4.3:

$$F = \left\langle \begin{matrix} (1,1,1) \\ (1,2,3) \\ (3,5,7) \end{matrix} \right\rangle \overset{F_2-F_1}{=} \left\langle \begin{matrix} (1,1,1) \\ (0,1,2) \\ (3,5,7) \end{matrix} \right\rangle \overset{F_3-3F_1}{=} \left\langle \begin{matrix} (1,1,1) \\ (0,1,2) \\ (0,2,4) \end{matrix} \right\rangle \overset{F_3-2F_2}{=} \left\langle \begin{matrix} (1,1,1) \\ (0,1,2) \\ (0,0,0) \end{matrix} \right\rangle = \left\langle \begin{matrix} (1,1,1) \\ (0,1,2) \end{matrix} \right\rangle$$

Tenemos que una base de F es $\{(1,1,1),(0,1,2)\}$. Por tanto, $\dim F = 2$ y los vectores $(1,1,1),(1,2,3)$ y $(3,5,7)$ no pueden ser linealmente independientes, pues si lo fueran formarían una base de F y la dimensión de F sería 3.

Una base ampliada de $\{(1,1,1),(0,1,2)\}$ en $\mathbb{R}^3$ es $\{(1,1,1),(0,1,2),(0,0,1)\}$.

Veamos si $(1,4,5) \in F$: Veamos si F coincide con $\langle (1,1,1),(1,2,3),(3,5,7),(1,4,5)\rangle = \langle (1,1,1),(0,1,2),(1,4,5)\rangle = \langle (1,1,1),(0,1,2),(0,3,4)\rangle = \langle (1,1,1),(0,1,2),(0,0,-2)\rangle$, que tiene dimensión 3, luego no coinciden y $(1,4,5) \notin F$.

2.5.2. Resolución de los sistemas de ecuaciones lineales

Resolvamos el sistema de ecuaciones lineales

$$\begin{aligned} x+y+z &= 3 \\ x+2y+3z &= 6 \\ 3x+5y+7z &= 15 \end{aligned}$$

Despejando el valor de x en la primera ecuación, obtenemos que $\boxed{x = 3-y-z}$. Sustituyamos x por su valor en las demás ecuaciones. Obtenemos

$$\begin{aligned} (3-y-z)+2y+3z = 6 &\equiv y+2z = 3 \\ 3(3-y-z)+5y+7z = 15 &\equiv 2y+4z = 6 \end{aligned}$$

Despejando el valor de y en la primera ecuación, obtenemos que $\boxed{y = 3-2z}$. Sustituyamos y por su valor en las demás ecuaciones. Obtenemos

$$2(3-2z)+4z = 6 \;\equiv\; 0 = 0.$$

En conclusión, $z = z$ (z puede tomar cualquier valor), de la ecuación encuadrada anterior $y = 3-2z$, y ahora de la ecuación encuadrada anterior a ésta $x = 3-y-z = 3-(3-2z)-z = z$. Variando z obtenemos todas las soluciones, por ejemplo, para $z = 1$ obtenemos la solución $x = 1$, $y = 1$, $z = 1$.

Hagamos esencialmente las mismas cuentas, usando el método de la cascada de Gauss. Evidentemente, el conjunto de soluciones del sistema de ecuaciones anterior no depende del orden el que se escriban cada una de las ecuaciones. Tampoco varía si una ecuación la multiplicamos por un escalar no nulo y no varía si a una ecuación le sumamos otra ecuación multiplicada por un escalar. Juguemos con estas ecuaciones como jugaríamos con los vectores $(1,1,1,3),(1,2,3,6)$, $(3,5,7,15)$ en el ejercicio anterior. Obtenemos

$$\left\langle \begin{matrix} (1,1,1,3) \\ (1,2,3,6) \\ (3,5,7,15) \end{matrix} \right\rangle \overset{F_2-F_1}{=} \left\langle \begin{matrix} (1,1,1,3) \\ (0,1,2,3) \\ (3,5,7,15) \end{matrix} \right\rangle \overset{F_3-3F_1}{=} \left\langle \begin{matrix} (1,1,1,3) \\ (0,1,2,3) \\ (0,2,4,6) \end{matrix} \right\rangle \overset{F_3-F_2}{=} \left\langle \begin{matrix} (1,1,1,3) \\ (0,1,2,3) \\ (0,0,0,0) \end{matrix} \right\rangle.$$

El sistema de ecuaciones lineales anterior tiene las mismas soluciones que el sistema con dos ecuaciones

$$\begin{aligned} \underline{x}+y+z &= 3 \\ \underline{y}+2z &= 3 \end{aligned}$$

Pasemos la variable z al otro lado de la igualdad (tratémosla como si fuese un número real) y obtenemos

$$\begin{aligned} x+y &= 3-z \\ y &= 3-2z \end{aligned}$$

Resolvamos este sistema de ecuaciones: $y = 3-2z$, $x = 3-z-y = 3-z-3+2z = z$ (y $z = z$). Variando z obtenemos todas las soluciones, por ejemplo, para $z = 5$ obtenemos la solución $x = 5$, $y = -7$, $z = 5$.

Resolvamos el sistema de ecuaciones lineales

$$\begin{aligned} x+y+z+t &= 3 \\ x+y+z-t &= 6 \\ x+2y+z+2t &= 0 \\ 2x+3y+2z+t &= 6 \end{aligned}$$

Tenemos que

$$\begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 1 & 1 & 1 & -1 & 6 \\ 1 & 2 & 1 & 2 & 0 \\ 2 & 3 & 2 & 1 & 6 \end{array} \overset{\substack{F_2-F_1\\F_3-F_1\\F_4-2F_1}}{\equiv} \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 1 & 0 & -1 & 0 \end{array} \overset{F_2\times F_3}{\equiv} \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 1 & 0 & -1 & 0 \end{array}$$

$$\overset{F_4-F_2}{\equiv} \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -2 & 3 \end{array} \equiv \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -2 & 3 \end{array}$$

El sistema de ecuaciones lineales anterior tiene las mismas soluciones que el sistema de tres ecuaciones

$$\begin{array}{rcr} \underline{x}+y+z+t & = & 3 \\ \underline{y}+0z+t & = & -3 \\ -2\underline{t} & = & 3 \end{array}$$

Pasemos la variable z al otro lado de la igualdad y obtenemos

$$\begin{array}{rcl} x+y+t & = & 3-z \\ y+t & = & -3 \\ -2t & = & 3 \end{array}$$

Luego $t = \frac{-3}{2}$, $y = -3-t = \frac{-3}{2}$, $x = 3-t-y-z = 6-z$ (y $z = z$).

Resolvamos el sistema de ecuaciones lineales

$$\begin{array}{rcl} x+y+z+t & = & 3 \\ x+y+z-t & = & 6 \\ x+2y+z+2t & = & 0 \\ 2x+3y+2z+t & = & 5 \end{array}$$

Tenemos que

$$\begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 1 & 1 & 1 & -1 & 6 \\ 1 & 2 & 1 & 2 & 0 \\ 2 & 3 & 2 & 1 & 5 \end{array} \equiv \cdots \equiv \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -2 & 4 \end{array} \equiv \begin{array}{ccccc} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{array}$$

El sistema de ecuaciones lineales anterior tiene las mismas soluciones que el sistema de ecuaciones

$$\begin{array}{rcr} x+y+z+t & = & 3 \\ y+t & = & -3 \\ -2t & = & 3 \\ 0 & = & 1 \end{array}$$

que no tiene ninguna solución ("es incompatible").

2. Ejemplo: (Tomado de [8]) Un líquido volátil (en un recipiente cerrado a cierta temperatura constante) se va evaporando hasta que se llega a un estado de saturación en el que el número de moléculas que se escapan del líquido es igual a las que vuelven al líquido. Denotemos la presión de vapor del benceno (en saturación) p_A y la del bromobenceno p_B. Si mezclamos los dos líquidos entonces la ley de Raoult nos dice que la presión de vapor de este líquido es $p = x_A \cdot p_A + x_B \cdot p_B$, donde x_A es la proporción de moléculas de benceno en el líquido y x_B es la proporción de moléculas de bromobenceno en el líquido (luego $x_A + x_B = 1$). Si p es igual a 72 kPa cuando $x_A = 0'3$ y $x_B = 0'7$ y p es igual a 80 kPa cuando $x_A = 0'5$ y $x_B = 0'5$, calculemos p_A y p_B. Tenemos que resolver el sistema de ecuaciones

$$\begin{array}{rcl} 0'3p_A + 0'7p_B & = & 72 \\ 0'5p_A + 0'5p_B & = & 80 \end{array}$$

Multipliquemos por 2 la segunda ecuación y coloquémosla en primer lugar

$$\begin{array}{ccc} 1 & 1 & 160 \\ 0'3 & 0'7 & 72 \end{array} \overset{F_2 - 0'3F_1}{\equiv} \begin{array}{ccc} 1 & 1 & 160 \\ 0 & 0'4 & 24 \end{array}$$

Es decir, equivale al sistema

$$\begin{array}{rcl} p_A + p_B & = & 160 \\ 0'4p_B & = & 24 \end{array}$$

Luego, $p_B = 60$ y $p_A = 100$.

3. Ejemplo: (Tomado de [8]) Tenemos dos muestras que contienen naftalina y antraceno. La primera contiene $0'16$ moles de naftalina y $0'06$ moles de antraceno y libera al arder $1197'3$ kJ. La segunda contiene $0'12$ moles de naftalina y $0'03$ moles de antraceno y libera al arder $830'7$ kJ. Calculemos cuántos x kJ libera al arder un mol de naftalina y cuántos y kJ libera un mol de antraceno. Tenemos que resolver el sistema de ecuaciones

$$\begin{array}{rcl} 0'16x + 0'06y & = & 1197'3 \\ 0'12x + 0'03y & = & 830'7 \end{array}$$

Multipliquemos por 100 las ecuaciones para eliminar los decimales,

$$\begin{array}{ccc} 16 & 6 & 119730 \\ 12 & 3 & 83070 \end{array} \overset{F_2 - \frac{3}{4}F_1}{\equiv} \begin{array}{ccc} 16 & 6 & 119730 \\ 0 & -3/2 & -13455/2 \end{array}$$

Luego, $y = 4485$ y $x = 5801'25$.

4. Ejercicio : Tres grupos de amigos toman unas bebidas en un bar. El primer grupo por tres cañas, dos cocacolas y tres tónicas pagó quince euros. El segundo por dos cañas, tres cocacolas y cuatro tónicas pagó veinte euros. El tercer grupo por cuatro cañas, dos cocacolas y dos tónicas pagó catorce euros ¿Cuánto cuesta cada caña, cocacola y tónica?

2.5.3. Programación lineal

5. Ejercicio : (Tomado de [2]) Una fábrica manufactura dos artículos y para producirlos usa tres máquinas. La primera puede usarse a lo sumo 80 horas, la segunda a lo más 50 horas y la tercera un máximo de 90 horas. El primer artículo usa dos horas en la primera máquina y una hora en cada una de las otras dos, mientras que el segundo requiere una hora de la primera, una hora de la segunda máquina y y dos horas de la tercera. La fábrica obtiene una ganancia de 500 euros por cada primer artículo producido y 700 euros por cada segundo artículo producido. Tiene unos gastos fijos de 10.000 euros ¿Cuántos unidades de cada artículo deben fabricarse para maximizar la ganancia?

Resolución: Sea x_1 y x_2 el número de unidades fabricadas del primer y segundo artículo, respectvamente. Sea $G = 500x_1 + 700x_2 - 10^4$ la ganancia obtenida. Puesto que la primera máquina se puede utilizar a lo más 80 horas, cada unidad del primer artículo requiere 2 horas y cada unidad del segundo articulo 1 hora, de esta primera máquina, se debe tener $2x_1 + x_2 \leq 80$. Igualmente $x_1 + x_2 \leq 50$ y $x_1 + 2x_2 \leq 90$. Por tanto, tenemos que maximizar $G = 500x_1 + 700x_2 - 10^4$ sujeto a las restricciones

$$\begin{array}{rcl} 2x_1 + x_2 & \leq & 80 \\ x_1 + x_2 & \leq & 50 \\ x_1 + 2x_2 & \leq & 90 \\ x_1, x_2 & \geq & 0 \end{array}$$

Este problema es equivalente a maximizar $G = 500x_1 + 700x_2 - 10^4$ sujeto a las restricciones

$$\begin{array}{rcrcrcrcrcrcl} 2x_1 & + & x_2 & + & x_3 & & & & & & & = & 80 \\ x_1 & + & x_2 & + & & & x_4 & & & & & = & 50 \\ x_1 & + & 2x_2 & + & & & & & & & x_5 & = & 90 \\ & & & & x_1, & & x_2, & & x_3, & & x_4, \; x_5 & \geq & 0 \end{array}$$

(a las variables x_3, x_4 y x_5 se las llama variables de holgura).

Como 80, 50 y 90 son mayores o iguales que cero, si G hubiese sido igual a la función $-100x_1 - 200x_2 - 3$, entonces G, sujeta a las restricciones, tomaría su valor máximo en $x_1 = x_2 = 0$ (y $x_3 = 80$, $x_4 = 50$, $x_5 = 90$). El conjunto de soluciones del sistema de ecuaciones no varía si a una ecuación la multiplico por un escalar no nulo o le sumo λ veces otra de las ecuaciones. Si defino $G' = G - \lambda\cdot(2x_1 + x_2 + x_3 - 80)$ el valor máximo de G es igual al valor máximo de G' sujeto a las mismas restricciones (lo mismo decimos si en vez de tomar la primera ecuación consideramos la segunda o la tercera).

Escribamos de modo sucinto el sistema de ecuaciones y "añadamos" la función G

$$\begin{array}{ccccc|c} 2 & 1 & 1 & 0 & 0 & 80 \\ 1 & 1 & 0 & 1 & 0 & 50 \\ 1 & 2 & 0 & 0 & 1 & 90 \\ \hline 500 & 700 & 0 & 0 & 0 & 10^4 \end{array} \qquad (*)$$

Por transformaciones elementales de las filas (de las que acabamos de mencionar) queremos obtener (salvo una permutación σ de las cinco primeras columnas) una matriz de la forma

$$\begin{array}{ccccc|c} a_{11} & a_{12} & a_1 & 0 & 0 & b_1 \\ a_{21} & a_{22} & 0 & a_2 & 0 & b_2 \\ a_{31} & a_{32} & 0 & 0 & a_3 & b_3 \\ \hline c_1 & c_2 & 0 & 0 & 0 & c \end{array}$$

con $a_1, a_2, a_3 > 0$, $b_1, b_2, b_3 \geq 0$, $c_1, c_2 \leq 0$, porque G alcanza el valor máximo en el mismo punto que la función $c_1 x_{\sigma(1)} + c_2 x_{\sigma(2)} - c$, y ésta lo alcanza en $x_{\sigma(1)} = x_{\sigma(2)} = 0$, $x_{\sigma(3)} = \frac{b_1}{a_1}$, $x_{\sigma(4)} = \frac{b_2}{a_2}$, $x_{\sigma(5)} = \frac{b_3}{a_3}$. Procedamos por el denominado método simplex.

Consideremos en $(*)$ la segunda columna (porque 700 es positivo y es mayor que 500) y veamos cuál es el menor de los cocientes de los coeficientes de la sexta columna $(80, 50, 90)$ por los coeficientes de la segunda columna $(1, 1, 2)$, que sean positivos, es decir, 80/1, 50/1 y 90/2. Nos sale 90/2. Pivotemos en el 2 (que está en la segunda columna y tercera fila)[2], obtenemos[3]

$$\begin{array}{ccccc|c} 2 & 1 & 1 & 0 & 0 & 80 \\ 1 & 1 & 0 & 1 & 0 & 50 \\ 1 & 2 & 0 & 0 & 1 & 90 \\ \hline 500 & 700 & 0 & 0 & 0 & c \end{array} \quad \begin{array}{c} F_1 - \frac{F_3}{2} \\ F_2 - \frac{F_3}{2} \\ \sim \\ F_4 - \frac{700F_3}{2} \end{array} \quad \begin{array}{ccccc|c} 3/2 & 0 & 1 & 0 & -1/2 & 35 \\ 1/2 & 0 & 0 & 1 & -1/2 & 5 \\ 1 & 2 & 0 & 0 & 1 & 90 \\ \hline 150 & 0 & 0 & 0 & -350 & c' \end{array}$$

Consideremos ahora la primera columna (porque 150 es positivo y -350 negativo) y veamos cuál es el menor de los cocientes de los coeficientes de la sexta columna

[2]Por el modo en el que lo hemos escogido, nos aseguramos de que los coeficientes de la última columna de la nueva matriz que vamos a obtener son no negativos.

[3]No necesitamos calcular c, ni c', etc. Recordemos que $-c$ es el valor de G en la solución $x_1 = x_2 = 0$ (y por tanto $x_3 = 80$, $x_4 = 50$, $x_5 = 90$), $-c'$ es el valor de G en la solución $x_1 = x_5 = 0$ (y por tanto $x_3 = 35$, $x_4 = 5$, $x_2 = 45$). Observemos que $c = 10^4 > 10^4 - \frac{700}{2}\cdot 90 = c'$, esto sucederá sucesivamente (en los ejemplos) por lo que este proceso terminará en un número finito de pasos.

$(35,5,90)$ por los coeficientes de la primera columna $(3/2,1/2,1)$, que sean no negativos, es decir, $\frac{35}{3/2}=\frac{70}{3}$, $\frac{5}{1/2}=10$ y $\frac{90}{1}=90$. Nos sale $\frac{5}{1/2}$. Pivotemos en el $1/2$ (que está en la primera columna y segunda fila):

$$\begin{array}{ccccc|c} 3/2 & 0 & 1 & 0 & -1/2 & 35 \\ 1/2 & 0 & 0 & 1 & -1/2 & 5 \\ 1 & 2 & 0 & 0 & 1 & 90 \\ \hline 150 & 0 & 0 & 0 & -350 & \end{array} \sim \begin{array}{ccccc|c} 0 & 0 & 1 & -3 & 1 & 20 \\ 1/2 & 0 & 0 & 1 & -1 & 5 \\ 0 & 2 & 0 & -2 & 2 & 80 \\ \hline 0 & 0 & 0 & -300 & -350 & \end{array}$$

Entonces, el valor máximo de G se alcanza cuando $x_4=x_5=0$, $x_3=20$, $x_1=10$ y $x_2=40$; luego $G=23000$.

6. Del mismo modo resolveremos el problema de maximizar una función

$$G=c_1x_1+\cdots+c_nx_n+c$$

sujeta a las restricciones

$$\begin{array}{ccc} a_{11}x_1+\cdots+a_{1n}x_n & \leq & b_1 \\ \cdots & \leq & \cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n & \leq & b_m \\ x_1,\ldots,x_n & \geq & 0 \end{array}$$

donde $b_1,\ldots,b_m\geq 0$.

7. Indiquemos cómo resolver problemas parecidos en una situación más general. Observemos lo que sigue:

1. Minimizar una función G equivale a maximizar $-G$.

2. El conjunto de soluciones de una inecuación $a_1x_1+\ldots+a_nx_n\geq b$ es igual al conjunto de soluciones de la inecuación $-a_1x_1-\ldots-a_nx_n\leq -b$.

3. Si tenemos $a_1x_1+\ldots+a_nx_n=b$, no será necesario añadir una variable de "holgura", pero imitando el ejemplo de abajo, añadiremos una variable "artificial".

Sabemos resolver el problema de maximizar una función $G=\sum_i c_ix_i+c$ en el conjunto de soluciones de un sistema de ecuaciones $A\cdot x^t=b^t$, de modo que $x\geq 0$ y suponiendo que $b=(b_1,\ldots,b_m)\geq 0$ y que la matriz A contiene una submatriz Id de orden m (reordenando las columnas si es necesario) ¿Pero cómo resolvemos este problema si algún b_i es negativo? Un ejemplo nos dará la respuesta.

8. Ejemplo : Resolved el problema de programación lineal: Maximizar $G=3x_1+5x_2$ sujeto a las restricciones

$$\begin{array}{rcl} -x_1+5x_2 & \leq & 29 \\ x_1+x_2 & \geq & 7 \\ 3x_1+2x_2 & \leq & 32 \\ x_1+3x_2 & \geq & 13 \\ x_1,x_2 & \geq & 0 \end{array}$$

El sistema de inecuaciones es equivalente a

$$\begin{array}{rcl} -x_1+5x_2 & \leq & 29 \\ -x_1-x_2 & \leq & -7 \\ 3x_1+2x_2 & \leq & 32 \\ -x_1-3x_2 & \leq & -13 \\ x_1,x_2 & \geq & 0 \end{array}$$

Tenemos que maximizar G sujeto a las restricciones

$$\begin{array}{rcl} -x_1+5x_2+x_3 & = & 29 \\ -x_1-x_2+x_4 & = & -7 \\ 3x_1+2x_2+x_5 & = & 32 \\ -x_1-3x_2+x_6 & = & -13 \\ x_1,\dots,x_6 & \geq & 0 \end{array} \sim \begin{array}{rcl} -x_1+5x_2+x_3 & = & 29 \\ x_1+x_2-x_4 & = & 7 \\ 3x_1+2x_2+x_5 & = & 32 \\ x_1+3x_2-x_6 & = & 13 \\ x_1,\dots,x_6 & \geq & 0 \end{array} \quad (*)$$

Desgraciadamente la matriz de coeficientes del sistema de ecuaciones (permutando las columnas) no contiene a la matriz identidad de orden 4. Añadamos las variables “artificiales” x_7, x_8 y consideremos el sistema de inecuaciones

$$\begin{array}{rcl} -x_1+5x_2+x_3 & = & 29 \\ x_1+x_2-x_4+x_7 & = & 7 \\ 3x_1+2x_2+x_5 & = & 32 \\ x_1+3x_2-x_6+x_8 & = & 13 \\ x_1,\dots,x_8 & \geq & 0 \end{array} \quad (**)$$

cuya matriz de coeficientes sí contiene a la matriz identidad. Si $N=-x_7-x_8$. sujeto a las restricciones $(**)$, no alcanza el máximo en $x_7=x_8=0$, esto implicaría que el sistema de inecuaciones $(*)$ no tiene sólución. En el caso de que N alcance el máximo en $x_7=x_8=0$, por el método simplex vamos a obtener un sistema de ecuaciones equivalente a $(*)$ que contiene a la matriz identidad:

Escribamos el sistema de ecuaciones $(**)$ y añadamos en la última fila la función N. Tenemos

$$\begin{array}{cccccccc|c} -1 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 29 \\ 1 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 7 \\ 3 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 32 \\ 1 & 3 & 0 & 0 & 0 & -1 & 0 & 1 & 13 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & \end{array} \overset{F_5+F_2+F_4}{\sim} \begin{array}{cccccccc|c} -1 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 29 \\ 1 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 7 \\ 3 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 32 \\ 1 & 3 & 0 & 0 & 0 & -1 & 0 & 1 & 13 \\ \hline 2 & 4 & 0 & -1 & 0 & -1 & 0 & 0 & \end{array}$$

Procedemos ya por el método simplex. Pivotamos por tanto en el coeficiente de la segunda columna y la cuarta fila:

$$
\begin{array}{cccccccc|c}
-1 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 29 \\
1 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 7 \\
3 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 32 \\
1/3 & 1 & 0 & 0 & 0 & -1/3 & 0 & 1/3 & 13/3 \\
\hline
2 & 4 & 0 & -1 & 0 & -1 & 0 & 0 &
\end{array}
\sim
\begin{array}{cccccccc|c}
-8/3 & 0 & 1 & 0 & 0 & 5/3 & 0 & -5/3 & 22/3 \\
2/3 & 0 & 0 & -1 & 0 & 1/3 & 1 & -1/3 & 8/3 \\
7/3 & 0 & 0 & 0 & 1 & 2/3 & 0 & -2/3 & 70/3 \\
1/3 & 1 & 0 & 0 & 0 & -1/3 & 0 & 1/3 & 13/3 \\
\hline
2/3 & 0 & 0 & -1 & 0 & 1/3 & 0 & -4/3 &
\end{array}
$$

Como me molestan las fracciones multiplico por 3 las filas que me convenga y luego pivoto en el coeficiente de la primera columna y segunda fila

$$
\begin{array}{cccccccc|c}
-8 & 0 & 3 & 0 & 0 & 5 & 0 & -5 & 22 \\
2 & 0 & 0 & -3 & 0 & 1 & 3 & -1 & 8 \\
7 & 0 & 0 & 0 & 3 & 2 & 0 & -2 & 70 \\
1 & 3 & 0 & 0 & 0 & -1 & 0 & 1 & 13 \\
\hline
2 & 0 & 0 & -3 & 0 & 1 & 0 & -4 &
\end{array}
\sim
\begin{array}{cccccc|cc|c}
0 & 0 & 3 & -12 & 0 & 9 & 12 & -9 & 54 \\
2 & 0 & 0 & -3 & 0 & 1 & 3 & -1 & 8 \\
0 & 0 & 0 & 21 & 6 & -3 & -21 & 3 & 84 \\
0 & 6 & 0 & 3 & 0 & -3 & -3 & 3 & 18 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & -3 & -4 &
\end{array}
$$

Luego, N alcanza el máximo en $x_7 = x_8 = 0$. Quedémosnos con las ecuaciones resultantes, hagamos $x_7 = x_8 = 0$ y maximicemos G,

$$
\begin{array}{cccccc|c}
0 & 0 & 3 & -12 & 0 & 9 & 54 \\
2 & 0 & 0 & -3 & 0 & 1 & 8 \\
0 & 0 & 0 & 21 & 6 & -3 & 84 \\
0 & 6 & 0 & 3 & 0 & -3 & 18 \\
\hline
3 & 5 & 0 & 0 & 0 & 0 &
\end{array}
\sim
\begin{array}{cccccc|c}
0 & 0 & 1 & -4 & 0 & 3 & 18 \\
1 & 0 & 0 & -3/2 & 0 & 1/2 & 4 \\
0 & 0 & 0 & 7 & 2 & -1 & 28 \\
0 & 1 & 0 & 1/2 & 0 & -1/2 & 3 \\
\hline
3 & 5 & 0 & 0 & 0 & 0 &
\end{array}
$$

$$
\sim
\begin{array}{cccccc|c}
0 & 0 & 1 & -4 & 0 & 3 & 18 \\
1 & 0 & 0 & -3/2 & 0 & 1/2 & 4 \\
0 & 0 & 0 & 7 & 2 & -1 & 28 \\
0 & 1 & 0 & 1/2 & 0 & -1/2 & 3 \\
\hline
0 & 0 & 0 & 2 & 0 & 1 &
\end{array}
\sim
\begin{array}{cccccc|c}
0 & 0 & 7 & 0 & 8 & 17 & 238 \\
14 & 0 & 0 & 0 & 6 & 4 & 140 \\
0 & 0 & 0 & 7 & 2 & -1 & 28 \\
0 & 14 & 0 & 0 & -2 & -6 & 14 \\
\hline
0 & 0 & 0 & 0 & -4 & 9 &
\end{array}
$$

$$
\sim
\begin{array}{cccccc|c}
0 & 0 & 7/17 & 0 & 8/17 & 1 & 14 \\
14 & 0 & - & 0 & - & 0 & 84 \\
0 & 0 & - & 7 & - & 0 & 42 \\
0 & 14 & - & 0 & - & 0 & 98 \\
\hline
0 & 0 & -63/14 & 0 & -140/17 & 0 &
\end{array}
$$

Luego, $x_1 = 6$, $x_2 = 7$ y $G = 53$.

En los problemas 13 y 14 se profundiza algo más en la Programación Lineal.

2.6. Coordenadas de un vector en una base

1. Coordenadas: Dada una base $B = \{e_1, \dots, e_n\}$ de un espacio vectorial E y un vector $e \in E$, existen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tales que

$$e = \lambda_1 e_1 + \dots + \lambda_n e_n$$

porque los e_i generan E. Además, estos λ_i son únicos, porque si

$$e = \lambda_1 e_1 + \dots + \lambda_n e_n = \lambda'_1 e_1 + \dots + \lambda'_n e_n,$$

entonces $(\lambda_1 - \lambda'_1)e_1 + \dots + (\lambda_n - \lambda'_n)e_n = 0$, luego $\lambda_1 - \lambda'_1 = \dots = \lambda_n - \lambda'_n = 0$, porque los e_i son linealmente independientes. Por tanto, $\lambda_i = \lambda'_i$, para todo i.

Diremos que $(\lambda_1, \dots, \lambda_n)$ son las coordenadas del vector e en la base B y escribiremos

$$\text{“}e_B = (\lambda_1, \dots, \lambda_n)\text{”}$$

Seguiremos también la notación $e^B = (\lambda_1, \dots, \lambda_n)^t := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$.

Dados $e, e' \in E$ y $\lambda \in \mathbb{R}$, si $e_B = (\lambda_1, \dots, \lambda_n)$ y $e'_B = (\lambda'_1, \dots, \lambda'_n)$, entonces

$$(e + e')_B = e_B + e'_B \quad \text{y} \quad (\lambda \cdot e)_B = \lambda \cdot e_B.$$

2. Ejercicio: Sea B la base estándar de $\mathbb{R}^n$. Probad que $(x_1, \dots, x_n)_B = (x_1, \dots, x_n)$.

2.7. Aplicaciones lineales. Matrices

Las aplicaciones que conservan la estructura de espacio vectorial son las aplicaciones lineales. Con precisión:

1. Definición: Sean E, E' dos $\mathbb{R}$-espacios vectoriales. Una aplicación $T: E \to E'$ se dice que es una *aplicación $\mathbb{R}$-lineal* si

$$T(e + v) = T(e) + T(v) \quad \text{y} \quad T(\lambda \cdot e) = \lambda \cdot T(e)$$

para cualesquiera $e, v \in E$, $\lambda \in \mathbb{R}$.

Observemos que $T(0) = 0$: $T(0) = T(0+0) = T(0) + T(0)$, y sumando $-T(0)$ en ambos lados de la igualdad, obtenemos $0 = T(0)$.

2. Ejemplos : 1. Si $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ es un giro de ángulo θ alrededor del origen entonces T es una aplicación lineal.

2. La aplicación $T\colon \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (2x+y, x-5y, 11y)$ es una aplicación lineal.

3. Consideremos el $\mathbb{R}$-espacio vectorial

$$\begin{aligned} E &:= \{\text{Aplicaciones } f\colon \mathbb{R} \to \mathbb{R} \text{ infinitamente derivables}\} \\ &= \{x^2-3, \cos x, \operatorname{sen} x + 7, e^x, \text{etc.}\} \end{aligned}$$

La aplicación $D\colon E \to E$, $D(f) = f'$ es una aplicación lineal

Una aplicación lineal está determinada por lo que vale en los vectores de una base. Efectivamente, si $\{e_1,\dots,e_n\}$ es una base de E, y $e \in E$, tendremos que $e = \sum_i x_i e_i$, para ciertos $x_i \in \mathbb{R}$ y tendremos que $T(e) = \sum_i x_i \cdot T(e_i)$. Por otra parte, si $\{e_1,\dots,e_n\}$ es una base de E, dados $v_1,\dots,v_n \in E'$ la aplicación $T\colon E \to E'$, definida por $T(\sum_i \lambda_i e_i) := \sum_i \lambda_i v_i$ es lineal y cumple que $T(e_i) = v_i$.

3. Matriz asociada a una aplicación lineal : Sean E y E' dos espacios vectoriales, $B = \{e_1,\dots,e_n\}$ una base de E y $B' = \{e'_1,\dots,e'_m\}$ una base de E'. Sea $T\colon E \to E'$ una aplicación lineal. Tenemos que

$$T(e_i) = \lambda_{1i}e'_1 + \cdots + \lambda_{mi}e'_m$$

para ciertas $\lambda_{1i},\dots,\lambda_{mi} \in \mathbb{R}$ únicas. Diremos que la caja de números

$$\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix}$$

de columnas $T(e_i)^{B'} = \begin{pmatrix} \lambda_{1i} \\ \vdots \\ \lambda_{mi} \end{pmatrix}$, es la matriz de T en las bases B y B'. Escribiremos

$$T_B^{B'} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix}.$$

Dado $e \in E$ de coordenadas $e_B = (x_1,\dots,x_n)$, calculemos las coordenadas de $T(e)$, es decir, $T(e)_{B'} = (x'_1,\dots,x'_m)$: Tenemos que

$$T(e) = T(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i T(e_i) = \sum_{i=1}^{n} x_i (\sum_{j=1}^{m} \lambda_{ji} e'_j) = \sum_{j=1}^{m} (\sum_{i=1}^{n} \lambda_{ji} x_i) e'_j.$$

Luego, $x'_j = \sum_{i=1}^{n} \lambda_{ji} x_i$ = "la fila j de la matriz (λ_{rs}) por la columna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$", y escribiremos

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o de modo abreviado $T(e)^{B'} = T_B^{B'} \cdot e^B$.

4. Ejemplo : Calculemos la matriz de $T\colon \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x+y, 2x, x-y)$ en las bases usuales $U_2 = \{(1,0),(0,1)\}$ de $\mathbb{R}^2$ y $U_3 = \{(1,0,0),(0,1,0),(0,0,1)\}$ de $\mathbb{R}^3$: 1. Tomemos el primer vector de la base de U_2 y calculemos las coordenadas en la base U_3 de su imagen por T:

$$T(1,0) = (1+0, 2\cdot 1, 1-0) = (1,2,1) = 1\cdot(1,0,0) + 2\cdot(0,1,0) + 1\cdot(0,0,1).$$

Luego, $(T(1,0))_{U_3} = (1,2,1)$ y la primera columna de $T_{U_2}^{U_3}$ es $(1,2,1)^t$.

2, Tomemos el segundo vector de la base de U_2 y calculemos las coordenadas en la base U_3 de su imagen por T:

$$T(0,1) = (0+1, 2\cdot 0, 0-1) = (1,0,-1) = 1\cdot(1,0,0) + 0\cdot(0,1,0) - 1\cdot(0,0,1).$$

Luego, $(T(0,1))_{U_3} = (1,0,-1)$ y la segunda columna de $T_{U_2}^{U_3}$ es $(1,0,-1)^t$.

3. Es decir, $T_{U_2}^{U_3} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}$.

5. Ejemplo : Sea E un espacio vectorial de base $B = \{e_1, \dots, e_n\}$ y $\mathrm{Id}\colon E \to E$ la aplicación lineal identidad, $\mathrm{Id}(e) = e$ para todo $e \in E$. La matriz de Id en las bases B y B es

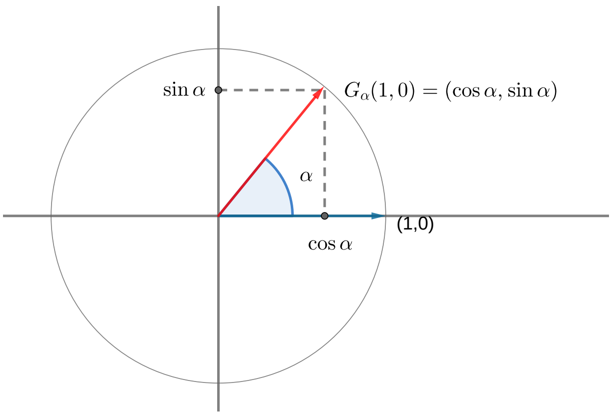
$$\mathrm{Id}_B^B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

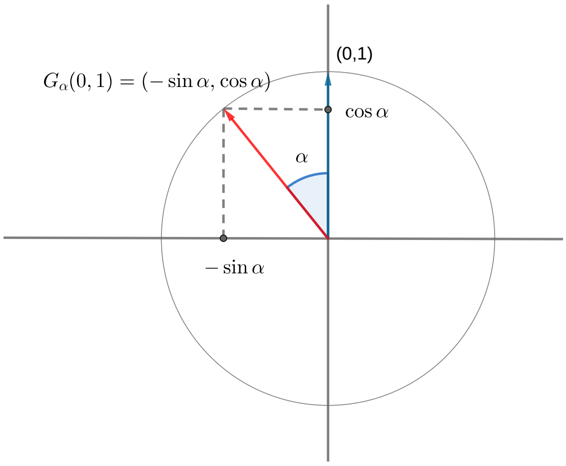
6. Ejemplo: Consideremos la matriz $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$. La aplicación $T\colon \mathbb{R}^n \to \mathbb{R}^m$ definida por $T(x_1,\dots,x_n) = (x'_1,\dots,x'_m)$, donde

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

es decir, $x'_j = \sum_{i=1}^n a_{ji}x_i$, para cada j, es una aplicación lineal. Además, la matriz asociada a T en las bases usuales U_n y U_m usuales de $\mathbb{R}^n$ y $\mathbb{R}^m$ es justamente (a_{ij}). Es usual escribir $T = (a_{ij})$ (en vez de $T_{U_n}^{U_m} = (a_{ij})$).

7. Ejemplo: Sea $G_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ el giro (antihorario) de ángulo α (radianes) alrededor del origen. Calculemos la matriz de G_α en las bases usuales de $\mathbb{R}^2$ y de $\mathbb{R}^2$. Tenemos que





Por lo tanto, la matriz de G_α en las bases usuales es $\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$. Calculemos $G_\alpha(x,y) = (x',y')$:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\alpha - y\operatorname{sen}\alpha \\ x\operatorname{sen}\alpha + y\cos\alpha \end{pmatrix}$$

8. Ejercicio: Sea $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ la simetría respecto del eje OX. Calcula la matriz de S en la base usual de $\mathbb{R}^2$ y en la base usual de $\mathbb{R}^2$.

2.7.1. Suma y producto de matrices

9. Definición: Diremos que $M_{m\times n}(\mathbb{R})$ es el conjunto de todas las matrices de m filas y n columnas con coeficientes números reales. Cuando $m = n$, seguiremos la siguiente notación $M_{n\times n}(\mathbb{R}) = M_n(\mathbb{R})$ y diremos que sus matrices son matrices cuadradas de orden n.

Si definimos para todo $(a_{ij}),(b_{ij}) \in M_{m\times n}(\mathbb{R})$ y todo $\lambda \in \mathbb{R}$

$$\begin{array}{l} (a_{ij})+(b_{ij}) := (c_{ij}), \text{ donde } c_{ij} = a_{ij}+b_{ij} \text{ para todo } i,j \\ \lambda\cdot(a_{ij}). = (d_{ij}), \text{ donde } d_{ij} = \lambda\cdot a_{ij} \text{ para todo } i,j \end{array}$$

entonces es fácil ver que $M_{m\times n}(\mathbb{R})$ es un $\mathbb{R}$-espacio vectorial.

10. Matriz de la suma de dos aplicaciones lineales: Sean $T,S\colon E \to E'$ dos aplicaciones lineales. La aplicación $T+S\colon E\to E'$ definida por

$$(T+S)(e) := T(e)+S(e), \qquad \forall e\in E$$

es lineal. Dada $\lambda\in\mathbb{R}$, la aplicación $\lambda\cdot T\colon E\to E'$ definida por $(\lambda\cdot T)(e) := \lambda\cdot T(e)$ es lineal. Sean B y B' bases de E y E' respectivamente. Es fácil probar que

$$(T+S)_B^{B'} = T_B^{B'} + S_B^{B'} \text{ y que } (\lambda\cdot T)_B^{B'} = \lambda\cdot T_B^{B'}.$$

11. Matriz de la composición de dos aplicaciones lineales: Sean E, E' y E'' espacios vectoriales de bases $B=\{e_1,\dots,e_n\}$, $B'=\{e'_1,\dots,e'_m\}$ y $B''=\{e''_1,\dots,e''_r\}$ respectivamente. Sea $T\colon E\to E'$ una aplicación lineal de matriz (λ_{ji}) en las bases B, B'. Sea $S\colon E'\to E''$ una aplicación lineal de matriz (μ_{kj}) en las bases B', B''. Es claro que la composición $S\circ T\colon E\to E''$, $(S\circ T)(e) := S(T(e))$, es una aplicación lineal ¿Cuál es la matriz (c_{ki}) de $S\circ T\colon E\to E''$ en las bases B, B''?

$$(S\circ T)(e_i) = S(\sum_j \lambda_{ji}e'_j) = \sum_j \lambda_{ji}S(e'_j) = \sum_j\sum_k \lambda_{ji}\cdot\mu_{kj}\cdot e''_k = \sum_k\sum_j \mu_{kj}\cdot\lambda_{ji}\cdot e''_k$$

Luego $c_{ki} = \sum_{1\le j\le m}\mu_{kj}\cdot\lambda_{ji}$ = "la fila k de (μ_{rs}) por la columna i de (λ_{uv})". Es decir, si por definición decimos que $(\mu_{rs})\cdot(\lambda_{ji})$ es la matriz cuyo coeficiente ki es la fila k de (μ_{rs}) por la columna i de (λ_{uv}), entonces

$$(c_{uv}) = (\mu_{rs})\cdot(\lambda_{ji}).$$

Con concisión, hemos probado

$$(S\circ T)_B^{B''} = S_{B'}^{B''}\cdot T_B^{B'}.$$

12. Ejemplo: Si $G_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ es el giro de ángulo α alrededor del origen y $G_\beta\colon \mathbb{R}^2 \to \mathbb{R}^2$ es el giro de ángulo β alrededor del origen, es obvio que $G_\alpha \circ G_\beta = G_{\alpha+\beta}$. Matricialmente tendremos que

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & -\operatorname{sen}\beta \\ \operatorname{sen}\beta & \cos\beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & -\operatorname{sen}(\alpha+\beta) \\ \operatorname{sen}(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}.$$

Ahora bien,

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & -\operatorname{sen}\beta \\ \operatorname{sen}\beta & \cos\beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha)\cos(\beta) - \operatorname{sen}(\alpha)\operatorname{sen}(\beta) & - \\ \operatorname{sen}(\alpha)\cos(\beta) + \cos(\alpha)\operatorname{sen}(\beta) & - \end{pmatrix}.$$

Luego, hemos obtenido las conocidas fórmulas trigonométricas

$$\begin{aligned} \cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \operatorname{sen}(\alpha)\operatorname{sen}(\beta) \\ \operatorname{sen}(\alpha+\beta) &= \operatorname{sen}(\alpha)\cos(\beta) + \cos(\alpha)\operatorname{sen}(\beta) \end{aligned}$$

13. Ejercicio: Calcula $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$.

2.8. Isomorfismos lineales

1. Definición: Una aplicación lineal $T\colon E \to E'$ se dice que es un isomorfismo lineal si existe una aplicación lineal $S\colon E' \to E$ tal que $T \circ S = \mathrm{Id}$ y $S \circ T = \mathrm{Id}$[4]. En este caso se dice que S es la inversa de T y escribiremos $S = T^{-1}$.

Si T es un isomorfismo lineal y $S \circ T = \mathrm{Id}$ entonces componiendo (por la derecha) con T^{-1} en ambos términos de la igualdad, tendremos que $S = T^{-1}$. Igualmente, si $T \circ S' = \mathrm{Id}$, componiendo (por la izquierda) con T^{-1} ambos términos de la igualdad, obtenemos que $S' = T^{-1}$.

Si T es un isomorfismo lineal entonces T^{-1} es un isomorfismo lineal, pues $(T^{-1})^{-1} = T$. La composición $T \circ S$ de isomorfismos lineales es un isomorfismo lineal, ya que $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$.

2. Ejemplos: La aplicación identidad $\mathrm{Id}\colon E \to E$ es evidentemente un isomorfismo y $\mathrm{Id}^{-1} = \mathrm{Id}$.

El giro $G_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ de ángulo α alrededor del origen es un isomorfismo lineal, cuyo inverso es $G_{-\alpha}$.

[4]Con otras palabras, T es isomorfismo lineal si y solo si es biyectiva, es decir, para cada $e' \in E'$ existe un único $e \in E$ tal que $T(e) = e'$.

Sea E un espacio vectorial y $B = \{e_1 \dots, e_n\}$ una base de E. Entonces, la aplicación $E \to \mathbb{R}^n$, $e \mapsto e_B$ es un isomorfismo lineal.

3. Proposición: *Sea E un espacio vectorial, $\{e_1, \dots, e_n\}$ una base de E y $T\colon E \to E'$ una aplicación lineal. Entonces, T es un isomorfismo si y solo si $\{T(e_1), \dots, T(e_n)\}$ es una base de E'.*

Demostración. $\Rightarrow$) $T(e_1), \dots, T(e_n)$ es un sistema generador de E': dado $e' \in E'$, $T^{-1}(e') = \sum_i \lambda_i e_i$ para ciertos λ_i, entonces $e' = T(T^{-1}(e')) = \sum_i \lambda_i T(e_i)$. Veamos que son linealmente independientes: Si $0 = \lambda_1 \cdot T(e_1) + \cdots + \lambda_n \cdot T(e_n)$, entonces aplicando T^{-1} a esta igualdad obtenemos que $0 = \lambda_1 \cdot e_1 + \cdots + \lambda_n \cdot e_n$, luego $\lambda_1 = \cdots = \lambda_n = 0$.

$\Leftarrow$) La aplicación lineal $S\colon E' \to E$ que sobre la base $\{T(e_1), \dots, T(e_n)\}$ cumple que $S(T(e_i)) := e_i$, para todo i, es la aplicación inversa de T. □

Esta proposición nos dice en particular que si $T\colon E \to E'$ es un isomorfismo lineal entonces $\dim E = \dim E'$. Si $\dim E = \dim E'$, entonces $\{T(e_1), \dots, T(e_n)\}$ es una base de E' si y solo si son linealmente independientes.

4. Corolario: *Supongamos que* $\dim E = \dim E'$. *Sean $T\colon E \to E'$ y $S\colon E' \to E$ aplicaciones lineales. Entonces,*

1. *S es la inversa de T si y solo si $S \circ T = \mathrm{Id}$.*

2. *S es la inversa de T si y solo si $T \circ S = \mathrm{Id}$.*

Demostración. 1. $\Leftarrow$) Sea $B = \{e_1, \dots, e_n\}$ una base de E. Los vectores $\{T(e_1), \dots, T(e_n)\}$ son linealmente independientes, porque $\{S(T(e_1)) = e_1, \dots, S(T(e_n)) = e_n\}$ son linealmente independientes. Por tanto, $\{T(e_1), \dots, T(e_n)\}$ es una base de E' y T tiene inversa. Como $S \circ T = \mathrm{Id}$, componiendo por la derecha con T^{-1}, obtenemos que $S = T^{-1}$.

2. $\Leftarrow$) Por 1. sabemos que T es la inversa de S, luego S es la inversa de T. □

Veamos que estamos diciendo en términos de las matrices asociadas. Sean $B = \{e_1, \dots, e_n\}$ y $B' = \{e'_1, \dots, e'_n\}$ bases de E y E' respectivamente. Sean $T\colon E \to E'$ y $S\colon E' \to E$ aplicaciones lineales. Entonces, $S \circ T = \mathrm{Id}$ si y solo si $(S \circ T)_B^B = S_{B'}^B \cdot T_B^{B'}$ es la matriz identidad. En este caso, diremos que $S_{B'}^B$ es la matriz inversa de $T_B^{B'}$ y la denotaremos $(T_B^{B'})^{-1}$. Es decir, si la matriz de $T_B^{B'}$ es (λ_{ij}) entonces

$$(\lambda_{ij})^{-1} \cdot (\lambda_{ij}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Observemos que $S = T^{-1}$, luego $(T^{-1})^B_{B'} = (T^{B'}_B)^{-1}$. Además, $(\lambda_{ij})\cdot(\lambda_{ij})^{-1}$ es la matriz identidad porque $T \circ S = \mathrm{Id}$.

Las columnas de la matriz $T^{B'}_B$ son los vectores $T(e_1)^{B'}, \dots, T(e_n)^{B'}$. Por la proposición anterior, una matriz (cuadrada) es invertible si y solo si sus columnas son linealmente independientes.

2.8.1. Transformaciones elementales

Sea $A \in M_{n\times m}(\mathbb{R})$ y sea $\delta_{ij} \in M_n(\mathbb{R})$ la matriz nula, salvo el coeficiente ij que es 1. Puede comprobarse que

$$B = (\mathrm{Id} + \lambda\cdot\delta_{ij})\cdot A$$

es igual a la matriz A salvo que la fila i de B es la suma de la fila i de A más λ-veces la fila j de A.

Si $i \neq j$, diremos que $\mathrm{Id} + \lambda\cdot\delta_{ij}$ es una transformación elemental especial. Es fácil comprobar que $(\mathrm{Id} + \lambda\cdot\delta_{ij})^{-1} = \mathrm{Id} - \lambda\cdot\delta_{ij}$. Puede comprobarse que

$$B' = (\mathrm{Id} + \delta_{ij})\cdot(\mathrm{Id} - \delta_{ji})\cdot(\mathrm{Id} + \delta_{ij})\cdot A$$

es igual a la matriz A salvo que la fila i de B' es igual a la fila j de A y que la fila j de B' es igual a menos la fila i de A.

Si $i = j$ y $1+\lambda \neq 0$, diremos que $\mathrm{Id} + \lambda\cdot\delta_{ii}$ es una transformación elemental homotética. Observemos que $(\mathrm{Id} + \lambda\cdot\delta_{ii})\cdot A$ es igual a A salvo que la fila i de A ha sido multiplicada por $1+\lambda$. Puede comprobarse que $(\mathrm{Id} + \lambda\cdot\delta_{ii})^{-1} = \mathrm{Id} + \frac{-\lambda}{1+\lambda}\cdot\delta_{ii}$.

Igualmente, dada $A \in M_{m\times n}(\mathbb{R})$ y $\delta_{ij} \in M_n(\mathbb{R})$ puede comprobarse que

$$C = A\cdot(\mathrm{Id} + \lambda\cdot\delta_{ij})$$

es igual a la matriz A salvo que la columna j de C es la suma de la columna j de A más λ-veces la columna i de A. En particular, si $i = j$ la columna i de C es $(1+\lambda)$-veces la columna i de A.

5. Proposición : *Sea $A \in M_{m\times n}(\mathbb{R})$. Existen transformaciones elementales especiales $T_1, \dots, T_r \in M_m(\mathbb{R})$ y $T'_1, \dots, T'_{r'} \in M_n(\mathbb{R})$ de modo que*

$$T_r \circ \cdots \circ T_1 \circ A \circ T'_1 \circ \cdots T'_{r'} = \begin{pmatrix} d_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & d_m & \dots & 0 \end{pmatrix}.$$

Luego, A es producto de transformaciones elementales especiales y una matriz "diagonal".

Demostración. Si $A = 0$ hemos acabado. Si $A \neq 0$ mediante transformaciones elementales especiales de filas y columnas podemos suponer que $a_{11} \neq 0$. Mediante transformaciones elementales especiales de filas podemos conseguir que $a_{i1} = 0$ para todo $i > 1$. Mediante transformaciones elementales especiales de columnas podemos conseguir que $a_{1i} = 0$ para todo $i > 1$. Es decir, hemos transformado A en la matriz

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Jugando del mismo modo con las filas $2, \dots, m$ y columnas $2, \dots, n$ y repitiendo este proceso de modo sucesivo, podemos transformar la matriz A en la matriz

$$\begin{pmatrix} d_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & d_m & \dots & 0 \end{pmatrix}.$$

□

6. Proposición: *Sea $A \in M_n(\mathbb{R})$ invertible. Entonces,*

1. *Existen transformaciones elementales especiales $T_1, \dots, T_r \in M_n(\mathbb{R})$ de modo que*

$$T_r \circ \cdots \circ T_1 \circ A = \begin{pmatrix} d_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_n \end{pmatrix}.$$

2. *Existen transformaciones elementales $T_1, \dots, T_r \in M_n(\mathbb{R})$ de modo que*

$$T_r \circ \cdots \circ T_1 \circ A = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

Luego A es producto de transformaciones elementales.

Demostración. 1. Como A es invertible, no puede ser la primera columna de A igual a cero. Mediante transformaciones elementales especiales de filas, podemos suponer que $a_{11} \neq 0$. Mediante transformaciones elementales de filas podemos conseguir que $a_{i1} = 0$ para todo $i > 1$. Es decir, hemos transformado A en la matriz

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

No pueden ser $a_{i2} = 0$ para todo $i > 1$, porque si no las dos primeras columnas serían linealmente dependientes. Mediante una transformación elemental especial de las filas $2, \dots, n$, podemos suponer que $a_{22} \neq 0$. Mediante transformaciones elementales especiales de filas podemos transformar la matriz A en la matriz

$$\begin{pmatrix} a_{11} & 0 & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Así sucesivamente, obtendremos la matriz diagonal buscada y habremos demostrado la proposición.

2. Una vez demostrado 1. es fácil obtener 2. □

2.8.2. Cálculo de la matriz inversa

Si A es una matriz cuadrada invertible sabemos que existen ciertas transformaciones elementales $T_1, \dots, T_m$ de modo que $T_m \circ T_{m-1} \circ \dots \circ T_1 \circ A = \mathrm{Id}$. Multipliquemos por la derecha por A^{-1} en esta igualdad, entonces

$$T_m \circ T_{m-1} \circ \dots \circ T_1 \circ \mathrm{Id} = A^{-1}.$$

Las mismas transformaciones elementales que transforman A en Id transforman Id en A^{-1}.

Calculemos la matriz inversa de la matriz $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$:

Después de las transformaciones elementales

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \overset{F_2-2F_1}{\underset{F_3-F_1}{\equiv}} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -2 & -4 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right) \overset{F_1+\frac{F_2}{2}}{\equiv} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & -2 & -4 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\overset{F_2-4F_3}{\equiv} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & -2 & 0 & 2 & 1 & -4 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right) \overset{\frac{-F_2}{2}}{\underset{-F_3}{\equiv}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 & \frac{-1}{2} & 2 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right)$$

obtenemos que $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ -1 & \frac{-1}{2} & 2 \\ 1 & 0 & -1 \end{pmatrix}$.

2.9. Fórmulas de cambio de base

Muchas veces al resolver problemas de Álgebra Lineal es necesario desechar las bases donde se expresan los vectores y aplicaciones lineales consideradas, escoger unas bases convenientes, hacer los cálculos en estas bases y finalmente expresar el resultado obtenido en las bases de partida. Para todo ello necesitaremos la fórmulas de cambio de base.

1. Matriz de cambio de base: Sea E un espacio vectorial y sean $B_1 = \{e_1, \dots, e_n\}$ y $B_2 = \{v_1, \dots, v_n\}$ dos bases de E. Supongamos que $e_i = \sum_{j=1}^n a_{ji} v_j$, para cada i. Supongamos que conocemos las coordenadas e_{B_1} de e en la base B_1, queremos calcular las coordenadas e_{B_2} de e en la base B_2.

Consideremos la aplicación lineal identidad $\mathrm{Id}\colon E \to E$, $\mathrm{Id}(e) := e$. Entonces, tenemos que $e^{B_2} = \mathrm{Id}(e)^{B_2} = \mathrm{Id}_{B_1}^{B_2} \cdot e^{B_1}$.

Observemos que $\mathrm{Id}_{B_1}^{B_2} = (a_{ji})$, porque $\mathrm{Id}(e_i) = e_i = \sum_{j=1}^n a_{ji} \cdot v_j$. Se dice que $\mathrm{Id}_{B_1}^{B_2} = (a_{ij})$ es la matriz de cambio de base de B_1 a B_2.

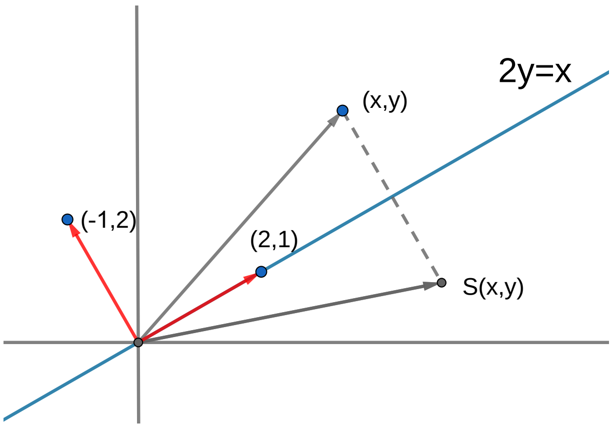
2. Observación: $\mathrm{Id}_{B_2}^{B_1} \cdot \mathrm{Id}_{B_1}^{B_2}$ es la matriz identidad $\mathrm{Id}_{B_1}^{B_1}$, luego $\mathrm{Id}_{B_2}^{B_1} = (\mathrm{Id}_{B_1}^{B_2})^{-1}$.

3. Matriz de una aplicación lineal después de cambios de bases: Sean B_1 y B_2 dos bases de E, y B'_1 y B'_2 dos bases de E'. Sea $T\colon E \to E'$ una aplicación lineal. Consideremos el diagrama conmutativo

$$\begin{tikzcd}
B_1 & E \arrow[r, "T"] & E' \arrow[d, "\mathrm{Id}"] & B'_1 \\
B_2 & E \arrow[r, "T"'] \arrow[u, "\mathrm{Id}"] & E' & B'_2
\end{tikzcd}$$

Entonces, $\boxed{T^{B'_2}_{B_2} = \mathrm{Id}^{B'_2}_{B'_1} \cdot T^{B'_1}_{B_1} \cdot \mathrm{Id}^{B_1}_{B_2}.}$

4. Ejemplo: Calculemos la matriz de la simetría $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ respecto de la recta $2y = x$ en las bases usuales de $\mathbb{R}^2$ y $\mathbb{R}^2$.



Sea $U = \{(1,0),(0,1)\}$ la base usual de $\mathbb{R}^2$. Nos piden que calculemos S^U_U. Sea $e_1 = (2,1)$ (vector que yace en la recta $2y = x$) y $e_2 = (-1,2)$ (vector perpendicular a la recta $2y = x$), y consideremos la base $B = \{e_1, e_2\}$. Observemos que $S(e_1) = e_1 = 1\cdot e_1 + 0\cdot e_2$ y $S(e_2) = -e_2 = 0\cdot e_1 - 1\cdot e_2$. Por lo tanto $S^B_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Observemos que $\mathrm{Id}^U_B = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ y del diagrama conmutativo

$$\begin{tikzcd}
U & \mathbb{R}^2 \arrow[r, "S^U_U", "S"'] & \mathbb{R}^2 & U \\
B & \mathbb{R}^2 \arrow[r, "S", "S^B_B=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}"'] \arrow[u, "\mathrm{Id}"', "\mathrm{Id}^U_B=\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}"] & \mathbb{R}^2 \arrow[u, "\mathrm{Id}", "\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}=\mathrm{Id}^U_B"'] & B
\end{tikzcd}$$

obtenemos que

$$S^U_U = \mathrm{Id}^U_B \cdot S^B_B \cdot (\mathrm{Id}^U_B)^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}^{-1} = \frac{1}{5} \cdot \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}.$$

Por tanto, $S(x,y) = \frac{1}{5}\cdot(3x+4y, 4x-3y)$.

Calculemos S_U^U de otro modo: Como $S(e_1) = e_1 = (2,1) = 2\cdot(1,0)+1\cdot(0,1)$ y $S(e_2) = -e_2 = (1,-2) = 1\cdot(1,0) - 2\cdot(0,1)$, tenemos que $S_B^U = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$. Consideremos el diagrama conmutativo

$$\begin{tikzcd}
U & \mathbb{R}^2 \arrow[r, "S_U^U", "S"'] & \mathbb{R}^2 & U \\
B & \mathbb{R}^2 \arrow[u, "\mathrm{Id}_B^U = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}", "\mathrm{Id}"'] \arrow[ur, "S", "S_B^U = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}"'] & &
\end{tikzcd}$$

Luego, $S_U^U = S_B^U \cdot (\mathrm{Id}_B^U)^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}^{-1} = \frac{1}{5} \cdot \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$.

5. Ejercicio: Sea $G_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ el giro de ángulo α alrededor del origen y $S : \mathbb{R}^2 \to \mathbb{R}^2$ la simetría respecto de la recta $2y = x$. Calcula la matriz de $S \circ G_\alpha$ en las bases usual de $\mathbb{R}^2$ y $\mathbb{R}^2$.

2.10. Imagen de una aplicación lineal. Rango

1. Definición: Sea $T : E \to E'$ una aplicación lineal. Se define la imagen de T, que denotamos $\mathrm{Im}\, T$, como sigue

$$\mathrm{Im}\, T := \{T(e) \in E', \forall\, e \in E\}$$

$\mathrm{Im}\, T$ es un subespacio vectorial de E' y si $E = \langle e_1, \dots, e_n \rangle$ entonces

$$\mathrm{Im}\, T = \langle T(e_1), \dots, T(e_n) \rangle,$$

pues dado $e \in E$, tenemos que $e = \sum_i \lambda_i e_i$ y $T(e) = \sum_i \lambda_i T(e_i)$.

2. Ejemplo: Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la aplicación lineal de matriz en las bases usuales $\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Entonces, $\mathrm{Im}\, T = \langle T(1,0), T(0,1) \rangle = \langle (1,3), (2,6) \rangle = \langle (1,3) \rangle$.

3. Definición: Sea $T : E \to E'$ una aplicación lineal. Llamaremos **rango** de T, que denotaremos $\mathrm{rango}(T)$, al siguiente número

$$\mathrm{rango}(T) := \dim(\mathrm{Im}\, T).$$

Si $E = \langle e_1, \dots, e_n \rangle$ entonces $\operatorname{rango}(T) = \dim \langle T(e_1), \dots, T(e_n) \rangle$, que es el número máximos de vectores $T(e_1), \dots, T(e_n)$ linealmente independientes. Si $A \in M_{m\times n}(\mathbb{R})$ es la matriz asociada a T (en unas bases dadas) llamaremos $\operatorname{rango}(A) := \operatorname{rango}(T)$, que es el número máximo de columnas linealmente independientes de A.

Si $S\colon E \to E$ es un isomorfismo lineal, entonces $\operatorname{rango}(T\circ S) = \operatorname{rango}(T)$, porque $\operatorname{Im}(T\circ S) = \operatorname{Im}(T)$. Si $S'\colon E' \to E'$ es un isomorfismo lineal, entonces $\operatorname{rango}(S'\circ T) = \operatorname{rango}(T)$, porque la aplicación lineal $S'\colon \operatorname{Im} T \to \operatorname{Im}(S'\circ T)$, $T(e) \mapsto S'(T(e))$ es un isomorfismo lineal, luego $\dim \operatorname{Im} T = \dim \operatorname{Im}(S'\circ T)$. En particular,

"El rango de una matriz no cambia por transformaciones elementales de filas y columnas".

4. Ejercicio: Sea $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, la aplicación lineal $T(x,y,z) = (x+y, x-y+2z, 2x+y+z)$. Calcula $\operatorname{rango}(T)$.

5. Definición: Sea $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$ una matriz. Llamaremos matriz traspuesta de A, que denotaremos $A^t = (a^t_{ij}) \in M_{n\times m}(\mathbb{R})$, a la matriz definida por $a^t_{ij} := a_{ji}$ (que es la matriz cuyas filas son las columnas de A).

6. Ejercicio: Calcula la matriz traspuesta de $\begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & -1 \end{pmatrix}$.

7. Proposición: *Se cumple que*

1. $(A^t)^t = A$, *para toda matriz* A.
2. $(A\cdot B)^t = B^t \cdot A^t$, *para todo* $A \in M_{r\times m}(\mathbb{R})$ *y* $B \in M_{m\times n}(\mathbb{R})$.
3. $(A^t)^{-1} = (A^{-1})^t$, *para toda matriz cuadrada invertible* A.

Demostración. 2. El término ij de $(A\cdot B)^t$ es el término ji de $A\cdot B$, que es la fila j de A por la columna i de B, que es la columna j de A^t por la fila i de B^t, que es el término ij de $B^t \cdot A^t$.

3. Como $A \cdot A^{-1} = \operatorname{Id}$, trasponiendo $(A^{-1})^t \cdot A^t = \operatorname{Id}^t = \operatorname{Id}$, luego $(A^t)^{-1} = (A^{-1})^t$.

□

8. Proposición: *Se cumple que* $\operatorname{rango}(A) = \operatorname{rango}(A^t)$.

Demostración. Después de diversas transformaciones elementales de filas y columnas A es igual a una matriz "diagonal"

$$D = \begin{pmatrix} d_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & d_n & \dots & 0 \end{pmatrix}$$

Igualmente, tras las correspondientes transformaciones elementales de columnas y filas A^t es igual a la matriz D^t. Luego, $\operatorname{rango}(A) = \operatorname{rango}(D) = \operatorname{rango}(D^t) = \operatorname{rango}(A^t)$. $\square$

Por tanto,

$$\operatorname{rango}(A) = \operatorname{rango}(A^t) = \begin{bmatrix} \text{Número máximo de columnas de} \\ A^t \text{ linealmente independientes} \end{bmatrix}$$
$$= \begin{bmatrix} \text{Número máximo de filas de} \\ A \text{ linealmente independientes} \end{bmatrix}$$

2.11. Núcleo de una aplicación lineal

1. Definición: Sea $T\colon E \to E'$ una aplicación lineal. Se define el núcleo de T, que denotamos $\operatorname{Ker} T$, como sigue

$$\operatorname{Ker} T := \{e \in E \text{ tales que } T(e) = 0\}.$$

Se cumple que $\operatorname{Ker} T$ es un subespacio vectorial de E.

2. Ejemplo: Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ la aplicación lineal $T(x, y) := (x + 2y, 2x + 4y)$. Calculemos $\operatorname{Ker} T$:

$$\begin{aligned} \operatorname{Ker} T &= \{(x, y) \in \mathbb{R}^2 : T(x, y) = (0,0)\} = \{(x, y) \in \mathbb{R}^2 \text{ tales que } (x + 2y, 2x + 4y) = (0,0)\} \\ &= \{(x, y) \in \mathbb{R}^2 \text{ tales que } x + 2y = 0\} = \{(-2\lambda, \lambda), \forall \lambda \in \mathbb{R}\} = \langle(-2,1)\rangle. \end{aligned}$$

Dado un subespacio vectorial $V \subset E$ y $e \in E$, denotaremos $e + V := \{e + v \in E$, para todo $v \in V\}$.

3. Proposición: *Sea $T\colon E \to E'$ una aplicación lineal y $e \in E$. Entonces,*

$$e + \operatorname{Ker} T = \{w \in E, \text{ tales que } T(w) = T(e)\}.$$

Demostración. Evidentemente, dado $v \in \operatorname{Ker} T$, $T(e + v) = T(e) + T(v) = T(e)$. Recíprocamente, si $T(w) = T(e)$, entonces $0 = T(w) - T(e) = T(w - e)$ y $w - e \in \operatorname{Ker} T$. Por tanto, $w = e + (w - e) \in e + \operatorname{Ker} T$. $\square$

4. Ejemplo: Dado un sistema de ecuaciones lineales

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n &=& b_1 \\ \cdots &=& \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &=& b_m \end{array} \qquad (*)$$

consideremos la aplicación lineal $A\colon \mathbb{R}^n \to \mathbb{R}^m$ de matriz $A = (a_{ij})$ en las bases usuales de $\mathbb{R}^n$ y $\mathbb{R}^m$. Denotemos $x = (x_1, \dots, x_n)$ y $b = (b_1, \dots, b_m)$. El sistema de ecuaciones $(*)$ es $(a_{ij}) \cdot x^t = b^t$. Si $p = (p_1, \dots, p_n)$ es una solución del sistema de ecuaciones $(*)$, es decir, $(a_{ij}) \cdot p^t = b^t$, entonces el conjunto de todas las soluciones del sistema de ecuaciones $(*)$ es $p + \operatorname{Ker} A$, y

$$\operatorname{Ker} A = \{x \in \mathbb{R}^n, \text{ tales que } A \cdot x^t = 0\} = \left\{ \begin{array}{c} \text{Sol. del sist. de ecuaciones homogéneo} \\ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n &=& 0 \\ \cdots &=& \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &=& 0 \end{array} \end{array} \right\}$$

5. Proposición: *Una aplicación $T\colon E \to E'$ es inyectiva*[5] *si y solo si* $\operatorname{Ker} T = \{0\}$.

Demostración. Es consecuencia inmediata de la proposición anterior. □

6. Lema: *Sea $\{e_1, \dots, e_s\}$ una base de* $\operatorname{Ker} T$ *y sea $\{e_1, \dots, e_s, \dots, e_n\}$ una base de E. Entonces, $T(e_{s+1}), \dots, T(e_n)$ es una base de* $\operatorname{Im} T$.

Demostración. $T(e_{s+1}), \dots, T(e_n)$ generan $\operatorname{Im} T$: En efecto,

$$\operatorname{Im} T = \langle T(e_1), \dots, T(e_n) \rangle = \langle T(e_{s+1}), \dots, T(e_n) \rangle.$$

$T(e_{s+1}), \dots, T(e_n)$ son linealmente independientes: Si $\lambda_{s+1} T(e_{s+1}) + \cdots + \lambda_n T(e_n) = 0$, entonces $T(\lambda_{s+1} e_{s+1} + \cdots + \lambda_n e_n) = 0$. Luego, $e = \lambda_{s+1} e_{s+1} + \cdots + \lambda_n e_n \in \operatorname{Ker} T$ y por tanto $e = \lambda_1 e_1 + \cdots + \lambda_s e_s$, para ciertos $\lambda_i \in \mathbb{R}$. Por tanto,

$$0 = \lambda_1 e_1 + \cdots + \lambda_s e_s - (\lambda_{s+1} e_{s+1} + \cdots + \lambda_n e_n)$$

y $\lambda_i = 0$, para todo i. □

7. Proposición: *Sea $T\colon E \to E'$ una aplicación lineal. Entonces*

$$\dim E = \dim \operatorname{Ker} T + \operatorname{rango}(T)$$

8. Ejercicio: Prueba que una aplicación lineal $T\colon E \to E'$ es un isomorfismo $\iff$ $\operatorname{Ker} T = 0$ y $\operatorname{rango}(T) = \dim E'$ $\iff$ $\dim E = \operatorname{rango}(T) = \dim E'$.

[5] Es decir, la aplicación T cumple que si $T(e) = T(v)$ entonces $e = v$.

2.12. Determinante de una matriz cuadrada

El determinante de una matriz cuadrada A es un número que vamos a denotarlo por $\det(A)$. Vamos a definir el determinante de una matriz cuadrada de modo recurrente: Dada una matriz $A=(a_{ij})$ denotaremos A^{ij} la matriz que es igual a A eliminando su fila i y su columna j. Si $A=(a_{11})$ es una matriz cuadrada de orden 1, se define el determinante de A como sigue

$$\det(A) := a_{11}.$$

Dada una matriz cuadrada $A=(a_{ij})$ de orden n, se define el determinante de A como sigue

$$\det(A) := a_{11}\cdot\det(A^{11}) - a_{12}\det(A^{12}) + a_{13}\det(A^{13}) + \cdots + (-1)^{n-1}a_{1n}\cdot\det(A^{1n}).$$

Seguiremos la notación $\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$.

1. Ejemplos: Tenemos que

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a\cdot\det(d) - b\cdot\det(c) = ad - bc.$$

Igualmente,

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} = a(ei-hf) - b(di-fg) + c(dh-eg)$$

$$= aei + bfg + dhc - (gec + dbi + hfa).$$

2. Ejercicio: Calcula $\begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 1 \end{vmatrix}$.

3. Ejercicio: Prueba que $\begin{vmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{vmatrix} = d_1\cdots d_n$.

Sean $v_i = (a_{i1},\dots,a_{in}) \in \mathbb{R}^n$, para $i=1,\dots,n$. Diremos que $(v_1,\dots,v_n)$ es la matriz de filas los v_i, es decir,

$$(v_1,\dots,v_n) = (a_{ij}).$$

4. Propiedades del determinante: *Se cumple que*

1. $\det\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = 1.$

2. $\det(v_1,\dots,\lambda\cdot v_i,\dots,v_n) = \lambda\cdot\det(v_1,\dots,v_i,\dots,v_n)$, *luego para* $\lambda = 0$,
$$\det(v_1,\dots,0,\dots,v_n) = 0.$$

3. $\det(v_1,\dots,v_i+v'_i,\dots,v_n) = \det(v_1,\dots,v_i,\dots,v_n)+\det(v_1,\dots,v'_i,\dots,v_n).$

4. $\det(v_1,\overset{i}{\dots},v_i,\dots,v_j,\dots,v_n) = -\det(v_1,\overset{i}{\dots},v_j,\dots,v_i,\dots,v_n)$, *para todo* $v_1,\dots,v_n$. *En particular,*
$$\det(v_1,\overset{i}{\dots},v,\dots,v,\dots,v_n) = 0$$
para todo $v_1,\dots,v,\dots,v_n$.

Demostración. 2. Para $n=1$ es cierto: $\det(\lambda\cdot a_{11}) = \lambda\cdot a_{11} = \lambda\cdot\det(a_{11})$. Supongamos que lo hemos probado para $1,2,\dots,n-1$. Veamos que lo podemos probar para n:

Para $i=1$, tenemos que

$$\begin{aligned}\det(\lambda\cdot v_1\dots,v_i,\dots,v_n) &= \lambda\cdot a_{11}\cdot|A^{11}|+\dots+(-1)^{i-1}\lambda\cdot a_{1i}\cdot|A^{1i}|+\dots+(-1)^{n-1}\lambda\cdot a_{1n}\cdot|A^{1n}|\\ &= \lambda\cdot\det(v_1\dots,v_i,\dots,v_n).\end{aligned}$$

Para $i>1$, sea $B=(v_1\dots,\lambda\cdot v_i,\dots,v_n)$. Entonces,

$$\begin{aligned}\det(v_1\dots,\lambda\cdot v_i,\dots,v_n) &= a_{11}\cdot|B^{11}|+\dots+(-1)^{n-1}a_{1n}\cdot|B^{1n}|\\ &= a_{11}\cdot\lambda\cdot|A^{11}|+\dots+(-1)^{n-1}a_{1n}\cdot\lambda\cdot|A^{1n}|\\ &= \lambda\cdot\det(v_1,v_2,\dots,v_n).\end{aligned}$$

3. Para $n=1$ es cierto: $\det(a_{11}+a'_{11}) = a_{11}+a'_{11} = \det(a_{11})+\det(a'_{11})$. Supongamos que lo hemos probado para $1,2,\dots,n-1$. Veamos que lo podemos probar para n:

Sea $v'_i = (a'_{i1},\dots,a'_{in})$. Si $i=1$, entonces

$$\begin{aligned}\det(v_1+v'_1,\dots,v_i,\dots,v_n) &= (a_{11}+a'_{11})\cdot A^{11}+\dots+(-1)^{n-1}(a_{1n}+a'_{1n})\cdot A^{1n}\\ &= \det(v_1,\dots,v_i,\dots,v_n)+\det(v'_1,\dots,v_i,\dots,v_n).\end{aligned}$$

Si $i > 1$, sea $B = (v_1, \dots, v'_i, \dots, v_n)$ y $C = (v_1, \dots, v_i + v'_i, \dots, v_n)$, entonces

$$\begin{aligned}\det(v_1, \dots, v_i + v'_i, \dots, v_n) &= a_{11} \cdot |C^{11}| + \cdots + (-1)^{n-1} a_{1n} \cdot |C^{1n}| \\ &= a_{11} \cdot (|A^{11}| + |B^{11}|) + \cdots + (-1)^{n-1} a_{1n} \cdot (|A^{1n}| + |B^{1n}|) \\ &= \det(v_1, \overset{i}{\ldots}, v_i, \dots, v_n) + \det(v_1, \overset{i}{\ldots}, v'_i, \dots, v_n).\end{aligned}$$

4. Probemos que $\det(v_1, \overset{i}{\ldots}, v_i, \dots, v_j, \dots, v_n) = -\det(v_1, \overset{i}{\ldots}, v_j, \dots, v_i, \dots, v_n)$. Para $n = 2$ es cierto: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = -\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}$. Supongamos que lo hemos probado para $2, \dots, n-1$. Probémoslo para n: Sea $B = (v_1, \overset{i}{\ldots}, v_j, \dots, v_i, \dots, v_n)$. Si $i, j \neq 1$,

$$\det(B) = b_{11} \cdot |B^{11}| + \cdots + (-1)^{n-1} b_{1n} \cdot |B^{1n}| = -a_{11} \cdot |A^{11}| + \cdots + (-1)^n a_{1n} \cdot |A^{1n}| = -\det(A).$$

Si $i = 1$ y $j = 2$, denotemos A^{1r2s} la matriz A eliminando las filas 1 y 2 y las columnas r y s,

$$\begin{aligned}\det(A) &= \sum_{r=1}^{n} (-1)^{r-1} a_{1r} \cdot |A^{1r}| = \sum_{r=1}^{n} (-1)^{r-1} a_{1r} \cdot \big(\sum_{s<r} (-1)^{s-1} a_{2s} |A^{1r2s}| + \sum_{s>r} (-1)^s a_{2s} |A^{1r2s}|\big) \\ &= \sum_{s<r} (-1)^{r+s} (a_{1r}a_{2s} - a_{1s}a_{2r}) |A^{1r2s}| = \sum_{s<r} (-1)^{r+s} (b_{2r}b_{1s} - b_{2s}b_{1r}) |B^{1r2s}| \\ &= -\sum_{s<r} (-1)^{r+s} (b_{1r}b_{2s} - b_{1s}b_{2r}) |B^{1r2s}| = -\det(B).\end{aligned}$$

Si $i = 1$, permutemos la fila j con la $j-1$, después la $j-1$ con la $j-2$ y así sucesivamente hasta que permutamos la fila 2 con la 1, obtenemos que

$$\det(v_1, \overset{j}{\ldots}, v_j, \dots, v_n) = (-1)^{j-1} \cdot \det(v_j, v_1, v_2, \dots, v_n).$$

Igualmente, $\det(v_j, v_1, v_2, \dots, v_n) = (-1)^{j-2} \det(v_j, v_2, \dots, \underset{j}{v_1}, \dots, v_n)$. Luego,

$$\det(v_1, \overset{j}{\ldots}, v_j, \dots, v_n) = -\det(v_j, \overset{j}{\ldots}, v_1, \dots, v_n).$$

En particular, se cumple que $\det(v_1, \overset{i}{\ldots}, v, \dots, v, \dots, v_n) = -\det(v_1, \overset{i}{\ldots}, v, \dots, v, \dots, v_n)$, luego $\det(v_1, \overset{i}{\ldots}, v, \dots, v, \dots, v_n) = 0$.

□

5. Determinante de una matriz cuadrada desarrollándolo por la fila i-ésima: Sea la matriz $A = (a_{ij}) = (v_1, \dots, v_n)$ y $B = (b_{ij}) := (v_i, v_1, v_2, \dots, v_n)$, entonces

$$\begin{aligned}\det(A) &= \det(v_1, \dots, v_{i-1}, v_i, \dots, v_n) = -\det(v_1, \dots, v_i, v_{i-1}, \dots, v_n) \\ &= (-1)^2 \det(v_1, \dots, v_i, v_{i-2}, v_{i-1}, \dots, v_n) = \cdots = (-1)^{i-1} \det(v_i, v_1, v_2, \dots, v_n) \\ &= (-1)^{i-1} \cdot \sum_{r=1}^{n} (-1)^{r-1} b_{1r} \cdot |B^{1r}| = \textstyle\sum_{r=1}^{n} (-1)^{i+r} a_{ir} \cdot |A^{ir}|.\end{aligned}$$

Como consecuencia de las propiedades 2.12.4 obtenemos el siguiente corolario.

6. Corolario: $\det(v_1, \overset{i}{\ldots}, v_i + \lambda_j v_j, \ldots, v_j, \ldots, v_n) = \det(v_1, \ldots, v_n)$ *(con $i \neq j$).*

Demostración. En efecto,

$$\begin{aligned}\det(v_1, \overset{i}{\ldots}, v_i + \lambda_j v_j, \ldots, v_j \ldots, v_n) &\overset{2.12.4}{=} \det(v_1, \ldots, v_n) + \lambda_j \cdot \det(v_1, \overset{i}{\ldots}, v_j, \ldots, v_j, \ldots v_n) \\ &\overset{2.12.4}{=} \det(v_1, \ldots, v_n).\end{aligned}$$

□

2.12.1. Determinante del producto de dos matrices

Una transformación elemental especial (de filas) es la transformación por la que pasamos de la matriz $(v_1, \ldots, v_n)$ a la matriz $(v_1, \ldots, v_i + \lambda \cdot v_j, \ldots, v_n)$ (con $i \neq j$). Entonces,

"El determinante no varía por transformaciones elementales especiales (de filas)".

7. Observación: Si T es una transformación elemental especial, entonces $\det(T) = \det(T \cdot \mathrm{Id}) = \det(\mathrm{Id}) = 1$.

Si D es una matriz diagonal, entonces por la propiedad 2. de 2.12.4 $\det(D \cdot B) = \det(D) \cdot \det(B)$.

8. Proposición: *Sean $A, B \in M_n(\mathbb{R})$ dos matrices cuadradas. Entonces,*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Demostración. Por la proposición 2.8.5, $A = T_1 \cdots T_r \cdot D \cdot S_1 \cdots S_s$ y $B = T'_1 \cdots T'_{r'} \cdot D' \cdot S'_1 \cdots S'_{s'}$, donde D, D' son matrices diagonales y T_i, T'_i, S_i, S'_i transformaciones elementales especiales. Ahora es fácil probar que

$$\det(A \cdot B) = \det(T_1 \cdots T_r \cdot D \cdot S_1 \cdots S_s \cdot T'_1 \cdots T'_{r'} \cdot D' \cdot S'_1 \cdots S'_{s'}) = \det(D) \cdot \det(D') = \det(A) \cdot (B).$$

□

Si T es una transformación elemental especial, entonces, $\det(A \cdot T) = \det(A) \cdot \det(T) = \det(A)$, luego:

"El determinante de una matriz cuadrada no varía
por transformaciones elementales especiales de filas y columnas."

9. Ejemplo: Calculemos mediante transformaciones elementales especiales el determinante del ejercicio 2.12.2:

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 1 \end{vmatrix} \overset{\substack{F_2-2F_1 \\ F_3-F_1 \\ F_4-F_1}}{=} \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & -2 & 2 & 1 \end{vmatrix} \overset{F_4+2F_2}{=} \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 3 \end{vmatrix} \overset{F_4-4F_3}{=} \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

que es igual por transformaciones elementales de columnas a $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} = -1.$

10. Ejercicio: Sea A una matriz cuadrada invertible. Probad que $\det(A^{-1}) = \det(A)^{-1}$.

Resolución: $1 = \det(\mathrm{Id}) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$, luego $\det(A^{-1}) = \det(A)^{-1}$.

11. Proposición: *Sea $A \in M_n(\mathbb{R})$ una matriz cuadrada. Entonces,*

$$\det(A) = \det(A^t).$$

Demostración. La transpuesta de una transformación elemental especial es una transformación elemental especial. Por la proposición 2.8.5, A es producto de transformaciones elementales especiales por una matriz diagonal D, luego A^t es producto de transformaciones elementales especiales por $D^t = D$ y

$$\det(A) = \det(D) = \det(D^t) = \det(A^t).$$

□

2.12.2. Cálculo del rango de una matriz

Teorema: Los vectores $v_1, \dots, v_n \in \mathbb{R}^n$ son linealmente independientes si y solo si $\det(v_1, \dots, v_n) \neq 0$.

Demostración. Observemos que $v_1, \dots, v_n$ son linealmente independientes si y sólo si el rango de la matriz $(v_1, \dots, v_n)$ es n. Tanto el rango como el determinante de una matriz cuadrada no varían por transformaciones elementales especiales de filas y columnas. Por la proposición 2.8.5, podemos suponer que $(v_1, \dots, v_n)$ es una matriz (cuadrada) diagonal D. Es claro que $\text{rango}(D) = n$ si y sólo si $\det(D) \neq 0$.

□

Sea $A \in M_{n\times m}(\mathbb{R})$ una matriz de rango r. Entonces, existen r columnas linealmente independiente. Sea $B \in M_{n\times r}$ la matriz formada por estas columnas. B es una matriz de rango r, luego existen r filas de B linealmente independientes. Sea $C \in M_{r\times r}(\mathbb{R})$ la matriz formada por estas r filas linealmente independientes. Entonces, $\det(C) \neq 0$. El determinante de toda matriz cuadrada C' de orden s, con $s > r$, que se obtenga eliminando $n-s$ filas y $m-s$ columnas de A es nulo, porque s filas de A son siempre linealmente dependientes, en particular las s filas de C' son linealmente dependientes. En conclusión:

$$\text{rango}(A) = \begin{bmatrix} \text{Máximo número } r \text{ para el que existe} \\ \text{una submatriz cuadrada de } A \text{ de orden} \\ r \text{ con determinante no nulo} \end{bmatrix}$$

12. Ejercicio: Calcula el rango de la matriz $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{pmatrix}$ ¿Son los vectores $(2,4,5)$ y $(1,2,2)$ linealmente independientes?

13. Corolario: *Una matriz cuadrada $A \in M_n(\mathbb{R})$ es invertible si y solo si $|A| \neq 0$.*

Demostración. Sea $\{e_1,\dots,e_n\}$ la base usual de $\mathbb{R}^n$. Por la proposición 2.8.3, una aplicación lineal $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es un isomorfismo si y solo si $\{T(e_1),\dots,T(e_n)\}$ son linealmente independientes. Por tanto, A es invertible si y solo rango$(A) = n$, es decir, $|A| \neq 0$. □

2.12.3. Regla de Cramer

Consideremos el sistema de m ecuaciones lineales[6].

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \cdots &= \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \qquad (*)$$

de variables $x_1,\dots,x_n$ y coeficientes $a_{ij}, b_i \in \mathbb{R}$. Queremos calcular las soluciones de este sistema de ecuaciones, es decir, los valores que han de tomar las variables $x_1,\dots,x_n$ para que se cumplan las m igualdades.

[6]Escribamos $A = (a_{ij})$, $x = (x_1,\dots,x_n)$ y $b = (b_1,\dots,b_m)$, entonces el sistema de ecuaciones $(*)$ se escribe de modo reducido $A\cdot x^t = b^t$.

14. Teorema de Rouché-Frobenius: *El sistema de ecuaciones* $(*)$ *tiene solución si y solo si*

$$\operatorname{rango}\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} = \operatorname{rango}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Demostración. Denotemos $v_1 := (a_{11}, \dots, a_{m1}), \dots, v_n := (a_{1n}, \dots, a_{mn})$ y $b := (b_1, \dots, b_m)$. El sistema $(*)$ tiene solución $\iff$ existen $x_1, \dots, x_n \in \mathbb{R}$ tales que $x_1 \cdot v_1 + \cdots + x_n v_n = b$ $\iff$ $b \in \langle v_1, \dots, v_n \rangle$ $\iff$ $\langle v_1, \dots, v_n, b \rangle = \langle v_1, \dots, v_n \rangle$ $\iff$ se cumple la igualdad del teorema de Rouché-Frobenius. □

15. Definición: Se dice que un sistema de ecuaciones es compatible si existen soluciones. Se dice que un sistema de ecuaciones es compatible determinado si existe una única solución.

16. Resolución por la regla de Cramer: 1. Supongamos en el sistema de ecuaciones $(*)$ que $n = m$ y que $\det(a_{ij}) \neq 0$ (es decir, $\operatorname{rango}(a_{ij}) = n$). Por el teorema de Rouché-Frobenius el sistema es compatible. Sea $x_1 = c_1, \dots, x_n = c_n$ solución del sistema de ecuaciones. Si añadimos al sistema $(*)$ la ecuación $x_i = c_i$, entonces el nuevo sistema de ecuaciones lineales es compatible y por el teorema de Rouché-Frobenius

$$0 = \begin{vmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} & b_n \\ 0 & \dots & 1 & \dots & 0 & c_i \end{vmatrix} = (-1)^{n+i+1} \cdot \begin{vmatrix} a_{11} & \dots & \cancel{a_{1i}} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & & \vdots & \\ a_{n1} & \dots & \cancel{a_{ni}} & \dots & a_{nn} & b_n \end{vmatrix} + c_i \cdot \det(a_{ij})$$

$$= -\begin{array}{c} \begin{array}{ccccc} \phantom{a_{11}} & & i & & \phantom{a_{1n}} \end{array} \\ \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} \end{array} + c_i \cdot \det(a_{ij}).$$

Luego,

$$c_i = \frac{\begin{array}{c} \begin{array}{ccccc} \phantom{a_{11}} & & i & & \phantom{a_{1n}} \end{array} \\ \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} \end{array}}{\det(a_{ij})}.$$

2. Para resolver el sistema de ecuaciones en general se procede como sigue: Sea $A_1 = (a_{i_p j_q})_{p,q\leq r}$ una submatriz cuadrada de orden r de (a_{ij}) de determinante no nulo, con $r = \text{rango}(a_{ij})$. Escribamos $\{1,\dots,m\} = \{i_1,\dots,i_m\}$, $\{1,\dots,n\} = \{j_1,\dots,j_n\}$ y $A_2 = (a_{i_p j_q})_{1\leq p\leq r<q\leq n}$. Tenemos

$$A_1\cdot\begin{pmatrix} x_{j_1}\\ \vdots\\ x_{j_r}\end{pmatrix} + A_2\cdot\begin{pmatrix} x_{j_{r+1}}\\ \vdots\\ x_{j_n}\end{pmatrix} = \begin{pmatrix} b_{i_1}\\ \vdots\\ b_{i_r}\end{pmatrix}$$

Denotemos $\begin{pmatrix} b'_{i_1}\\ \vdots\\ b'_{i_r}\end{pmatrix} = \begin{pmatrix} b_{i_1}\\ \vdots\\ b_{i_r}\end{pmatrix} - A_2\cdot\begin{pmatrix} x_{j_{r+1}}\\ \vdots\\ x_{j_n}\end{pmatrix}$. Entonces, $A_1\cdot\begin{pmatrix} x_{j_1}\\ \vdots\\ x_{j_r}\end{pmatrix} = \begin{pmatrix} b'_{i_1}\\ \vdots\\ b'_{i_r}\end{pmatrix}$ y procedemos como en el apartado 1., para el cálculo de $x_{j_1},\dots,x_{j_r}$ (en función de $x_{j_{r+1}},\dots,x_{j_n}$).

17. Ejemplo : Resolvamos el siguiente sistema de ecuaciones lineales

$$\begin{aligned} x+y+z &= 3\\ x+y+2z &= 4\\ 3x+3y+5z &= 11 \end{aligned}$$

El rango de $\begin{pmatrix} \underline{1} & 1 & \underline{1}\\ \underline{1} & 1 & \underline{2}\\ 3 & 3 & 5\end{pmatrix}$ es 2: la tercera fila es una vez la primera más dos veces la segunda, luego el rango es menor o igual que dos; ahora bien la submatriz cuadrada donde se cortan la primera fila y segunda fila con la primera y tercera columna tiene determinante no nulo

$$\begin{vmatrix} 1 & 1\\ 1 & 2\end{vmatrix} = 1 \neq 0.$$

Igualmente, el rango de la matriz ampliada

$$\text{rango}\begin{pmatrix} 1 & 1 & 1 & 3\\ 1 & 1 & 2 & 4\\ 3 & 3 & 5 & 11\end{pmatrix} = 2.$$

Luego el sistema de ecuaciones es compatible. Tenemos que resolver el sistema

$$\begin{aligned} \underline{x}+y+\underline{z} &= 3\\ \underline{x}+y+\underline{2z} &= 4 \end{aligned}$$

que equivale a resolver

$$\begin{aligned} x+z &= 3-y\\ x+2z &= 4-y \end{aligned}$$

y por la regla de Cramer

$$x = \frac{\begin{vmatrix} 3-y & 1 \\ 4-y & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}} = 2-y, \qquad z = \frac{\begin{vmatrix} 1 & 3-y \\ 1 & 4-y \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}} = 1, \qquad y = y.$$

18. Ejemplo : Resolvamos por la regla de Cramer el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcrcr} x & + & y & + & z & + & t & = & 3 \\ x & + & y & + & z & - & t & = & 6 \\ x & + & 2y & + & z & + & 2t & = & 0 \\ 2x & + & 3y & + & 2z & + & t & = & 6 \end{array}$$

La cuarta ecuación es suma de la segunda más la tercera. Podemos eliminarla. Consideremos la submatriz formada por los coeficientes de las columnas $2,3,4$. Tenemos que

$$\text{rango}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \end{pmatrix} = 3, \text{ porque } \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \end{vmatrix} = -2 \neq 0.$$

El rango de la matriz ampliada es 3 (no puede ser mayor), luego el sistema es compatible. Tenemos

$$\begin{array}{rcl} y+z+t & = & 3-x \\ y+z-t & = & 6-x \\ 2y+z+2t & = & -x \end{array}$$

Luego,

$$y = \frac{\begin{vmatrix} 3-x & 1 & 1 \\ 6-x & 1 & -1 \\ -x & 1 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \end{vmatrix}}, \quad z = \frac{\begin{vmatrix} 1 & 3-x & 1 \\ 1 & 6-x & -1 \\ 2 & -x & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \end{vmatrix}}, \quad t = \frac{\begin{vmatrix} 1 & 1 & 3-x \\ 1 & 1 & 6-x \\ 2 & 1 & -x \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \end{vmatrix}}, \quad x = x.$$

2.12.4. Cálculo de la matriz inversa

Sea $A = (a_{ij}) \in M_n(\mathbb{R})$ una matriz cuadrada invertible con coeficientes reales y escribamos $(a_{ij})^{-1} = (b_{ij})$. Como $(a_{ij}) \cdot (b_{ij}) = \mathrm{Id}$, fijándonos en la columna j tenemos

$$\begin{array}{rcl} a_{11} \cdot b_{1j} + \cdots + a_{1n} \cdot b_{nj} & = & 0 \\ \vdots & & \vdots \\ a_{j1} \cdot b_{1j} + \cdots + a_{jn} \cdot b_{nj} & = & 1 \\ \vdots & & \vdots \\ a_{n1} \cdot b_{1j} + \cdots + a_{nn} \cdot b_{nj} & = & 0 \end{array}$$

para todo j. Por la regla de Cramer,

$$b_{ij} = \frac{\begin{array}{c} \phantom{a_{11} \dots{}} i \phantom{{}\dots a_{1n}} \\ \begin{vmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{vmatrix} \end{array}}{\det(a_{ij})} = \frac{(-1)^{i+j}|A^{ji}|}{|A|}.$$

Sea $Ad_{ij} := (-1)^{i+j}|A^{ji}|$ y $Ad := (Ad_{ij})$. Entonces, $A^{-1} = \frac{1}{|A|} \cdot Ad$. Es decir,

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} |A^{11}| & -|A^{12}| & |A^{13}| & \dots \\ -|A^{21}| & |A^{22}| & -|A^{23}| & \dots \\ |A^{31}| & -|A^{32}| & |A^{33}| & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^t$$

19. Ejemplo : Calculemos la matriz inversa de la matriz $A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 0 \\ 1 & 1 & 1 \end{pmatrix}$: Tenemos que $\det(A) = 4 + 6 - (8 + 6) = -4$, además

$$Ad = ((-1)^{i+j}|A^{ij}|)^t = \begin{pmatrix} 4 & -3 & -1 \\ 0 & -1 & 1 \\ -8 & 6 & -2 \end{pmatrix}^t = \begin{pmatrix} 4 & 0 & -8 \\ -3 & -1 & 6 \\ -1 & 1 & -2 \end{pmatrix} \text{ y } A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3/4 & 1/4 & -3/2 \\ 1/4 & -1/4 & 1/2 \end{pmatrix}.$$

2.12.5. Volumen de un paralelepípedo

Para determinar el volumen de un paralelepípedo es necesario fijar primero una unidad de volumen. Digamos que el volumen del volumen del paralelepípedo definido por los vectores de la base usual de $\mathbb{R}^n$ (y sus trasladados) es 1. Consideremos la aplicación

$$V : M_n(\mathbb{R}) \to \mathbb{R},\ V(v_1,\dots,v_n) := \left[\begin{array}{l}\text{Volumen del paralelepípedo}\\ \text{generado por } v_1,\dots,v_n\end{array}\right]$$

No hemos definido qué es volumen de un paralelepípedo, pero tenemos cierta intuición de qué es y qué propiedades debe cumplir.

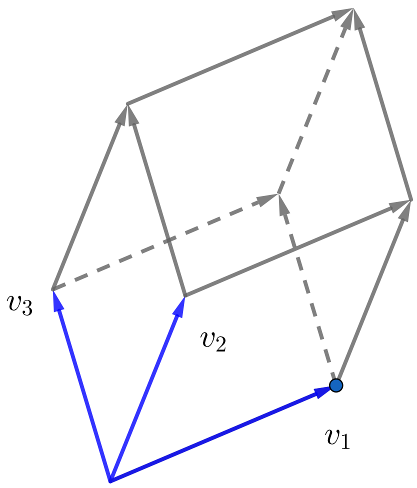
20. Proposición: *Dados $v_1,\dots,v_n \in \mathbb{R}^n$, se cumple que*[7]

$$V(v_1,\dots,v_n) = |\det(v_1,\dots,v_n)|.$$

Demostración. Tenemos la intuición de que

1. $V(v_1,\dots,\lambda v_i,\dots,v_n) = |\lambda| \cdot V(v_1,\dots,v_n)$.
2. $V(v_1,\overset{i}{\dots},v_i + \lambda_j v_j,\dots,v_j,\dots,v_n) = V(v_1,\dots,v_n)$ (con $i \neq j$).

Sea $A \in M_n(\mathbb{R})$. Si D es una matriz diagonal entonces $V(D\cdot A) = |\det(D)| \cdot V(A)$ y si T es una transformación elemental especial $V(T\cdot A) = V(A)$.



$(n = 3)$

Por la proposición 2.8.5, $A = T_1 \cdots T_r \cdot D \cdot S_1 \cdots S_s$, donde T_i, S_i son transformaciones elementales especiales y D es una matriz diagonal. Entonces,

$$V(A) = V(D) = |\det(D)| = |\det(A)|.$$

□

21. Proposición: *Sea $A : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación $\mathbb{R}$-lineal de matriz (a_{ij}) en las bases usuales de $\mathbb{R}^n$ y $\mathbb{R}^n$. Para cualesquiera $v_1,\dots,v_n \in \mathbb{R}^n$, se cumple que*

$$V(A(v_1),\dots,A(v_n)) = |\det(a_{ij})| \cdot V(v_1,\dots,v_n).$$

"La aplicación lineal A transforma paralelepípedos de volumen V en paralelepípedos de volumen $|\det(A)| \cdot V$."

[7] Los matemáticos, que somos muy escrupulosos en cuanto al rigor y sentido de las frases, diríamos que esta proposición es en realidad la definición de volumen.

Demostración. Escribamos $(v_1,\dots,v_n) = (b_{ij}) \in M_n(\mathbb{R})$. Entonces, $(A(v_1),\dots,A(v_n)) = ((a_{ij})\cdot(b_{ij})^t)^t = (b_{ij})\cdot(a_{ij})^t$ y

$$V(A(v_1),\dots,A(v_n)) = |\det((b_{ij})\cdot(a_{ij})^t)| = |\det(b_{ij})|\cdot|det(a_{ij})^t)| = |\det(a_{ij})|\cdot V(v_1,\dots,v_n).$$

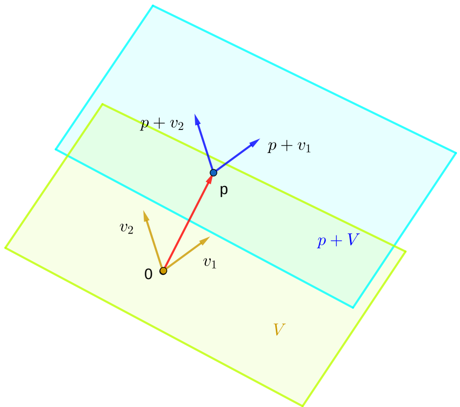
□

2.13. Subvariedades lineales

1. Definición: Sea $V \subseteq \mathbb{R}^n$ un subespacio vectorial y $p \in \mathbb{R}^n$. Diremos que el subconjunto de $\mathbb{R}^n$:

$$p+V := \{p+v \in \mathbb{R}^n, \text{ para todo } v \in V\}$$

es la subvariedad lineal de $\mathbb{R}^n$ que pasa por p y de subespacio vectorial director V.



Supongamos $V = \langle v_1,\dots,v_r\rangle$ y escribamos $v_1 = (a_{11},\dots,a_{1n}),\dots,v_r = (a_{r1},\dots,a_{rn})$ y $p = (p_1,\dots,p_n)$. .

2. Ecuaciones paramétricas: Dado $x = (x_1,\dots,x_n) \in \mathbb{R}^n$, tenemos que $x \in p+V$ si y solo si existen $\lambda_1,\dots,\lambda_r$ de modo que

$$(x_1,\dots,x_n) = (p_1,\dots,p_n) + \lambda_1(a_{11},\dots,a_{1n}) + \cdots + \lambda_r(a_{r1},\dots,a_{rn}).$$

Es decir, $x = (x_1,\dots,x_n) \in p+V$ si y solo si para ciertos $\lambda_1,\dots,\lambda_r$

$$\begin{array}{rcl} x_1 & = & p_1 + \lambda_1 a_{11} + \cdots + \lambda_r a_{r1} \\ \cdots & = & \cdots \\ x_n & = & p_n + \lambda_1 a_{1n} + \cdots + \lambda_r a_{rn} \end{array}$$

y se dice que son las "ecuaciones paramétricas" de la subvariedad lineal $p+V$.

3. Ecuaciones implícitas: Tenemos que $x \in p+V \iff x-p \in V \iff \langle x-p, v_1,\dots,v_r\rangle = \langle v_1,\dots,v_r\rangle \iff \dim\langle x-p, v_1,\dots,v_r\rangle = \dim\langle v_1,\dots,v_r\rangle$. Entonces, $x \in p+V$ si y solo si

$$\text{rango}\begin{pmatrix} x_1-p_1 & \cdots & x_n-p_n \\ a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix} = \text{rango}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix}.$$

Supongamos por sencillez que $v_1,\dots,v_r$ es una base de V y que las r-primeras columnas de (a_{ij}) son linealmente independientes, entonces $(x_1,\dots,x_n)\in p+V$ si y solo si

$$\begin{vmatrix} x_1-p_1 & \cdots & x_r-p_r & x_{r+1}-p_{r+1}\\ a_{11} & \cdots & a_{1r} & a_{1r+1}\\ \vdots & & \vdots & \vdots\\ a_{r1} & \cdots & a_{rr} & a_{rr+1}\end{vmatrix}=0,\dots,\begin{vmatrix} x_1-p_1 & \cdots & x_r-p_r & x_n-p_n\\ a_{11} & \cdots & a_{1r} & a_{1n}\\ \vdots & & \vdots & \vdots\\ a_{r1} & \cdots & a_{rr} & a_{rn}\end{vmatrix}=0$$

Es decir, $p+V$ es el conjunto de soluciones de este sistema de ecuaciones lineales, que se dice que son las "ecuaciones implícitas" de $p+V$.

El conjunto de soluciones de un sistema de ecuaciones compatible $Ax^t=b^t$ es una subvariedad lineal que pasa por p, donde p es una solución particular del sistema de ecuaciones, y de subespacio vectorial director $\operatorname{Ker} A=\{x\in\mathbb{R}^n$ tales que $Ax^t=(0,\dots,0)^t\}$,

Al calcular las soluciones del sistema de ecuaciones obtendremos las ecuaciones paramétricas de esta subvariedad lineal. Recíprocamente, si nos dan las ecuaciones paramétricas de una subvariedad lineal y "eliminamos" los parámetros, obtendremos las ecuaciones implícitas de la subvariedad lineal.

Si S y S' son subvariedades lineales (y $S\cap S'\neq\emptyset$), entonces $S\cap S'$ es una subvariedad lineal de ecuaciones implícitas la unión de las ecuaciones implicitas de S y S'.

4. Definición: Diremos que la dimensión de la subvariedad lineal $p+V\subset\mathbb{R}^n$ es $\dim V$. Diremos que $p+V$ es una recta si tiene dimensión 1, que es un plano si tiene dimensión 2 y que es un hiperplano de $\mathbb{R}^n$ si tiene dimensión $n-1$.

Si $A\cdot(x_1,\dots,x_n)^t=(b_1,\dots,b_m)^t$ son las ecuaciones (compatibles) implícitas de una subvariedad lineal S, entonces la dimensión de S es igual a la dimensión de $\operatorname{Ker} A$, y como A es una matriz de m filas y n columnas recordemos que

$$\dim S=\dim\operatorname{Ker} A=n-\operatorname{rango}(A).$$

5. Ejemplo: Calculemos las ecuaciones paramétricas e implicítas del plano $(1,1,1)+\langle(1,2,3),(1,-1,1)\rangle\subset\mathbb{R}^3$: Las ecuaciones paramétricas son

$$(x_1,x_2,x_3)=(1,1,1)+\lambda(1,2,3)+\mu(1,-1,1).$$

Es decir,

$$\begin{aligned} x_1 &= 1+\lambda+\mu\\ x_2 &= 1+2\lambda-\mu\\ x_3 &= 1+3\lambda+\mu\end{aligned}$$

Las ecuaciones implícitas son

$$\begin{vmatrix} x_1-1 & x_2-1 & x_3-1 \\ 1 & 2 & 3 \\ 1 & -1 & 1 \end{vmatrix} = 0.$$

Es decir, $5x_1 + 2x_2 - 3x_3 = 4$.

6. Ejemplo : Sea $V = \langle (a,b,c) \rangle$ $(a,b,c \neq 0)$ y $p = (p_1,p_2,p_3)$. Entonces, $(x_1,x_2,x_3) \in p+V$ si y solo si

$$\operatorname{rango}\begin{pmatrix} x_1-p_1 & x_2-p_2 & x_3-p_3 \\ a & b & c \end{pmatrix} = \dim V = 1.$$

Es decir, $\frac{x_1-p_1}{a} = \frac{x_2-p_2}{b} = \frac{x_3-p_3}{c}$.

7. Ejercicio : Probad que si $p' \in p+V$ entonces $p+V = p'+V$.

8. Ejercicio : Sea $q \in (p+V)\cap(p'+W)$. Probad que $(p+V)\cap(p'+W) = q+(V\cap W)$.

9. Ejercicio : Probad que la mínima subvariedad lineal que pasa por $p+V$ y $q+W$ es $p+(V+W+\langle q-p\rangle)$.

10. Definición : Se dice que dos subvariedades lineales $p+V$ y $q+W$ son paralelas si $V \subseteq W$ o $W \subseteq V$.

Sea S una subvariedad lineal de ecuaciones implícitas $A\cdot x^t = b^t$. Las ecuaciones implícitas de la subvariedad lineal paralela a S, de la misma dimensión que S, que pasa por un punto p, son $A\cdot x^t = c^t$, donde $c^t = A\cdot p^t$.

11. Ejemplo : Calculemos la recta paralela a los planos $x+y+z=3$ y $2x+y-z=5$, que pase por el punto $(1,1,2)$: El plano paralelo a $x+y+z=3$ que pasa por $(1,1,2)$ es el plano $x+y+z=4$; y el plano paralelo a $2x+y-z=5$ es $2x+y-z=1$. La recta buscada es la recta de ecuaciones implícitas

$$\begin{aligned} x+y+z &= 4 \\ 2x+y-z &= 1 \end{aligned}$$

2.14. Biografía de Arthur Cayley

CAYLEY BIOGRAPHY

Arthur Cayley's father Henry Cayley (1768-1850), although from a family who had lived for many generations in Yorkshire, England, worked as a merchant in St Petersburg, Russia. The family, although living in St Petersburg, returned to England for the summers and it was on such a summer visit in 1821 that Arthur Cayley was born. The family returned to live permanently in England in 1828 and took up residence in a fine house in London.

Arthur showed great skill in numerical calculations at a private school in Blackheath and, after he moved to King's College School in 1835, at age 14 rather than the usual age of entry of 16, his aptitude for advanced mathematics became apparent. He won the Chemistry Prize in each of his final two years despite not specialising in science.

In 1838 Arthur began his studies at Trinity College, Cambridge. His favourite mathematical topics were linear transformations and analytical geometry and while still an undergraduate he had three papers published in the newly founded Cambridge Mathematical Journal edited by Duncan Gregory. Cayley graduated as Senior Wrangler in 1842 and won the first Smith's prize.

For four years Cayley taught at Cambridge having won a Fellowship and, during this period, he published 28 papers in the Cambridge Mathematical Journal. He worked on a large variety of mathematical topics including algebraic curves and surfaces, elliptic functions, determinants and the theory of integration. For example he published On a theory of determinants in 1843 in which he extended the idea of a 2-dimensional determinant (rows and columns) to multidimensional arrays. Two papers he published in 1845 and 1846 are regarded as laying the foundations for invariant theory.

The Cambridge fellowship had a limited tenure, since Cayley was not prepared to take Holy Orders, so he had to find a profession. He chose law and began training in April 1846. He spent 14 years as a lawyer but Cayley, although very skilled in conveyancing (his legal speciality), always considered it as a means to make money so that he could pursue mathematics. The superabundant verbiage of legal forms was always distasteful to him. He once remarked that "the object of law was to say a thing in the greatest number of words, of mathematics to say it in the fewest."During these 14 years as a lawyer Cayley published about 250 mathematical papers. During the years 1853-1856 he averaged ten published papers per year, but as he sought to raise his profile, he published on average thirty papers per year during 1857-1860. He had not been successful in his application in 1856 for the chair of natural philosophy at Marischal College, Aberdeen. He was also interested in the Lowndean Chair of Geometry and Astronomy at Cambridge in 1858 and the chair of astronomy at Glasgow University in the following year. However, he failed to obtain both these and other positions. Perhaps one reason was that, despite a remarkable research record, he had little experience of teaching.

In 1863 Cayley was appointed Sadleirian professor of Pure Mathematics at Cambridge. . This appointment involved a very large decrease in income for Cayley. He married in 1863, He published over 900 papers and notes covering nearly every aspect of modern mathematics. The most important of his work is in developing the algebra of matrices, work on non-euclidean geometry and n-dimensional geometry. Of his work on matrices, Richard Feldmann writes: "Although the term "matrix"was introduced into mathematical literature by James Joseph Sylvester in 1850, the credit for founding the theory of matrices must be given to Arthur Cayley, since he published the first expository articles on the subject. He introduces, although quite sketchily, the ideas of inverse matrix and of matrix multiplication, or "compounding" as Cayley called it."

As early as 1849 Cayley had written a paper linking his ideas on permutations with Cauchy's. In 1854 Cayley wrote two papers which are remarkable for the insight they have of abstract groups. At that time the only known groups were permutation groups and even this was a radically new area, yet Cayley defines an abstract group and gives a table to display the group multiplication. He gives the "Cayley tables" of some special permutation groups but, much more significantly for the introduction of the abstract group concept, he realised that matrices and quaternions were groups. Cayley developed the theory of algebraic invariance, and his development of n-dimensional geometry has been applied in physics to the study of the space-time continuum. His work on matrices served as a foundation for quantum mechanics, which was developed by Werner Heisenberg in 1925. Cayley also suggested that euclidean and non-euclidean geometry are special types of geometry. He united projective geometry and metrical geometry which is dependent on sizes of angles and lengths of lines.

Article by: J.J. O'Connor and E.F. Robertson: http://mathshistory.st-andrews.ac.uk/Biographies/Cayley.html

2.15. Cálculos con ordenador

En Mathematica el vector $(a,b,c,d)\in\mathbb{C}^4$ se escribe $\{a,b,c,d\}$. La matriz $\begin{pmatrix} a & b & c \\ e & f & g \end{pmatrix}$ se escribe $\{\{a,b,c\},\{e,f,g\}\}$. Con el símbolo $*$ calcularemos el producto de matrices.

1. Para calcular una base "sencilla" de $\langle(1,1,1),(1,2,3),(3,5,7)\rangle$ por el método de la cascada de Gauss, escribe $RowReduce[\{1,1,1\},\{1,2,3\},\{3,5,7\}]$.

2. Escribe $Solve[\{x+y+z==3, x+2y+3z==6, 3x+5y+7z==15\},\{x,y,z\}]$. Escribe $NullSpace[\{\{1,2\},\{2,4\}\}]$ para calcular $\operatorname{Ker}\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$. Por otra parte, escribe en wolframalpha $x+y+z=3, x+2y+3z=6, 3x+5y+7z=15$.

3. Escribe MatrixPower[{{1,2},{1,1}},3] para calcular $\begin{pmatrix}1 & 2\\ 1 & 1\end{pmatrix}^3$. Escribe en wolframalpha {{1,2},{1,1}}^3.

4. Escribe

$$\{\{Cos[\backslash[Alpha]], -Sin[\backslash[Alpha]]\},\{Sin[\backslash[Alpha]], Cos[\backslash[Alpha]]\}\}.\{\{x\},\{y\}\}$$

   y comprueba que es igual a $\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$.

5. Usa los comandos MatrixRank[], Det[], Inverse[] y Transpose[] para calcular el rango, el determinante, la matriz inversa y la matriz transpuesta de la matriz $\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 2\\ 2 & 1 & 1\end{pmatrix}$.

2.16. Problemas

1. Sea $E' = \langle(1,2,3,4),(2,3,1,1),(5,8,5,6),(0,0,0,1)\rangle \subseteq \mathbb{R}^4$. Calcula una base de E'.

2. El arsénico AsH_3, se obtiene mediante la fórmula (no balanceada)

$$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \to HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

   Balancea la fórmula.

3. Un paciente tiene un tumor. Si el tumor es maligno y no se opera la esperanza de vida del paciente es 4 años y si se opera 22 años. Si el tumor es benigno y no se opera la esperanza de vida es 32 años y si se opera 26 años. Si la probabilidad de que el tumor sea maligno es $q = 1/3$ ¿cuál es la esperanza de vida si se opera? y ¿si no se opera? ¿A partir de qué probabilidad q es conveniente operar? (Tomado de [2]).

4. Consideremos n puntos $p_1,\ldots,p_n \in \mathbb{R}^3$ y consideremos en cada punto p_i una masa de m_i gramos. El centro de gravedad $p \in \mathbb{R}^3$ de estas n masas viene dada por la fórmula $p = \frac{m_1 \cdot p_1 + \cdots + m_n \cdot p_n}{m_1 + \cdots + m_n}$. Si $n = 3$, $p_1 = (1,0,0)$, $p_2 = (0,1,-10)$, y $p_3 = (0,0,1)$, calcula m_1, m_2 y m_3 para que el centro de gravedad de estas masas esté en el punto $(3/28, 1/14, 3/28)$ y $m_1 + m_2 + m_3 = 6$.

5. Una mezcla de dos compuestos, $NaBr$ y Na_2SO_4, contiene un 30% de Na en masa. Calcula el porcentaje en masa de cada compuesto en la mezcla.

6. Cierto compuesto orgánico contiene C, H y O. Cuando se produce la combustión de $1'57$ g del mismo se obtienen $3'00$ g de dióxido de carbono y $1'84$ g de agua. Una muestra (gaseosa) de $0'41$ g de este compuesto a 360 K y $0'97$ atm ocupa un volumen de $270'6$ cm^3. Determina la fórmula molecular del compuesto.

7. Usa la siguiente tabla (tomada de [5])

	Cantidad en gramos proporcionada		
Nutrimento	100g Leche desgrasada	100g Harina de soja	100g Suero
Proteínas	36	51	13
Carbohidratos	52	34	74
Grasas	0	7	1’1

para responder a la pregunta: ¿Cúantos gramos de leche desgrasada, harina de soja y suero he de tomar para obtener una dieta de 33 gramos de proteínas, 45 de carbohidratos y 3 de grasas?

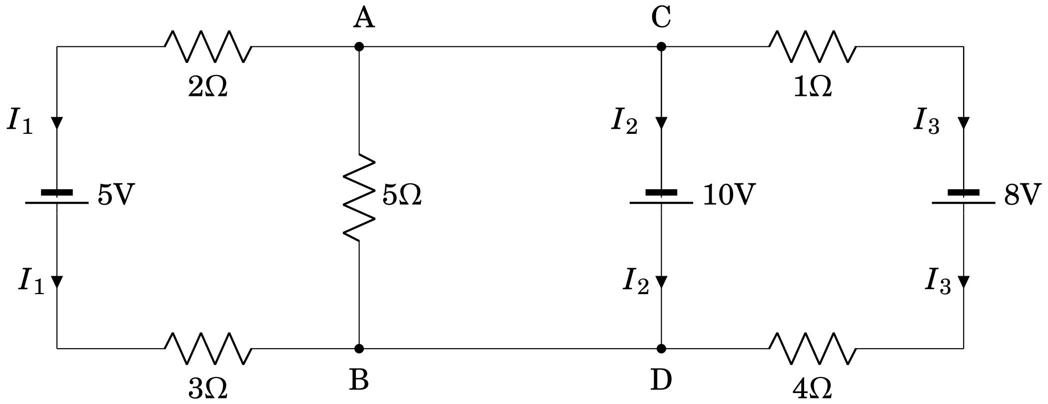
8. **Criptografía**. Consideremos la siguiente tabla, que hace corresponder a letras, signos ortográficos y palabras, números del 0 al 30,

A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
O	Q	R	S	T	U	V	W	X	Y	Z	.	,	ANA	
16	17	18	19	20	21	22	23	24	25	26	27	28	29	0

Juan escribe un mensaje, lo transcribe a números, y por último lo encripta como sigue: cada tres números consecutivos x, y, z obtiene otros tres x', y', z', donde $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Si Juan finalmente escribe $39, 5, 29, 56, 24, 13, 2, 13, -12, 10, 18, -13, 46, 34, -6$ ¿qué mensaje escribió al principio?

9. La primera ley de Kirchhoff afirma que la suma de las intensidades de las corrientes que entran en un nodo es igual a la suma de las intensidades de las

corrientes que salen. La segunda ley de Kirchhoff afirma que la suma de las diferencias de potencial eléctrico en un circuito cerrado es igual a cero. Calcula las intensidades de corriente en cada tramo del circuito



10. Diremos que una matriz N cuadrada de orden n es **nilpotente** si existe un número natural $r \geq 1$ tal que $N^r = 0$. Prueba que si N es nilpotente, entonces la matriz $\mathrm{Id} - N$ es invertible y, además:

$$(\mathrm{Id} - N)^{-1} = \mathrm{Id} + N + N^2 + \ldots + N^{r-1}.$$

Como aplicación, calcula la matriz inversa de la matriz siguiente:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

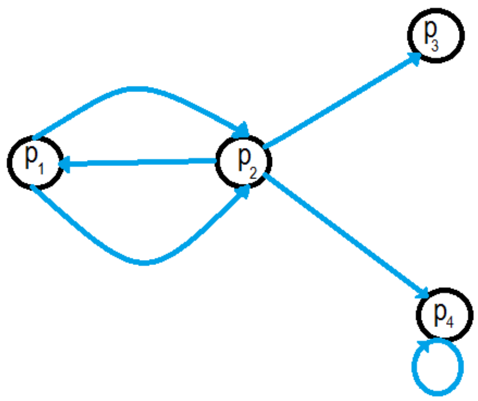
11. Modelo de Leontief (tomado de [5]) Se divide la economía de Estados Unidos de 1958, en 7 sectores: (A) productos no metálicos personales y domésticos, (B) productos metálicos finales (como vehículos de motor), (C) productos básicos de metal y minería, (D) productos básicos no metálicos y de agricultura, (E) energía, (F) servicios, y (G) entretenimiento y productos diversos. Se ha constatado que el sector A para producir bienes por un millón de dólares gasta $0'1588$ millones en bienes del sector A, $0'0057$ en el sector B, $0'0264$ en el sector C, $0'3299$ en el sector D, $0'0089$ en el sector E, $0'119$ en el sector F y $0'0063$ en el sector G,

según se muestra en la primera columna de la matriz P.

$$P=\begin{pmatrix} 0'1588 & 0'0064 & 0'0025 & 0'0304 & 0'0014 & 0'0083 & 0'1594 \\ 0'0057 & 0'2645 & 0'0436 & 0'0099 & 0'0083 & 0'0201 & 0'3413 \\ 0'0264 & 0'1506 & 0'3557 & 0'0139 & 0'0142 & 0'007 & 0'0236 \\ 0'3299 & 0'0565 & 0'0495 & 0'3636 & 0'0204 & 0'0483 & 0'0649 \\ 0'0089 & 0'0081 & 0'0333 & 0'0295 & 0'3412 & 0'0237 & 0'002 \\ 0'119 & 0'0901 & 0'0996 & 0'126 & 0'1722 & 0'2368 & 0'3369 \\ 0'0063 & 0'0126 & 0'0196 & 0'0098 & 0'0064 & 0'0132 & 0'0012 \end{pmatrix}$$

Lo equivalente expresan las demás columnas de P para los demás sectores. ¿Cuánta producción en millones de dólares tendrá que producir cada sector para satisfacer una demanda de los consumidores de 74000 millones en bienes de A, 56000 en B, 10500 en C, 25000 en D, 17500 en E, 19000, en F y 5000 en G?

12. Multidigrafo. Consideremos un conjunto de "vértices" $\{p_1,\ldots,p_n\}$ y un cierto conjunto de "flechas" que van de los vértices p_i a los p_j. Sea (a_{ij}) la matriz cuadrada de orden n de "adyacencia" definida por a_{ij} = número de flechas que van de p_j a p_i (véase el ejemplo de abajo)



Matriz de adyacencia

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Sea $(b_{ij})=(a_{ij})^r$. Demuestra que b_{ij} = número de caminos formados por r flechas que van de p_j a p_i. En el ejemplo ¿cuántos caminos formados por 4 flechas van de p_1 a p_4?

13. Una planta química fabrica dos productos A y B mediante dos máquinas I y II. La máquina I produce 2 Kg de A cada hora y produce 3 Kg del producto B cada hora. La máquina II produce 3 Kg del producto A cada hora y 4 Kg del producto B cada hora. Se dispone de 16 horas de uso de la máquina I y de 24 horas de uso de la máquina II. Por cada Kg de A fabricado se obtiene un beneficio de 40 euros y por cada Kg de B 100 euros.

La producción de B, da además un subproducto C, de modo que por cada Kg de B se obtienen 2 Kg de C. El producto C se puede vender a 30 euros el Kg, sin embargo no se puede vender más de 5 Kg y el sobrante ha de destruirse a un costo de 20 euros Kg.

¿Cuál será el plan de producción para obtener el máximo beneficio?

Problema dual en programación lineal: Para resolver el problema de maximizar $G = c\cdot x^t$ ($c \in \mathbb{R}^n$, $x = (x_1,\dots,x_n)$) sujeto a m-restricciones $A\cdot x^t \leq b^t$ (con $b \geq 0$) y $x \geq 0$, añadimos m variables de holgura y, y probamos que es equivalente a maximizar G sujeto a las ecuaciones

$$A\cdot x^t + \mathrm{Id}\cdot y^t = b^t, \text{ (con } x, y \geq 0) \qquad (*)$$

Vimos (cuando era posible) que este sistema es equivalente a un sistema de ecuaciones

$$A'\cdot x'^t + \mathrm{Id}\cdot y'^t = b'^t, \text{ (con } x', y' \geq 0) \qquad (**)$$

donde x', y' es una permutación de las variables x, y y $b' \geq 0$, y que existe una combinación lineal de las ecuaciones del sistema $(*)$ con coeficientes $l = (l_1,\dots,l_m)$, de modo que restadas a G, obtenemos $G' = (-lAx^t - ly^t + lb^t) + cx^t = c'x'^t + lb^t$ con $c' \leq 0$, es decir, $-lA + c \leq 0$, $l \geq 0$. Esto nos dice que el valor máximo de G (o G') en el conjunto de restricciones es lb^t y que los coeficientes en las variables y de G' son $-l$.
Restemos a G una combinación lineal l' de las ecuaciones del sistema $*$ (que es igual a restar a G' una combinación lineal l'' de las ecuaciones del sistema $(**)$) de modo que $(-l'\cdot A\cdot x^t - l'\cdot y^t + l'\cdot b^t) + G = d\cdot(x,y)^t + l'\cdot b^t$, con $d \leq 0$. Es decir, $-l'\cdot A + c \leq 0$ y $l' \geq 0$; por ejemplo podríamos tomar l' igual a l. Tenemos que

$$d\cdot(x,y)^t + l'\cdot b^t = (-l''\cdot A'\cdot x'^t - l''\cdot y'^t + l''\cdot b'^t) + c'\cdot x'^t + l\cdot b^t$$

luego $l'' \geq 0$ y por tanto $l'b^t = lb^t + l''b'^t \geq lb^t$. Luego, lb^t es el valor mínimo de $y\cdot b^t$ sujeto a las restricciones $y\cdot A \geq c$ e $y \geq 0$.

En conclusión, el máximo de $c\cdot x^t$ sujeto a $A\cdot x^t \leq b^t$ y $x \geq 0$ es igual al mínimo de $b\cdot y^t$ sujeto a $A^t\cdot y^t \geq c^t$ e $y \geq 0$. Además la solución óptima $y = l$ del problema dual, es el valor de las variables de holgura y que hemos obtenido al calcular la solución óptima del problema de partida.
Por último, el dual del problema dual es el problema de partida.

14. Tomado de [2]. La empresa Martín tiene dos compañías I y II. En promedio paga por ellas 5 y 10 millones de euros de impuestos, respectivamente. Para cada una de estas empresas puede declarar los ingresos reales y pagar los impuestos correspondientes, o falsificar la contabilidad y evadir el pago de impuestos. Hacienda tiene únicamente medios para investigar una de las dos compañías cada año.

Si investiga a una compañía y descubre el fraude, ésta tiene que pagar lo que corresponde de impuestos, más una multa del doble de lo defraudado. Hacienda investigará la compañía I o la compañía II con probabilidades p_1, p_2 respectivamente (con $p_1 + p_2 = 1$). La empresa no defraudará, o defraudará solo en I, o defraudará solo en II, o defraudará en I y II con probabilidades $q = (q_1, q_2, q_3, q_4)$ que desconoce Hacienda. Calcula p_1 para maximizar el mínimo de los ingresos promedio (para cada q) por recaudación.

15. Calcula las coordenadas de un vector de $\mathbb{R}^3$ respecto de la base

$$\mathscr{B}_1 = \{(1,2,3),\ (3,4,0),\ (1,1,0)\}$$

sabiendo que sus coordenadas respecto de la base $\mathscr{B}_2 = \{(1,1,0),\ (0,1,1),\ (1,0,1)\}$ son $(1,1,1)$.

16. Sea $\mathscr{B}$ una base de $\mathbb{R}^3$ y $T, S : \mathbb{R}^3 \to \mathbb{R}^3$ dos aplicaciones lineales tales que

$$T_B^B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \qquad S_B^B = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Calcula las matrices asociadas a las aplicaciones $S \circ T$ y $T \circ S$ respecto de $\mathscr{B}$ y $\mathscr{B}$.

17. Sean $\mathscr{B}_1 = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathscr{B}_2 = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $\mathscr{B}_3 = \{\mathbf{v}_1, \mathbf{v}_2\}$ tres bases de $\mathbb{R}^2$ tales que $\mathbf{u}_1 = \mathbf{e}_1$, $\mathbf{u}_2 = 2\mathbf{e}_1 + \mathbf{e}_2$, $\mathbf{v}_1 = \mathbf{e}_1$ y $\mathbf{v}_2 = \mathbf{e}_1 + 4\mathbf{e}_2$. Usando las matrices de cambio de bases, calcula las coordenadas del vector $\mathbf{u} = 2\mathbf{u}_1 + 5\mathbf{u}_2$ respecto de la base $\mathscr{B}_3$.

18. Dada la aplicación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $T(x,y,z) = (2x + y, y - z)$, calcula la matriz asociada a T

    1. en las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^2$;
    2. en las bases $\mathscr{B} = \{(1,1,1), (0,1,2), (0,2,1)\}$ de $\mathbb{R}^3$ y $\mathscr{B}' = \{(2,1), (1,0)\}$ de $\mathbb{R}^2$.

19. Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ una aplicación lineal tal que $T(1,2) = (1,1,2)$, $T(2,3) = (2,10,1)$. Calcula la matriz asociada a T en las bases $\mathscr{B} = \{(1,1), (2,3)\}$ de $\mathbb{R}^2$ y $\mathscr{B}' = \{(1,0,1), (1,1,0), (0,0,2)\}$ de $\mathbb{R}^3$.

20. Calcula el rango de la matriz

$$\begin{pmatrix} 2 & 2 & 2 & 1 & 1 & 4 \\ -1 & -1 & -3 & 0 & 2 & -1 \\ 1 & 2 & 1 & 1 & 1 & 3 \\ 3 & 1 & 2 & -2 & -1 & -1 \\ 4 & -2 & -2 & -6 & 0 & 8 \end{pmatrix}.$$

21. Sea $T:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la aplicación lineal definida como $T(x,y)=(x+y,x+y,x+y)$.

    1. Halla la matriz asociada a T en las bases usuales.
    2. Calcula bases de $\mathrm{Ker}(T)$ e $\mathrm{Im}(T)$.

22. Consideremos la aplicación lineal $T:\mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T(x,y,z)=(x+z,y+z,x+z,y+z)$$

    1. Calcula la matriz A de T respecto de las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^4$.
    2. Calcula el rango r de A y determina matrices P y Q tales que

$$Q^{-1}AP=\begin{pmatrix} \mathrm{Id}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

    3. Calcula una base de $\mathrm{Ker}(T)$ y una base de $\mathrm{Im}(T)$.

23. Sean $A \in M_n(\mathbb{R})$, $C \in M_m(\mathbb{R})$, $B \in M_{n\times m}(\mathbb{R})$ y la matriz cero $0 \in M_{m\times n}(\mathbb{R})$. Prueba que $\begin{vmatrix} A & B \\ 0 & C \end{vmatrix}=|A|\cdot|C|$.

24. Se llama **determinante de Vandermonde** de unos ciertos escalares $(x_1,\ldots,x_n)$ al determinante definido por la igualdad

$$V(x_1,\ldots,x_n)=\begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}.$$

Prueba que $V(x_1,\ldots,x_n)=\prod_{i<j}(x_j-x_i)$. Como aplicación de lo anterior prueba que se satisface la igualdad

$$\begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 2^2 & \ldots & 2^{n-1} \\ 1 & 3 & 3^2 & \ldots & 3^{n-1} \\ \ldots & \ldots & \ldots & \ddots & \ldots \\ 1 & n & n^2 & \ldots & n^{n-1} \end{vmatrix}=1!\cdot 2!\cdot\cdots\cdot(n-1)!.$$

25. Prueba que el cuadrilátero de vértices los puntos medios de los lados de otro cuadrilátero es un paralelogramo.

26. Escribe las ecuaciones paramétricas e implícitas del plano que pasa por los puntos $(1,0,0)$, $(0,0,1)$ y por el punto de intersección del plano y la recta

$$\pi \equiv x+y+z=2 \text{ y } r \equiv \left\{ \begin{array}{rcr} x & = & 1+t \\ y & = & t \\ z & = & 1+t \end{array} \right.$$

27. Sea S la subvariedad (no vacía) de $\mathbb{R}^n$ de ecuaciones implícitas

$$\begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & b_1 \\ \cdots & = & \cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n & = & b_m \end{array}$$

Prueba que un hiperplano $\pi \equiv c_1x_1+\cdots+c_nx_n = d$ pasa por S si y solo si la ecuación del hiperplano es combinación lineal de las ecuaciones de la subvariedad S. Calcula el plano de $\mathbb{R}^3$ que pasa por el punto $(1,1,1)$ y la recta de ecuaciones implícitas

$$r \equiv \left\{ \begin{array}{rcl} x+y+z & = & 2 \\ x-y-z & = & 0 \end{array} \right.$$

28. Escribe las ecuaciones implícitas de la subvariedad paralela a la subvariedad S del problema 27, de la misma dimensión que S que pase por $(p_1,\ldots,p_n)$. Calcula el plano paralelo a $x+y+z=3$ que pasa por el punto $(1,0,1)$.

29. Halla las ecuaciones de la recta que pasa por el punto $(1,1,1)$, es paralela al plano $-x+2y+z=5$ y es coplanaria con la recta $\frac{x-1}{1}=\frac{y}{2}=\frac{z}{3}$.

2.17. Solución de los problemas

P1. Tenemos

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 1 \\ 5 & 8 & 5 & 6 \\ 0 & 0 & 0 & 1 \end{array} \equiv \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 0 & -1 & -5 & -7 \\ 0 & -2 & -10 & -14 \\ 0 & 0 & 0 & 1 \end{array} \equiv \begin{array}{cccc} 1 & 0 & -7 & -10 \\ 0 & 1 & 5 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \equiv \begin{array}{cccc} 1 & 0 & -7 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}$$

Luego una base es $\{(1,0,-7,0),(0,1,5,0),(0,0,0,1)\}$.

P2. Tenemos que calcular x,y,z,t,u,v,w de modo que

$$x\cdot MnS+y\cdot As_2Cr_{10}O_{35}+z\cdot H_2SO_4 = t\cdot HMnO_4+u\cdot AsH_3+v\cdot CrS_3O_{12}+w\cdot H_2O.$$

Podemos suponer que $\boxed{t=1}$ y de la ecuación obtenemos que $\boxed{x=1}$. Igualando los moles de azufre, $1+z=3v$, los de arsénico $2y=u$, los de cromo $10y=v$, los de oxígeno $35y+4z=4+12v+w$, los de hidrógeno $2z=1+3u+2w$. Sustituyendo $\boxed{u=2y}$ y $\boxed{v=10y}$ en las demás ecuaciones tenemos

$$\begin{array}{rcl} 30y-z+0\cdot w &=& 1 \\ -85y+4z-w &=& 4 \\ -6y+2z-2w &=& 1 \end{array}$$

Sustituyendo $\boxed{z=30y-1}$. Obtenemos,

$$\begin{array}{rcl} 35y-w &=& 8 \\ 54y-2w &=& 3 \end{array}$$

Restando a la segunda fila dos veces la primera $-16y=-13$, luego $\boxed{y=\dfrac{13}{16}}$, $\boxed{w=\dfrac{327}{16}}$. Multiplicando todas las variables por 16. Podemos suponer que $x=t=16$, $y=13$, $w=327$, $u=26$, $v=130$, $z=374$.

P3. Consideremos la tabla

	Benigno	Maligno
Opera	26	22
No Opera	32	4

. Si tuviéramos n pacientes con tumor benigno y m pacientes con tumor maligno y operamos los $n+m$ pacientes, entonces el promedio de años de esperanza de vida sería

$$\frac{26\cdot n+22\cdot m}{n+m}=26\cdot\frac{n}{n+m}+22\cdot\frac{m}{n+m}.$$

Si no los operamos, el promedio de años de esperanza de vida promedio sería

$$\frac{32\cdot n+4\cdot m}{n+m}=32\cdot\frac{n}{n+m}+4\cdot\frac{m}{n+m}.$$

Datos que podríamos haberlos expresado con la operación: $\begin{pmatrix}26 & 22\\ 32 & 4\end{pmatrix}\begin{pmatrix}n/(n+m)\\ m/(n+m)\end{pmatrix}$. Observemos que $\frac{n}{n+m}$ es la probabilidad de que uno de los pacientes tenga un tumor benigno y $\frac{m}{n+m}$ es la probabilidad de que uno de los pacientes tenga un tumor maligno. Tenemos que $\begin{pmatrix}26 & 22\\ 32 & 4\end{pmatrix}\begin{pmatrix}2/3\\ 1/3\end{pmatrix}=\begin{pmatrix}24'6\\ 22'6\end{pmatrix}$, es decir $24'6$ años es la esperanza de vida si se opera y $22'6$ años si no se opera. $\begin{pmatrix}26 & 22\\ 32 & 4\end{pmatrix}\begin{pmatrix}1-q\\ q\end{pmatrix}=\begin{pmatrix}26-4q\\ 32-28q\end{pmatrix}$ y $26-4q\geq 32-28q$ si y solo si $q\geq 1/4$.

P4. El enunciado nos dice que $p^t = \frac{1}{m_1+\cdots+m_m}\cdot(p_1^t,\ldots,p_n^t)\cdot(m_1,\ldots,m_n)^t$. En el caso pedido

$$\begin{pmatrix}3/28\\1/14\\3/28\end{pmatrix} = \frac{1}{6}\cdot\begin{pmatrix}1&0&0\\0&1&0\\0&-10&1\end{pmatrix}\cdot\begin{pmatrix}m_1\\m_2\\m_3\end{pmatrix}.$$

Luego, $\begin{pmatrix}m_1\\m_2\\m_3\end{pmatrix} = 6\cdot\begin{pmatrix}1&0&0\\0&1&0\\0&-10&1\end{pmatrix}^{-1}\cdot\begin{pmatrix}3/28\\1/14\\3/28\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1&0\\0&10&1\end{pmatrix}\cdot\begin{pmatrix}9/14\\3/7\\9/14\end{pmatrix} = \begin{pmatrix}9/14\\3/7\\69/14\end{pmatrix}.$

P5. Podemos suponer que tenemos 100 gramos de mezcla. Sea x el porcentaje (o gramos) de $NaBr$ en la mezcla, e y el porcentaje (o gramos) de Na_2SO_4 en la mezcla. Entonces, $x+y=100$ y hay 30 gramos de Na en la mezcla. El peso atómico de Na es 23, el de Br es 80, el de O es 16 y el de S es 32. El número de moles de $NaBr$ es $\frac{x}{23+80}$, el de Na_2SO_4 es $\frac{y}{2\cdot 23+32+4\cdot 16}$ y el de Na es $\frac{30}{23}$. Por tanto, $\frac{x}{23+80}+\frac{2y}{2\cdot 23+32+4\cdot 16}=\frac{30}{23}$. Tenemos que resolver el sistema de ecuaciones lineales

$$\begin{array}{rcl} x+y &=& 100\\ \frac{x}{103}+\frac{y}{71} &=& \frac{30}{23}\end{array}$$

Por tanto, $y(\frac{1}{71}-\frac{1}{103})=\frac{30}{23}-\frac{100}{103}$, luego $y=76'2$ y $x=23'8$.

P6. La fórmula química de la combustión del compuesto $C_aH_bO_c$ (con $a,b,c\in\mathbb{N}$) es

$$C_aH_bO_c+\frac{4a+b-2c}{4}O_2\longrightarrow a\cdot CO_2+\frac{b}{2}H_2O$$

En $1'57$ gramos de $C_aH_bO_c$ hay $1'57/(12a+b+16c)$ moles de $C_aH_bO_c$. Por tanto,

$$\frac{1'57}{12a+b+16c}\cdot a\cdot(12+32)=3 \;\text{ y }\; \frac{1'57}{12a+b+16c}\cdot\frac{b}{2}\cdot(2+16)=1'84$$

Por otra parte, los $270'6$ cm^3 de gas son en condiciones normales (0 grados centígrados y una atmósfera) $270'6\cdot\frac{273}{360}\cdot\frac{0'97}{1}$ cm^3, que son $270'6\cdot\frac{273}{360}\cdot\frac{0'97}{1}\frac{1}{22140}$ moles de $C_aH_bO_c$ cuya masa es $270'6\cdot\frac{273}{360}\cdot\frac{0'97}{1}\frac{12a+b+16c}{22140}=0'41$. Tenemos que resolver el sistema de ecuaciones lineales

$$\begin{array}{rcl} -11'02a+b+16c &=& 0\\ 12a-6'6b+16c &=& 0\\ 12a+b+16c &=& 45'60\end{array}$$

Restando a la primera y segunda ecuación la tercera obtenemos $23'02a=45'60$ y $7'6b=45'60$, luego $a\simeq 2$, $b\simeq 6$ y $c\simeq 1$.

P7. Tenemos que resolver el sistema de ecuaciones lineales

$$\begin{aligned} 36x+51y+13z &= 33\\ 52x+34y+74z &= 45\\ 7y+1'1z &= 3 \end{aligned}$$

Entonces, $x\cdot 100$ serán los gramos de leche desgrasada, $y\cdot 100$ los de harina de soja y $z\cdot 100$ los de suero.

$$\begin{matrix} 36 & 51 & 13 & 33\\ 52 & 34 & 74 & 45\\ 0 & 7 & 1'1 & 3 \end{matrix} \equiv \begin{matrix} 36 & 51 & 13 & 33\\ 0 & -39'6 & 55'2 & -2'6\\ 0 & 7 & 1'1 & 3 \end{matrix} \equiv \begin{matrix} 36 & 51 & 13 & 33\\ 0 & -39'6 & 55'2 & -2'6\\ 0 & 0 & 10'8 & 2'54 \end{matrix}$$

Luego $z=0'23$, $y=0'39$, $x=0'27$.

P8. Desencriptemos y traduzcamos:

$$\begin{pmatrix} 1 & 0 & 2\\ 0 & 1 & 1\\ 1 & -1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 39 & 56 & 2 & 10 & 46\\ 5 & 24 & 13 & 18 & 34\\ 29 & 13 & -12 & -13 & -6 \end{pmatrix} = \begin{pmatrix} 29 & 18 & 0 & 0 & 10\\ 0 & 5 & 12 & 13 & 16\\ 5 & 19 & 1 & 5 & 18 \end{pmatrix}$$

"29, 0, 5, 18, 5, 19, 0, 12, 1, 0, 13, 5, 10, 16, 18" = "ANA ERES LA MEJOR".

P9. Del generador izquierdo sale una corriente de intensidad I_1 (y llega con intensidad I_1). Del generador central sale una corriente de intensidad I_2 (y llega con intensidad I_2). Del generador derecho sale una corriente de intensidad I_3 (y llega con intensidad I_3). La intensidad de corriente que llega a D es I_2+I_3, luego del tramo DB es I_2+I_3 y en el tramo BA es $I_1+I_2+I_3$. Si aplicamos la segunda ley de Kirchhoff (y la ley de Ohm) a los tres "recuadros" cerrados obtenemos

$$\begin{aligned} (2+3+5)I_1+5I_2+5I_3 &= 5\\ 5I_1+5I_2+5I_3 &= 10\\ (4+1)I_3 &= 8-10 \end{aligned}$$

La solución es $I_1=-1$, $I_2=\frac{17}{15}$ e $I_3=\frac{-2}{5}$.

P10. Observemos que $\mathrm{Id}-N$ es una matriz cuadrada y que

$$(\mathrm{Id}-N)\cdot(\mathrm{Id}+N+\ldots+N^{r-1}) = \mathrm{Id}+N+\ldots+N^{r-1}-(N+N^2+\ldots+N^{r-1}+N^r)=\mathrm{Id}.$$

Por último, $N=A-\mathrm{Id}$ es nilpotente y

$$A^{-1}=(\mathrm{Id}+N)^{-1}=\mathrm{Id}-N+N^2-N^3+N^4=\begin{pmatrix} 1 & -2 & 1 & 0 & 0\\ 0 & 1 & -2 & 1 & 0\\ 0 & 0 & 1 & -2 & 1\\ 0 & 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

P11. Sea $x=(x_1,\dots,x_7)\in\mathbb{R}^7$ la solución, es decir, la economía ha de producir x_1 millones de dólares en bienes de A, x_2 millones de dólares de bienes de B, etc. Entonces, Px^t expresa los millones de dólares que se han gastado en el proceso de producción en los distintos sectores Sea $d=(d_1,\dots,d_7)\in\mathbb{R}^7$ la demanda de los consumidores, es decir, $d_1=74000$, $d_2=56000$, etc. Luego $x^t=Px^t+d^t$ y $(\mathrm{Id}-P)x^t=d^t$. Realizando los cálculos con ayuda del ordenador obtenemos que

$$x^t=(\mathrm{Id}-P)^{-1}d^t=(95565'7,88564'5,45455'9,109155,39308'3,87531'1,10097'7)^t.$$

P12. Supóngase por hipótesis de inducción que si $(c_{ij})=(a_{ij})^{r-1}$ entonces $c_{ij}=$ número de caminos formados por $r-1$ flechas que van de p_j a p_i. Conclúyase.

En el ejemplo, $(a_{ij})^4=\begin{pmatrix}4&0&0&0\\0&4&0&0\\4&0&0&0\\6&3&0&1\end{pmatrix}$. Luego la respuesta es 6.

P13. Sean a_1 y b_1 los Kg de A y B producidos por I, respectivamente. Sean a_2 y b_2 los Kg de de A y B producidos por II, respectivamente. Sea

$$z:=\left\{\begin{array}{ll}0, & \text{si } 5-2(b_1+b_2)\geq 0\\ -5+2(b_1+b_2), & \text{si } -5+2(b_1+b_2)\geq 0\end{array}\right\}=\min\{z:\, z\geq 0 \text{ y } z\geq -5+2(b_1+b_2)\}.$$

El beneficio G obtenido es

$$G=40(a_1+a_2)+100(b_1+b_2)+30\cdot 2\cdot(b_1+b_2)-50\cdot z=40a_1+40a_2+160b_1+160b_2-50z$$

Tenemos que maximizar G sujeto a las restricciones

$$\begin{array}{rcl}2a_1+3b_1&\leq&16\\3a_2+4b_2&\leq&24\\z+5-2(b_1+b_2)&\geq&0\\a_1,b_1,a_2,b_2,z&\geq&0\end{array}\sim\begin{array}{rcl}2a_1+3b_1&\leq&16\\3a_2+4b_2&\leq&24\\-z+2b_1+2b_2&\leq&5\\a_1,b_1,a_2,b_2,z&\geq&0\end{array}\sim\begin{array}{rcl}2a_1+3b_1+t_1&=&16\\3a_2+4b_2+t_2&=&24\\-z+2b_1+2b_2+t_3&=&5\\a_1,\dots,t_1,t_2,t_3&\geq&0\end{array}$$

Por programación lineal (ordenamos las variables: $a_1,b_1,a_2,b_2,z,t_1,t_2,t_3$):

$$\begin{array}{cccccccc|c}2&3&0&0&0&1&0&0&16\\0&0&3&4&0&0&1&0&24\\0&2&0&2&-1&0&0&1&5\\\hline 40&160&40&160&-50&0&0&0&\end{array}\sim\begin{array}{cccccccc|c}4&0&0&-6&3&2&0&-3&17\\0&0&3&4&0&0&1&0&24\\0&2&0&2&-1&0&0&1&5\\\hline 40&0&40&0&30&0&0&-80&\end{array}$$

$$\begin{array}{cccccccc|c}4&0&0&-6&3&2&0&-3&17\\0&0&3&4&0&0&1&0&24\\0&2&0&2&-1&0&0&1&5\\\hline 0&0&40&60&0&-20&0&-50&\end{array}\sim\begin{array}{cccccccc|c}4&6&0&0&0&2&0&0&32\\0&-4&3&0&2&0&1&-2&14\\0&2&0&2&-1&0&0&1&5\\\hline 0&-60&40&0&30&-20&0&-80&\end{array}$$

$$
\begin{array}{cccccccc|c}
4 & 6 & 0 & 0 & 0 & 2 & 0 & 0 & 32\\
0 & -4 & 3 & 0 & 2 & 0 & 1 & -2 & 14\\
0 & 2 & 0 & 2 & -1 & 0 & 0 & 1 & 5\\
\hline
0 & -20 & 0 & 0 & 10 & -60 & -40 & -160 &
\end{array}
\sim
\begin{array}{cccccccc|c}
4 & 6 & 0 & 0 & 0 & 2 & 0 & 0 & 32\\
0 & -4 & 3 & 0 & 2 & 0 & 1 & -2 & 14\\
0 & 0 & 3 & 4 & 0 & 0 & 1 & 0 & 24\\
\hline
0 & 0 & -15 & 0 & 0 & -60 & -45 & -150 &
\end{array}
$$

El máximo lo alcanza en $a_2 = t_1 = t_2 = t_2 = 0$ y cualesquiera valores de las demás variables que cumplan el sistema de ecuaciones: $a_1 = 8 - 3/2 \cdot b_1$, $z = 7 + 2b_1$, $b_2 = 6$. Por ejemplo, $b_1 = 0$, $a_1 = 8$, $z = 7$ y $b_2 = 6$. En todos, los casos $G = 930$.

P14. Escribamos los ingresos de Hacienda en cada uno de los cuatro casos en los que puede incurrir la empresa:

	No Defrauda	Defrauda I	Defrauda II	Defrauda I y II
Hacienda investiga I	15	25	5	15
Hacienda investiga II	15	10	35	30

Si A es la matriz de números de la tabla de encima y $p = (p_1, p_2)$, entonces $p \cdot A$ son los ingresos de Hacienda en cada uno de los casos. Si la empresa defrauda con probabilidades $q = (q_1, q_2, q_3, q_4)$ en cada uno de los cuatro casos, entonces los ingresos de Hacienda serían en promedio $Ing(p,q) = p \cdot A \cdot q^t$. Sea $v(p) = \text{mín}\{Ing(p,q), \forall\, q\}$. Buscamos $p = (p_1, p_2)$ (con $p_1, p_2 \geq 0$ y $p_1 + p_2 = 1$) tal que $v(p)$ sea máximo.

Observemos que

$$Ing(p,q) = (15p_1 + 15p_2)q_1 + (25p_1 + 10p_2)q_2 + (5p_1 + 35p_2)q_3 + (15p_1 + 30p_2)q_4.$$

Por tanto, $v(p) = \text{mín}\{15p_1 + 15p_2, 25p_1 + 10p_2, 5p_1 + 35p_2, 15p_1 + 30p_2\}$. En conclusión, tenemos que hallar el máximo de v sujeto a las restricciones

$$
\left.\begin{array}{rcl}
v & \leq & 15p_1 + 15p_2\\
v & \leq & 25p_1 + 10p_2\\
v & \leq & 5p_1 + 35p_2\\
v & \leq & 15p_1 + 30p_2
\end{array}\right.
\qquad (p_1 + p_2 = 1,\ p_1, p_2 \geq 0).
$$

Que equivale a hallar el mínimo de $1/v$ sujeto a las restricciones

$$
\left.\begin{array}{rcl}
1 & \leq & 15p_1/v + 15p_2/v\\
1 & \leq & 25p_1/v + 10p_2/v\\
1 & \leq & 5p_1/v + 35p_2/v\\
1 & \leq & 15p_1/v + 30p_2/v
\end{array}\right.
\qquad (p_1 + p_2 = 1, p_1, p_2 \geq 0).
$$

Definamos $x_1 := p_1/v$ y $x_2 := p_2/v$. Entonces, $1/v = x_1 + x_2$ y tenemos que hallar su mínimo sujeto a las restricciones

$$x \cdot A^t \geq (1,1,1,1),\ (x \geq 0). \qquad \text{Donde } A = \begin{pmatrix} 15 & 25 & 5 & 15\\ 15 & 10 & 35 & 30 \end{pmatrix}.$$

Consideremos el problema dual. Calculemos el máximo de $G = y_1 + y_2 + y_3 + y_4$ sujeto a las restricciones $A \cdot y^t \leq (1,1)^t$, $y \geq 0$. Consideremos las variables de holgura $z = (z_1, z_2)$ y calculemos el máximo de $G = y_1 + y_2 + y_3 + y_4$ sujeto a las restricciones $A \cdot y^t + z^t = (1,1)^t$, $y \geq 0$. Por el método simplex obtenemos

$$\begin{array}{cccccc|c} 15 & 25 & 5 & 15 & 1 & 0 & 1 \\ 15 & 10 & 35 & 30 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 0 & 0 & \end{array} \sim \begin{array}{cccccc|c} 1 & 5/3 & 1/3 & 1 & 1/15 & 0 & 1/15 \\ 0 & -15 & 30 & 15 & -1 & 1 & 0 \\ \hline 0 & -2/3 & 2/3 & 0 & -1/15 & 0 & \end{array}$$

$$\sim \begin{array}{cccccc|c} 1 & 11/6 & 0 & 5/6 & 7/90 & -1/90 & 1/15 \\ 0 & -1/2 & 1 & 1/2 & -1/30 & 1/30 & 0 \\ \hline 0 & -1/3 & 0 & -1/3 & -2/45 & -1/45 & \end{array}$$

Por tanto, $y = (1/15, 0, 0, 0)$ y el valor máximo de G es $1/15 + 0 + 0 + 0$, y el valor mínimo de $1/v$ es $1/15$, es decir, el valor máximo de v es 15. Por otra parte, $x_1 = 2/45$ y $x_2 = 1/45$, es decir, $p_1 = v \cdot x_1 = 2/3$ y $p_2 = v \cdot x_2 = 1/3$.

Si el problema que nos planteásemos ahora fuese calcular q para minimizar el máximo de los pagos promedio (para cada p) a Hacienda de la empresa Martín, tendremos que éste es justamente el problema dual: El pago promedio es $Ing(p,q) = p \cdot A \cdot q^t$ y tenemos que minimizar $w(q) = \max\{Ing(p,q), \forall p\}$. Razonando del mismo modo que antes, tendremos que hallar el máximo de $1/w = y_1 + y_2 + y_3 + y_4$ sujeto a las restricciones $(1,1) \geq y \cdot A$, $y \geq 0$. El máximo de $1/w$ coincide con el mínimo de $1/v$. Si q es donde w alcanza el mínimo y p donde v alcanza el máximo, entonces $\max\{Ing(p', q), \forall p'\} = w(q) = v(p) = \min\{Ing(p, q'), \forall q'\}$.

P15. Sea B la base usual de $\mathbb{R}^3$. Tenemos $v_{B_2} = (1,1,1)$. Entonces,

$$v^{B_1} = \mathrm{Id}_B^{B_1} \cdot \mathrm{Id}_{B_2}^{B} \cdot v^{B_2} = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -2/3 \\ 10/3 \end{pmatrix}.$$

P16. Tenemos $(S \circ T)_B^B = S_B^B \cdot T_B^B = \begin{pmatrix} 5 & 4 & 3 \\ 8 & 6 & 6 \\ 3 & 2 & 3 \end{pmatrix}$ y $(T \circ S)_B^B = T_B^B \cdot S_B^B = \begin{pmatrix} 3 & 7 & 2 \\ 2 & 8 & 3 \\ 3 & 9 & 3 \end{pmatrix}$.

P17. Tenemos $\mathbf{u}_{B_2} = (2,5)$. La matriz de cambio de base de $\mathscr{B}_2$ a $\mathscr{B}_1$ es $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ y la matriz de cambio de base de $\mathscr{B}_3$ a $\mathscr{B}_1$ es $\begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix}$. Entonces,

$$\mathbf{u}^{B_3} = \mathrm{Id}_{B_1}^{B_3} \cdot \mathrm{Id}_{B_2}^{B_1} \cdot \mathbf{u}^{B_2} = \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 43/4 \\ 5/4 \end{pmatrix}.$$

P18. La matriz asociada a T en las bases usuales U_1, U_2 es $T_{U_1}^{U_2} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$. Entonces,

$$T_B^{B'} = \mathrm{Id}_{U_2}^{B'} \cdot T_{U_1}^{U_2} \cdot \mathrm{Id}_B^{U_1} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ 3 & 3 & 0 \end{pmatrix}.$$

P19. Sea $C = \{(1,2),(2,3)\}$ una base de $\mathbb{R}^2$, U, U' las bases usuales de $\mathbb{R}^2$ y $\mathbb{R}^3$. Tenemos $T_C^{U'} = \begin{pmatrix} 1 & 2 \\ 1 & 10 \\ 2 & 1 \end{pmatrix}$. Entonces $T_U^{U'} = T_C^{U'} \cdot \mathrm{Id}_U^C = \begin{pmatrix} 1 & 2 \\ 1 & 10 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1}$ y

$$T_B^{B'} = \mathrm{Id}_{U'}^{B'} \cdot T_U^{U'} \cdot \mathrm{Id}_B^U = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 \\ 1 & 10 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -8 & 8 \\ 9 & -7 \\ 7/2 & -3/2 \end{pmatrix}.$$

P20. Tenemos

$$\begin{array}{cccccc} 2 & 2 & 2 & 1 & 1 & 4 \\ -1 & -1 & -3 & 0 & 2 & -1 \\ 1 & 2 & 1 & 1 & 1 & 3 \\ 3 & 1 & 2 & -2 & -1 & -1 \\ 4 & -2 & -2 & -6 & 0 & 8 \end{array} \equiv \begin{array}{cccccc} 0 & 0 & -4 & 1 & 5 & 2 \\ -1 & -1 & -3 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 3 & 2 \\ 0 & -2 & -7 & -2 & 5 & -4 \\ 0 & -6 & -14 & -6 & 8 & 4 \end{array} \equiv \begin{array}{ccccc} 0 & -4 & 1 & 5 & 2 \\ 1 & -2 & 1 & 3 & 2 \\ -2 & -7 & -2 & 5 & -4 \\ -6 & -14 & -6 & 8 & 4 \end{array}$$

$$\begin{array}{ccccc} 0 & -4 & 1 & 5 & 2 \\ 1 & -2 & 1 & 3 & 2 \\ 0 & -11 & 0 & 11 & 0 \\ 0 & -26 & 0 & 26 & 16 \end{array} \equiv \begin{array}{cccc} -4 & 1 & 5 & 2 \\ -11 & 0 & 11 & 0 \\ -26 & 0 & 26 & 16 \end{array} \equiv \begin{array}{ccc} -11 & 11 & 0 \\ -26 & 26 & 16 \end{array}$$

Luego el rango es $(1+1+1)+2 = 5$.

P21. La matriz asociada a T es $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$. $\operatorname{Ker} T = \langle(1,-1)\rangle$, $\operatorname{Im} T = \langle(1,1,1)\rangle$.

P22. Tenemos $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. $\operatorname{Ker} T = \langle(-1,-1,1)\rangle$, $\operatorname{Im} T = \langle(1,0,1,0),(0,1,0,1)\rangle$. Por lo tanto, $r = 2$ y $P = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ y $Q = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$.

P23. $\begin{vmatrix} A & B \\ 0 & C \end{vmatrix} = \left|\begin{pmatrix} \mathrm{Id} & B \\ 0 & C \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & \mathrm{Id} \end{pmatrix}\right| = \begin{vmatrix} \mathrm{Id} & B \\ 0 & C \end{vmatrix} \cdot \begin{vmatrix} A & 0 \\ 0 & \mathrm{Id} \end{vmatrix} = \begin{vmatrix} \mathrm{Id} & 0 \\ 0 & C \end{vmatrix} \cdot |A| = |C| \cdot |A|$.

P24. Tras sucesivas transformaciones elementales

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} \overset{F_n - x_1 F_{n-1}}{=} \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

$$\overset{F_{n-1} - x_1 F_{n-2}}{=} \dots \overset{F_2 - x_1 F_1}{=} \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & x_2 - x_1 & \dots & x_n - x_1 \\ 0 & x_2(x_2 - x_1) & \dots & x_n(x_n - x_1) \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

$$= (x_2 - x_1)\cdots(x_n - x_1) V(x_2, \dots, x_n).$$

Por recurrencia $V(x_2,\dots,x_n) = \prod_{1<i<j}(x_j - x_i)$, luego $V(x_1,\dots,x_n) = \prod_{i<j}(x_j - x_i)$. Si utilizamos esta fórmula en el caso que sigue, obtenemos

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^{n-1} \\ 1 & 3 & 3^2 & \dots & 3^{n-1} \\ \dots & \dots & \dots & \ddots & \dots \\ 1 & n & n^2 & \dots & n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \\ 1^2 & 2^2 & \dots & n^2 \\ \vdots & \vdots & & \vdots \\ 1^{n-1} & 2^{n-1} & \dots & n^{n-1} \end{vmatrix} = n!(n-1)!\cdots 2!1.$$

P26. Sean e_1, e_2, e_3, e_4 los cuatro vértices consecutivos de un cuadrilátero. Entonces los puntos medios consecutivos de los lados del cuadrilátero son $v_1 = \frac{e_1+e_2}{2}$, $v_2 = \frac{e_2+e_3}{2}$, $v_3 = \frac{e_3+e_4}{2}$, $v_4 = \frac{e_4+e_1}{2}$. Obviamente, $v_2 - v_1 = \frac{e_3-e_1}{2} = v_3 - v_4$ y $v_3 - v_2 = \frac{e_4-e_1}{2} = v_4 - v_1$.

P26. Calculemos $\pi \cap r$: Buscamos t tal que $(1+t)+(t)+(1+t)=2$, que es $t=0$. Luego el punto de corte es $(1,0,1)$. La ecuación del plano es

$$\begin{vmatrix} x-1 & y & z \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 0,$$

que es $y=0$. Las ecuaciones paramétricas del plano son: $x=\lambda_1$, $y=0$, $z=\lambda_2$.

P27. Si la ecuación de π es combinación lineal de las de S, obviamente π pasa por S. Si la ecuación de π no es combinación lineal de las de S (y $S \cap \pi \neq \emptyset$), entonces

$$\dim S = n - \text{rango}(a_{ij}|b_i) > n - \text{rango}\left(\begin{array}{c|c} a_{ij} & b_i \\ c_j & d \end{array}\right) = \dim(S \cap \pi).$$

Luego, $S \cap \pi \underset{\neq}{\subset} S$ y π no pasa por S.

Tenemos que calcular λ y μ de modo que el plano $\lambda \cdot (x+y+z-2) + \mu \cdot (x-y-z) = 0$ pase por $(1,1,1)$, es decir, $\lambda - \mu = 0$ y podemos tomar $\lambda = \mu = 1$. El plano, $x-1=0$ pasa por r y $(1,1,1)$.

P28. El espacio vectorial director de S son las soluciones del sistema de ecuaciones homogéneo

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & 0 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & 0 \end{array}$$

Si S' es una variedad paralela a S, de la misma dimensión, que pasa por $(p_1, \ldots, p_n)$ entonces sus ecuaciones implícitas son

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b'_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b'_m \end{array}$$

donde $b'_i = a_{i1}p_1 + \cdots + a_{in}p_n$, para todo i. El plano paralelo a $x+y+z=3$ que pasa por $(1,0,1)$ es $x+y+z=1+0+1=2$.

P29. Sea $\pi \equiv \lambda(2(x-1)-y) + \mu(3y-2z) = 0$ un plano que pasa por la recta dada y el punto $(1,1,1)$, luego $-\lambda + \mu = 0$ y $\pi \equiv 2x+2y-2z=2$. Y la recta buscada es de ecuaciones implícitas $x+y-z=1$, $-x+2y+z=2$.

Capítulo 3

Operadores lineales. Diagonalización

3.1. Introducción

Los giros y simetrías del plano, derivar en el conjunto de las funciones infinitamente derivables y otras muchas aplicaciones son aplicaciones lineales de un espacio vectorial en sí mismo, y se denominan operadores lineales. El estudio de los operadores se aplicará al estudio de los movimientos en $\mathbb{R}^n$, a la resolución de ecuaciones diferenciales, a la resolución de ecuaciones en diferencias finitas, etc.

Para el estudio de los operadores lineales será fundamental el encontrar bases en las que la matriz asociada al operador sea muy sencilla. Para cada operador T buscaremos aquellas bases $\{e_1, \dots, e_n\}$ tales que $T(e_1) = \lambda_1 \cdot e_1, \dots, T(e_n) = \lambda_n \cdot e_n$. Para ello hablaremos del polinomio característico de T, de autovectores y autovalores.

Muchos procesos o transformaciones son operadores lineales, y la reiteración n-ésima de estas transformaciones se corresponde con la potencia n-ésima del operador lineal. De nuevo, para su cálculo y para la predicción de lo que sucederá cuando $n \to \infty$ será conveniente encontrar bases donde la matriz asociada al operador lineal sea sencilla.

3.2. Máximo común divisor de dos polinomios

Diremos que el polinomio $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{R}[x]$, con $a_n \neq 0$, tiene grado n y escribiremos

$$\operatorname{gr}(p(x)) = n.$$

Si $a_n = 1$ diremos que $p(x)$ es mónico.

1. Definición: Llamaremos máximo común divisor de dos polinomios $p(x), q(x) \neq 0 \in \mathbb{R}[x]$, y lo denotaremos por $m.c.d.(p(x), q(x))$, al polinomio mónico de grado máximo que divide a $p(x)$ y $q(x)$.

Vamos a desarrollar el algoritmo de Euclides para calcular el máximo común divisor de dos polinomios $p_1(x), p_2(x) \neq 0$: Sea $p_3(x)$ el resto de dividir $p_1(x)$ por $p_2(x)$. Es decir, tenemos que

$$p_1(x) = c_1(x) \cdot p_2(x) + p_3(x), \quad \text{con } \mathrm{gr}(p_3(x)) < \mathrm{gr}(p_2(x)) \text{ (ó } p_3(x) = 0).$$

Observemos que un polinomio divide a $p_1(x)$ y a $p_2(x)$ si y solo si divide a $p_2(x)$ y $p_3(x)$. Luego, $m.c.d.(p_1(x), p_2(x)) = m.c.d.(p_2(x), p_3(x))$. Si $p_3(x) \neq 0$, repitamos la división con $p_2(x)$ y $p_3(x)$: tenemos

$$p_2(x) = c_2(x) \cdot p_3(x) + p_4(x), \quad \text{con } \mathrm{gr}(p_4(x)) < \mathrm{gr}(p_3(x)) \text{ (ó } p_4(x) = 0).$$

De nuevo, $m.c.d.(p_2(x), p_3(x)) = m.c.d.(p_3(x), p_4(x))$. Así sucesivamente, vamos obteniendo polinomios $p_1(x), p_2(x), p_3(x), \ldots, p_n(x)$. Este proceso termina cuando $p_n(x) = 0$ (recordemos que $\mathrm{gr}(p_2(x)) > \mathrm{gr}(p_3(x)) > \mathrm{gr}(p_4(x)) > \ldots$), tenemos

$$\begin{aligned} p_{n-3}(x) &= c_{n-3}(x) \cdot p_{n-2}(x) + p_{n-1}(x) \\ p_{n-2}(x) &= c_{n-2}(x) \cdot p_{n-1}(x) + 0 \end{aligned}$$

Por tanto,

$$\begin{aligned} m.c.d.(p_1(x), p_2(x)) = m.c.d.(p_2(x), p_3(x)) = \cdots &= m.c.d.(p_{n-1}(x), p_n(x)) \\ &= m.c.d.(p_{n-1}(x), 0) = p_{n-1}(x). \end{aligned}$$

Observemos que

$$\begin{aligned} m.c.d.(p_1(x), p_2(x)) &= p_{n-1}(x) = p_{n-3}(x) - c_{n-3}(x) p_{n-2}(x) \\ &= p_{n-3}(x) - c_{n-3}(x)(p_{n-4}(x) - c_{n-4}(x) p_{n-3}(x)) \\ &= -c_{n-3}(x)) p_{n-4}(x) + (1 + c_{n-3}(x) c_{n-4}(x)) p_{n-3}(x) = \cdots = \lambda(x) p_1(x) + \mu(x) p_2(x). \end{aligned}$$

2. Identidad de Bézout: *Sean $p(x), q(x) \neq 0$ dos polinomios. Entonces, existen otros dos polinomios $\lambda(x), \mu(x)$, que sabemos calcular por el algoritmo de Euclides, de modo que*

$$m.c.d.(p(x), q(x)) = \lambda(x) \cdot p(x) + \mu(x) \cdot q(x).$$

3. Ejemplo : Calculemos el máximo común divisor de x^3+3x^2+3x+1 y x^2-1. Tenemos

$$\underline{x^3+3x^2+3x+1} = (x+3)\cdot\underline{(x^2-1)} + \underline{4x+4}$$
$$\underline{x^2-1} = \tfrac{1}{4}(x-1)\cdot\underline{(4x+4)} + 0$$

Luego, $4x+4$ es el máximo común divisor, o mejor dicho (porque ha de ser mónico) $x+1$ es el máximo común divisor de x^3-3x^2+3x+1 y x^2-1.

4. Ejemplo : Calculemos el máximo común divisor de x^3+3x^2+3x+1 y x^3-x^2-x+1. Tenemos

$$\underline{x^3+3x^2+3x+1} = \underline{x^3-x^2-x+1} + \underline{4x^2+4x}$$
$$\underline{x^3-x^2-x+1} = (\tfrac{1}{4}x-\tfrac{1}{2})\underline{(4x^2+4x)} + \underline{x+1}$$
$$\underline{4x^2+4x} = 4x\cdot\underline{(x+1)} + 0$$

Por tanto, $m.c.d(x^3+3x^2+3x+1, x^3-x^2-x+1) = x+1$.

Se dice que dos polinomios de grados mayor que cero, $p(x)$ y $q(x)$, son primos entre sí si el máximo común divisor de $p(x)$ y $q(x)$ es 1. Si $p(x)$ y $q(x)$ son primos entre sí existen dos polinomios $\lambda(x), \mu(x)$ tales que $\lambda(x)\cdot p(x) + \mu(x)\cdot q(x) = 1$.

El teorema de D'Alambert afirma que si $p(x) \in \mathbb{C}[x]$ entonces existen $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ (denominadas las raíces de $p(x)$) de modo que

$$p(x) = a\cdot(x-\alpha_1)\cdots(x-\alpha_n).$$

Si $p(x) = a\cdot(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$ y $q(x) = b\cdot(x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}$, con $n_i, m_i \geq 0$ y $\alpha_i \neq \alpha_j$ siempre que $i \neq j$, entonces

$$m.c.d.(p(x), q(x)) = (x-\alpha_1)^{min(n_1,m_1)}\cdots(x-\alpha_r)^{min(n_r,m_r)}.$$

Por tanto, $p(x)$ y $q(x)$ son primos entre sí si y solo si no tienen raíces comunes.

3.3. Operadores lineales

1. Definición : Una aplicación lineal $T: E \to E$ de un espacio vectorial E en sí mismo diremos que es un operador o un endomorfismo lineal de E.

Sea $T: E \to E$ un operador $\mathbb{R}$-lineal. Denotamos por T^2 el operador $T\circ T$, es decir $T^2: E \to E$ es el operador lineal definido por $T^2(e) = T(T(e))$. Denotamos por T^3 el operador $T\circ T\circ T$, es decir $T^3: E \to E$ es el operador lineal definido por $T^3(e) = T(T(T(e)))$.

Dado un polinomio $p(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$, denotaremos

$$p(T) := a_n\cdot T^n + \cdots + a_1\cdot T + a_0 := a_n\cdot(T\circ \overset{n}{\cdots} \circ T) + \cdots + a_1\cdot T + a_0\cdot \mathrm{Id}.$$

Dados $p(x), q(x) \in \mathbb{R}[x]$, si $h(x) = p(x) \cdot q(x)$, entonces $h(T) = p(T) \circ q(T)$. Además, como $q(x) \cdot p(x) = h(x)$, entonces $q(T) \circ p(T) = h(T) = p(T) \circ q(T)$. Para simplificar notaciones, $p(T) \cdot q(T)$ querrá decir $p(T) \circ q(T)$.

2. Ejemplo : Sea $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ la aplicación lineal de matriz, en las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^3$,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{pmatrix}$$

Sea $p(x) = x^2 + 3x - 1$. Calculemos la matriz de $p(T)$ en las bases usuales de $\mathbb{R}^3$ y $\mathbb{R}^3$:

$$p(T) = T^2 + 3T - \mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{pmatrix}^2 + 3\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 20 & 24 & 27 \\ 54 & 59 & 66 \\ 23 & 24 & 24 \end{pmatrix}.$$

3. Nota : Si $E_1, E_2 \subset E$ son dos subespacios vectoriales denotamos

$$E_1 + E_2 := \{e_1 + e_2 \in E, \text{ para todo } e_1 \in E_1 \text{ y todo } e_2 \in E_2\}.$$

Supongamos $E_1 \cap E_2 = 0$. Si $e_1 + e_2 = 0$ con $e_1 \in E_1$ y $e_2 \in E_2$ entonces $e_1 = 0$ y $e_2 = 0$, porque $e_1 = -e_2 \in E_1 \cap E_2 = 0$. Si $\{u_1, \dots, u_n\}$ es una base de E_1 y $\{v_1, \dots, v_m\}$ es una base de E_2, entonces $\{u_1, \dots, u_n, v_1, \dots, v_m\}$ es una base de $E_1 + E_2$, porque obviamente es un sistema generador y son linealmente independientes porque si

$$\lambda_1 \cdot u_1 + \cdots + \lambda_n \cdot u_n + \mu_1 \cdot v_1 + \cdots + \mu_m \cdot v_m = 0$$

entonces $\lambda_1 \cdot u_1 + \cdots + \lambda_n \cdot u_n = 0 = \mu_1 \cdot v_1 + \cdots + \mu_m \cdot v_m$, luego $\lambda_1 = \cdots = \lambda_n = 0 = \mu_1 = \cdots = \mu_m$.

Si $E_1, E_2 \subset E$ son dos subespacios vectoriales y $E_1 \cap E_2 = 0$ escribiremos $E_1 + E_2 = E_1 \oplus E_2$. Hemos probado que $\dim(E \oplus E') = \dim E + \dim E'$.

4. Proposición : *Sea $T \colon E \to E$ un operador lineal y $p_1(x), p_2(x)$ dos polinomios primos entre sí. Entonces,*

$$\mathrm{Ker}(p_1(T) \circ p_2(T)) = \mathrm{Ker}\, p_1(T) \oplus \mathrm{Ker}\, p_2(T).$$

Demostración. Sean $\lambda(x), \mu(x)$ tales que $\lambda(x) \cdot p_1(x) + \mu(x) \cdot p_2(x) = 1$. Luego,

$$\mathrm{Id} = \lambda(T) \circ p_1(T) + \mu(T) \circ p_2(T).$$

Dado $e \in \operatorname{Ker}(p_1(T) \circ p_2(T))$, tenemos

$$e = \operatorname{Id}(e) = (\lambda(T) \circ p_1(T) + \mu(T) \circ p_2(T))(e) = (\lambda(T) \circ p_1(T))(e) +)(\mu(T) \circ p_2(T))(e).$$

Ahora bien, $(\lambda(T) \circ p_1(T))(e) \in \operatorname{Ker} p_2(T)$, porque

$$p_2(T)((\lambda(T) \circ p_1(T))(e)) = (\lambda(T) \circ p_1(T) \circ p_2(T))(e) = \lambda(T)(0) = 0,$$

Igualmente $\mu(T)(p_2(T)(e)) \in \operatorname{Ker} p_1(T)$. Por lo tanto, hemos probado que $\operatorname{Ker}(p_1(T) \circ p_2(T)) \subseteq \operatorname{Ker} p_1(T) + \operatorname{Ker} p_2(T)$.

Evidentemente, $\operatorname{Ker} p_1(T) + \operatorname{Ker} p_2(T) \subseteq \operatorname{Ker}(p_1(T) \circ p_2(T))$, luego son iguales.

Por último, $\operatorname{Ker} p_1(T) \cap \operatorname{Ker} p_2(T) = 0$, porque si $e \in \operatorname{Ker} p_1(T) \cap \operatorname{Ker} p_2(T)$, entonces

$$e = \operatorname{Id}(e) = (\lambda(T) \circ p_1(T) + \mu(T) \circ p_2(T))(e) = 0 + 0 = 0.$$

□

3.4. Ecuaciones diferenciales lineales de orden n con coeficientes constantes

Sea $F = \{e^{ix}, \cos x + ix^2$, etc.$\}$ el conjunto de todas las funciones de los números reales a valores complejos infinitamente diferenciables. F es un $\mathbb{C}$-espacio vectorial. Sea

$$D : F \to F,\ D(f(x)) := f'(x)$$

el "operador derivada". Es claro que D es un operador $\mathbb{C}$-lineal de F. Dado $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, entonces $P(D) \colon F \to F$ es el operador definido por

$$P(D)(f) = a_n D^n(f) + a_{n-1}D^{n-1}(f) + \cdots + a_0 \cdot f,$$

donde $D^r(f)$ es la derivada r-ésima de f.

1. Ejercicio : Calcula $(D^2 + \operatorname{Id})(\cos x)$.

Queremos resolver ecuaciones diferenciales del tipo

$$a_n \cdot f^{n)} + \cdots + a_2 f'' + a_1 f' + a_o \cdot f = 0, \qquad \text{(donde los } a_i \text{ son constantes).}$$

Es decir, buscamos aquellas funciones $f \in F$. que cumplen que $P(D)(f) = 0$, donde $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Tenemos que calcular $\operatorname{Ker} P(D)$.

Veamos que

$$\operatorname{Ker} D^r = \{\text{Polinomios de grado estrictamente menor que } r\}.$$

En efecto,

$D(f)=0 \iff f = cte$
$D^2(f)=0 \iff D(D(f))=0 \iff D(f)=cte \iff f = cte\cdot x + cte'$
$D^3(f)=0 \iff D^2(Df)=0 \iff Df = cte\cdot x+cte' \iff f=\frac{cte}{2}\cdot x^2+cte'\cdot x+cte''$
Etcétera.

2. Movimiento uniformemente acelerado: Supongamos que un objeto se mueve a lo largo de la recta (real) con una aceleración constante a. Digamos que $f(t)$ es la posición del móvil en el instante t. La velocidad del móvil es $f'(t)$ en cada instante t y la aceleración es $f''(t)=a$ en todo instante t. Por tanto, $f'''=0$, es decir, $f(t)\in \operatorname{Ker} D^3$. Luego, $f(t)=\lambda+\mu t+\gamma t^2$. Observemos que $f(0)=\lambda$, $f'(0)=\mu$ y $f''(0)=2\cdot\gamma=a$. Por tanto,

$$f(t)=f(0)+f'(0)\cdot t+\frac{a}{2}\cdot t^2 \quad \text{y} \quad f'(t)=f'(0)+a\cdot t.$$

3. Fórmula de conmutación: *Sea $P(x)\in\mathbb{C}[x]$. Para toda $f\in F$ y $\alpha\in\mathbb{C}$, se cumple que*

$$P(D)(e^{\alpha x}\cdot f)=e^{\alpha x}\cdot P(D+\alpha\cdot \mathrm{Id})(f).$$

Demostración. $D(e^{\alpha x}\cdot f)=\alpha\cdot e^{\alpha x}\cdot f+e^{\alpha x}\cdot D(f)=e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})(f)$, por la regla de Leibniz de derivación de un producto de funciones. Entonces,
$D^2(e^{\alpha x}\cdot f)=D(D(e^{\alpha x}\cdot f))=D(e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})(f))=e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})((D+\alpha\cdot\mathrm{Id})(f))=e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})^2(f)$. Así sucesivamente, $D^n(e^{\alpha x}\cdot f)=e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})^n(f)$. Por último, si $P(D)=\sum_i a_i D^i$ tenemos que

$$P(D)(e^{\alpha x}\cdot f)=\sum_i a_i D^i(e^{\alpha x}\cdot f)=\sum_i a_i\cdot e^{\alpha x}\cdot(D+\alpha\cdot\mathrm{Id})^i(f)=e^{\alpha x}\cdot P(D+\alpha\cdot\mathrm{Id})(f).$$

□

4. Teorema: *Se cumple que*

1. $\operatorname{Ker}(D-\alpha\cdot\mathrm{Id})^r=e^{\alpha x}\cdot\{\text{Polinomios de grado estrictamente menor que } r\}$.
2. *Si* $P(x)=(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, *entonces*

$$\operatorname{Ker} P(D)=e^{\alpha_1 x}\cdot\{\text{Pol. de grado } <n_1\}\oplus\cdots\oplus e^{\alpha_r x}\cdot\{\text{Pol. de grado } <n_r\}$$

Demostración. 1. En efecto,

$$\begin{aligned}(D-\alpha\cdot \mathrm{Id})^r f(x)=0 &\iff 0=(D-\alpha\cdot \mathrm{Id})^r(e^{\alpha x}\cdot e^{-\alpha x}\cdot f(x))=e^{\alpha x}\cdot D^r(e^{-\alpha x}\cdot f(x))\\ &\iff 0=D^r(e^{-\alpha x}\cdot f(x)) \iff e^{-\alpha x}\cdot f(x) \text{ es un polinomio de grado menor que } r\\ &\iff f(x) \text{ es } e^{\alpha x} \text{ multiplicado por un polinomio de grado menor que } r.\end{aligned}$$

2. Es consecuencia de que $\operatorname{Ker} P(D)=\operatorname{Ker}(D-\alpha_1\cdot \mathrm{Id})^{n_1}\oplus\cdots\oplus\operatorname{Ker}(D-\alpha_r\cdot \mathrm{Id})^{n_r}$ y de 1.

□

5. Nota: Supongamos $P(x)\in\mathbb{R}[x]$ y $\alpha=a+bi$ es una raíz compleja de $P(x)$, con $b\neq 0$. Entonces $\bar\alpha=a-bi$ también es una raíz compleja de $P(x)$. Más aún, la multiplicidad con la que aparece α es la misma con la aparece $\bar\alpha$, es decir, $P(x)=(x-\alpha)^n\cdot(x-\bar\alpha)^n\cdot Q(x)$, con $Q(\alpha),Q(\bar\alpha)\neq 0$. Es fácil probar, a partir del teorema 3.4.4, que

$$\left\{\begin{array}{l} f\in\operatorname{Ker}(D-\alpha\,\mathrm{Id})^n\oplus\operatorname{Ker}(D-\bar\alpha\,\mathrm{Id})^n\\ \text{tales que } f(x)\in\mathbb{R} \text{ para todo } x\end{array}\right\}=e^{ax}\cdot\{q(x)\cdot\cos bx+r(x)\cdot\operatorname{sen} bx\}_{q(x),r(x)\in P_n}$$

donde P_n es el conjunto de todos los polinomios con coeficientes reales de grado $<n$.

6. Ejemplo: Resolvamos la ecuación diferencial: $f''''-2f'''+2f''=0$. Tenemos que resolver $(D^4-2D^3+2D^2)(f)=0$, es decir, calcular $\operatorname{Ker}(D^4-2D^3+2D^2)$. Observemos que $x^4-2x^3+2x^2=x^2(x^2-2x+2)=x^2(x-(1+i))(x-(1-i))$. Luego,

$$\begin{aligned}\operatorname{Ker}(D^4-2D^3+2D^2)&=\operatorname{Ker} D^2\oplus\operatorname{Ker}(D-(1+i)\cdot\mathrm{Id})\oplus\operatorname{Ker}(D-(1-i)\cdot\mathrm{Id})\\ &=\{a+bx+c\cdot e^{(1+i)\cdot x}+d\cdot e^{(1-i)\cdot x}\}.\end{aligned}$$

Luego,

$$\left\{\begin{array}{l}\text{Las funciones con valores en } \mathbb{R} \text{ solución de}\\ \text{la ecuación diferencial } f''''-2f'''+2f''=0\end{array}\right\}=\{a+bx+e^x\cdot(\lambda\cos x+\mu\operatorname{sen} x)\}.$$

7. Ejercicio: Resuelve la ecuación diferencial: $f''+f=0$.

Ecuaciones diferenciales no homogéneas:

Hasta ahora hemos resuelto ecuaciones diferenciales del tipo $P(D)(f)=0$. Consideremos ahora una ecuación diferencial del tipo $P(D)(f)=g$, con $g\in F$. Sea f_0 una solución particular. Entonces, $f\in F$ cumple que $P(D)(f)=g$ si y solo si

$$f=f_0+f_1,\ \text{con } f_1\in\operatorname{Ker} P(D)$$

Con palabras:

$$\begin{bmatrix}\text{Todas las soluciones}\\ \text{de } P(D)(f)=g\end{bmatrix} = \begin{bmatrix}\text{Una solución particular}\\ \text{de } P(D)(f)=g\end{bmatrix} + \begin{bmatrix}\text{Todas las soluciones de la}\\ \text{“homogénea” } P(D)(f)=0\end{bmatrix}$$

8. Cálculo de una solución particular: *Consideremos una ecuación diferencial $P(D)f = g$. Si tenemos un subespacio vectorial $E = \langle f_1,\dots,f_n\rangle \subset F$, tal que $g \in E$ y $D(E) \subseteq E$, consideremos la aplicación lineal $P(D)\colon E \to E$, $h \mapsto P(D)(h)$. Si encontramos $f \in E$ tal que $P(D)(f) = (g)$, obtenemos así una solución particular.*

9. Ejemplo: Resolvamos la ecuación diferencial $f''' - f = x^3$: Tenemos que resolver la ecuación $(D^3 - \mathrm{Id})(f) = x^3$. Sea $E = \langle x^3, x^2, x, 1\rangle$ y consideremos la aplicación lineal $D^3 - \mathrm{Id}\colon E \to E$. Entonces, $(D^3 - \mathrm{Id})(ax^3 + bx^2 + cx + d) = x^3$ si $6a - (ax^3 + bx^2 + cx + d) = x^3$. Luego, $a = -1, b = 0, c = 0, d = -6$ y una solución particular de la ecuación diferencial es $-6 - x^3$.

Todas las soluciones son

$$\begin{aligned} f_0 + \mathrm{Ker}(D^3 - \mathrm{Id}) &= f_0 + \mathrm{Ker}(D - \mathrm{Id}) + \mathrm{Ker}(D - \frac{-1-\sqrt{-3}}{2}\cdot \mathrm{Id}) + \mathrm{Ker}(D - \frac{-1+\sqrt{-3}}{2}\cdot \mathrm{Id}) \\ &= \{-6 - x^3 + a\cdot e^x + b\cdot e^{\frac{-1-\sqrt{-3}}{2}\cdot x} + c\cdot e^{\frac{-1+\sqrt{-3}}{2}\cdot x},\ a,b,c \in \mathbb{C}\}. \end{aligned}$$

Todas las soluciones que son funciones con valores reales son

$$\{-6 - x^3 + a\cdot e^x + e^{\frac{-1}{2}x}\cdot (b\cdot \cos\frac{\sqrt{3}}{2}x + c\cdot \mathrm{sen}\,\frac{\sqrt{3}}{2}x),\ a,b,c \in \mathbb{R}\}.$$

10. Ejercicio: Resuelve la ecuación diferencial $f'' - f = sen x$.

11. Sea $P(x) = (x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$. Existen polinomios $Q_i(x)$ de grados menores que n_i tales que

$$\frac{1}{(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}} = \frac{Q_1(x)}{(x-\alpha_1)^{n_1}} + \overset{r}{\cdots} + \frac{Q_r(x)}{(x-\alpha_r)^{n_r}}.$$

Resolvamos la ecuación diferencial $P(D)f = g$. Vamos a denotar $\int f = \frac{1}{D}f$ y en general $\int \overset{n}{\cdots} \int f = \frac{1}{D^n}f$. Entonces,

$$\begin{aligned} f = \frac{1}{P(D)}g &= (\frac{Q_1(D)}{(D-\alpha_1\mathrm{Id})^{n_1}} + \overset{r}{\cdots} + \frac{Q_r(D)}{(D-\alpha_r\mathrm{Id})^{n_r}})g = \sum_i \frac{Q_i(D)}{(D-\alpha_i\mathrm{Id})^{n_i}}g \\ &\overset{3.4.3}{=} \sum_i e^{\alpha_i x}\cdot \frac{Q_i(D+\alpha_i\mathrm{Id})}{D^{n_i}} e^{-\alpha_i x}\cdot g = \sum_i e^{\alpha_i x}\cdot Q_i(D+\alpha_i\mathrm{Id}) \int \overset{n_i}{\cdots} \int e^{-\alpha_i x}\cdot g. \end{aligned}$$

3.4.1. Ejemplos

Demos diversos ejemplos de la Física, Química y Economía donde las ecuaciones diferenciales ayudan en su estudio.

12. Ley de desintegración radiactiva: Consideremos que tenemos una cierta cantidad $U(t)$ de gramos de uranio en el instante t de tiempo. Suponemos que la cantidad de uranio que se desintegra en un intervalo de tiempo (muy pequeño) t_1, $U(t)-U(t+t_1)$, es proporcional al tiempo t_1 transcurrido y a la cantidad $U(t)$ de gramos que había en el instante t. Es decir, tenemos

$$U(t+t_1)-U(t) = -cte\cdot t_1\cdot U(t) \qquad (cte > 0).$$

Por tanto,

$$\frac{U(t+t_1)-U(t)}{t_1} = -cte\cdot U(t).$$

Tomando límite $t_1 \to 0$, obtenemos

$$U'(t) = -cte\cdot U(t).$$

Es decir, $U(t)$ verifica la ecuación diferencial $U' + cte\cdot U = 0$. Tenemos $(D+cte\cdot \mathrm{Id})(U) = 0$, luego $U(t) = a\cdot e^{-cte\cdot t}$, para cierta constante a. Observemos que $U(0) = a\cdot e^0 = a$. Luego,

$$\boxed{U(t) = U(0)\cdot e^{-cte\cdot t}}$$

Veamos cuál es la semivida del uranio, es decir, cuánto tiempo s ha de transcurrir para que se desintegre la mitad del uranio:

$$U(0)\cdot e^{-cte\cdot s} = U(s) = \frac{U(0)}{2}.$$

Luego, $e^{-cte\cdot s} = \frac{1}{2}$. Tomando logaritmo neperiano, $-cte\cdot s = \ln 2^{-1} = -\ln 2$, luego

$$s = \frac{\ln 2}{cte} \quad \text{y} \quad U(t) = U(0)\cdot e^{-\frac{ln2}{s}\cdot t} = U(0)\cdot 2^{-\frac{t}{s}}.$$

El carbono 14 que hay en la atmósfera aunque se desintegra en carbono no radiactivo (carbono 12 y 13), también se crea continuamente debido a las colisiones de los neutrones generados por los rayos cósmicos con el nitrógeno de la atmósfera superior. Resulta que la proporción de carbono 14 y carbono no radiactivo permanece en un nivel casi constante en la atmósfera a lo largo del tiempo. Las plantas adquieren el carbono atmosférico mediante la fotosíntesis, y los animales, mediante el consumo de plantas y de otros animales. Cuando un organismo muere el carbono 14 existente en el organismo va desintegrándose. La proporción de carbono 14 y carbono no radiactivo cuando se examinan los restos del organismo proporciona una indicación del tiempo transcurrido desde su muerte.

Ejercicio: El cobalto-60 (^{60}Co) es un isótopo radiactivo sintético del cobalto que tiene una semivida de 5.27 años. El cobalto-60 de una unidad de radioterapia debe ser reemplazado cuando su masa desciende un 25 %. Si una muestra de cobalto-60 se adquirió en septiembre de 2016 ¿cuándo será necesario reemplazarla?

13. Partícula en un pozo Consideremos una partícula que se mueve a lo largo del eje 0X. La ecuación de Schrödinger para una partícula de masa m que se mueve en el eje 0X es

$$-\frac{\hbar^2}{2m}\psi'' + V\psi = E\psi.$$

Donde $\hbar = \frac{h}{2\pi}$ es la constante de Planck normalizada, V es la energía potencial de la partícula, E es la energía total (constante) de la partícula y $\psi(x)$ la función de onda de la partícula, cuyo módulo al cuadrado multiplicado por Δx indica la probabilidad de que la partícula esté en el intervalo $(x, x+\Delta x)$. Suponemos que en el intervalo $(0,a)$ no interviene ninguna fuerza. Es más, que $V = 0$ en el intervalo $(0,a)$ y $V = \infty$ fuera del intervalo, lo que en mecánica cuántica significa que $\psi(x) = 0$ fuera del intervalo. Calculemos ψ:

$$(D^2 + w^2\,\mathrm{Id})\psi = 0$$

(donde $w = \frac{\sqrt{2mE}}{\hbar}$). Luego,

$$\psi(x) \in \operatorname{Ker}(D^2 + w^2\,\mathrm{Id}) = \{\lambda \cos wx + \mu \operatorname{sen} wx\}.$$

Como $\psi(0) = 0$ entonces $\lambda = 0$ y $\psi(a) = 0$, luego $\operatorname{sen} wa = 0$ y $w = \frac{n\pi}{a}$. En conclusión las soluciones son $\psi_n(x) = \mu \operatorname{sen}\frac{n\pi x}{a}$ y para cada una de estas funciones de onda la energía total de la partícula es $E_n = \frac{n^2h^2}{8ma^2}$. Como la probabilidad que la partícula esté en el intervalo $(0,a)$ es 1, entonces

$$1 = \int_0^a \psi_n(x)^2 dx = \int_0^a \mu^2 \operatorname{sen}^2\frac{n\pi x}{a}\,dx = \cdots = \mu^2\frac{a}{2}.$$

Luego, $\mu = \sqrt{2/a}$ y $\psi_n(x) = \sqrt{2/a}\cdot \operatorname{sen}\frac{n\pi x}{a}$.

14. Presión atmosférica: Sea $P(h)$ la presión atmosférica a una altura h del suelo. Suponemos que la temperatura y la fuerza de la gravedad no varían con la altura (considerada). La diferencia de presión $P(h+t) - P(h)$ es proporcional a t y a la densidad del aire en h (que es proporcional a $P(h)$). Entonces,

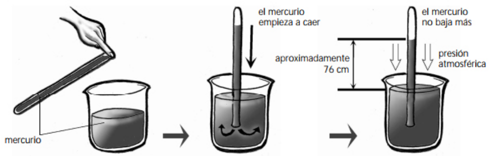
$$\frac{P(h+t)-P(h)}{t} = -K\cdot P(h).$$

Luego, $P'(h) = -K \cdot P(h)$, es decir,

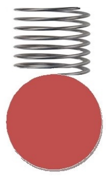
$$(D + K \cdot \mathrm{Id})(P) = 0.$$

Por tanto, $P(h) = a \cdot e^{-K \cdot h}$, donde $a = P(0)$. Puede medirse la altura en términos de la presión: $h = \frac{\ln P(0) - \ln P(h)}{K}$.

Problema: En 1643, Torricelli al realizar su experimento, al nivel del mar, obtuvo que la columna de mercurio medía 760 mm. (por encima del nivel de mercurio de la cubeta). Si a una altura de 100 metros sobre el nivel del mar la columna de mercurio mide 751 mm. ¿a qué altura sobre el nivel del mar estaríamos si la columna de mercurio midiese 600 mm.?



Experimento de Torricelli



15. Movimiento armónico simple: Consideremos un muelle de longitud a en la recta real, con un extremo fijo en $-a$ y el otro situado en el origen. Sea $f(t)$ la posición del extremo libre del muelle en el instante t. Si el extremo del muelle está en la posición $f(t)$ en el instante t, entonces el muelle ejerce una fuerza (luego aceleración) proporcional a $f(t)$ con sentido hacia el origen. Entonces,

$$f''(t) = -cte \cdot f(t), \qquad (cte > 0).$$

Es decir, $f(t)$ cumple la ecuación diferencial

$$(D^2 + cte)(f) = 0.$$

Como $x^2 + cte = (x - \sqrt{cte} \cdot i)(x + \sqrt{cte} \cdot i)$, tenemos que $f(t) = a\cos(cte \cdot t) + b\,\mathrm{sen}(cte \cdot t)$. Si $f(0) = 0$, entonces $a = 0$ y

$$f(t) = b \cdot \mathrm{sen}(cte \cdot t).$$

Observemos que $b = f'(0)/cte$ es la máxima elongación del muelle y $2\pi/cte$ es el periodo del movimiento armónico.

16. Interés compuesto continuo: Supongamos que tenemos un capital de 10^6 euros invertidos en un banco. El banco nos da por la inversión un interés del 2 por ciento anual y nos permite retirar el dinero en cualquier momento sin penalización y con el pago de los intereses del capital por el tiempo exacto transcurrido. Si retiramos el capital, con los intereses generados, al año y medio ¿cuánto dinero nos llevaremos?: Sea $f(t)$ el capital más los intereses generados que tenemos en el banco en el momento

t. Observemos que $f(t+h)-f(t)$ es proporcional a $f(t)$ y al tiempo h transcurrido ("cuando h es muy pequeño"), es decir,

$$f(t+h)-f(t)=K\cdot h\cdot f(t), \quad \text{y} \quad \frac{f(t+h)-f(t)}{h}=K\cdot f(t).$$

Luego,

$$f'(t)=\lim_{h\to 0}\frac{f(t+h)-f(t)}{h}=K\cdot f(t).$$

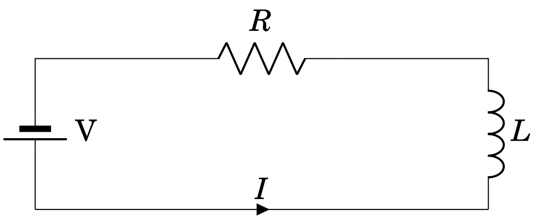
Es decir, $(D-K\cdot \mathrm{Id})(f)=0$, luego $f(t)=a\cdot e^{Kt}$. Sabemos que $f(0)=10^6$, luego $a=10^6$; y $10^6\cdot e^K=f(1)=10^6\cdot(1+0'02)$, luego $e^K=1'02$. Por tanto,

$$f(t)=10^6\cdot(1'02)^t$$

y $f(1'5)=10^6\cdot(1'02)^{1'5}$.

Demos ahora ejemplos de ecuaciones diferenciales no homogéneas.

17. Circuito eléctrico Consideremos el circuito eléctrico que sigue:



La segunda ley de Kirchhoff afirma que la suma de las diferencias de potencial eléctrico en un circuito cerrado es igual a cero. La caída de potencial en una bobina es proporcional a la tasa de variación de la intensidad de corriente: $V_L=L\cdot\frac{dI}{dt}$, la caída de potencial en una resistencia es proporcional a la intensidad: $V_R=R\cdot I$ (ley de Ohm). Por tanto,

$$V=L\cdot\frac{dI}{dt}+R\cdot I.$$

Resolvamos la ecuación diferencial $LI'+RI=(LD+R\,\mathrm{Id})(I)=V$, cuando V es constante. Una solución particular es $I=V/R$ y la general es

$$I(t)=V/R+\mathrm{Ker}(LD+R\,\mathrm{Id})=V/R+\mathrm{Ker}(D+\frac{R}{L}\mathrm{Id})=V/R+\{\lambda\cdot e^{-\frac{Rt}{L}}\}.$$

Si suponemos que $I(0)=0$, entonces $\lambda=-V/R$ y $I(t)=V/R\cdot(1-e^{-\frac{Rt}{L}})$.

Supongamos ahora que $V=A\cos(wt)$. Tenemos que resolver la ecuación diferencial

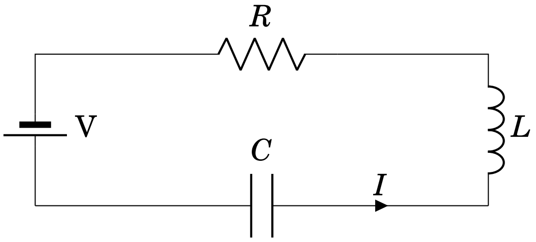
$$(LD+R\,\mathrm{Id})(I)=A\cdot\cos(wt).$$

Sea $E = \langle \cos(wt), \operatorname{sen}(wt)\rangle$. Buscamos una solución $\lambda \cdot \cos(wt) + \mu \cdot \operatorname{sen}(wt) \in E$:

$$\begin{aligned} A \cdot \cos(wt) &= (LD + R\,\mathrm{Id})(\lambda \cdot \cos(wt) + \mu \cdot \operatorname{sen}(wt)) \\ &= (R\mu - Lw\lambda)\operatorname{sen}(wt) + (R\lambda + Lw\mu)\cos(wt). \end{aligned}$$

Obtenemos $\lambda = \frac{AR}{L^2w^2+R^2}$ y $\mu = \frac{ALw}{L^2w^2+R^2}$. Por tanto, $I(t) = \frac{AR\cos(wt)}{L^2w^2+R^2} + \frac{ALw\operatorname{sen}(wt)}{L^2w^2+R^2} + \gamma \cdot e^{-\frac{Rt}{L}}$. Si suponemos que $I(0) = 0$, entonces

$$\begin{aligned} I(t) &= \frac{A}{L^2w^2+R^2} \cdot (R \cdot (\cos(wt) - e^{-\frac{Rt}{L}}) + Lw \cdot \operatorname{sen}(wt)) \\ &\overset{t \gg 0}{\simeq} \frac{A}{L^2w^2+R^2} \cdot (R \cdot \cos(wt) + Lw \cdot \operatorname{sen}(wt)). \end{aligned}$$



Añadamos un condensador. La diferencia de potencial causada por un condensador es proporcional a la carga, $V_C = \frac{Q}{C}$, luego derivando respecto del tiempo $V_C' = \frac{I}{C}$. Por tanto, $V = L \cdot \frac{dI}{dt} + R \cdot I + V_C$ y derivando

$$V' = LI'' + RI' + \frac{I}{C}.$$

Supongamos $V(t) = V_0 \cdot \cos(wt)$. Si $\tilde{V} := V_0 \cdot e^{iwt}$, entonces I es la parte real de las soluciones de la ecuación diferencial

$$\tilde{V}' = L\tilde{I}'' + R\tilde{I}' + \frac{\tilde{I}}{C}.$$

Busquemos una solución particular en $\langle e^{iwt}\rangle$, es decir, $\tilde{I} = Y \cdot e^{iwt}$, con $Y \in \mathbb{C}$. Tenemos que $iwV_0 \cdot e^{iwt} = (-Lw^2 + Rwi + \frac{1}{C}) \cdot Y \cdot e^{iwt}$, luego $\tilde{I} = \frac{iwV_0}{(-Lw^2+\frac{1}{C})+Rwi} \cdot e^{iwt}$. Por lo tanto,

$$\tilde{I} = \frac{\tilde{V}}{Z},$$

donde $Z = Z_0 \cdot e^{i\phi}$ (denominada impedancia), $\phi = \arctan \frac{Lw - \frac{1}{wc}}{R}$ (denominada desfase) y $Z_0 = \sqrt{(Lw - \frac{1}{wC})^2 + R^2}$. En conclusión,,

$$\boxed{I(t) = \frac{V_0}{Z_0} \cdot \cos(wt - \phi)}$$ [1]

Observemos que la amplitud del voltaje es V_0 y el de la intensidad $I_0 := \frac{V_0}{Z_0}$, luego

$$\boxed{V_0 = Z_0 \cdot I_0}$$

18. Propagación de un virus: Consideremos que tenemos una población con N personas. Un virus aparece e infecta a las personas. Sea $I(t)$ el número de infectados en el instante t de tiempo. Supongamos que el incremento de nuevos casos de infectados, en un incremento de tiempo h, es proporcional al número de infectados por la proporción de sanos en toda la población (y el incremento del tiempo)[2]. Tenemos

$$I(t+h) - I(t) = R \cdot I(t) \cdot \frac{N - I(t)}{N} \cdot h$$

y R mide lo infeccioso que es el virus. Entonces, $I' = R \cdot I \cdot (1 - I/N)$, que no es una ecuación diferencial de las estudiadas. Hay mucho más que estudiar. Nos resistimos a dejarla sin resolver. Tenemos $\frac{I'}{R \cdot I \cdot (1-I/N)} = 1$. Integrando

$$\begin{aligned} t + cte &= \int 1 \cdot dt = \int \frac{I'}{R \cdot I \cdot (1 - I/N)} \cdot dt = \int \frac{1}{R \cdot I \cdot (1 - I/N)} \cdot dI = \int \frac{1}{R} \cdot (\frac{1}{I} + \frac{1}{N - I}) \cdot dI \\ &= \frac{\ln I - \ln(N - I)}{R} = \frac{\ln(I/(N - I))}{R}. \end{aligned}$$

Multiplicando por R y tomando exponenciales, obtenemos $e^{(t+cte)R} = I/(N - I)$ (luego $e^{cte \cdot R} = I(0)/(N - I(0))$. Despejando I obtenemos

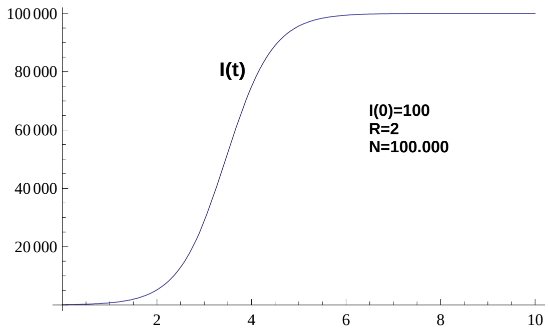
$$I(t) = \frac{N \cdot e^{(t+cte)R}}{1 + e^{(t+cte)R}} = \cdots = \frac{I(0) \cdot e^{tR}}{1 + \frac{I(0)}{N} \cdot (e^{tR} - 1)}.$$

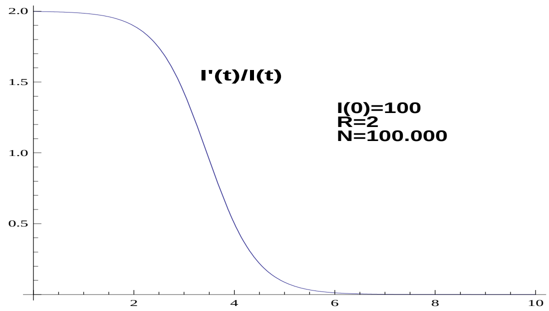
[1]Podemos decir que es la solución, porque si $R^2 \neq \frac{4L}{C}$ entonces

$$\begin{aligned} \operatorname{Ker}(LD^2 + RD + \frac{\mathrm{Id}}{C}) &= \operatorname{Ker}(D - \frac{-R + \sqrt{R^2 - 4L/C}}{2L}) \oplus \operatorname{Ker}(D - \frac{-R - \sqrt{R^2 - 4L/C}}{2L}) \\ &= \{a \cdot e^{\frac{-R+\sqrt{R^2-4L/C}}{2L}t} + be^{\frac{-R-\sqrt{R^2-4L/C}}{2L}t} \overset{t \gg 0}{\simeq} 0\}, \end{aligned}$$

y si $R^2 = 4L/C$, $\operatorname{Ker}(D - \frac{-R-\sqrt{R^2-4L/C}}{2L}) = \{e^{\frac{-R}{2L}t} \cdot (a + bt) \overset{t \gg 0}{\simeq} 0\}$.

[2]¡Estamos suponiendo que los infectados permanecen infecciosos con el paso del tiempo!





$I'(t)/I(t)$ mide lo infecciosa que es la enfermedad (en la condiciones dadas en t). Obsérvese que $I'(0)/I(0) = R(1-\frac{I(0)}{N})$. Cuando la población es muy grande en comparación con $I(0)$, entonces $I'(0)/I(0) \simeq R$; además si t es pequeño $I(t) \simeq I(0)\cdot e^{tR}$.

3.5. Ecuaciones en diferencias finitas

Sea $S = \{(a_n)_n\}$ el $\mathbb{C}$-espacio vectorial de las sucesiones de números complejos. La sucesión $(a_n)_n$ muchas veces la denotaremos simplemente (a_n). Consideremos el "operador siguiente" $\nabla\colon S \to S$ que es el operador $\mathbb{C}$-lineal definido por

$$\nabla(a_n) = (a'_n), \text{ donde } a'_n = a_{n+1}.$$

Diremos que $\Delta := \nabla - \mathrm{Id}$ es el "operador diferencia".

Como sabemos, dado un polinomio $P(x) = c_r x^r + c_{r-1}x^{r-1} + \cdots + c_0 \in \mathbb{C}[x]$, podemos considerar el operador lineal $P(\nabla)\colon S \to S$ definido por

$$\begin{aligned} P(\nabla)(a_n) &= (c_r\nabla^r + c_{r-1}\nabla^{r-1} + \cdots + c_0\,\mathrm{Id})(a_n) \\ &= (c_r a_{n+r} + c_{r-1}a_{n+r-1} + \cdots + c_0 a_n). \end{aligned}$$

Observemos que $\operatorname{Ker}\Delta = \{(c)$, para cualquier $c \in \mathbb{C}\}$. En efecto, $\Delta(a_n) = (0)$ si y solo si $(a_{n+1} - a_n) = (0)$, que equivale a $a_{n+1} - a_n = 0$ para todo n, es decir, la sucesión (a_n) es constante.

Observemos que $\Delta(n^r) = ((n+1)^r - n^r) = (n^r + \binom{r}{1}n^{r-1} + \cdots + 1 - n^r) = (q_r(n))$, donde $q_r(n)$ es un polinomio en n de grado menor estricto que r. Por tanto, dado un polinomio $p_r(n) = c_r n^r + c_{r-1}n^{r-1} + \cdots + c_0$ de grado r, se cumple que

$$\begin{aligned} \Delta(p_r(n)) &= (c_r\Delta(n^r) + c_{r-1}\Delta(n^{r-1}) + \cdots + \Delta(c_0)) = (c_r q_r(n) + c_{r-1}q_{r-1}(n) + \cdots + 0) \\ &= (q(n)), \end{aligned}$$

donde $q(n)$ es un polinomio de grado menor estricto que r. Por lo tanto,

$$\Delta^{r+1}(p_r(n)) = 0.$$

Es decir, $S_r := \{(c_r n^r + c_{r-1} n^{r-1} + \cdots + a_0)$, para todo $c_r, \dots, c_0 \in \mathbb{C}\} \subseteq \operatorname{Ker} \Delta^{r+1}$. Probemos que que $S_r = \operatorname{Ker} \Delta^{r+1}$. Lo hemos probado para $r = 0$. Supongamos que lo hemos probado para $r = 0, \dots, m-1$. Nos basta ver que es cierto para $r = m$. S_r es un $\mathbb{C}$-espacio vectorial de base las sucesiones $(1), (n), (n^2), \dots, (n^r)$, luego es un espacio vectorial de dimensión $r+1$. Por tanto, $\dim \operatorname{Ker} \Delta^r \geq \dim S_{r-1} = m$. Consideremos la aplicación $\bar{\Delta}\colon \operatorname{Ker} \Delta^{m+1} \to \operatorname{Ker} \Delta^m$ $\bar{\Delta}(p(n)) := \Delta(p(n))$. Sabemos que

$$\dim \operatorname{Ker} \Delta^{m+1} = \dim \operatorname{Ker} \bar{\Delta} + \dim \operatorname{Im} \bar{\Delta} \leq \dim \operatorname{Ker} \Delta + \dim \operatorname{Ker} \Delta^m = 1 + \dim S_{m-1} = m+1.$$

Por tanto, $m+1 \leq \dim \operatorname{Ker} \Delta^{m+1} \leq m+1$, luego $\dim \operatorname{Ker} \Delta^{m+1} = m+1$ y $\operatorname{Ker} \Delta^{m+1} = S_m$.

Hemos probado que

$$\operatorname{Ker} \Delta^m = \{(p(n)), \text{ donde } p(n) \text{ es cualquier polinomio de grado} < m\}.$$

1. Fórmula de conmutación: *Sea $P(x) \in \mathbb{C}[x]$. Entonces para toda sucesión $(a_n) \in Suc(\mathbb{C})$ y $\alpha \in \mathbb{C}$, se cumple que*

$$P(\nabla)((\alpha^n) \cdot (a_n)) = (\alpha^n) \cdot P(\alpha \nabla)(a_n).$$

Demostración. $\nabla((\alpha^n) \cdot (a_n)) = (\alpha^{n+1}) \cdot (a_{n+1}) = \alpha \cdot (\alpha^n) \cdot \nabla(a_n) = (\alpha^n) \cdot (\alpha \nabla)(a_n)$. Luego,

$$\begin{aligned} \nabla^2((\alpha^n) \cdot (a_n)) &= \nabla(\nabla((\alpha^n) \cdot (a_n))) = \nabla((\alpha^n) \cdot (\alpha \nabla)(a_n)) = (\alpha^n) \cdot (\alpha \nabla)((\alpha \nabla)(a_n)) \\ &= (\alpha^n) \cdot (\alpha \nabla)^2 (a_n). \end{aligned}$$

Así sucesivamente, $\nabla^r((\alpha^n) \cdot (a_n)) = (\alpha^n) \cdot (\alpha \nabla)^r (a_n)$. Por último, si $P(x) = \sum_i c_i x^i$, entonces

$$P(\nabla)((\alpha^n) \cdot (a_n)) = \sum_i c_i \nabla^i ((\alpha^n) \cdot (a_n)) = \sum_i c_i \cdot (\alpha^n) \cdot (\alpha \nabla)^i (a_n) = (\alpha^n) \cdot P(\alpha \nabla)(a_n)$$

□

2. Teorema: *Se cumple que*

1. $\operatorname{Ker}(\nabla - \alpha \cdot \operatorname{Id})^r = (\alpha^n) \cdot \{(\textit{Pol. } q(n) \textit{ de grado menor que } r)\}$ *(suponemos $\alpha \neq 0$).*
2. *Si $P(x) = (x - \alpha_1)^{r_1} \cdots (x - \alpha_s)^{r_s}$ (suponemos $\alpha_i \neq 0$, $\forall i$), entonces*

$$\operatorname{Ker} P(\nabla) = (\alpha_1{}^n) \cdot \{(\textit{Pol. } q(n) \textit{ de grado} < r_1)\} \oplus \overset{r}{\cdots} \oplus (\alpha_s^n) \cdot \{\textit{Pol. } q(n) \textit{ de grado} < r_s\}.$$

Demostración. 1. En efecto,

$$
\begin{aligned}
(s(n)) \in \operatorname{Ker}(\nabla-\alpha)^r &\iff 0=(\nabla-\alpha)^r(s(n)) \iff 0=(\nabla-\alpha)^r((\alpha^n)\cdot(\alpha^{-n})\cdot(s(n)))\\
&\iff 0=(\alpha^n)\cdot(\alpha\Delta)^r((\alpha^{-n})\cdot(s(n))) \iff 0=\Delta^r((\alpha^{-n})\cdot(s(n)))\\
&\iff \text{ existe un polinomio } q(n) \text{ de grado menor que } r \text{ tal que } \alpha^{-n}\cdot s(n)=q(n)\\
&\iff s(n)=\alpha^n\cdot q(n).
\end{aligned}
$$

2. Es consecuencia de que $\operatorname{Ker} P(\nabla)=\operatorname{Ker}(\nabla-\alpha_1\operatorname{Id})^{n_1}\oplus\cdots\oplus\operatorname{Ker}(\nabla-\alpha_r\operatorname{Id})^{n_r}$ y de 1. □

3. Nota : Supongamos $P(x)\in\mathbb{R}[x]$ y $\alpha=\rho\cdot(\cos\beta+i\cdot\operatorname{sen}\beta)$ es una raíz compleja de $P(x)$, con $\beta\neq 0,\pi$ y $\rho>0$. Entonces, $P(x)=(x-\alpha)^r\cdot(x-\bar\alpha)^r\cdot Q(x)$, con $Q(\alpha),Q(\bar\alpha)\neq 0$. Es fácil probar, a partir del teorema 3.5.2, que

$$
\left\{\begin{array}{l} (s(n))\in\operatorname{Ker}(\nabla-\alpha\operatorname{Id})^n\oplus\operatorname{Ker}(\nabla-\bar\alpha\operatorname{Id})^r\\ \text{tales que } s(n)\in\mathbb{R} \text{ para todo } n\end{array}\right\}=
$$

$$
\left\{(\rho^n\cdot\cos(\beta\cdot n)\cdot\begin{array}{l}\text{Pol. } p(n) \text{ con coef.}\\ \text{reales de grado } <r\end{array}+\rho^n\cdot\operatorname{sen}(\beta\cdot n)\cdot\begin{array}{l}\text{Pol. } q(n) \text{ con coef.}\\ \text{reales de grado } <r\end{array})\right\}.
$$

Ecuaciones en diferencias finitas no homogéneas.

Consideremos una ecuación en diferencias $P(\nabla)(s(n))=(z(n))$. Sea $(s_0(n))$ una solución particular. Entonces, $(s(n))$ es una solución de la ecuación en diferencias si y solo si

$$s(n)=s_0(n)+t(n),\ \text{con } t(n)\in\operatorname{Ker}P(\nabla).$$

Con palabras:

$$
\left[\begin{array}{l}\text{Todas las soluciones}\\ \text{de } P(\nabla)(s(n))=(z(n))\end{array}\right]=\left[\begin{array}{l}\text{Solución particular}\\ \text{de } P(\nabla)(s(n))=(z(n))\end{array}\right]+\left[\begin{array}{l}\text{Todas las soluciones de la}\\ \text{“homogénea” } P(\nabla)(s(n))=0\end{array}\right]
$$

4. Solución particular : *Consideremos la ecuación en diferencias $P(\nabla)(s(n))=(z(n))$. Si tenemos $E=\langle(z_1(n)),\dots,(z_r(n))\rangle$ tal que $(z(n))\in E$ y $\nabla(E)\subseteq E$, consideremos la aplicación lineal $P(\nabla)\colon E\to E$, $(t(n))\mapsto P(\nabla)(t(n))$. Si existe $(s(n))\in E$ tal que $P(\nabla)(s(n))=(z(n))$, obtenemos así una solución particular.*

3.5.1. Ejemplos

5. Sucesión de Fibonacci: Resolvamos la ecuación $a_{n+2} = a_{n+1} + a_n$, con las condiciones iniciales $a_0 = 0, a_1 = 1$. "Esta sucesión fue descrita por Leonardo de Pisa, matemático italiano del siglo XIII también conocido como Fibonacci. Tiene numerosas aplicaciones en ciencias de la computación, matemática y teoría de juegos. También aparece en configuraciones biológicas, como por ejemplo en las ramas de los árboles, en la disposición de las hojas en el tallo, en las flores de alcachofas y girasoles, en las inflorescencias del brécol romanesco, en la configuración de las piñas de las coníferas, en la reproducción de los conejos y en cómo el ADN codifica el crecimiento de formas orgánicas complejas. De igual manera, se encuentra en la estructura espiral del caparazón de algunos moluscos, como el nautilus." (Wikipedia). Tenemos que $a_{n+2} - a_{n+1} - a_n = 0$, luego

$$(\nabla^2 - \nabla - \mathrm{Id})(a_n) = (0).$$

Observemos que $x^2 - x - 1 = (x - \frac{1+\sqrt{5}}{2}) \cdot (x - \frac{1-\sqrt{5}}{2})$, por tanto

$$(a_n) \in \mathrm{Ker}(\nabla^2 - \nabla - \mathrm{Id}) = \mathrm{Ker}(\nabla - \tfrac{1+\sqrt{5}}{2} \cdot \mathrm{Id}) \oplus (\nabla - \tfrac{1-\sqrt{5}}{2} \cdot \mathrm{Id}) = \{(a \cdot (\tfrac{1+\sqrt{5}}{2})^n + b \cdot (\tfrac{1-\sqrt{5}}{2})^n), \forall\, a, b\}.$$

Recordemos que

$$\begin{aligned} 0 &= a_0 = a + b \\ 1 &= a_1 = a \cdot (\tfrac{1+\sqrt{5}}{2}) + b \cdot (\tfrac{1-\sqrt{5}}{2}) \end{aligned}$$

Resulta que $a = \frac{1}{\sqrt{5}}$ y $b = -\frac{1}{\sqrt{5}}$. Luego,

$$a_n = \frac{1}{\sqrt{5}} \cdot (\frac{1+\sqrt{5}}{2})^n - \frac{1}{\sqrt{5}} \cdot (\frac{1-\sqrt{5}}{2})^n.$$

6. Ejercicio : Calcula cuántos números de longitud n se pueden escribir con ceros y unos, de modo que nunca aparezcan dos ceros seguidos (ejemplo: los números de tres cifras cumpliendo lo dicho son $010, 011, 101, 110, 111$, que son cinco distintos).

7. Resolvamos $a_{n+2} + 2a_{n+1} - 6a_n = 2^n$: Tenemos que resolver

$$(\nabla^2 + 2\nabla - 6\,\mathrm{Id})(a_n) = (2^n).$$

Calculemos una solución particular. Sea $E = \langle (2^n) \rangle$ y consideremos la aplicación lineal $\nabla^2 + 2\nabla - 6\,\mathrm{Id} \colon E \to E$. Entonces, si $(\nabla^2 + 2\nabla - 6\,\mathrm{Id})(a2^n) = (2^n)$ entonces $(4 + 4 - 6) \cdot a = 1$ y $a = 1/2$. Una solución particular es $a_n = \frac{1}{2} \cdot 2^n$. Todas las soluciones son

$$\begin{aligned} a_n = 2^{n-1} + \mathrm{Ker}(\nabla^2 + 2\nabla - 6\,\mathrm{Id}) &= 2^{n-1} + \mathrm{Ker}(\nabla - (-1 + \sqrt{7})) + \mathrm{Ker}(\nabla - (-1 - \sqrt{7})) \\ &= 2^{n-1} + a \cdot (-1 + \sqrt{7})^n + b \cdot (-1 - \sqrt{7})^n. \end{aligned}$$

8. Hipoteca: Un banco nos presta un capital K, a devolver en N años, a un tipo de interés anual I. ¿Cuánto dinero D deberemos pagar al año, de modo que todos los años paguemos la misma cantidad y en los N años hayamos saldado nuestra deuda con el banco?

Resolución: Sea i_n el dinero que pagamos en el año n por los intereses del capital que tenemos prestado durante el año n y a_n el dinero que amortizamos en el año n (por el capital K que nos han prestado). Entonces $D = a_n + i_n$. Además, $i_n = I\cdot(K - \sum_{r=1}^{n-1} a_r)$. Por tanto, $D = a_n + I\cdot(K - \sum_{r=1}^{n-1} a_r)$. Si aplicamos el operador diferencia Δ entonces

$$0 = \Delta(a_n) - I\cdot a_n = (\nabla - (1+I))(a_n).$$

Por tanto, $a_n = (1+I)^n \cdot \lambda$ y

$$D = a_1 + IK = (1+I)\cdot\lambda + IK. \qquad (*)$$

Tenemos que calcular λ. Nos falta decir que amortizamos la hipoteca en N años, es decir, $K = \sum_{r=1}^{N} a_r$, que equivale a decir que $D = a_{N+1} = (1+I)^{N+1}\cdot\lambda$. Despejando λ y sustituyendo su valor en $(*)$ obtendremos que

$$D = \frac{IK}{1 - \frac{1}{(1+I)^N}}.$$

9. Préstamos con gradiente lineal: Por la compra de un coche en un concesionario pagaremos cada año n un dinero d_n de modo que $d_n = A + G\cdot(n-1)$ (con $A = 1000$ y $G = 100$), durante $N = 20$ años. Se supone que el tipo de interés anual es $I = 5\%$. Determina cuál es el valor K del coche (en la actualidad).

Resolución: Podemos decir que nos han prestado un capital K. Sea i_n es el dinero que pagamos en el año n por los intereses del capital que tenemos prestado durante el año n y a_n el dinero que amortizamos en el año n por el capital prestado. Entonces, $d_n = i_n + a_n$. Tenemos que $i_n = I\cdot(K - \sum_{r=1}^{n-1} a_r)$. Por tanto,

$$A + G\cdot(n-1) = d_n = a_n + I\cdot(K - \sum_{r=1}^{n-1} a_r). \qquad (*)$$

Para $n = 1$, tenemos que

$$A = a_1 + I\cdot K. \qquad (**)$$

Tenemos que determinar a_1. Aplicando Δ en $(*)$ obtenemos

$$G = \Delta(a_n) - I\cdot(a_n) = (\Delta - I)(a_n) = (\nabla - (1+I))(a_n).$$

Una solución particular, es $a_n = \frac{-G}{I}$ y todas las soluciones son $a_n = \frac{-G}{I} + cte \cdot (1+I)^n$. Nos falta imponer que $\sum_{i=1}^N a_n = K$. Puede comprobarse que

$$K = \sum_{i=1}^{N} a_n = \frac{-G}{I} \cdot N + cte \cdot \frac{(1+I)^{N+1} - (1+I)}{I}.$$

Luego, $cte = \frac{IK+GN}{(1+I)\cdot((1+I)^N-1)}$ y $a_1 = \frac{-G}{I} + \frac{IK+GN}{(1+I)^N-1}$. Sustituyendo el valor de a_1 en $(**)$, y despejando K obtenemos que

$$K = \frac{A}{I} \cdot \frac{(1+I)^N - 1}{(1+I)^N} + \frac{G}{I^2} \cdot \frac{(1+I)^N - 1 - IN}{(1+I)^N} = 22311'1.$$

Supongamos que voy a un banco que me ofrece un interés anual $I = 1\%$ por mi dinero. Tendría que depositar $K = 34592$ euros para que el banco me fuese dando cada año lo que el concesionario me pide (y al final el banco quedase en paz conmigo). En fin, un coche de 22311 euros me ha costado 34592 euros.

10. Sumatorios Dada una sucesión de números complejos $(s(n))$, definamos la sucesión $S(n) := \sum_{i=0}^{n-1} s(i)$ (y $S(0) := 0$). Entonces, $\Delta((S(n)) = (S(n+1) - S(n)) = (s(n))$. Denotaremos

$$\frac{1}{\Delta'}(s(n)) := (\sum_{i=0}^{n-1} s(i)).$$

Hemos dicho que $\Delta(\frac{1}{\Delta'}(s(n))) = (s(n))$. Luego, $\Delta^{-1}(s(n)) = \frac{1}{\Delta'}(s(n)) + \{(cte)$, para toda $cte \in \mathbb{C}\}$.

Se define $\binom{n}{i} := \frac{n\cdot(n-1)\cdots(n-i+1)}{i\cdot(i-1)\cdots 2\cdot 1}$. Es fácil comprobar que $\Delta(\binom{n}{r}) = (\binom{n}{r-1})$. Por tanto, $\frac{1}{\Delta'}(\binom{n}{r}) = \binom{n}{r+1}$.

Es fácil probar la igualdad $\langle (1), (n), \dots, (n^r) \rangle = \langle (\binom{n}{0}), (\binom{n}{1}), \dots, (\binom{n}{r}) \rangle$. Si escribimos $p(n) = \sum_{i=0}^r \lambda_i \binom{n}{i}$, entonces

$$\lambda_0 = p(0), \dots, \ \lambda_i = p(i) - \lambda_{i-1} \binom{i}{i-1} - \cdots - \lambda_0, \text{ etc.}$$

Por tanto, $\sum_{i=0}^{n-1} p(i) = \frac{1}{\Delta'} p(n) = \frac{1}{\Delta'} \sum_{i=0}^r \lambda_i \binom{n}{i} = \sum_{i=0}^r \lambda_i \binom{n}{i+1}$ y $\sum_{i=0}^n p(i) = \sum_{i=0}^r \lambda_i \binom{n+1}{i+1}$.

11. Ejercicio: Calcula $\sum_{i=0}^n (i^2 + i - 3)$.

3.6. Cambio de base en operadores lineales. Determinante

Si B es una base de E y $T\colon E \to E$ un operador lineal, diremos que la matriz de T en las bases B, B es la matriz de T en la base B, y por brevedad escribiremos

$$T_B := T_B^B.$$

1. Proposición: *Sea $T\colon E \to E$ un operador lineal. Sean B y B' dos bases de E. Entonces,*

$$T_{B'} = \mathrm{Id}_B^{B'} \cdot T_B \cdot \mathrm{Id}_{B'}^{B} = \mathrm{Id}_B^{B'} \cdot T_B \cdot (\mathrm{Id}_B^{B'})^{-1}.$$

Demostración. Es consecuencia del párrafo 2.9.3.

□

2. Proposición: *Sea $T\colon E \to E$ un operador lineal. Sean B y B' dos bases de E. Entonces,*

$$\det(T_B) = \det(T_{B'}).$$

Se dice que $\det(T) := \det(T_B)$ *es el determinante del operador lineal T.*

Demostración. Sea C la matriz de cambio de base de B a B'. Por la proposición 3.6.1

$$T_{B'} = C \cdot T_B \cdot C^{-1}$$

Luego, $\det(T_{B'}) = \det(C \cdot T_B \cdot C^{-1}) = \det(C) \cdot \det(T_B) \cdot \det(C)^{-1} = \det(T_{B'})$.

□

3. Ejemplo: Calculemos el determinante del giro de α radianes alrededor del origen de $\mathbb{R}^2$ (en sentido horario): La matriz del giro en la base usual es

$$\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

cuyo determinante es $(\cos\alpha)^2 + (\operatorname{sen}\alpha)^2 = 1$.

3.7. Autovectores y autovalores

Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal. Nuestro objetivo es encontrar (si existe) una base $B = \{e_1, \dots, e_n\}$ de E tal que $T(e_i) = \lambda_i \cdot e_i$, es decir, $T_B = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.

1. Definición: Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal. Diremos que $0 \neq e \in E$ es un autovector de T si $T(e) = \lambda \cdot e$, para cierto $\lambda \in \mathbb{R}$, en este caso diremos que λ es un autovalor de T (o que es el autovalor asociado al autovector e).

Sea $(a_{ij}) = T_B$ la matriz asociada a T en una base $B = \{e_1, \dots, e_n\}$. Entonces, $e \in E$ es un autovector de T de autovalor λ si y solo si $T_B \cdot e^B = \lambda \cdot e^B$, es decir, si $e_B = (x_1, \dots, x_n)$

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Diremos que $(x_1, \dots, x_n)$ es un autovector de A de autovalor λ.

2. Ejemplo: Calculemos los autovectores del giro de α radianes alrededor del origen de $\mathbb{R}^2$ (en sentido horario). La matriz del giro en la base usual es

$$\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

Buscamos vectores $(x, y) \in \mathbb{R}^2$ no nulos tales que $T(x, y) = \lambda \cdot (x, y)$. Tenemos que $T(x, y) = (x\cos\alpha - y\operatorname{sen}\alpha, x\operatorname{sen}\alpha + y\cos\alpha)$. Entonces,

$$\begin{aligned} x\cos\alpha - y\operatorname{sen}\alpha &= \lambda \cdot x \\ x\operatorname{sen}\alpha + y\cos\alpha &= \lambda \cdot y \end{aligned}$$

Es decir,

$$\begin{aligned} x \cdot (\cos\alpha - \lambda) - y \cdot \operatorname{sen}\alpha &= 0 \\ x \cdot \operatorname{sen}\alpha + y \cdot (\cos\alpha - \lambda) &= 0 \end{aligned}$$

El sistema es compatible y $(x, y) = (0, 0)$ es una solución. Tiene más soluciones si y solo si

$$0 = \begin{vmatrix} \cos\alpha - \lambda & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha - \lambda \end{vmatrix} = \lambda^2 - (2\cos\alpha) \cdot \lambda + 1$$

Es decir, si y solo si $\lambda = \frac{2\cos\alpha \pm \sqrt{4\cos\alpha - 4}}{2} \in \mathbb{R}$. Lo cual solo sucede si $\alpha = 0, \pi$. Si $\alpha = 0$, entonces todos los vectores (no nulos) son autovectores de autovalor 1. Si $\alpha = \pi$ todos los vectores (no nulos) son autovectores de autovalor -1. Para $\alpha \neq 0, \pi$ no hay autovectores.

3. Ejercicio : Sea $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ el operador lineal de matriz en la base usual

$$T = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix}.$$

Comprueba que los vectores $(1,-1,0)$ y $(0,0,1)$ son autovectores de T.

4. Proposición : *Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal. Entonces,*

1. *{Autovectores de T de autovalor λ}* $= \operatorname{Ker}(T - \lambda \cdot \operatorname{Id})\backslash\{0\}$.
2. *$\lambda \in \mathbb{R}$ es un autovalor de T si y solo si* $\det(T - \lambda \cdot \operatorname{Id}) = 0$.

Demostración. 1. En efecto, $0 \neq e \in E$ es un autovector de autovalor λ si y solo si $T(e) = \lambda \cdot e$, que equivale a $(T - \lambda \cdot \operatorname{Id})(e) = 0$, es decir, si y solo si $e \in \operatorname{Ker}(T - \lambda \cdot \operatorname{Id})$.

2. En efecto, $\lambda \in \mathbb{R}$ es un autovalor de T si y solo si existe un autovector $0 \neq e \in E$ de autovalor λ, es decir, $\operatorname{Ker}(T - \lambda \cdot \operatorname{Id}) \neq 0$. Ahora bien, $\operatorname{Ker}(T - \lambda \cdot \operatorname{Id}) = 0$ si y solo si $\operatorname{rango}(T - \lambda \cdot \operatorname{Id}) = \dim E$ y esto sucede si y solo si $\det(T - \lambda \cdot \operatorname{Id}) \neq 0$. Luego, $\operatorname{Ker}(T - \lambda \cdot \operatorname{Id}) \neq 0$ si y solo si $\det(T - \lambda \cdot \operatorname{Id}) = 0$ y hemos concluido. □

5. Definición : Sea $T\colon E \to E$ un operador lineal. Llamaremos polinomio característico de T, que denotaremos $c_T(x)$, al polinomio

$$c_T(x) := \det(T - x \cdot \operatorname{Id}).$$

Si $A = (a_{ij})$ es la matriz asociada a T en una base, entonces

$$c_T(x) = \det(A - x \cdot \operatorname{Id}) = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{vmatrix} =: c_A(x),$$

que es un polinomio mónico de grado $n = \dim E$. Si $A' = C \cdot A \cdot C^{-1}$, entonces

$$c_{A'}(x) := |A' - x \cdot \operatorname{Id}| = C \cdot A \cdot C^{-1} - x \cdot \operatorname{Id}| = |C \cdot (A - x \cdot \operatorname{Id}) \cdot C^{-1}| = |A - x \cdot \operatorname{Id}| = c_A(x).$$

6. Proposición : *Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal. Entonces, $\lambda \in \mathbb{R}$ es un autovalor de T si y solo si λ es una raíz del polinomio característico de T.*

7. Nota: Diremos que $\lambda \in \mathbb{C}$ es un autovalor imaginario de T, si es una raíz del polinomio característico de T.

8. Ejemplo: Calculemos los autovectores y autovalores del operador lineal del ejercicio 3.7.3: Calculemos primero el polinomio característico de T

$$c_T(x) = \begin{vmatrix} 1-x & 1 & 0 \\ 2 & 2-x & 0 \\ 1 & 1 & 2-x \end{vmatrix} = -(x^3 - 5x^2 + 6x) = -x(x-2)(x-3).$$

Los autovalores son las raíces $0,2,3$ de $c_T(x)$.

Calculemos los autovectores de autovalor 0, es decir, $\operatorname{Ker} T$,

$$\begin{aligned} \operatorname{Ker} T &= \{(x,y,z) \in \mathbb{R}^3 : T(x,y,z) = (0,0,0)\} \\ &= \{(x,y,z) \in \mathbb{R}^3 : (x+y, 2x+2y, x+y+2z) = (0,0,0)\} = \langle (1,-1,0) \rangle. \end{aligned}$$

Calculemos los autovectores de autovalor 2, es decir, $\operatorname{Ker}(T - 2\,\mathrm{Id})$,

$$\begin{aligned} \operatorname{Ker}(T - 2\,\mathrm{Id}) &= \{(x,y,z) \in \mathbb{R}^3 : (T - 2\,\mathrm{Id})(x,y,z) = (0,0,0)\} \\ &= \{(x,y,z) \in \mathbb{R}^3 : (-x+y, 2x, x+y) = (0,0,0)\} = \langle (0,0,1) \rangle. \end{aligned}$$

Calculemos los autovectores de autovalor 3, es decir, $\operatorname{Ker}(T - 3\,\mathrm{Id})$,

$$\begin{aligned} \operatorname{Ker}(T - 3\,\mathrm{Id}) &= \{(x,y,z) \in \mathbb{R}^3 : (T - 3\,\mathrm{Id})(x,y,z) = (0,0,0)\} \\ &= \{(x,y,z) \in \mathbb{R}^3 : (-2x+y, 2x-y, x+y-z) = (0,0,0)\} = \langle (1,2,3) \rangle. \end{aligned}$$

La matriz de T en la base $B = \{(1,-1,0),(0,0,1),(1,2,3)\}$ es

$$T_B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

y podemos comprobar la fórmula de la proposición 3.6.1:

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}^{-1}.$$

3.8. Operadores diagonalizables

Buscamos bases en las que la matriz asociada a un operador lineal sea muy simple.

1. Definición: Se dice que un operador lineal $T\colon E \to E$ es diagonalizable si existe una base de E, $B = \{e_1, \dots, e_n\}$, tal que $T(e_i) = \lambda_i \cdot e_i$, para todo i, es decir, si la matriz asociada a T en la base B es

$$T_B = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Demos condiciones computables que caractericen los operadores diagonalizables.

2. Proposición: *Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal y escribamos el polinomio característico $c_T(x) = \pm(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j \in \mathbb{C}$ si $i \neq j$. Entonces, el operador lineal T es diagonalizable $\iff$ $\alpha_i \in \mathbb{R}$ y $n_i = \dim \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id})$, para todo i.*

Demostración. $\Rightarrow$) Si T es diagonalizable existe una base en la que la matriz asociada a T es diagonal, reordenando los vectores de esta base podemos suponer que

$$T \equiv \begin{pmatrix} \alpha_1 & & & & & & \\ & \overset{n_1}{\ddots} & & & & & \\ & & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r & & \\ & & & & & \overset{n_r}{\ddots} & \\ & & & & & & \alpha_r \end{pmatrix}$$

con $\alpha_i \neq \alpha_j \in \mathbb{R}$, si $i \neq j$. En este caso, $c_T(x) = \pm(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, con $\alpha_i \neq \alpha_j \in \mathbb{R}$, para $i \neq j$, y $\dim \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) = n_i$, para todo i.

$\Leftarrow$) Observemos que $\sum_i n_i = \operatorname{gr}(c_T(x)) = \dim E$. Además,

$$\operatorname{Ker}((T-\alpha_1 \mathrm{Id})\cdots(T-\alpha_r \mathrm{Id})) = \operatorname{Ker}(T-\alpha_1 \mathrm{Id}) \oplus \cdots \oplus \operatorname{Ker}(T-\alpha_r \mathrm{Id}).$$

Por dimensiones, $E = \operatorname{Ker}(T-\alpha_1 \mathrm{Id}) \oplus \cdots \oplus \operatorname{Ker}(T-\alpha_r \mathrm{Id})$. Si consideramos una base en cada $\operatorname{Ker}((T-\alpha_i \mathrm{Id})$, obtendremos una base de E en la que T diagonaliza.

□

3. Proposición: *Sea $T\colon E \to E$ un operador $\mathbb{R}$-lineal. Si el polinomio característico de T tiene todas sus raíces en $\mathbb{R}$ y son distintas, entonces T diagonaliza.*

Demostración. Tenemos $c_T(x) = \pm(x-\alpha_1)\cdots(x-\alpha_n)$, con $\alpha_i \neq \alpha_j$ para $i \neq j$ y $n = \dim E$. Por tanto, $\dim \operatorname{Ker}(T-\alpha_i \operatorname{Id}) \geq 1$ y por dimensiones tenemos que

$$\operatorname{Ker}((T-\alpha_1\cdot\operatorname{Id})\cdots(T-\alpha_n \operatorname{Id})) = \operatorname{Ker}(T-\alpha_1\cdot\operatorname{Id})\oplus\cdots\oplus\operatorname{Ker}(T-\alpha_n \operatorname{Id}) = E.$$

Si consideramos una base en cada $\operatorname{Ker}(T-\alpha_i \operatorname{Id})$, obtendremos una base de E en la que T diagonaliza.

□

4. Ejemplo: Comprobemos si la matriz $A = \begin{pmatrix} 0'5 & 0'4 & 0'6 \\ 0'3 & 0'4 & 0'3 \\ 0'2 & 0'2 & 0'1 \end{pmatrix}$ diagonaliza. Tenemos que $c_A(x) = -(x-1)\cdot(x-0'1)\cdot(x+0'1)$, luego diagonaliza (en una base de autovectores). Calculemos la base.

$$\operatorname{Ker}(A-\operatorname{Id}) = \langle(16,11,6)\rangle, \quad \operatorname{Ker}(A-0'1\operatorname{Id}) = \langle(1,-1,0)\rangle, \quad \operatorname{Ker}(A+0'1\operatorname{Id}) = \langle(1,0,-1)\rangle.$$

Luego, $\{(16,11,6),(1,-1,0),(1,0,-1)\}$ es una base donde A diagonaliza y

$$A = \begin{pmatrix} 16 & 1 & 1 \\ 11 & -1 & 0 \\ 6 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0'1 & 0 \\ 0 & 0 & -0'1 \end{pmatrix} \cdot \begin{pmatrix} 16 & 1 & 1 \\ 11 & -1 & 0 \\ 6 & 0 & -1 \end{pmatrix}^{-1}.$$

5. Ejemplo: Comprobemos que la matriz $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ diagonaliza. Tenemos que $c_A(x) = -(x-1)^3$ y $\operatorname{Ker}(A-\operatorname{Id}) = \langle(1,0,0)\rangle$. Por la proposición 3.8.2, no diagonaliza.

3.9. Operadores triangulables

1. Definición: Diremos que un operador lineal $T: E \to E$ es triangulable si existe una base $B = \{e_1,\ldots,e_n\}$ de E de modo que

$$T_B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

2. Proposición: *Un operador $\mathbb{R}$-lineal $T\colon E \to E$ es triangulable si y solo si todos sus autovalores (imaginarios o no) son reales.*

Demostración. $\Rightarrow$) Sea B una base de E en la que T triangule

$$T_B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Entonces,

$$c_T(x) = \begin{vmatrix} a_{11}-x & a_{12} & \dots & a_{1n} \\ 0 & a_{22}-x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}-x \end{vmatrix} = (a_{11}-x)\cdots(a_{nn}-x).$$

Luego, los autovalores son $a_{11}, a_{22}, \dots, a_{nn} \in \mathbb{R}$.

$\Leftarrow$) Si $\dim_{\mathbb{R}} E = 1$ entonces es obvio. Supongamos que hemos demostrado la proposición siempre que $\dim_{\mathbb{R}} E = 1, \dots, n-1$. Tenemos que probarla para $\dim_{\mathbb{R}} E = n$.

Sea $\lambda_1 \in \mathbb{R}$ una raíz del polinomio característico de T y $e_1 \in E$ un autovector de autovalor λ_1. Sea $e_1, \dots, e_n$ una base de E. La matriz de T en esta base es

$$\begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Sea $E' = \langle e_2, \dots, e_n \rangle \subset E$ y $T'\colon E' \to E'$ la aplicación lineal cuya matriz asociada en la base $B' = \{e_2, \dots, e_n\}$ de E' es

$$T'_{B'} = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Observemos que $c_T(x) = (\lambda_1 - x) \cdot c_{T'}(x)$, luego todos las raíces de $c_{T'}(x)$ son raíces de $c_T(x)$ y son reales. Entonces, existe una base $B' = \{e'_2, \dots, e'_n\}$ de E' en la que la matriz de T' es triangular,

$$T'_{B'} = \begin{pmatrix} a'_{22} & \cdots & a'_{2n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a'_{nn} \end{pmatrix}.$$

Observemos que $T(e)-T'(e)\in\langle e_1\rangle$, para todo $e\in E'$. Por tanto, la matriz de T en la base $\{e_1,e'_2,\dots,e'_n\}$ es

$$\begin{pmatrix} \lambda_1 & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a'_{nn} \end{pmatrix}.$$

□

Igualmente, obtenemos la proposición que sigue.

3. Proposición: *Sea E un $\mathbb{C}$-espacio vectorial y $T\colon E\to E$ un operador $\mathbb{C}$-lineal. Entonces, T es triangulable.*

4. Teorema de Cayley-Hamilton: *Sea $T\colon E\to E$ un operador lineal y $c_T(x)$ su polinomio característico. Entonces,*

$$\boxed{c_T(T)=0}$$

Demostración. Sea $A\in M_n(\mathbb{R})$ la matriz asociada a T en una base B y sea $c_A(x)=\det(A-x\cdot\mathrm{Id})$. Tenemos que probar que $c_A(A)=0$. Consideremos el operador $\mathbb{C}$-lineal $T'\colon\mathbb{C}^n\to\mathbb{C}^n$ de matriz A, en la base usual. Si $c_{T'}(T')=0$, entonces $c_A(A)=0$. En conclusión, podemos suponer que T es un operador $\mathbb{C}$-lineal.

Por la proposición anterior sabemos que existe una base $B=\{e_1,\dots,e_n\}$ de E en la que la matriz asociada a T es triangular

$$\begin{pmatrix} \alpha_1 & a_{12} & \cdots & a_{1n} \\ 0 & \alpha_2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

Entonces, el polinomio característico $c_T(x)=\pm(x-\alpha_1)\cdots(x-\alpha_n)$. Sea $E_i:=\langle e_1,\dots,e_i\rangle$. Observemos que $(T-\alpha_i\cdot\mathrm{Id})(e_j)=(\alpha_j-\alpha_i)\cdot e_j+a_{j-1j}\cdot e_{j-1}+\cdots+a_{1j}\cdot e_1$. Por tanto,

$$(T-\alpha_i\cdot\mathrm{Id})(E_i)\subseteq E_{i-1},$$

para todo i. Observemos que $E=E_n$ y

$$\begin{aligned} c_T(T)(E_n) &= (T-\alpha_1\,\mathrm{Id})\cdots(T-\alpha_n\,\mathrm{Id})(E_n)\subseteq(T-\alpha_1\,\mathrm{Id})\cdots(T-\alpha_{n-1}\,\mathrm{Id})(E_{n-1}) \\ &\subseteq(T-\alpha_1\,\mathrm{Id})\cdots(T-\alpha_{n-2}\,\mathrm{Id})(E_{n-2})\subseteq\cdots\subseteq(T-\alpha_1\,\mathrm{Id})(E_1)=0. \end{aligned}$$

Es decir, $c_T(T)=0$.

□

5. Lema : *Sean E_1 y E_2 dos subespacios vectoriales de E de modo que $E = E_1 \oplus E_2$. Sean $T_1: E_1 \to E_1$ y $T_2: E_1 \to E_2$ dos operadores lineales y $T: E_1 \oplus E_2 \to E_1 \oplus E_2$ el operador lineal definido por $T(e_1+e_2) = T_1(e_1)+T_2(e_2)$, para todo $e_1 \in E_1$ y todo $e_2 \in E_2$. Entonces, $c_T(x) = c_{T_1}(x) \cdot c_{T_2}(x)$.*

Demostración. Sea $B_1 = \{e_1,\dots,e_r\}$ una base de E_1 y $B_2 = \{v_1,\dots,v_s\}$ una base de E_2. Entonces, $B = \{e_1,\dots,e_r,v_1,\dots,v_s\}$ es una base de $E_1 \oplus E_2$. Si $(a_{ij}) \in M_r(\mathbb{R})$ es la matriz asociada a T_1 en la base B_1 y $(b_{uv}) \in M_s(\mathbb{R})$ es la matriz asociada a T_2 en la base B_2, entonces la matriz asociada a T en la base B es $\left(\begin{array}{c|c} (a_{ij}) & 0 \\ \hline 0 & (b_{uv}) \end{array}\right)$ y

$$c_T(x) = \left|\begin{array}{c|c} (a_{ij}) - x\,\mathrm{Id} & 0 \\ \hline 0 & (b_{uv}) - x\,\mathrm{Id} \end{array}\right| = |(a_{ij}) - x\,\mathrm{Id}| \cdot |(b_{uv}) - x\,\mathrm{Id}| = c_{T_1}(x) \cdot c_{T_2}(x).$$

□

6. Lema : *Sea E un espacio vectorial de dimensión n y $T: E \to E$ un operador lineal tal que $(T - \alpha\,\mathrm{Id})^m = 0$, para cierto m. Entonces, $c_T(x) = \pm(x-\alpha)^n$.*

Demostración. Sea $A \in M_n(\mathbb{R}) \subset M_n(\mathbb{C})$ la matriz asociada a T en cierta base. Sabemos que $(A - \alpha\,\mathrm{Id})^m = 0$ y tenemos que probar que $c_A(x) = \pm(x-\alpha)^n$. Sea $\beta \in \mathbb{C}$ distinto de α. Sabemos que $\mathrm{Ker}((A - \alpha\,\mathrm{Id})^m \cap \mathrm{Ker}(A - \beta\,\mathrm{Id}) = 0$ porque $(x-\alpha)^m$ es primo con $(x-\beta)$. Como $\mathrm{Ker}((A - \alpha\,\mathrm{Id})^m) = \mathbb{C}^n$, entonces $\mathrm{Ker}(A - \beta\,\mathrm{Id}) = 0$ y β no es un autovalor (imaginario o no) de A.

□

7. Corolario : *Sea $T: E \to E$ un operador lineal. Si $c_T(x) = \pm(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$ con $\alpha_i \neq \alpha_j$ si $i \neq j$. Entonces,*

1. $E = \mathrm{Ker}(T - \alpha_1\,\mathrm{Id})^{n_1} \oplus \cdots \oplus \mathrm{Ker}(T - \alpha_r\,\mathrm{Id})^{n_r}$.
2. $\dim \mathrm{Ker}(T - \alpha_i\,\mathrm{Id})^{n_i} = n_i$, *para todo i.*

Demostración. Por el teorema de Cayley-Hamilton

$$E = \mathrm{Ker}\, c_T(T) = \mathrm{Ker}((T - \alpha_1\,\mathrm{Id})^{n_1}\cdots(T - \alpha_r\,\mathrm{Id})^{n_r}) = \mathrm{Ker}(T - \alpha_1\,\mathrm{Id})^{n_1} \oplus \cdots \oplus \mathrm{Ker}(T - \alpha_r\,\mathrm{Id})^{n_r}.$$

Denotemos $E_i = \mathrm{Ker}(T - \alpha_i\,\mathrm{Id})^{n_i}$. Observemos que si $e \in \mathrm{Ker}(T - \alpha_i\,\mathrm{Id})^{n_i}$ entonces $T(e) \in \mathrm{Ker}(T - \alpha_i\,\mathrm{Id})^{n_i}$, porque $(T - \alpha_i)^{n_i}(T(e)) = T((T - \alpha_i)^{n_i}(e)) = T(0) = 0$.

Sea $T_i: E_i \to E_i$ el operador lineal definido por $T_i(e) = T(e)$, para todo $e \in E_i$. Observemos que $T(e_1 + \cdots + e_r) = T(e_1) + \cdots + T(e_r) = T_1(e_1) + \cdots + T_r(e_r)$, para $e_1 \in E_1, \dots, e_r \in E_r$. Por el lema 3.9.5, $c_T(x) = c_{T_1}(x) \cdots c_{T_r}(x)$.

Observemos que $(T_i - \alpha_i \,\mathrm{Id})^{n_i}(e) = (T - \alpha_i \,\mathrm{Id})^{n_i}(e) = 0$ para todo $e \in E_i$. Por el lema 3.9.6, $c_{T_i}(x) = \pm(x - \alpha_i)^{m_i}$, donde $m_i = \dim E_i$. Entonces,

$$(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r} = \pm c_T(x) = (x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}.$$

Luego, $m_i = n_i$, para todo i, es decir, $\dim \mathrm{Ker}(T - \alpha_i \,\mathrm{Id})^{n_i} = n_i$.

□

3.10. Potencias de una matriz cuadrada

Sea $T : E \to E$ un operador lineal diagonalizable. Supongamos que nos piden calcular T^{100}. Si T_B es la matriz asociada a T en una base B dada, entonces la matriz asociada a T^{100} en la base B es $(T_B)^{100}$, cuyo cálculo es muy arduo. Existe una base $B' = \{e_1, e_2, \dots, e_n\}$ tal que $T(e_1) = \lambda_1 e_1$, $T(e_2) = \lambda_2 e_2, \dots$, $T(e_n) = \lambda_n e_n$. Por tanto, $T^{100}(e_1) = \lambda_1^{100} \cdot e_1$, $T^{100}(e_2) = \lambda_2^{100} \cdot e_2, \dots$, y $T^{100}(e_n) = \lambda_n^{100} \cdot e_3$ y

$$(T^{100})_{B'} = \begin{pmatrix} \lambda_1^{100} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^{100} \end{pmatrix}.$$

Luego,

$$(T^{100})_B = \mathrm{Id}_{B'}^{B} \cdot (T^{100})_{B'} \cdot (\mathrm{Id}_{B'}^{B})^{-1}.$$

1. Ejemplo : Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ el operador lineal de matriz $A = \begin{pmatrix} -1 & 1 & 1 \\ -3 & 3 & 1 \\ -3 & 1 & 3 \end{pmatrix}$ en la base usual. Para calcular A^{100} calculemos primero el polinomio característico del operador T, que es $c_T(x) = -(x-1)\cdot(x-2)^2$, luego los autovalores, que son 1 y 2, y después los autovectores:

$$\mathrm{Ker}(T - \mathrm{Id}) = \{(x,y,z) \in \mathbb{R}^3 \text{ tales que } (A - \mathrm{Id})\cdot(x,y,z)^t = (0,0,0)^t\} = \langle (1,1,1) \rangle.$$

$$\mathrm{Ker}(T - 2\,\mathrm{Id}) = \{(x,y,z) \in \mathbb{R}^3 \text{ tales que } (A - 2\,\mathrm{Id})\cdot(x,y,z)^t = (0,0,0)^t\} = \langle (1,0,3), (0,1,-1) \rangle.$$

Una base de $\mathbb{R}^3$, formada por autovectores de T, es $B = \{(1,1,1), (1,0,3), (0,1,-1)\}$. Entonces,

$$T_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad T_B^{100} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{100} & 0 \\ 0 & 0 & 2^{100} \end{pmatrix}$$

y la matriz de T^{100} en la base usual es

$$A^{100} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{100} & 0 \\ 0 & 0 & 2^{100} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 3 & -1 \end{pmatrix}^{-1}.$$

3.10.1. Comportamiento asintótico de las potencias de una matriz

Dado $e = (x_1, \dots, x_m) \in \mathbb{R}^m$ definamos $||e||_1 := |x_1| + \cdots + |x_n|$. Observemos que e está "cerca" de e' si $||e - e'||_1$ es "pequeño". Decimos que una sucesión de vectores $\{v_n \in \mathbb{R}^m\}$ converge a $v \in \mathbb{R}^m$, y lo denotamos $\lim\limits_{n\to\infty} v_n = v$, si $\lim\limits_{n\to\infty} ||v - v_n||_1 = 0$.[3] Si $\phi\colon \mathbb{R}^m \to \mathbb{R}^s$ es una aplicación $\mathbb{R}$-lineal es fácil probar que si una sucesión de vectores $\{v_n \in \mathbb{R}^m\}$ converge a $v \in \mathbb{R}^m$ entonces la sucesión de vectores $\{\phi(v_n) \in \mathbb{R}^s\}$ converge a $\phi(v) \in \mathbb{R}^s$.

Sea E un $\mathbb{R}$-espacio vectorial. Dada una base B en E tenemos el isomorfismo $E \simeq \mathbb{R}^m$, $e \mapsto e_B$. Diremos que una sucesión de vectores $\{e_n \in E\}$ converge a $e \in E$ si $\{(e_n)_B \in \mathbb{R}^m\}$ converge a e_B. Puede probarse que esta definición no depende de la base B escogida.

2. Ejemplo: Sea $T\colon E \to E$ un operador lineal diagonalizable, tal que todos sus autovalores son de módulo menor que 1. Sea $\{e_1, \dots, e_m\}$ una base de autovectores. Entonces, $T(e_i) = \lambda_i \cdot e_i$, con $|\lambda_i| < 1$. Dado un vector cualquiera $e = \sum\limits_{i=1}^m \mu_i \cdot e_i$, tendremos que

$$\lim_{n\to\infty} T^n(e) = \lim_{n\to\infty} (T^n(\sum_{i=1}^m \mu_i \cdot e_i)) = \lim_{n\to\infty} (\sum_{i=1}^m \lambda_i^n \cdot \mu_i \cdot e_i) = 0.$$

3. Lema: *Sea $T\colon E \to E$ un operador lineal, que tiene un único autovalor (imaginario o no) α. Si $|\alpha| < 1$, entonces*

$$\lim_{m\to\infty} T^m = 0,$$

es decir, $\lim\limits_{m\to\infty} T^m(e) = 0$, para todo $e \in E$.

Demostración. El polinomio característico de T es $(x-\alpha)^n$. Por el teorema de Cayley-Hamilton $(T - \alpha \cdot \mathrm{Id})^n = 0$. Escribamos $S = T - \alpha \cdot \mathrm{Id}$. Entonces, $S^n = 0$, $T = \alpha \cdot \mathrm{Id} + S$ y S conmuta con $\alpha \cdot \mathrm{Id}$. Para $m > n$,

$$T^m = (\alpha \cdot \mathrm{Id} + S)^m = \alpha^m \cdot \mathrm{Id} + \binom{m}{1} \alpha^{m-1} \cdot S + \cdots + \binom{m}{n-1} \alpha^{m-n+1} \cdot S^{n-1}$$

[3] Un teorema matemático bonito dice que todas las normas de un espacio vectorial real (de dimensión finita) son topológicamente equivalentes, y que por ello en el estudio de la convergencia de una sucesión de vectores puede escogerse la norma que más convenga.

y $\lim_{m\to\infty} T^m = 0\cdot \mathrm{Id} + 0\cdot S + \cdots + 0\cdot S^{n-1} = 0.$

□

4. Definiciones: Diremos que un autovalor $\lambda \in \mathbb{R}$ de un operador lineal $T\colon E \to E$ es dominante, si todo otro autovalor $\alpha \in \mathbb{C}$ de T cumple que $|\alpha| < |\lambda|$.

Diremos que un autovalor $\lambda \in \mathbb{R}$ de un operador lineal T es de multiplicidad r si $c_T(x) = (x-\lambda)^r \cdot q(x)$, con $q(\lambda) \neq 0$.

5. Teorema: *Sea $T\colon E \to E$ un operador lineal.*

1. *Si todos los autovalores de T (imaginarios o no) son de módulo menor estricto que* 1*, entonces* $\lim_{m\to\infty} T^m = 0$.

2. *Si* 1 *es un autovalor dominante de multiplicidad* 1 *entonces* $\operatorname{Ker}(T-\mathrm{Id}) = \langle e\rangle$ *y para cada $v \in E$*

$$\lim_{m\to\infty} T^m(v) = \mu\cdot e,$$

para cierto $\mu \in \mathbb{R}$.

Demostración. 2. Escribamos $c_T(x) = \pm(x-1)\cdot(x-\alpha_1)^{n_1}\cdots(x-\alpha_r)^{n_r}$, con $|\alpha_i| < 1$ para todo i y $\alpha_i \neq \alpha_j$ si $i \neq j$. Tenemos que

$$E = \operatorname{Ker}(T-\mathrm{Id}) \oplus \operatorname{Ker}(T-\alpha_1\mathrm{Id})^{n_1} \oplus \cdots \oplus \operatorname{Ker}(T-\alpha_r\mathrm{Id})^{n_r}.$$

Escribamos $E_i = \operatorname{Ker}(T-\alpha_i\mathrm{Id})^{n_i}$ y sea $T_i\colon E_i \to E_i$, $T_i(e_i) := T(e_i)$, para todo $e_i \in E_i$. Como el único autovalor (imaginario o no) de T restringido a E_i es α_i, entonces el único autovalor (imaginario o no) de T_i es α_i. Dado $v \in E$, escribamos $v = \mu\cdot e + e'_1 + \cdots + e'_r$, con $e'_i \in E_i$. Entonces,

$$\begin{aligned}
\lim_{m\to\infty} T^m(v) &= \lim_{m\to\infty} T^m(\mu\cdot e + e'_1 + \cdots + e'_n)\\
&= \lim_{m\to\infty} \mu\cdot e + \lim_{m\to\infty} T^m(e'_1) + \cdots + \lim_{m\to\infty} T^m(e'_r)\\
&= \mu\cdot e + \lim_{m\to\infty} T_1^m(e'_1) + \cdots + \lim_{m\to\infty} T_r^m(e'_r) \underset{3.10.3}{=} \mu\cdot e + 0 + \cdots + 0 = \mu\cdot e.
\end{aligned}$$

1. Se prueba igual que 2. □

6. Nota: Sean $\{\alpha_i\}$ los autovalores de una matriz cuadrada $A \in M_n(\mathbb{R})$. Los autovalores de $\frac{1}{\lambda}\cdot A$ son $\{\frac{\alpha_i}{\lambda}\}$ y v es un autovector de autovalor α_i de A si y sólo si v es un autovector de autovalor $\frac{\alpha_i}{\lambda}$ de $\frac{1}{\lambda}\cdot A$. Por tanto, si $|\alpha_i| < |\lambda|$ para todo i, entonces $\lim_{m\to\infty}(\frac{1}{\lambda}\cdot A)^m = 0$.

Supongamos que $\lambda \in \mathbb{R}$ es un autovalor dominante de multiplicidad 1 de A. Entonces 1 es un autovalor dominante de multiplicidad 1 de $\frac{1}{\lambda}\cdot A$. Tenemos que $c_A(x) = (x-\lambda)\cdot q(x)$, con $q(\lambda)\neq 0$ y escribamos $\operatorname{Ker}(A-\lambda\cdot \operatorname{Id}) = \langle e\rangle$ y podemos suponer que $||e||_1 = 1$. $\mathbb{R}^n = \operatorname{Ker}(A-\lambda\cdot\operatorname{Id})\oplus \operatorname{Ker} q(A)$. Sea $v\in\mathbb{R}^n$, tenemos que $v = \mu\cdot e + e'$, con $e'\in \operatorname{Ker} q(T)$. En el teorema anterior hemos probado que

$$\lim_{m\to\infty} (\tfrac{1}{\lambda}\cdot A)^m(v) = \mu\cdot e.$$

Si $\lambda > 0$ y $v\notin \operatorname{Ker} q(A)$ (luego $\mu\neq 0$), entonces

$$\lim_{m\to\infty}\frac{A^m(v)}{||A^m(v)||_1} = \lim_{m\to\infty}\frac{(\frac{1}{\lambda}\cdot A)^m(v)}{||(\frac{1}{\lambda}\cdot A)^m(v)||_1} = \frac{\mu\cdot e}{||\mu\cdot e||_1} = \frac{\mu\cdot e}{|\mu|} = \pm e.$$

Si $\lambda < 0$ y $v\notin \operatorname{Ker} q(A)$, entonces $\lim_{m\to\infty}\frac{(-A)^m(v)}{||(-A)^m(v)||_1} = \pm e$.

7. Ejemplo: (Tomado de [4]) Consideremos una población de mochuelos moteados, que dividimos en jóvenes de menores de un año, semiadultos entre uno y dos años y adultos mayores de dos años. Se ha observado que después de un año, sobrevive el 18 por ciento de los jóvenes, el 71 por ciento de los semiadultos y el 94 por ciento de los adultos, además el 33 por ciento de los adultos tienen una cría ¿Cuál será la evolución de la población de mochuelos a largo plazo?¿Cuánto habrá que aumentar la supervivencia de los jóvenes para que la población no se extinga y con tal aumento cuál será el porcentaje de población adulta a largo plazo?

Si x_1 es el número de jóvenes, x_2 el de semiadultos y x_3 el de adultos y después de un año y_1 es el número de jóvenes, y_2 el de semiadultos y y_3 el de adultos, tenemos que $y_1 = 0'33\cdot x_3$, $y_2 = 0'18\cdot x_1$, $y_3 = 0'71\cdot x_2 + 0'94\cdot x_3$. Es decir,

$$\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0'33\\ 0'18 & 0 & 0\\ 0 & 0'71 & 0'94\end{pmatrix}\cdot\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}.$$

Denotemos $A = \begin{pmatrix} 0 & 0 & 0'33\\ 0'18 & 0 & 0\\ 0 & 0'71 & 0'94\end{pmatrix}$. La población de mochuelos después de n años será $A^n\cdot(x_1,x_2,x_3)^t$. Tenemos que calcular $\lim_{m\to\infty} A^m(x_1,x_2,x_3)^t$. Las raíces del polinomio característico de A, es decir, los autovalores de A son de módulo menor que $0'99$, entonces por el teorema 3.10.5 $\lim_{m\to\infty} A^m(x_1,x_2,x_3)^t = (0,0,0)^t$ y la población se extinguirá. Para que la población no se extinga aumentemos la supervivencia de los jóvenes. Calculemos a tal que el 1 sea un autovalor (dominante de multiplicidad 1) de

$B = \begin{pmatrix} 0 & 0 & 0'33 \\ 0'18+a & 0 & 0 \\ 0 & 0'71 & 0'94 \end{pmatrix}$. Tenemos que $0 = c_B(1) = 0'042 + 0'2343 \cdot a - 0'06$, luego a ha de ser aproximadamente $0'078$. El vector $(4,1,12)$ es aproximadamente un autovector de autovalor 1 de B, luego la población adulta a largo plazo será un $\frac{1200}{17} = 70'5$ por ciento de toda la población

3.10.2. Matrices positivas. Autovector de Perron

Nuestro objetivo es dar condiciones sencillas que nos aseguren que las hipótesis del teorema 3.10.5 se cumplen.

8. Definiciones: Sean $(a_{ij}),(b_{ij}) \in M_{n\times m}(\mathbb{R})$. Diremos que $(a_{ij}) \geq (b_{ij})$ si $a_{ij} \geq b_{ij}$, para todo i y j. Diremos que $(a_{ij}) \geq 0$ si $a_{ij} \geq 0$, para todo i y j. Diremos que $(a_{ij}) \gg 0$ si $a_{ij} > 0$, para todo i y j.

Sea $A \in M_n(\mathbb{R})$ y $x, x' \in \mathbb{R}^n$. Supongamos que $A \geq 0$. Si $x \geq x'$ entonces $Ax \geq Ax'$.

9. Definiciones: Dado $(x_1,\dots,x_n) \in \mathbb{C}^n$, denotaremos $|(x_1,\dots,x_n)| = (|x_1|,\dots,|x_n|)$. Sea $A \in M_n(\mathbb{R})$ y supongamos que $A \geq 0$. Dado $0 \neq x \in \mathbb{C}^n$, definimos

$$\begin{aligned} \rho_x &:= \sup\{\lambda \in \mathbb{R} \colon |Ax^t| \geq \lambda \cdot |x|^t\} \\ \rho_A &= \sup\{\rho_x,\ \text{para todo } 0 \neq x \in \mathbb{C}^n\} \end{aligned}$$

Para aligerar un poco las notaciones dada una matriz $A \in M_{m\times n}(\mathbb{R})$ denotaremos también con la letra A la aplicación lineal inducida $A\colon \mathbb{R}^n \to \mathbb{R}^m$ por esta matriz, además mucha veces escribiremos Ax en vez de $A(x)$.

10. Lema: *Sea $A \in M_n(\mathbb{R})$ con $A \geq 0$ y sea $x \in \mathbb{C}^n$. Se cumple que*

1. $\rho_{\mu\cdot x} = \rho_x$, *para todo* $\mu \in \mathbb{C}$.
2. $\rho_{A|x|} \geq \rho_{|x|} \geq \rho_x$.
3. $\rho_{A|x|} > \rho_{|x|}$ *si* $A \gg 0$ *y* $A|x| \neq \rho_{|x|} \cdot |x|$.
4. $\rho_x = \inf\{\frac{[A|x|]_i}{|x_i|}$, *para todos los* $x_i \neq 0\}$.

Demostración. 2. Como $A|x| \geq |Ax| \geq \rho_x|x|$, entonces $\rho_{|x|} \geq \rho_x$. Como $A|x| \geq \rho_{|x|} \cdot |x|$, aplicando A,

$$AA|x| \geq \rho_{|x|} \cdot A|x|$$

y $\rho_{A|x|} \geq \rho_{|x|}$.

3. Si $(A - \rho_{|x|}\,\mathrm{Id})|x| \neq 0$ entonces

$$(A - \rho_{|x|} \cdot \mathrm{Id})A|x| = A(A - \rho_{|x|} \cdot \mathrm{Id})|x| \gg 0.$$

Luego, $\rho_{A|x|} > \rho_{|x|}$.

□

11. Teorema de Perron-Frobenius: *Sea $A \in M_n(\mathbb{R})$ y supongamos que $A \gg 0$. Entonces, ρ_A es un autovalor dominante y de multiplicidad 1. Además, si $x \in \mathbb{R}^n$ es el autovector de autovalor ρ_A (único salvo multiplicación por escalares) entonces $x \gg 0$ ó $-x \gg 0$.*

Demostración. Sea $S = \{x \in \mathbb{C}^n$ tales que $|x_1| + \cdots + |x_n| = 1\}$. Por el lema 3.10.10.1, es fácil de probar que $\rho_A = \sup\{\rho_x$ para todo $x \in S\}$. Por el lema 3.10.10.2, es fácil de probar que $\rho_A = \sup\{\rho_{A|x|}$ para todo $x \in S\}$. La aplicación,

$$\begin{array}{ccl} S & \to & \mathbb{R} \\ x & \mapsto & \rho_{A|x|} = \inf\{\frac{[A^2|x|]_i}{[A|x|]_i}, \text{ para } i = 1, \ldots, n\} \end{array}$$

es continua, luego alcanza el máximo, ρ_A, en un cierto $y \in S$. Sea $z = A|y| \gg 0$, entonces $\rho_A = \rho_z$. Si $Az \neq \rho_A \cdot |z|$, por el lema 3.10.10.3 $\rho_{Az} > \rho_z = \rho_A$, lo que es contradictorio. Luego, z es un autovector de autovalor ρ_A.

Sea $v \in \mathbb{C}^n$ un autovector de autovalor $\alpha \in \mathbb{C}$. Entonces, $|Av| = |\alpha v| = |\alpha||v|$. Por tanto, $\rho_v = |\alpha|$ y $|\alpha| \leq \rho_A$. Supongamos $|\alpha| = \rho_A$. Entonces $A|v| \geq |Av| = |\alpha||v| = \rho_A|v|$, luego $\rho_{|v|} = \rho_A$. Además, si $A|v| \neq \rho_A \cdot |v|$, por el lema 3.10.10.3 $\rho_{A|v|} > \rho_{|v|} = \rho_A$, lo que es contradictorio. Luego, $A|v| = \rho_A \cdot |v|$, y como $A|v| \gg 0$, entonces $|v| \gg 0$. Como $|Av| = |\alpha| \cdot |v| = \rho_A \cdot |v| = A|v|$, tenemos que $|\sum_i a_{1j} v_j| = \sum_i a_{1j}|v_j|$ y esto solo puede suceder si todos los números complejos v_j están en la misma semirrecta con origen 0. Luego, $v = \frac{v_1}{|v_1|}|v|$, v es un autovector de autovalor ρ_A y $v >> 0$ ó $-v >> 0$.

Si v_1, v_2 son dos autovectores de autovalor ρ_A, entonces existe $\lambda \in \mathbb{C}$ de modo que el vector $v_1 + \lambda \cdot v_2$ tiene una coordenada cero, pero si no es nulo sabemos que $|v_1 + \lambda \cdot v_2| \gg 0$, contradicción salvo que $v_1 + \lambda \cdot v_2 = 0$. Por tanto, $\dim \operatorname{Ker}(A - \rho_A \operatorname{Id}) = 1$.

Escribamos $c_A(x) = (x - \rho_A)^n \cdot q(x)$, con $q(\rho_A) \neq 0$. Supongamos que $n > 1$. Sabemos que $\operatorname{Ker}(A - \rho_A \cdot \operatorname{Id}) = \langle v \rangle$ y $\dim \operatorname{Ker}(A - \rho_A \cdot \operatorname{Id})^n = n$. Existe $w \in \operatorname{Ker}(A - \rho_A \cdot \operatorname{Id})^n$ que no pertenece a $\operatorname{Ker}(A - \rho_A \cdot \operatorname{Id})$. Sea $m > 1$ mínimo tal que $(A - \rho_A \cdot \operatorname{Id})^m w^t = 0$, entonces $0 \neq (A - \rho_A \cdot \operatorname{Id})^{m-1} w^t \in \operatorname{Ker}(A - \rho_A \cdot \operatorname{Id})$. Es decir, $(A - \rho_A \cdot \operatorname{Id})^{m-1} w^t = \lambda \cdot v^t$, con $\lambda \neq 0$ y $v \gg 0$. Cambiando w por $-w$, si es necesario, podemos suponer que $\lambda > 0$. Entonces, $(A - \rho_A \cdot \operatorname{Id})(A - \rho_A \cdot \operatorname{Id})^{m-2} w^t \gg 0$ y $\rho_{(A-\rho_A \cdot \operatorname{Id})^{m-2} w^t} > \rho_A$ y hemos llegado a contradicción.

□

12. Notación: Dado $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ denotaremos $||x||_1 = |x_1| + \cdots + |x_n|$.

13. Definición: Sea $A \in M_n(\mathbb{R})$ una matriz tal que $A \gg 0$. El autovector $x \gg 0$ de A de autovalor máximo tal que $||x|||_1 = 1$ se denomina autovector de Perron de A.

14. Corolario: *Sea $A \in M_n(\mathbb{R})$ y supongamos que $A^m \gg 0$ para algún número natural m impar. Si λ es un autovalor de A de módulo máximo, entonces es un autovalor dominante de multiplicidad 1, es un número real positivo y*

$$\lim_{n\to\infty} \frac{A^n y^t}{||A^n y^t||_1} = v^t,$$

para todo $y \geq 0$, donde $v \in \mathbb{R}^n$ es el único autovector de autovalor λ, tal que $||v||_1 = 1$ y $v \gg 0$.

Demostración. Sea $c_A(x) = \pm(x-\alpha_1)\cdots(x-\alpha_n)$, entonces $c_{A^m}(x) = \pm(x-\alpha_1^m)\cdots(x-\alpha_n^m)$. Por el teorema de Perron-Fröbenius podemos suponer que α_1^m es el autovalor dominante de multiplicidad 1 de A^m, luego $\lambda = \alpha_1$ es el autovalor dominante de multiplicidad 1 de A. Como el conjugado de α_1 es raíz de $c_A(x)$ y tiene el mismo módulo que α_1, entonces $\alpha_1 \in \mathbb{R}$ y es positivo porque $\alpha_1^m > 0$. Por la nota 3.10.6, solo tenemos que probar que $y \notin \operatorname{Ker} q(A)$. $\operatorname{Ker} q(A) = \operatorname{Ker} q(A^m)$, porque $\operatorname{Ker} q(A) \subseteq \operatorname{Ker} q(A^m)$ y ambos son de dimensión $n-1$. Observemos que $\operatorname{Ker} q(A^m) = \operatorname{Im}(A^m - \lambda^m \cdot \mathrm{Id})$. Si $y \in \operatorname{Ker} q(A^m)$, entonces $y^t = (A^m - \lambda \mathrm{Id})(x^t)$, para cierto $x \in \mathbb{R}^n$. Si aplicamos A^m tenemos que $(A^m - \lambda^m \mathrm{Id})(A^m x^t) = A^m y^t \gg 0$, luego $\rho_{A^m x^t} > \lambda^m$ y hemos llegado a contradicción. □

15. Observaciones: 1. En la demostración del corolario anterior, solo hemos usado que m es impar para afirmar que $\lambda > 0$. Si m es par y $\lambda < 0$ el autovalor dominante de A, entonces $(-A)^m \gg 0$ y el autovalor dominante $-\lambda$ de $-A$ es positivo, y

$$\lim_{n\to\infty} \frac{(-A)^n y^t}{||(-A)^n y^t||_1} = v^t,$$

para todo $y \geq 0$ no nulo, donde v es el único autovector de autovalor λ de A, con $||v||_1 = 1$ (y $v \gg 0$).

2. Si $A \geq 0$ y $A^m \gg 0$, entonces $A^r \gg 0$ para todo $r \geq m$.

3. Si $A \geq 0$ y una fila de A es $\gg 0$, se puede probar que $A^m \gg 0$ para cierto m si y solo si $(A + \mathrm{Id})^{n-1} \gg 0$ (ver problema 37).

3.10.3. Matrices de Leslie

Consideremos una población de hembras de una misma especie entre los 0 años hasta el fin de su edad fértil. Dividamos la población en sucesivos grupos de edad $G_1, G_2, \dots, G_n$, donde cada grupo tiene la misma amplitud L de edad.

Denotemos por f_i el número promedio de hijas de cada hembra del grupo G_i cada L años y por s_i la fracción de individuos del grupo G_i que sobreviven pasados L años.

Sea $p_i(m)$ el número de hembras del grupo G_i después de $m \cdot L$ años y denotemos $\mathbb{P}(m) := (p_1(m), \dots, p_n(m))$. Sea

$$A = \begin{pmatrix} f_1 & f_2 & \dots & f_{n-1} & f_n \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{n-1} & 0 \end{pmatrix}$$

Entonces, $\mathbb{P}(m)^t = A \cdot \mathbb{P}(m-1)^t$. La matriz A se llama **matriz de Leslie** en honor de P.H. Leslie que introdujo este modelo en 1945.

La matriz A es una matriz no negativa, pues $s_i \geq 0$, $i = 1, \dots, n-1$ y $f_i \geq 0$, $i = 1, \dots, n$. Además, si $n \geq 2$ y f_{n-1}, f_n y $s_1, \dots, s_{n-1}$ son no nulos, entonces $A^{n(n-1)} \gg 0$: Sea B la matriz A, pero con $s_1 = \ldots = s_{n-1} = 1 = f_{n-1} = f_n$ y $f_1 = \cdots = f_{n-2} = 0$. Es fácil comprobar que $B^n = B + \mathrm{Id}$ y que $(B^n)^{n-1} = \mathrm{Id} + \binom{n-1}{1} B + \binom{n-1}{2} B^2 + \cdots + B^{n-1} \gg 0$. Por tanto, para todo $N > 0$, $(\frac{1}{N} B)^{n \cdot (n-1)} \gg 0$. Para N muy grande $A \geq \frac{1}{N} B$, luego $A^{n(n-1)} \gg 0$.

3.10.4. Matrices estocásticas

16. Definiciones : Diremos que $(x_1, \dots, x_n) \in \mathbb{R}^n$ es un vector de probabilidad si cumple $x_1, \dots, x_n \geq 0$ y $x_1 + \cdots + x_n = 1$. Diremos que $A \in M_n(\mathbb{R})$ es una matriz estocástica si las columnas de A son vectores de probabilidad.

Si x es un vector de probabilidad y A una matriz estocástica entonces $A \cdot x^t$ es un vector de probabilidad, ya que $Ax^t = (\sum_{j=1}^n a_{1j} x_j, \dots, \sum_{j=1}^n a_{nj} x_j)$ y

$$\sum_{i=1}^n (\sum_{j=1}^n a_{ij} x_j) = \sum_{j=1}^n (\sum_{i=1}^n a_{ij} x_j) = \sum_{j=1}^n x_j = 1.$$

Demos una demostración directa del siguiente teorema sin usar el teorema de Perron-Frobenius.

17. Teorema : *Sea $A \in M_n(\mathbb{R})$ una matriz estocástica, con una fila de coeficientes todos positivos. Entonces, para todo vector de probabilidad p,*

$$\lim_{m \to \infty} A^m p^t = w^t,$$

donde w es el único autovector de autovalor 1 de A que es de probabilidad.

Demostración. Probemos que 1 es un autovalor de multiplicidad 1 y es dominante.

Observemos que $(1,\dots,1)\cdot A = (1,\dots,1)$, luego $A^t(1,\dots,1)^t = (1,\dots,1)^t$ y el 1 es un autovalor de A^t. Por tanto, el 1 es un autovalor de A.

Sea $v = (v_1,\dots,v_n) \in \mathbb{C}^n$ un autovector de A, de autovalor ρ. Entonces, $A|v|^t \geq |Av^t| = |\rho||v|^t$. Multiplicando por $(1,\dots,1)$ por la izquierda, obtenemos que $\sum_i |v_i| \geq |\rho|\cdot\sum_i |v_i|$. Por tanto, $|\rho| \leq 1$. Supongamos $|\rho| = 1$. Si $A|v^t| \neq |v|^t$ entonces multiplicando por $(1,\dots,1)$ por la izquierda, obtenemos que $\sum_i |v_i| > \sum_i |v_i|$, lo que es contradictorio. Luego, $A|v|^t = |v|^t$. Además, $A|v|^t = |v^t| = |Av^t|$, luego $\sum_j a_{ij}|v_j| = |\sum_j a_{ij}v_j|$, lo que implica que $v = \lambda\cdot|v|$ para un cierto $\lambda \in \mathbb{C}$, y podemos suponer que $v \geq 0$ (y obviamente $\rho = 1$). Si $v_1, v_2 \geq 0$ son autovectores de autovalor 1, linealmente independientes, entonces existe $\lambda \in \mathbb{R}$ tal que $v_1 + \lambda\cdot v_2$ es un autovector de autovalor 1 que tiene coeficientes positivos y negativos, lo que es contradictorio.

Escribamos $c_A(x) = (x-1)^n\cdot q(x)$, con $q(1) \neq 0$. Supongamos $n > 1$. Sabemos que $\operatorname{Ker}(A - \mathrm{Id}) = \langle v\rangle$, con $v \geq 0$ y $\dim \operatorname{Ker}(A-\mathrm{Id})^n = n$. Sea $u \in \operatorname{Ker}(A-\mathrm{Id})^n$ un vector que no pertenezca a $\operatorname{Ker}(A-\mathrm{Id})$. Sea $m > 1$ mínimo tal que $(A-\mathrm{Id})^m u^t = 0$, entonces $0 \neq (A-\mathrm{Id})^{m-1}u^t \in \operatorname{Ker}(A-\mathrm{Id})$. Por tanto, $(A-\mathrm{Id})^{m-1}u^t = \lambda\cdot v^t$, con $\lambda \neq 0$. Multiplicando por $(1,\dots,1)$ por la izquierda, se obtiene

$$0 = (1,\dots,1)\cdot(A-\mathrm{Id})^{m-1}u^t = \lambda(1,\dots,1)\cdot v^t \neq 0.$$

Hemos llegado a contradicción. Luego, $n = 1$.

Por el teorema 3.10.5, $\lim\limits_{m\to\infty} A^m p^t \in \langle w^t\rangle$ y como $\lim\limits_{m\to\infty} A^m p^t$ es un vector de probabilidad ha de coincidir con w^t.

□

18. Ejemplo : Un estudio ha determinado que el sector de ocupación de un niño, cuando sea adulto, depende del sector en que trabaje su padre, y está dado por la siguiente matriz de transición, con los sectores de ocupación P = sector primario, S = sector secundario, T = sector terciario.

$$\begin{array}{cc} & \text{Sector del padre} \\ & \begin{array}{ccc} \text{P} & \text{S} & \text{T} \end{array} \\ \text{Sector del hijo} \begin{array}{c} \text{P} \\ \text{S} \\ \text{T} \end{array} & \begin{pmatrix} 0'6 & 0'2 & 0'1 \\ 0'2 & 0'5 & 0'1 \\ 0'2 & 0'3 & 0'8 \end{pmatrix} \end{array} =: A.$$

Así, la probabilidad de que el hijo de alguien que trabaja en el sector secundario trabaje en el primario es $0'2$, que lo haga en el sector secundario es $0'5$ y en el terciario $0'3$.

Si tuviésemos una población de 100 trabajadores de modo que 50 trabajan en el primario, 40 en el secundario y 10 en el terciario y cada trabajador tuviese un hijo,

$50\cdot 0'6+40\cdot 0'2+10\cdot 0'1$ hijos trabajarán en el primario, $50\cdot 0'2+40\cdot 0'5+10\cdot 0'1$ en el secundario y $50\cdot 0'2+40\cdot 0'3+10\cdot 0'8$ en el terciario. Es decir, si los padres trabajan con probabilidad $0'5$ en el primario, $0'4$ en el secundario y 0,1 en el terciario, los hijos trabajan con probabilidad $0'5\cdot 0'6+0'4\cdot 0'2+0'1\cdot 0'1$ en el primario, 0'5·0'2+0'4·0'5+0'1·0'1 en el secundario y $0'5\cdot 0'2+0$'4$\cdot 0'3+0'1\cdot 0'8$ en el terciario. La Teoría de la Probabilidad nos dice (en el caso de que la fecundidad sea en todos los sectores la misma) que si un individuo trabaja con probabilidad x en el sector primario, probabilidad y en el secundario y probabilidad z en el terciario (datos que representamos por el "vector de probabilidad" $(x,y,z)^t$) entonces un hijo suyo trabajará en los diferentes sectores con vector de probabilidad $A\cdot(x,y,z)^t$.

1. ¿Cuál es la probabilidad de que el nieto de un trabajador del sector terciario trabaje en ese sector? Tenemos que calcular

$$A^2\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}=A\cdot\begin{pmatrix}0'1\\0'1\\0'8\end{pmatrix}=\begin{pmatrix}0'16\\0'15\\0'69\end{pmatrix}.$$

   El nieto trabajará en el sector terciario con probabilidad $0'69$.

2. A largo plazo ¿qué proporción de la población trabajará en el sector secundario?

   Tenemos que calcular $\lim\limits_{n\to\infty} A^n(x,y,z)^t$. Por el teorema 3.10.17, sabemos que

$$\lim_{n\to\infty} A^n(x,y,z)^t=w^t,$$

   donde w es el vector de probabilidad que es un autovector de A de autovalor 1. $\operatorname{Ker}(A-\operatorname{Id})=\langle(7,6,16)\rangle$, luego $w=(7/29,6/29,16/29)$. El $\frac{600}{29}=20'6$ por ciento de la población trabajará a largo plazo en el sector secundario.

Compliquemos el ejemplo. Supongamos que cada trabajador del sector secundario tiene un 20% más de hijos que cada trabajador del primario y los del terciario un 30% más de hijos que los del primario, a largo plazo ¿qué proporción de la población trabajará en el sector secundario?

Sea D la matriz diagonal de coeficientes 1, $1'2$ y $1'3$ en la diagonal y $B=A\cdot D$. Tenemos que calcular $\lim\limits_{n\to\infty}\frac{B^n\cdot(x,y,z)^t}{||B^n\cdot(x,y,z)^t||_1}$ y para ello aplicaremos el corolario 3.10.14.

3.10.5. Sistemas de ecuaciones diferenciales lineales

19. Consideremos el sistema de ecuaciones diferenciales lineales con coeficientes constantes

$$\begin{array}{rcl} x_1' & = & a_{11}x_1 + \cdots + a_{1n}x_n \\ \cdots & = & \cdots \\ x_n' & = & a_{n1}x_1 + \cdots + a_{nn}x_n \end{array}$$

donde $A = (a_{ij})$ es una matriz cuadrada de orden n, con coeficientes complejos. De modo conciso escribiremos el sistema anterior $X' = AX$. Por cambio lineal de coordenadas. $Y = BX$, tenemos $Y' = BX' = BAX = BAB^{-1}Y$ y para B conveniente podemos conseguir que $J = BAB^{-1}$ sea una matriz triangular. Ahora ya es fácil calcular Y y por tanto podemos calcular X.

De otro modo: Se define $e^{At} := \sum_{n=0}^{\infty} A^n \cdot \frac{t^n}{n!}$. Las soluciones del sistema $X' = AX$ son $\{X = e^{At} \cdot C\}$, siendo C una matriz columna de constantes cualesquiera. Entonces,

$$X = e^{At} \cdot C = e^{BJB^{-1}t} \cdot C = Be^{Jt}B^{-1} \cdot C = Be^{Jt} \cdot C'.$$

siendo C' una matriz columna de constantes cualesquiera.

Si A es diagonalizable, tomemos $B = (v_1, \dots, v_n)^t$ donde $v_1, \dots, v_n$ es una base (de autovectores) donde A diagonaliza, entonces J es diagonal y el coeficiente i-ésimo de la diagonal de J es el autovalor λ_i de v_i. Entonces,

$$X = B \cdot e^{Jt} \cdot C' = c_1' e^{\lambda_1 t} v_1 + \cdots + c_n' e^{\lambda_n t} v_n.$$

En general, sea D la matriz diagonal cuyos coeficientes en la diagonal son los de J . Entonces, $J = D + N$ y $N^n = 0$ (puede tomarse J de modo que D y N conmuten). Entonces,

$$e^{Jt} = e^{Dt}e^{Nt} = e^{Dt} \cdot (\mathrm{Id} + Nt + \frac{N^2t^2}{2!} + \cdots + \frac{N^{n-1}t^{n-1}}{(n-1)!}).$$

20. Sea $X' = AX + B(t)$ un sistema lineal de ecuaciones diferenciales. Tenemos que $(D - A)X = B(t)$, luego una solución particular es

$$X = \frac{1}{D-A}B(t) = \frac{1}{D-A}(e^{At}e^{-At}B(t)) = e^{At}\frac{1}{D}(e^{-At}B(t))) = e^{At}\int e^{-At}B(t)dt.$$

21. Sea $p(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{R}[x]$ un polinomio de grado n. La ecuación diferencial

$$p(D)y = f(x)$$

es equivalente al sistema de ecuaciones diferenciales lineales con coeficientes constantes de primer orden de n variables:

$$\begin{array}{lcl} x_1' & = & x_2 \\ x_2' & = & x_3 \\ \cdots & = & \cdots \\ x_{n-1}' & = & x_n \\ x_n' & = & \frac{-1}{a_n}\cdot(\sum_{i=0}^{n-1} a_i x_{i+1}) + f(t) \end{array}$$

22. Ejercicio : Resuelve el siguiente sistema de ecuaciones diferenciales

$$\begin{array}{lcl} x' & = & -11x - 4y \\ y' & = & 15x + 6y \end{array}$$

23. Ejemplo : Resolvamos el siguiente sistema de ecuaciones diferenciales

$$\begin{array}{lcl} x' & = & x - 3y + 3z \\ y' & = & -2x - 6y + 13z \\ z' & = & -x - 4y + 8z \end{array}$$

Sea $A = \begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$. Calculemos una base donde A triangula. Tenemos que $c_A(x) = 1 - 3x + 3x^2 - x^3 = -(x-1)^3$. Sea $v = (1,0,0)$ y consideremos la base

$$\{v^t, (A - \mathrm{Id})v^t, (A - \mathrm{Id})^2 v^t\} = \{(1,0,0)^t, (0,-2,-1)^t, (3,1,1)^t\}.$$

Entonces, $A = BJB^{-1}$, con

$$B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} \quad \text{y} \quad J = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

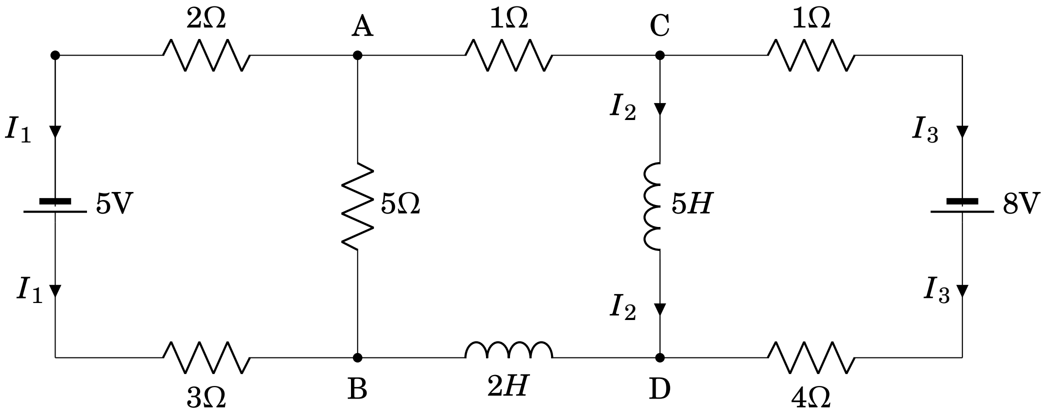
Las soluciones son $X = e^{At}\cdot C = Be^{Jt}B^{-1}C = Be^{Jt}C'$. Tenemos que $J = \mathrm{Id} + N$, con $N = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ y

$$e^{Jt} = e^{(\mathrm{Id}+N)t} = e^{Idt}\cdot e^{Nt} = \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^t \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ t^2/2 & t & 1 \end{pmatrix}.$$

Luego

$$X = e^t \cdot \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ t^2/2 & t & 1 \end{pmatrix} \cdot C' = e^t \cdot \begin{pmatrix} c_1'(\frac{3t^2}{2}+1)+3c_2't+3c_3' \\ c_1'(\frac{t^2}{2}-2t)+c_2'(t-2)+c_3' \\ c_1'(\frac{t^2}{2}-t)+c_2'(t-1)+c_3' \end{pmatrix}.$$

24. Circuito eléctrico: Calculemos las intensidades de corriente en los diferentes tramos del siguiente circuito



De D a B la corriente es de intensidad I_2+I_3. De B a A la corriente es de intensidad $I_1+I_2+I_3$. De A a C la corriente es de intensidad I_2+I_3. La suma de diferencia de potenciales en todo circuito cerrado es nula. Luego,

$$\begin{aligned} 3I_1+5(I_1+I_2+I_3)+2I_1 &= 5 \\ 2(I_2'+I_3')+5(I_1+I_2+I_3)+1(I_2+I_3)+5I_2' &= 0 \\ 4I_3-5I_2'+I_3 &= 8 \end{aligned}$$

$I_1=\frac{1-I_2-I_3}{2}$. Tenemos que resolver el sistema de ecuaciones

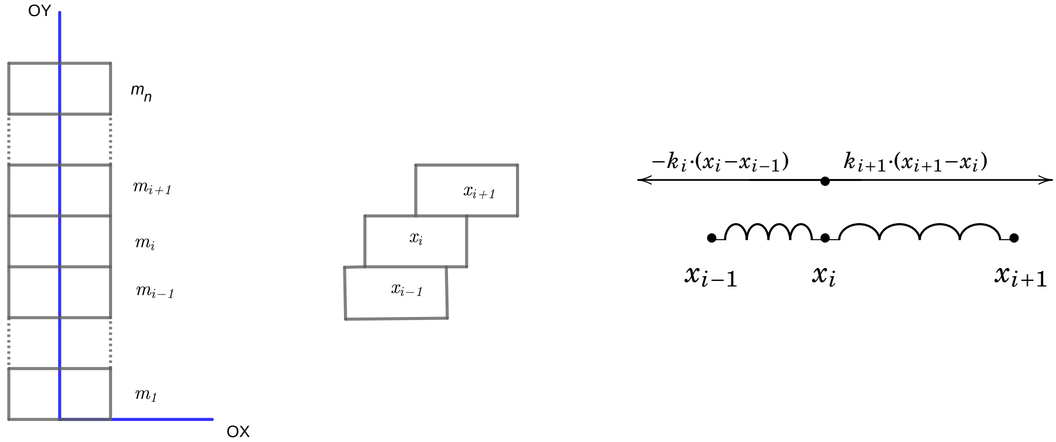
$$\begin{aligned} 7I_2'+2I_3' &= -\tfrac{7}{2}I_2-\tfrac{7}{2}I_3-\tfrac{5}{2} \\ -5I_2' &= -5I_3+8 \end{aligned}$$

que es equivalente al sistema

$$\begin{aligned} I_2' &= I_3-\tfrac{8}{5} \\ I_3' &= \tfrac{-7}{4}I_2-\tfrac{21}{4}I_3+\tfrac{87}{20} \end{aligned}$$

Una solución particular es $\begin{pmatrix} I_2 \\ I_3 \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ -7/4 & -21/4 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -8/5 \\ 87/20 \end{pmatrix} = \begin{pmatrix} -81/35 \\ 8/5 \end{pmatrix}$. Dejamos que el lector calcule la solución general e imponga $I_2(0)=0$, $I_3(0)=0$.

25. Modelo de desplazamiento de las plantas de un edificio por un terremoto[4]: Consideremos un edificio de n plantas. Suponemos que la i-ésima planta tiene masa m_i y que las plantas sucesivas están conectadas por un conector elástico cuyo efecto se asemeja al de un muelle (los elementos estructurales de los grandes edificios son de acero, que es un material muy elástico). Supongamos que sucede un terremoto, que mueve el suelo según la función $f(t)$. Sigamos el siguiente dibujo esquemático:



Denotemos por x_i la coordenada sobre el eje de las abscisas del centro de la planta i-ésima. Por tanto, $x_0(t) = f(t)$. El desplazamiento horizontal de la planta i-ésima respecto de la $i+1$-ésima es $x_{i+1}-x_i$, por tanto la planta $i+1$-ésima "tira" horizontalmente de la planta i-ésima con una fuerza $k_{i+1}\cdot(x_{i+1}-x_i)$. Tenemos el siguiente sistema de ecuaciones diferenciales

$$\begin{array}{rcl} m_1x_1'' & = & -k_1x_1 + k_2(x_2-x_1) + k_1 f(t) \\ m_2x_2'' & = & -k_2(x_2-x_1) + k_3(x_3-x_2) \\ \cdots & = & \cdots \\ m_nx_n'' & = & -k_n\cdot(x_n-x_{n-1}) \end{array}$$

Podemos escribirlo abreviadamente como $MX'' = KX + F(t)$, donde M y K son las matrices

$$\begin{pmatrix} m_1 & 0 & \ldots & 0 \\ 0 & m_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & m_n \end{pmatrix}, \quad \begin{pmatrix} -k_1-k_2 & k_2 & 0 & 0 & \ldots & 0 & 0 \\ k_2 & -k_2-k_3 & k_1 & 0 & \ldots & 0 & 0 \\ 0 & k_3 & -k_3-k_4 & k_4 & \ldots & 0 & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & k_n & -k_n \end{pmatrix}$$

[4] Tomado de [10].

y $F(t)^t = (k_1 f(t), 0, \ldots, 0)$. Escribamos $A = M^{-1}K$ (que es una matriz simétrica tridiagonal) y $G(t) = (\frac{k_1}{m_1} f(t), 0, \ldots, 0)^t$, entonces

$$X'' = AX + G.$$

$A = P\Delta P^{-1}$, donde P es la matriz cuyas columnas son los autovectores de A y Δ es la matriz diagonal cuyos coeficientes de la diagonal son los autovalores de A. Entonces, $X'' = P\Delta P^{-1}X + G$, luego $(P^{-1}X)'' = \Delta(P^{-1}X) + P^{-1}G$. Si escribimos $Y = P^{-1}X$, entonces $Y'' = \Delta Y + P^{-1}G$ que es más fácil de resolver y $X = PY$.

Resolvamos un caso sencillo $n = 2$, $m_i = 5000\ kg$, $k_i = 10000\ kg/s^2$ para todo i y $f(t) = c \cdot \operatorname{sen}(wt)$.

Tenemos que resolver el sistema de ecuaciones diferenciales

$$\begin{array}{rcl} x_1'' & = & -4x_1 + 2x_2 + 2c\operatorname{sen}(wt) \\ x_2'' & = & 2x_1 - 2x_2 \end{array}$$

Sea $A = \begin{pmatrix} -4 & 2 \\ 2 & -2 \end{pmatrix}$, entonces $c_A(x) = (x + 3 + \sqrt{5})(x + 3 - \sqrt{5})$ y los autovectores son $(\frac{-1-\sqrt{5}}{2}, 1)$, $(\frac{-1+\sqrt{5}}{2}, 1)$. Sea $P = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & \frac{-1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$ e $Y = P^{-1}X$. Entonces,

$$\begin{pmatrix} y_1'' \\ y_2'' \end{pmatrix} = \begin{pmatrix} -3-\sqrt{5} & 0 \\ 0 & -3+\sqrt{5} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \frac{c\operatorname{sen}(wt)}{\sqrt{5}} \begin{pmatrix} -2 \\ -1+\sqrt{5} \end{pmatrix}$$

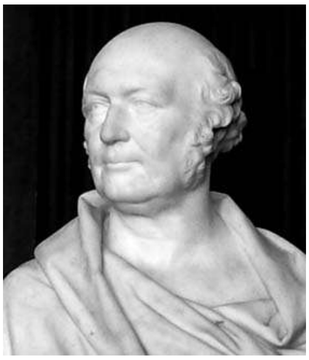
Una solución particular es $y_1 = \frac{2c}{\sqrt{5}(w^2-3-\sqrt{5})}\operatorname{sen}(wt)$ e $y_2 = \frac{c(1-\sqrt{5})}{\sqrt{5}(w^2-3+\sqrt{5})}\operatorname{sen}(wt)$. Luego,

$$\begin{array}{rcl} y_1 & = & \lambda_1 \cdot \cos(\sqrt{3+\sqrt{5}} \cdot t) + \mu_1 \cdot \operatorname{sen}(\sqrt{3+\sqrt{5}} \cdot t) + \frac{2c}{\sqrt{5}(w^2-3-\sqrt{5})}\operatorname{sen}(wt) \\ y_2 & = & \lambda_2 \cdot \cos(\sqrt{3-\sqrt{5}} \cdot t) + \mu_2 \cdot \operatorname{sen}(\sqrt{3-\sqrt{5}} \cdot t) + \frac{c(1-\sqrt{5})}{\sqrt{5}(w^2-3+\sqrt{5})}\operatorname{sen}(wt) \end{array}$$

Podemos suponer que $X = 0$ para $t = 0$, entonces $Y = 0$, para $t = 0$ y $\lambda_1 = \lambda_2 = 0$. Podemos suponer también que $X' = 0$ para $t = 0$, entonces $Y' = 0$, para $t = 0$, y obtenemos que $\mu_1 = \frac{-2cw}{\sqrt{5}(w^2-3-\sqrt{5})\sqrt{3+\sqrt{5}}}$ y $\mu_2 = \frac{-cw(1-\sqrt{5})}{\sqrt{5}(w^2-3+\sqrt{5})\sqrt{3-\sqrt{5}}}$. Por último, $X = PY$.

Observemos que $w^2 = 3+\sqrt{5}, 3-\sqrt{5}$ son valores críticos. Cuando el opuesto de algún valor propio de A esté muy próximo a w^2 el edificio corre grave riesgo de derrumbamiento en un terremoto.

3.11. Biografía de Sir William Hamilton



HAMILTON BIOGRAPHY

William Hamilton had not had a university education and it is thought that Hamilton's genius came from his mother, Sarah Hutton. By the age of five, William had already learned Latin, Greek, and Hebrew. Hamilton's introduction to mathematics came at the age of 13 when he studied Clairaut's Algebra, a task made somewhat easier as Hamilton was fluent in French by this time. At age 15 he started studying the works of Newton and Laplace. In 1822 Hamilton found an error in Laplace's Mécanique céleste and, as a result of this, he came to the attention of John Brinkley, the Royal Astronomer of Ireland, who said: "This young man, I do not say will be, but is, the first mathematician of his age." Hamilton entered Trinity College, Dublin at the age of 18 and in his first year he obtained an 'optime' in Classics, a distinction only awarded once in 20 years.

In August 1824, Uncle James took Hamilton to Summerhill to meet the Disney family. It was at this point that William first met their daughter Catherine and immediately fell hopelessly in love with her. Unfortunately, as he had three years left at Trinity College, Hamilton was not in a position to propose marriage. The following February, Catherine's mother informed William that her daughter was to marry a clergyman, who was fifteen years her senior. He was affluent and could offer more to Catherine than Hamilton. He became ill and at one point he even considered suicide. In this period he turned to poetry, which was a habit that he pursued for the rest of his life in times of anguish.

In 1826 Hamilton received an 'optime' in both science and Classics, which was unheard of, while in his final year as an undergraduate he presented a memoir Theory of Systems of Rays to the Royal Irish Academy. It is in this paper that Hamilton introduced the characteristic function for optics.

Hamilton's finals examiner, Boyton, persuaded him to apply for the post of Royal Astronomer at Dunsink observatory. It turned out that Hamilton had made a poor choice as he lost interest in astronomy and spend all time on mathematics. Before beginning his duties in this prestigious position, Hamilton toured England and Scotland (from where the Hamilton family originated). He met the poet Wordsworth and they became friends. The two men had long debates over science versus poetry. Hamilton liked to compare the two, suggesting that mathematical language was as artistic as poetry. William also took the opportunity to visit Catherine, as she was living relatively nearby, which she then reciprocated by coming to the observatory. Hamilton was so nervous in her presence that he broke the eyepiece of the telescope whilst trying to give her a demonstration. This episode inspired another interval of misery and poem writing. Hamilton seemed quite fickle when it came to relationships with women. Perhaps this was because he thought that he ought to marry and so, if he could not have Catherine, then it did not really matter whom he married. In the end he married Helen Maria Bayly who lived just across the fields from the observatory. Unfortunately, the marriage was fated from the start.

In 1832 Hamilton published this third supplement to Theory of Systems of Rays which is essentially a treatise on the characteristic function applied to optics. Near the end of the work he applied the characteristic function to study Fresnel's wave surface. From this he predicted conical refraction and asked the Professor of Physics at Trinity College, Humphrey Lloyd, to try to verify his theoretical prediction experimentally. This Lloyd did two months later and this theoretical prediction brought great fame to Hamilton.

On 4 November 1833 Hamilton read a paper to the Royal Irish Academy expressing complex numbers as algebraic couples, or ordered pairs of real numbers. He used algebra in treating dynamics in On a General Method in Dynamics in 1834. In this paper Hamilton gave his first statement of the characteristic function applied to dynamics and wrote a second paper on the topic the following year.

Hamilton was knighted in 1835. After the discovery of algebraic couples, he tried to extend the theory to triplets, and this became an obsession that plagued him for many years. Hamilton was so preoccupied with the triplets that even his children were aware of it. Every morning they would inquire: "Well, Papa can you multiply triplets?" On 16 October 1843 (a Monday) Hamilton was walking in along the Royal Canal with his wife to preside at a Council meeting of the Royal Irish Academy. Although his wife talked to him now and again Hamilton hardly heard, for the discovery of the quaternions. He could not resist the impulse to carve the formulae for the quaternions $i^2 = j^2 = k^2 = ijk = -1$ in the stone of Broome Bridge as he and his wife passed it. In 1958 the Royal Irish Academy erected a plaque commemorating this.

Hamilton felt this discovery would revolutionise mathematical physics and he spent the rest of his life working on quaternions. Shortly after Hamilton's discovery of the quaternions his personal life started to prey on his mind again. In 1845, Thomas Disney visited Hamilton at the observatory and brought Catherine with him. This must have upset William as his alcohol dependency took a turn for the worse. In 1848, Catherine began writing to Hamilton, which cannot have helped at this time of depression. The correspondence continued for six weeks and became more informal and personal until Catherine felt so guilty that she confessed to her husband. However, Catherine wrote once more and this time attempted suicide (unsuccessfully) as her remorse was so great. She then spent the rest of her life living with her mother or siblings, although there was no official separation from her husband. Hamilton persisted in his correspondence to Catherine, which he sent through her relatives. He published Lectures on Quaternions in 1853. Hamilton went straight to Catherine and gave her a copy of Lectures on Quaternions. She died two weeks later.

Determined to produce a work of lasting quality, Hamilton began to write another book Elements of Quaternions which he estimated would be 400 pages long and take 2 years to write. The title suggests that Hamilton modelled his work on Euclid's Elements and indeed this was the case. The book ended up double its intended length and took seven years to write.

Hamilton died from a severe attack of gout shortly after receiving the news that he had been elected the first foreign member of the National Academy of Sciences of the USA.

Article by: J.J. O'Connor and E.F. Robertson: http://mathshistory.st-andrews.ac.uk/Biographies/Hamilton.html

3.12. Cálculos con ordenador

1. Escribe Roots[x^3 $-5x+4==0,x$], para calcular las raíces de x^3-5x+4. Escribe PolynomialGCD[x^3 $+3x$^2 $+3x+1, x$^3 $-x$^2 $-x+1$] para calcular el máximo común divisor de los polinomios x^3+3x^2+3x+1 y x^3-x^2-x+1. Escribe también, PolynomialExtendedGCD[x^3 $+3x$^2 $+3x+1, x$^3 $-x$^2 $-x+1$] para calcular además los polinomios $\lambda(x)$ y $\mu(x)$ de la identidad de Bézout.

2. Escribe $DSolve[f''[x]-f[x]==Sin[x],f[x],x]$ para resolver la ecuación diferencial $f''-f=\operatorname{sen} x$.

3. Escribe Eigenvalues[A], Eigenvectors[A] y CharacteristicPolynomial[A,x], donde A es la matriz asociada a T del ejercicio 3.7.3, para calcular los autovalores, autovectores y polinomio característico de A.

4. Usa el comando MatrixExp[] para calcular e^{At} con $A=\begin{pmatrix}-11 & -4\\ 15 & 6\end{pmatrix}$. Escribe

$$DSolve[\{x'[t]==-11*x[t]-4*y[t],y'[t]==15*x[t]+6*y[t]\},\{x[t],y[t]\},t]$$

para resolver el ejercicio 3.10.22.

3.13. Problemas

1. La ley del enfriamiento de Newton establece que la tasa de pérdida de calor de un cuerpo es proporcional a la diferencia de temperatura entre el cuerpo y sus alrededores. Un sólido a 20º centígrados es introducido en un lago de agua a temperatura de 5º. Si tarda dos minutos en enfriarse diez grados ¿Cuántos minutos tardará en enfriarse diez grados?

2. Un depósito contiene 10000 litros de agua salada con una concentración de sal del 5 % (50-gramos de sal por litro). Un grifo va vertiendo 50-litros de agua pura por minuto. Por otro lado se desagua 50 litros del depósito por minutos ¿Cuánto tiempo transcurrirá hasta que la concentración salina del agua salada del depósito baje al 1 %?

3. Sea F el espacio vectorial formado por las funciones de $\mathbb{R}$ a $\mathbb{C}$ infinitamente derivables y $D: F \to F$ el operador derivada. Prueba que

   a) $D(e^f \cdot g) = e^f \cdot (D + f' \cdot \mathrm{Id})(g)$.

   b) Calcula las soluciones de la ecuación diferencial lineal $y' + fy = g$.

4. Sea F el espacio vectorial formado por las funciones de $\mathbb{R}^+$ a $\mathbb{C}$ infinitamente derivables y $\Theta: F \to F$ el operador $\mathbb{C}$-lineal definido por $\Theta(f) := xf'$.

   a) Prueba que $\mathrm{Ker}(\Theta - \alpha)^r = x^\alpha \cdot \{\sum_{i=0}^{r-1} \lambda_i (\ln x)^i : \; \forall \lambda_i \in \mathbb{C}\}$.

   b) Resuelve la ecuación de Euler-Cauchy

   $$x^2y'' + bxy' + cy = 0,$$

   para $b, c \in \mathbb{C}$ y $x > 0$.

5. Conducimos un coche con un consumo constante de gasolina a lo largo del tiempo. Supongamos que no hay más fuerza de rozamiento que la producida por el aire por causa de la velocidad del coche. Supongamos también que la fuerza de rozamiento es proporcional a la velocidad. Calcula la velocidad del coche en cada instante

6. Resuelve la ecuación diferencial del movimiento armónico amortiguado

$$f'' + af' + bf = 0, \;\; (\text{con } a^2 - 4b < 0 \text{ y } a > 0).$$

7. Resuelve la ecuación diferencial del movimiento armónico forzado

$$f'' + af' + bf = c\cos(wx), \;\; (\text{con } a^2 - 4b < 0 \text{ y } a > 0).$$

8. Resuelve la ecuación diferencial $f''' - 2f'' + f = xe^x$.

9. Sea $p(x) \in \mathbb{R}[x]$ un polinomio mónico de grado n. Sean $s_1(x), \ldots, s_n(x)$ soluciones, linealmente independientes, de la ecuación diferencial $p(D)y = 0$. Pruébese que si las funciones $c_1(x), \ldots, c_n(x)$ cumplen las ecuaciones

$$\begin{array}{rcl} c_1(x)'s_1(x) + \ldots + c_n(x)'s_n(x) & = & 0 \\ \ldots & = & 0 \\ c_1(x)'s_1(x)^{n-2)} + \ldots + c_n(x)'s_n(x)^{n-2)} & = & 0 \\ c_1(x)'s_1(x)^{n-1)} + \ldots + c_n(x)'s_n(x)^{n-1)} & = & f(x) \end{array}$$

entonces $c_1(x)s_1(x) + \ldots + c_n(x)s_n(x)$ es una solución particular de $p(D)y = f(x)$.

10. Sea $p(x) \in \mathbb{R}[x]$ un polinomio mónico de grado r. Sean $s_1(n), \ldots, s_r(n)$ soluciones, linealmente independientes, de la ecuación en diferencias $p(\nabla)y = 0$. Pruébese que si las sucesiones $c_1(n), \ldots, c_r(n)$ cumplen las ecuaciones

$$\begin{array}{rcl} \Delta(c_1)\nabla(s_1) + \ldots + \Delta(c_r)\nabla(s_r) & = & 0 \\ \ldots & = & 0 \\ \Delta(c_1)\nabla^{r-1}(s_1) + \ldots + \Delta(c_r)\nabla^{r-1}(s_r) & = & 0 \\ \Delta(c_1)\nabla^{r}(s_1) + \ldots + \Delta(c_r)\nabla^{r}(s_r) & = & f \end{array}$$

entonces $c_1s_1 + \ldots + c_rs_r$ es una solución particular de $p(\nabla)y = f$.

11. Calcula los coeficientes a_i de la serie $s(x) = \sum_{i=0}^{\infty} a_i x^i$ tal que $s'' - s = \operatorname{sen} x$.

12. Préstamos de gradiente exponencial: Un préstamo de $K = 10^5$ euros se quiere devolver durante $N = 20$ años, pagando cada año n una anualidad d_n de modo que $d_n = I'd_{n-1}$ ($I' = 1 + 2\%$). Se supone que nos prestan el dinero a un tipo de interés anual $I = 5\%$. Determina d_1.

13. Sea $A \in M_n(\mathbb{R})$ una matriz cuadrada. Prueba que $c_A(x) = c_{A^t}(x)$.

14. Sea $A \in M_n(\mathbb{R})$ una matriz cuadrada y $p(x) \in \mathbb{R}[x]$ un polinomio. Prueba que si $c_A(x) = \pm(x-\alpha_1)\cdots(x-\alpha_n)$ entonces $c_{p(A)}(x) = \pm(x - p(\alpha_1))\cdots(x - p(\alpha_n))$.

15. Dada la matriz $A = \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}$, determina los valores del parámetro a para que se cumpla cada una de las siguientes condiciones:

a) A sea invertible.

b) Un autovalor de A sea 2.

c) A tenga dos autovalores positivos.

16. Escribe una matriz que tenga como autovectores $(1,0)$ y $(1,1)$ y como autovalores asociados 1 y 2 respectivamente.

17. Escribe una matriz A que tenga los siguientes autovectores y autovalores:

    $v_1 = (1,1,1)$ autovector asociado al autovalor $\lambda_1 = 1$, $v_2 = (0,1,0)$ autovector asociado al autovalor $\lambda_2 = 3$, $v_3 = (0,1,1)$ autovector asociado al autovalor $\lambda_3 = -1$.

18. Calcula los autovalores y los autovectores de las siguientes matrices

$$A = \begin{pmatrix} 10 & -3 & 14 \\ 6 & -1 & 12 \\ -3 & 1 & -3 \end{pmatrix} \text{ y } B = \begin{pmatrix} -2 & 0 & -4 \\ 6 & 2 & 6 \\ 6 & -1 & 9 \end{pmatrix}$$

19. Sea $A = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ 2 & 2 & -3 \end{pmatrix}$. Calcula A^{50}.

20. Explica por qué no es diagonalizable la matriz $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

21. Indica cuáles de las siguientes matrices tienen como autovalores $\lambda_1 = 1$ y $\lambda_2 = -1$ y como autovectores asociados $u_1 = (1,1)$ y $u_2 = (-1,0)$ respectivamente:

$$A = \begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} \quad C = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}.$$

22. Calcula el valor del parámetro α para que el vector $v = (12,2,1)$ sea autovector de la matriz

$$A = \begin{pmatrix} 0 & \alpha & 4 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{3}{4} & 0 \end{pmatrix}.$$

23. Sea A una matriz cuadrada tal que la suma de los elementos de cada una de sus filas es igual a 1. Prueba que 1 es autovalor de A.

24. Sea A una matriz cuadrada y sea λ un autovalor de A. Prueba:

    a) λ^m es autovalor de A^m, para cualquier $m \in \mathbb{N}$.

    b) Si A es invertible, entonces $\lambda \neq 0$ y $1/\lambda$ es autovalor de A^{-1}

25. Calcula los autovalores y una base de autovectores de la matriz $\begin{pmatrix} 0,5 & 0,4 & 0,6 \\ 0,3 & 0,4 & 0,3 \\ 0,2 & 0,2 & 0,1 \end{pmatrix}$.

26. Calcula los autovalores de la matriz $A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ \alpha & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & \beta & 3 & 0 \\ 0 & 0 & 0 & \gamma & 3 \end{pmatrix}$ y determina para qué valores de los parámetros α, β y γ es diagonalizable.

27. Dada la matriz $A = \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}$, determina los valores del parámetro a para que se cumpla cada una de las siguientes condiciones:

    a) El vector $(1,1)$ sea autovector de A.

    b) El vector $(1,-1)$ sea autovector de A.

    c) El vector $(2,1)$ sea autovector de A.

28. Calcula el valor de los parámetros α y β que aparecen en la matriz A para que el vector $v = (10,5,40)$ sea autovector de A:

$$A = \begin{pmatrix} 0 & 0 & \alpha \\ \beta & 0 & 0 \\ 0 & 0{,}8 & 0{,}9 \end{pmatrix}.$$

29. Calcula los autovalores y autovectores de la matriz $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ y determina si es diagonalizable.

30. Dada la matriz $A = \begin{pmatrix} 0,5 & 0,4 & a \\ 0,3 & 0,4 & 0,3 \\ 0,2 & 0,2 & 0,1 \end{pmatrix}$,

a) Calcula el valor del parámetro a sabiendo que $\lambda = 1$ es un autovalor de A.

b) Calcula el valor del parámetro b sabiendo que el vector $v = (16, 11, b)$ es un autovector asociado al autovalor $\lambda = 1$.

31. Indica razonadamente si cada una de las siguientes afirmaciones es verdadera o falsa:

    a) Si λ es un autovalor de una matriz A y u es un autovector asociado a λ, entonces $2u$ es también un autovector asociado a λ.

    b) Si λ es un autovalor de una matriz A y u es un autovector asociado a λ, entonces 2λ es un autovalor de la matriz $2A$ y tiene a $2u$ como autovector asociado.

    c) Si λ es un autovalor de una matriz A y u es un autovector asociado a λ, entonces 2λ es un autovalor de la matriz $2A$ y tiene a u como autovector asociado.

32. Sea A una matriz cuadrada. Indica razonadamente si cada una de las siguientes afirmaciones es verdadera:

    a) Si $\lambda = 0$ es autovalor de A entonces A no es invertible.

    b) Si $\lambda = 0$ es autovalor de A entonces A no es diagonalizable.

    c) Si $\lambda = 0$ es autovalor de A, el vector cero es el único autovector asociado a $\lambda = 0$.

    d) $\lambda = 0$ es autovalor de A, entonces A no tiene autovalores positivos ni negativos.

33. Determina si cada una de las siguientes afirmaciones es cierta. Justifica la respuesta.

    a) Si la segunda columna de una matriz cuadrada A de orden 3×3 es cero, entonces $(0, 1, 0)$ es un autovector de A.

    b) Si una matriz cuadrada A tiene dos filas iguales, entonces 0 es un autovalor de A.

    c) La suma de dos autovectores de una matriz cuadrada A es un autovector de A.

34. Sean $T, S : E \to E$ dos operadores lineales diagonalizables que conmutan, es decir, $T \circ S = S \circ T$. Prueba que existe una base de E en los que T y S diagonalizan simultáneamente.

35. Sea $(a_{ij}) \in M_n(\mathbb{C})$ tal que $\sum_i |a_{ij}| < 1$ para todo j. Prueba que el módulo de todos los autovalores de (a_{ij}) son menores estrictos que 1. Prueba que $\lim\limits_{n\to\infty} A^n = 0$.

36. **Esperanza de vida.** Un automóvil tiene dos sistemas de frenado independientes: el freno de pedal y el freno de mano. El automóvil pasa un control cada año y si al menos uno de los sistemas de frenado está operativo será reparado, pero si ambos sistemas fallan simultáneamente se supone que se destruirá.

    En el año n en un punto de reparación, si uno de los frenos falla pero el otro funciona, entonces el freno defectuoso es reemplazado. Si un freno funciona en el año n entonces se estima fiable en un 90 % en el año $n+1$. Sin embargo, si un freno falla en el año n, entonces su recambio no probado se estima fiable en un 60 % en el año $n+1$.

    ¿Cuánto tiempo se espera que el sistema funcione antes de la destrucción?

37. Sea $A \in M_n(\mathbb{R})$ tal que $A \geq 0$, una fila de A es estrictamente mayor que cero y $(A + \mathrm{Id})^m \gg 0$ para algún número natural m. Prueba que existe un número natural N tal que $A^N \gg 0$. Prueba que $(A+\mathrm{Id})^{n-1} \gg 0$.

38. En un experimento controlado, se cruzan sementales de genotipo AB respecto a un gen que presenta dos alelos A y B, con hembras escogidas al azar. Las hembras de la descendencia se aparean después con sementales AB y así sucesivamente.

    Si al comenzar el experimento partimos de una población en la que el 60 % es de tipo AA, el 20 % de tipo AB y el otro 20 % de tipo BB ¿Cuál será el porcentaje de individuos de cada tipo en la tercera generación?¿Y después de muchas generaciones?

39. En una gran población de plantas hay tres posibles genotipos respecto de un gen: AA, Aa y aa. Llamamos x_n, y_n y z_n a la proporción de plantas con genotipo AA, Aa y aa, respectivamente, en la generación n-ésima y denotamos por g_n al vector $g_n = (x_n, y_n, z_n)$. Al realizarse un programa de fecundación controlada, podemos calcular el vector g_n en función de g_{n-1}, utilizando una matriz de transición A: $g_n^t = A g_{n-1}^t = \cdots = A^n g_0^t$.

    a) Si cada planta se fecunda con una de su mismo genotipo, di cuál de las siguientes es la matriz de transición A. Razona la respuesta:

$$A = \begin{pmatrix} 1 & 1/4 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1/4 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/4 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}$$

b) Calcula los autovalores y autovectores de la matriz A y escribe las matrices P y D tales que $A = PDP^{-1}$.

c) Calcula $\lim_{n\to\infty} D^n$ y comprueba que

$$\lim_{n\to\infty} g_n^t = (x_0 + \frac{1}{2}y_0, 0, \frac{1}{2}y_0 + z_0)^t.$$

40. Un psicólogo sitúa cada día un grupo de ratones en una jaula con dos puertas, A y B. Cada ratón puede ir a través de A donde se encuentra una cortina o a través de B donde recibe comida.

Si un ratón pasa por A, al día siguiente la probabilidad de volver a pasar por A es de $0'5$. Si un ratón pasa por B, al día siguiente la probabilidad de volver a pasar por B es de $0'5$.

Denominamos p_k a la probabilidad de que un ratón pase por A el día k-ésimo y q_k a la probabilidad de que pase por B, y llamamos A a la matriz de transición que permite calcular (p_{k+1}, q_{k+1}) a partir de (p_k, q_k), es decir, tal que

$$\begin{pmatrix} p_{k+1} \\ q_{k+1} \end{pmatrix} = A \begin{pmatrix} p_k \\ q_k \end{pmatrix}.$$

a) Indica razonadamente cual de las siguientes matrices es la matriz de transición A:

$$E = \begin{pmatrix} 0'5 & 0'5 \\ 0'2 & 0'8 \end{pmatrix} \qquad F = \begin{pmatrix} 0'5 & 0'2 \\ 0'5 & 0'8 \end{pmatrix} \qquad G = \begin{pmatrix} 0'5 & 0'8 \\ 0'5 & 0'2 \end{pmatrix}$$

b) Sabemos que el primer día del experimento, la probabilidad de que un ratón pase por A es igual a la de que pase por B, es decir, $0'5$ en cada caso. Calcula la probabilidad de pasar por la puerta A y por la puerta B los días segundo y tercero del experimento.

c) Calcula los autovalores y los autovectores de la matriz A.

d) Al cabo de un tiempo largo, calcula la probabilidad de que un ratón pase por A.

41. Compliquemos el problema 40. Supongamos que la probabilidad de que el ratón pase por la puerta A o la B, depende de por dónde haya pasado los dos días anteriores según el sistema de ecuaciones en diferencias siguiente:

$$\begin{pmatrix} p_{n+2} \\ q_{n+2} \end{pmatrix} = A \cdot \begin{pmatrix} p_{n+1} \\ q_{n+1} \end{pmatrix} + B \cdot \begin{pmatrix} p_n \\ q_n \end{pmatrix}$$

con $A = \frac{3}{4} \cdot \begin{pmatrix} 1/2 & 2/10 \\ 1/2 & 8/10 \end{pmatrix}$ y $B = \frac{1}{4} \cdot \begin{pmatrix} 1/2 & 1/10 \\ 1/2 & 9/10 \end{pmatrix}$. Calcula (p_n, q_n) y $\lim\limits_{n\to\infty} (p_n, q_n)$.

Ecuaciones diofánticas. Sea $A = (a_{ij}) \in M_{n\times m}(\mathbb{Z})$ una matriz con coeficientes números enteros. Sea $\delta_{ij} \in M_n(\mathbb{Z})$ la matriz cuyos coeficientes son todos nulos salvo el ij que es igual a 1. Sea $\lambda \in \mathbb{Z}$. Recordemos (ver 2.8.1) que $(\mathrm{Id} + \lambda \cdot \delta_{ij}) \cdot A$ es igual a la matriz A salvo que la fila i es igual a la fila i de A más λ veces la fila j de A. Decíamos que $\mathrm{Id} + \lambda \cdot \delta_{ij}$ es una transformación elemental especial. Igualmente, si $A \in M_{m\times n}$, $A \cdot (\mathrm{Id} + \lambda \cdot \delta_{ij})$ es igual a la matriz A salvo que la columna j es igual a la columna j de A más λ veces la columna i de A. Mediante transformaciones elementales especiales [5]de filas y columnas de A podemos obtener una matriz "diagonal" (d_{ij}) (con $d_{ij} = 0$ si $i \neq j$). Veámoslo con un ejemplo:

Diagonalicemos la matriz $\begin{pmatrix} 3 & 2 & 4 \\ 8 & 4 & 10 \end{pmatrix}$. Permutando la primera fila con otra (o la primera columna con otra) coloquemos en el lugar $1,1$ el número entero no nulo más pequeño en valor absoluto.

Admitamos también, multiplicar una fila o una columna por -1, con esto, componiendo transformaciones elementales podemos intercambiar dos filas o dos columnas.

$$\begin{pmatrix} 3 & 2 & 4 \\ 8 & 4 & 10 \end{pmatrix} \overset{C_1\times C_2}{\rightsquigarrow} \begin{pmatrix} 2 & 3 & 4 \\ 4 & 8 & 10 \end{pmatrix}$$

Si a la segunda columna le restamos una vez la primera columna obtendremos (en el lugar $1,2$) un número no nulo más pequeño en valor absoluto que el 2, que a continuación colocaremos en el lugar $1,1$:

$$\begin{pmatrix} 2 & 3 & 4 \\ 4 & 8 & 10 \end{pmatrix} \overset{C_2-C_1}{\rightsquigarrow} \begin{pmatrix} 2 & 1 & 4 \\ 4 & 4 & 10 \end{pmatrix} \overset{C_1\times C_2}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & 4 \\ 4 & 4 & 10 \end{pmatrix}$$

Repetiríamos el proceso, pero ya no es posible, ahora pivotando con el coeficiente $1,1$, obtenemos ceros en la primera columna y en la primera fila:

$$\begin{pmatrix} 1 & 2 & 4 \\ 4 & 4 & 10 \end{pmatrix} \underset{C_3-4\cdot C_1}{\overset{C_2-2\cdot C_1}{\rightsquigarrow}} \begin{pmatrix} 1 & 0 & 0 \\ 4 & -4 & -6 \end{pmatrix} \overset{F_2-4F_1}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -6 \end{pmatrix}$$

Ahora repitamos el proceso con la matriz A^{11} (jugando del mismo modo con las columnas 2 y 3 y la fila 2):

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -6 \end{pmatrix} \overset{C_3-C_2}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -2 \end{pmatrix} \overset{C_2\times C_3}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & -4 \end{pmatrix} \overset{C_3-2\cdot C_2}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \end{pmatrix}.$$

Consideremos ahora un sistema de ecuaciones lineales

$$\begin{aligned} a_{11}x_1+\cdots+a_{1n}x_n &= b_1 \\ &\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n &= b_m \end{aligned}$$

con $a_{ij}, b_k \in \mathbb{Z}$ para todo $i,j.k$, que escribimos abreviadamente $A\cdot x = b$. Mediante transformaciones elementales (en columnas y filas) diagonalizamos la matriz A. Es decir, sabemos calcular matrices cuadradas invertibles F y C de modo que $F\cdot A\cdot C = D$, donde $D=(d_{ij})$, con $d_{ij}=0$ para todo $i\neq j$. Sea $x' = C^{-1}\cdot x$. Entonces,

$$D\cdot x' = FACC^{-1}\cdot x = FA\cdot x = F\cdot b$$

Este sistema de ecuaciones es sencillo de resolver; y una vez obtenido x', tenemos que $x = C\cdot x'$.
Igualmente podemos diagonalizar matrices con coeficientes polinomios, sólo tenemos que hablar del polinomio de grado mínimo donde antes hablábamos del número entero de valor absoluto mínimo.

42. Calcula las soluciones enteras del sistema de ecuaciones diofánticas

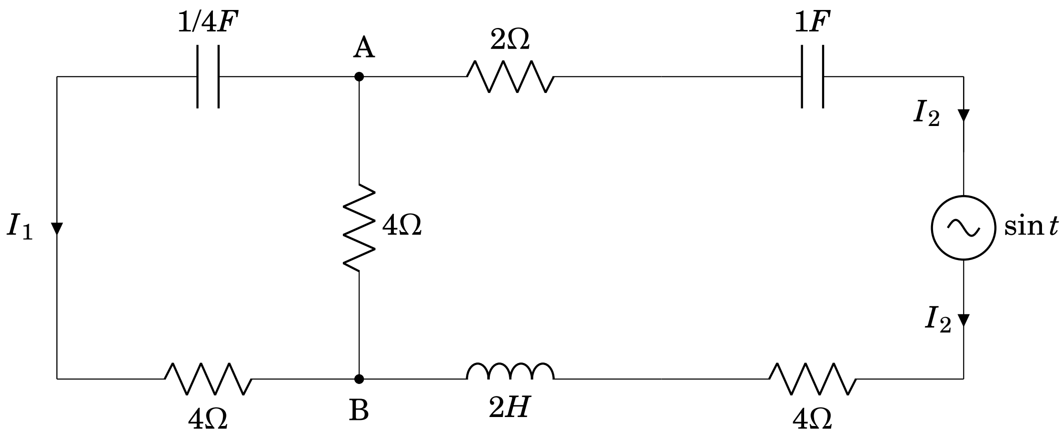
$$\begin{aligned} 2x+3y+z&=6 \\ 4x+2y+z&=5 \end{aligned}$$

43. Compliquemos el ejemplo 3.10.18. Supongamos que la ocupación de los hijos depende de la ocupación de los padres y abuelos según el sistema de ecuaciones en diferencias finitas

$$\begin{pmatrix} p_{n+2} \\ s_{n+2} \\ t_{n+2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0'6 & 0'2 & 0'1 \\ 0'2 & 0'5 & 0'1 \\ 0'2 & 0'3 & 0'8 \end{pmatrix}\cdot\begin{pmatrix} p_{n+1} \\ s_{n+1} \\ t_{n+1} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0'4 & 0'2 & 0'1 \\ 0'3 & 0'5 & 0'1 \\ 0'3 & 0'3 & 0'8 \end{pmatrix}\cdot\begin{pmatrix} p_n \\ s_n \\ t_n \end{pmatrix}$$

($p_n+s_n+t_n=1$, para todo n). Resuelve el sistema de ecuaciones y calcula $\lim\limits_{n\to\infty} p_n$.

44. Calcula las intensidades de corriente en los diferentes tramos del siguiente **circuito eléctrico.**



3.14. Solución de los problemas

P1. Sea $T(t)$ la temperatura del sólido en el instante t. Por la ley de enfriamiento de Newton $\frac{dT}{dt} = K \cdot (5 - T)$, es decir, $T' + KT = 5K$, que escribimos

$$(D + K\,\mathrm{Id})(T) = 5K$$

Una solución particular T_0 de esta ecuación diferencial es $T_0 = 5$. Todas las soluciones son $5 + \mathrm{Ker}(D + K\,\mathrm{Id}) = 5 + e^{-Kt} \cdot \mu$. Tenemos que $T(t) = 5 + e^{-Kt} \cdot \mu$, $T(0) = 20$ y $T(2) = 10$. Luego $\mu = 15$ y $K = \ln\sqrt{3}$.

P2. Sea $s(t)$ los gramos de sal que hay en el depósito en el minuto t, luego $c(t) = \frac{s(t)}{1000}$ es la concentración salina del agua salada del depósito en el minuto t. Tenemos que $\Delta s = s(t+h) - s(t)$ es igual a los gramos de sal que hay en 50 litros de agua salada en el instante t por el incremento de tiempo Δt transcurrido (h) cuando el incremento es pequeño, multiplicado por -1, es decir, $\Delta s = -50 \cdot \frac{s(t)}{10000} \cdot \Delta t$. Luego, $s'(t) = -\frac{50}{10000} \cdot s(t)$. Dividiendo por 1000, tenemos $c'(t) = \frac{-50}{10000} \cdot c(t)$. Es decir,

$$(D + \frac{50}{10000})(c(t)) = 0$$

Luego $c(t) = a \cdot e^{\frac{-50}{10000}t}$. Como $c(0) = 5\,\%$, entonces $a = 5$. Por tanto, $t = -200 \cdot \ln(\frac{c(t)}{5})$. Después de $-200 \cdot \ln(\frac{1}{5}) = 322$ minutos la concentración baja al $1\,\%$.

P3. El apartado a) es una comprobación inmediata. Resolvamos b). Tenemos

$$g = y' + fy = (D + f\,\mathrm{Id})(y) = (D + f\,\mathrm{Id})(e^{-\int f} e^{\int f} y) = e^{-\int f} D(e^{\int f} y).$$

Luego, $D(e^{\int f} y) = e^{\int f} g$. Entonces, $e^{\int f} y = cte + \int e^{\int f} g$ e $y = e^{-\int f}(cte + \int e^{\int f} g)$.

P4. a) Consideremos el cambio de variable $x = e^t$ (o $t = \ln x$). Tenemos que $\frac{\partial}{\partial t} = \Theta$, por la regla de la cadena:

$$\frac{\partial f(x)}{\partial t} = \frac{\partial f(x)}{\partial x} \cdot \frac{\partial x}{\partial t} = \frac{\partial f(x)}{\partial x} \cdot x = \Theta(f(x)).$$

Por tanto,

$$\operatorname{Ker}(\Theta - \alpha \operatorname{Id})^r = \operatorname{Ker}(\frac{\partial}{\partial t} - \alpha \operatorname{Id})^r = e^{\alpha t} \cdot \{\sum_{i=0}^{r-1} \lambda_i t^i\} = x^\alpha \cdot \{\sum_{i=0}^{r-1} \lambda_i (\ln x)^i\}.$$

b) $\Theta^2 = x^2 \frac{\partial^2}{\partial x^2} + x \frac{\partial}{\partial x}$. Por tanto,

$$0 = x^2 y'' + bxy' + cy = (\Theta^2 + (b-1)\Theta + c \operatorname{Id})(y).$$

Entonces, $y \in \operatorname{Ker}(\Theta^2 + (b-1)\Theta + c \operatorname{Id})$. Sean $\alpha, \beta = \frac{-b+1 \pm \sqrt{(b-1)^2 - 4c}}{2}$. Entonces,

$$y = \begin{cases} \lambda x^\alpha + \mu x^\beta, & \text{si } \alpha \neq \beta. \\ x^\alpha (\lambda + \mu \ln x), & \text{si } \alpha = \beta. \end{cases}$$

P5. El gasto de energía E es constante, es decir $E = c \cdot t$ (donde t es el tiempo y c constante). E es la suma del gasto de energía por la fuerza de rozamiento, $\int_0^l Kmv \cdot dl$ (donde l es el espacio que recorre el coche, $v = \frac{dl}{dt}$ la velocidad del coche y m la masa del coche y K una constante), más la energía por la aceleración del coche $\int_0^l ma \cdot dl$. Luego,

$$c \cdot t = E = \int_0^l Kmv \cdot dl + \int_0^l ma \cdot dl = \int_0^l (Kmv + ma) \cdot dl$$

Derivando respecto al tiempo obtenemos

$$c = \frac{d \int_0^l (Kmv + ma) \cdot dl}{dt} = \frac{d \int_0^l m(Kv + a) \cdot dl}{dl} \cdot \frac{dl}{dt} = m(Kv + a)v = Kmv^2 + \frac{1}{2} m (v^2)'.$$

Luego, $(D + 2K)(v^2) = \frac{2c}{m}$. Una solución particular es $v^2 = \frac{c}{Km}$ y la solución general es $v^2 = \frac{c}{Km} + \lambda \cdot e^{-2Kt}$. Supongamos que $v(0) = 0$, entonces $\lambda = -\frac{c}{Km}$ y

$$v(t) = \sqrt{\frac{c}{Km} - \frac{c}{Km} \cdot e^{-2Kt}} = \sqrt{\frac{c}{Km}} \cdot \sqrt{1 - e^{-2Kt}}.$$

P6. Tenemos $(D^2 + aD + b \operatorname{Id})(f) = 0$. Las raíces de $x^2 + ax + b$ son $\frac{-a \pm \sqrt{a^2 - 4b}}{2}$. Luego,

$$f \in \operatorname{Ker}(D^2 + aD + b \operatorname{Id}) = e^{\frac{-ax}{2}} \cdot (\lambda \cos(\frac{\sqrt{-a^2 + 4b}}{2} x) + \mu \operatorname{sen}(\frac{\sqrt{-a^2 + 4b}}{2} x))$$

P7. En el problema 6 hemos calculado la solución de la homogénea. Calculemos una solución particular. Sea $E = \langle \cos wt, \operatorname{sen} wt\rangle$. Busquemos en E una solución particular. Resolvamos $(D^2 + aD + b\,\mathrm{Id})(\lambda \cos wt + \mu \operatorname{sen} wt) = c \cos wt$, es decir,

$$\begin{aligned} (b-w^2)\cdot\lambda + aw\cdot\mu = c \\ -aw\cdot\lambda + (b-w^2)\cdot\mu = 0 \end{aligned}$$

Luego, $\lambda = \frac{c(b-w^2)}{(b-w^2)^2+a^2w^2}$ y $\mu = \frac{acw}{(b-w^2)^2+a^2w^2}$.

P8. Sigamos otro método para calcular una solución particular de la ecuación diferencial: Tenemos $(D^3 - 2D^2 + \mathrm{Id})(f) = xe^x$, luego una solución particular es

$$\begin{aligned} f &= \frac{1}{D^3-2D^2+\mathrm{Id}} xe^x = \frac{1}{(D^2-D-\mathrm{Id})(D-\mathrm{Id})} xe^x = e^x \frac{1}{(D^2+D-\mathrm{Id})D} x \\ &\overset{*}{=} e^x(-\mathrm{Id} - D - 2D^2)\frac{1}{D} x = e^x\Big(\frac{-x^2}{2} + cte - x - 2\Big) = e^x\Big(\frac{-x^2}{2} - x + a\Big) \end{aligned}$$

($*$ el desarrollo de Taylor de $\frac{1}{x^2+x-1}$ en $x = 0$ hasta orden tres es $-1 - x - 2x^2$).

P9. Observemos que $D^i(\sum_i c_i(x)s_i(x)) = \sum_i c_i(x) D^i s_i(x)$, para $i < n$. Entonces,

$$D^n\Big(\sum_i c_i(x)s_i(x)\Big) = D\Big(\sum_i c_i(x)D^{n-1}s_i(x)\Big) = \sum_i c_i(x)D^n s_i(x) + \sum_i c_i(x)' D^{n-1} s_i(x).$$

Luego, $p(D)(\sum_i c_i(x)s_i(x)) = \sum_i c_i(x)p(D)(s_i(x)) + \sum_i c_i(x)' s_i(x)^{n-1)} = 0 + f(x)$.

P10. Observemos que

$$\nabla(gh) = g\nabla(h) + \Delta(g)\nabla(h).$$

Por tanto, por esta fórmula y la primera ecuación del sistema de ecuaciones

$$\nabla\Big(\sum_i c_i s_i\Big) = \sum_i c_i \nabla(g_i) + \sum_i \Delta(c_i)\nabla(s_i) = \sum_i c_i \nabla(g_i).$$

Luego, por la fórmula y la segunda ecuación del sistema de ecuaciones

$$\nabla^2\Big(\sum_i c_i s_i\Big) = \nabla\Big(\sum_i c_i \nabla(g_i)\Big) = \sum_i c_i \nabla^2(g_i)$$

y recurrentemente

$$\nabla^{r-1}\Big(\sum_i c_i s_i\Big) = \nabla\Big(\sum_i c_i \nabla^{r-2}(g_i)\Big) = \sum_i c_i \nabla^{r-1}(g_i).$$

Escribamos $p(\nabla) = \nabla^r + q(\nabla)$. Entonces,

$$\begin{aligned} p(\nabla)(\sum_i c_i s_i) &= \nabla^r(\sum_i c_i s_i) + \sum_i c_i q(\nabla)s_i = \nabla(\sum_i c_i \nabla^{r-1} s_i) + \sum_i c_i q(\nabla)s_i \\ &= (\sum_i \Delta(c_i)\nabla^r s_i) + (\sum_i c_i \nabla^r s_i) + \sum_i c_i q(\nabla)s_i \\ &= (\sum_i \Delta(c_i)\nabla^r s_i) + \sum_i c_i p(\nabla)s_i = f + 0 = f. \end{aligned}$$

P11. Definamos $t(x) = \sum_{n=0}^{\infty} b_n \cdot \frac{x^n}{n!}$ y supongamos que $t'' - t = e^{ix}$. Entonces, $b_{n+2} - b_n = i^n$, es decir, $(\nabla^2 - \mathrm{Id})(b_n) = (i^n)$. Busquemos una solución particular en $\langle (i^n) \rangle$. Resulta que $b_n = -\frac{i^n}{2}$ es una solución particular. Como

$$\mathrm{Ker}(\nabla^2 - \mathrm{Id}) = \mathrm{Ker}(\nabla - \mathrm{Id}) \oplus \mathrm{Ker}(\nabla + \mathrm{Id}) = \{\lambda + \mu(-1)^n\},$$

entonces

$$b_n = \frac{i^n}{-2} + \lambda + \mu \cdot (-1)^n.$$

Luego, $t(x) = \frac{e^{ix}}{-2} + \lambda e^x + \mu e^{-x}$. La parte real de $t(x)$ es $s(x) = \frac{\cos x}{-2} + \lambda e^x + \mu e^{-x}$ $(\lambda, \mu \in \mathbb{R})$ y la parte real de $\frac{b_n}{n!}$ es $a_n = \frac{\frac{(-1)^{\frac{n}{2}}}{-2} + \lambda + \mu}{n!}$ para n par y $a_n = \frac{\lambda - \mu}{n!}$ para n impar.

P12. Sea i_n es el dinero que pagamos en el año n por los intereses del capital que tenemos prestado durante el año n y a_n el dinero que amortizamos en el año n por el capital prestado. Entonces, $d_n = i_n + a_n$. Tenemos que $i_n = I \cdot (K - \sum_{r=1}^{n-1} a_r)$. Por tanto,

$$d_n = a_n + I \cdot (K - \sum_{r=1}^{n-1} a_r).$$

Si aplicamos el operador diferencia, $\Delta(d_n) = \Delta(a_n) - I \cdot a_n = (\nabla - (1+I))(a_n)$. Por otra parte, $(\nabla - I')(d_n) = 0$ (luego $d_n = \lambda' I'^n$). Por tanto, si aplicamos $\nabla - I'$, obtenemos que

$$(\nabla - I')(\nabla - (1+I))(a_n) = 0.$$

Por tanto, $a_n = \lambda I'^n + \mu(1+I)^n$. Sabemos que $d_{N+1} = a_{N+1}$, de lo que se deduce que $\lambda' = \lambda + \mu(\frac{1+I}{I'})^{N+1}$. De las ecuaciones

$$\begin{aligned} \lambda' I' &= d_1 = a_1 + IK = \lambda I' + \mu(1+I) + IK \\ \lambda' I'^2 &= d_2 = a_2 + I(K - a_1) = \lambda I'(I' - I) + \mu(1+I) + IK \end{aligned}$$

se obtiene que $d_1 = \frac{K(1 - I' + I)}{1 - (\frac{I'}{1+I})^N}$.

P13. En efecto, $c_A(x) = |A - x\,\mathrm{Id}| = |(A - x\,\mathrm{Id})^t| = |A^t - x\,\mathrm{Id}| = c_{A^t}(x)$.

P14. Sea $T\colon \mathbb{C}^n \to \mathbb{C}^n$ el operador lineal de matriz A en la base usual. Sea B una base de $\mathbb{C}^n$ donde la matriz asociada a T sea triangular. Los coeficientes de la diagonal son $\alpha_1, \dots, \alpha_n$ porque $c_T(x) = \pm(x-\alpha_1)\cdots(x-\alpha_n)$. La matriz asociada a $p(T)$ en la base B es triangular y los coeficientes de la diagonal son $p(\alpha_1), \dots, p(\alpha_n)$, luego $c_{p(T)}(x) = \pm(x - p(\alpha_1))\cdots(x - p(\alpha_n))$.

P15. Tenemos que $c_A(x) = |\begin{pmatrix} a-x & 1 \\ 1 & a-x \end{pmatrix}| = (x-a)^2 - 1 = x^2 - 2ax + (a^2-1)$ y $|A| = a^2 - 1$. Para que A sea invertible, debe ser $a \neq \pm 1$. Los autovalores de A son $a \pm 1$, que son las raíces de $c_A(x)$. 2 es un autovalor de A si y solo si $a = 1, 3$. Con todo, $a = 3$.

P16. Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ un operador lineal de autovectores $(1,0), (1,1)$ y autovalores asociados 1 y 2. Consideremos la base $B = \{(1,0),(1,1)\}$, entonces $T_B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Entonces, la matriz de T en la base usual es

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

P17. Sea P la matriz de columnas los vectores v_1, v_2, v_3 y D la matriz diagonal con coeficientes $1, 3, -1$ en la diagonal. Entonces,

$$A = PDP^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & -4 \\ 2 & 0 & -1 \end{pmatrix}.$$

P18. Tenemos $c_A(x) = -6 - 11x - 6 - x^2 - x^3 = -(x-3)(x-2)(x-1)$. Luego los autovalores son $3, 2, 1$. Los autovectores de autovalor 3: $\mathrm{Ker}(A - 3\,\mathrm{Id}) = \langle(-2,0,1)\rangle$. Los autovectores de autovalor 2: $\mathrm{Ker}(A - 2\,\mathrm{Id}) = \langle(-1,2,1)\rangle$. Los autovectores de autovalor 1: $\mathrm{Ker}(A - \mathrm{Id}) = \langle(1,3,0)\rangle$.

Tenemos que $c_B(x) = 24 + 26x + 9x^2 + x^3 = (x-4)(x-3)(x-2)$. Luego los autovalores son $4, 3, 2$. Los autovectores de autovalor 4: $\mathrm{Ker}(B - 4\,\mathrm{Id}) = \langle(-2,3,3)\rangle$. Los autovectores de autovalor 3: $\mathrm{Ker}(B - 3\,\mathrm{Id}) = \langle(-4,6,5)\rangle$. Los autovectores de autovalor 1: $\mathrm{Ker}(B - 2\,\mathrm{Id}) = \langle(-1,1,1)\rangle$.

P19. Calculemos los valores propios de A: $c_A(x) = -(x+1)^2(x-1)$, luego los valores propios son $-1, 1$. Calculemos los vectores propios $\mathrm{Ker}(A + \mathrm{Id}) = \langle(1,0,1),(-1,1,0)\rangle$ y $\mathrm{Ker}(A - \mathrm{Id}) = \langle(1,1,1)\rangle$. A diagonaliza en la base $B = \{(1,0,1),(-1,1,0),(1,1,1)\}$.

Si P es la matriz de columnas los vectores de la base B, y D es la matriz diagonal con coeficientes $-1,-1,1$ en la diagonal, entonces $A = PDP^{-1}$ y

$$A^{50} = (PDP^{-1})^{50} = PD^{50}P^{-1} = PP^{-1} = \mathrm{Id}.$$

P20. La matriz A no es diagonalizable porque $\dim \mathrm{Ker}(A-\mathrm{Id}) = 1$ es distinto de $2 = \dim \mathrm{Ker}(A-\mathrm{Id})^2$.

P21. Sólo C cumple que $C\cdot(1,1)^t = (1,1)^t$ y $C\cdot(-1,0)^t = -1\cdot(-1,0)^t$.

P22. Tenemos que $Av^t = (2\alpha+4,3,\frac{3}{2})^t$. Entonces, $\lambda\cdot v = (2\alpha+4,3,\frac{3}{2})$, si $\lambda = \frac{3}{2}$ y $\alpha = 7$.

P23. El vector $(1,1,\dots,1)$ es un autovector de autovalor 1.

P24. Sea v un autovector de autovalor λ. Entonces, $A^m v^t = A^{m-1}(\lambda v^t) = \cdots = \lambda^m\cdot v^t$. Luego v es un autovector de A de autovalor λ^m. Si A es invertible, entonces $\mathrm{Ker}\,A = 0$, luego $0 \neq Av^t = \lambda v^t$ y $\lambda \neq 0$. Además, $A^{-1}(Av^t) = v^t$, luego $A(\lambda v^t) = v^t$ y $Av^t = \frac{1}{\lambda}v^t$ y $\frac{1}{\lambda}$ es un autovalor de A^{-1}.

P25. Tenemos $c_A(x) = -(x-1)(x-0'1)(x+0{,}1)$. $\mathrm{Ker}(A-\mathrm{Id}) = \langle(\frac{8}{3},\frac{11}{6},1)\rangle$, $\mathrm{Ker}(A-0'1\,\mathrm{Id}) = \langle(-1,1,0)\rangle$ y $\mathrm{Ker}(A+0'1\,\mathrm{Id}) = \langle(-1,0,1)\rangle$.

P26. Tenemos que $c_A(x) = -(x-2)^2(x-3)^3$. $\dim \mathrm{Ker}(A-2\,\mathrm{Id}) = 2$ si y solo si $\alpha = 0$ y $\dim \mathrm{Ker}(A-3\,\mathrm{Id}) = 3$ si y solo si $\beta = \gamma = 0$. Entonces, A es diagonalizable si y solo si $\alpha = \beta = \gamma = 0$.

P27. Tenemos que $A(1,1)^t = (1+a,1+a)^t$, luego $(1,1)$ es autovector de autovalor $1+a$. $A(1,-1)^t = (-1+a,1-a)^t$, luego $(-1,1)$ es autovector de autovalor $-1+a$. Observemos que $1+a \neq -1+a$ y que $(2,1) \notin \langle(1,1)\rangle \cup \langle(-1,1)\rangle$, luego no es un autovector.

P28. Tenemos que $\lambda v^t = Av^t = (40\alpha,10\beta,40)^t$. Luego, $\lambda = 1$ y $\alpha = \frac{1}{4}$ y $\beta = \frac{1}{2}$.

P29. Tenemos que $c_A(x) = -(x+1)^2(x-2)$. $\mathrm{Ker}(A-\mathrm{Id}) = \langle(0,1,0)\rangle$ y $\mathrm{Ker}(A-2\,\mathrm{Id}) = \langle(-1,-1,1)\rangle$. Como $\dim \mathrm{Ker}(A-\mathrm{Id}) = 1 \leq 2$, A no es diagonalizable.

P30. a) Tenemos que $0 = |A-\mathrm{Id}| = 0'002 - 0'02\cdot a$, luego $a = 0'6$.

b) $A(16,11,b)^t = (16,11,b)^t$, luego $8 + 4'4 + 0'6b = 16$ y $b = 6$.

P31. a) $Au^t = \lambda u^t$, luego $A(2u^t) = 2Au^t = 2\lambda u^t = \lambda(2u^t)$, luego $2u$ es un autovector de autovalor λ.

b),c) $(2A)u^t = 2(Au^t) = 2\lambda u^t$, luego u es un autovector de $2A$ de autovalor 2λ. Por a), $2u$ es autovector de $2A$.

P32. a) Si $\lambda = 0$, entonces $|A| = |A - 0 \cdot \mathrm{Id}| = 0$ y A no es invertible.

b) Si $A = 0$ entonces es diagonalizable y $\lambda = 0$ es un autovalor de A.

c) Existe un autovector (no nulo) con autovalor 0.

d) La matriz diagonal con $1, 0 - 1$ en la diagonal tiene autovalores 1, -1 y 0.

P33. a) $A(0,1,0)^t = (0,0,0) = 0 \cdot (0,1,0)$, luego $(0,1,0)$ es un autovector de autovalor 0.

b) $|A| = 0$ entonces existe $0 \neq v \in \operatorname{Ker} A$, luego v es un autovector de autovalor 0.

c) Si u es un autovector de autovalor 1 y v es un autovector de autovalor 0, entonces $A(u+v)^t = u^t$ y $u+v$ no es un autovector.

P34. Si $F : E \to E$ es un operador lineal diagonalizable y $V \subset E$ es un subespacio vectorial tal que $F(V) \subseteq V$, entonces $F_{|V}$ es diagonalizable: Dado $v \in V$ podemos escribir $v = e_1 + \ldots + e_r$, donde los e_i son autovectores de F de autovalor λ_i y $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces,

$$(\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_r) \cdot e_1 = (F - \lambda_2 \cdot \mathrm{Id}) \circ \cdots \circ (F - \lambda_r \cdot \mathrm{Id})(v) \in V,$$

luego $e_1 \in V$. Igualmente $e_2, \ldots, e_r \in V$. Es fácil probar que existe un sistema generador en V formado por autovectores de F, luego existe una base en V formada por autovectores y F diagonaliza en V.

Sean $\{\alpha_i\}_{i \in I}$ los autovalores de T y $\{\beta_j\}_{j \in J}$ los de S. Si $e \in \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id})$ entonces $S(e) \in \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id})$: $(T - \alpha_i \cdot \mathrm{Id})(S(e)) = S((T - \alpha_i \cdot \mathrm{Id})(e)) = S(0) = 0$. $S_{|\operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id})}$ diagonaliza, luego $\operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) = \oplus_j (\operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) \cap \operatorname{Ker}(S - \beta_j \cdot \mathrm{Id})$ y

$$E = \bigoplus_i \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) = \bigoplus_{i,j} \operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) \cap \operatorname{Ker}(S - \beta_j \cdot \mathrm{Id}).$$

Tomando una base en cada sumando directo $\operatorname{Ker}(T - \alpha_i \cdot \mathrm{Id}) \cap \operatorname{Ker}(S - \beta_j \cdot \mathrm{Id})$ obtenemos una base de E donde T y S diagonalizan.

P35. Sea $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$ un autovector de autovalor $\lambda \in \mathbb{C}$. Entonces, $(a_{ij})v^t = \lambda \cdot v^t$. Tomando $|| \ ||_1$ obtenemos $|\lambda|(|v_1| + \ldots + |v_n|) = ||(a_{ij})v^t||_1 \leq \sum_{ij} |a_{ij}||v_j| < \sum_j |v_j|$, luego $|\lambda| < 1$. Por el teorema 3.10.5, $\lim_{n \to \infty} A^n = 0$.

P36. El coche puede estar en cuatro estados distintos: Estado BB: los frenos de pedal y mano funcionan, estado BM: el freno de pedal funciona, el freno de mano no funciona, estado MB: el freno de pedal no funciona, el freno de mano funciona y estado MM: el coche está destruido. Si el coche está en el estado BB en el año n, entonces en el año $n+1$ estará en el estado BB con probabilidad $0'9 \cdot 0'9$, en el estado BM con probabilidad $0'9 \cdot 0'1$, en el estado MB con probabilidad $0'1 \cdot 0'9$ y en estado MM con probabilidad $0'1 \cdot 0'1$, etc. Obtenemos la llamada matriz de transición del año n al año $n+1$:

$$P = \begin{pmatrix} 0'9\cdot 0'9 & 0'9\cdot 0'6 & 0'6\cdot 0'9 & 0 \\ 0{,}9\cdot 0'1 & 0'9\cdot 0'4 & 0'6\cdot 0'1 & 0 \\ 0'1\cdot 0'9 & 0'1\cdot 0'6 & 0'4\cdot 0'9 & 0 \\ 0'1\cdot 0'1 & 0'1\cdot 0'4 & 0'1\cdot 0'4 & 1 \end{pmatrix}.$$

Si tenemos en el año n, 10 coches en estado BB, 20 en estado BM, 30 en estado MB y 40 en estado MM, y esta información la codificamos con el vector $v_n = (10,20,30,40)^t$, entonces al año siguiente los coches en cada estado vienen codificados por el vector $v_{n+1} = P \cdot v_n$, como el lector puede comprobar. Si dividimos ahora por el número de coches, que es 100, tendremos que en el año n, las proporciones de coches en cada estado (o probabilidad de que un coche esté en cada estado) es $\tilde{v}_n = (0'1,0'2,0'3,0'4)^t$ y al año siguiente es $\tilde{v}_{n+1} = P \cdot \tilde{v}_n$.

Partamos de un estado concreto, por ejemplo $(1,0,0,0)$. La probabilidad m_n de que en el año n no estemos en estado de desastre es

$$m_n = (1,1,1,0)\cdot P^n \cdot (1,0,0,0)^t.$$

Sea Q la submatriz de P tal que $P = \begin{pmatrix} Q & \mathbf{0} \\ \mathbf{p} & 1 \end{pmatrix}$, donde $\mathbf{p} = (0'01, 0'04, 0'04)$ y $\mathbf{0} = (0,0,0)^t$. Escribamos $u = (1,1,1)$. Entonces, $m_n = u \cdot Q^n \cdot (1,0,0)^t$.

Por tanto, nuestra esperanza de vida[6] es

$$\begin{aligned} m &= 1 + u\cdot Q\cdot(1,0,0)^t + \cdots + u\cdot Q^n\cdot(1,0,0)^t + \cdots \\ &= u\cdot(\mathrm{Id} + Q + \cdots + Q^n + \cdots)\cdot(1,0,0)^t \overset{*}{=} u\cdot(\mathrm{Id} - Q)^{-1}\cdot(1,0,0)^t. \end{aligned}$$

($\overset{*}{=}$ Observemos que $(\mathrm{Id}-Q)\cdot(\mathrm{Id}+Q+\cdots+Q^n) = \mathrm{Id}-Q^{n+1}$ y que $\lim\limits_{n\to\infty} Q^n = 0$ porque la suma de los coeficientes de las columnas es menor que 1, ver problema 35. Por lo tanto, $(\mathrm{Id}-Q)\cdot(\mathrm{Id}+Q+\cdots+Q^n+\cdots) = \mathrm{Id}$ y $(\mathrm{Id}+Q+\cdots+Q^n+\cdots) = (\mathrm{Id}-Q)^{-1}$).

En nuestro caso, calculamos $(\mathrm{Id}-Q)^{-1} = \begin{pmatrix} 44'615 & 41'538 & 41'538 \\ 6'9231 & 8'022 & 6'5934 \\ 6'9231 & 6'5934 & 8'022 \end{pmatrix}$, y obtenemos $u\cdot(\mathrm{Id}-Q)^{-1} = (58'462, 56'154, 56'154)$.

[6] Dada una población, denotemos por s_n la probabilidad de que en el año n cada individuo de la población esté vivo. Promediando podríamos decir que todos los individuos han vivido una s_n parte del año n, por tanto, la esperanza de vida de los individuos de la población es $\sum_{i=1}^{\infty} s_i$ años.

Interpretemos los resultados. El tiempo medio para el desastre si partimos con los dos controles probados es algo más de 58 años, mientras que el tiempo medio para el desastre si partimos con uno de los frenos no probado está alrededor de los 56 años. Cuando solo tenemos un sistema de frenado, solamente hay dos estados: B el freno funciona y M coche destruido. La matriz de transición queda $P = \begin{pmatrix} 0'9 & 0 \\ 0'1 & 1 \end{pmatrix}$, por lo que el tiempo medio de fallo es solo $u(\mathrm{Id}-Q)^{-1} = 10$ años.

P37. Sin más que transponer A, podemos suponer que una columna (en vez de una fila) de A es estrictamente mayor que cero, es decir, $Ae_s \gg 0$ para cierto s, donde denotamos $e_s = (0,\dots,\underset{s}{1},\dots,0)^t$. Luego $A^t e_s \gg 0$, para todo $t > 0$. Observemos que $(A+\mathrm{Id})^m = A^m + \binom{m}{1}A^{m-1} + \cdots + \binom{m}{m-1}A + \mathrm{Id}$, luego $(A+\mathrm{Id})^m \gg 0$ si y sólo si $A^m + A^{m-1} + \cdots + A + \mathrm{Id} \gg 0$. Por tanto, para cada e_i y cada e_j, existe un $0 \le r_{ij} \le m$ tal que $A^{r_{ij}} e_i = \sum_l a_{li} e_l$ con $a_{ji} \neq 0$. Por tanto, $A^t \cdot A^{r_{is}} e_i \gg 0$, para todo $t \geq 0$. Sea $N = \text{máx}\{r_{is}, \forall\, i\} + 1$. Entonces, $A^N e_i \gg 0$ para todo i, luego $A^N \gg 0$.

Dado $x = (x_1,\dots,x_n)$, denotemos $[x]_i = x_i$. Sea $B_i^r := \{e_j : [(A+\mathrm{Id})^r(e_i)]_j \neq 0\}$. Observemos que $\{e_i\} \subseteq B_i^1 \subseteq B_i^2 \subseteq \cdots \subseteq B_i^r \subseteq \cdots$ y que si $B_i^r = B_i^{r+1}$ entonces $B_i^r = B_i^{r+1} = B_i^{r+2} = \cdots$. Por tanto, como no pueden existir más de $n-1$ desigualdades estrictas, tenemos que $B_i^{n-1} = B_i^n = B_i^{n+1} = \cdots$. Como $(A+\mathrm{Id})^N \gg 0$ entonces $B_i^N = \{e_1,\dots,e_n\}$ para todo i. Entonces, $B_i^{n-1} = \{e_1,\dots,e_n\}$ para todo i, luego $(A+\mathrm{Id})^{n-1} \gg 0$.

P38. La descendencia es de genotipo AA, AB y BB. De una hembra AA y un macho AB la descendencia es AA, AB, AA, AB. Es decir, $\frac{1}{2}$ AA y $\frac{1}{2}$ AB. De una hembra AB y un macho AB la descendencia es AA, AB, BA, BB. Es decir, $\frac{1}{4}$ AA y $\frac{1}{2}$ AB y $\frac{1}{4}$ BB. De una hembra BB y un macho AB la descendencia es BA, BB, BA, BB. Es decir, $\frac{1}{2}$ AB y $\frac{1}{2}$ BB. Luego la matriz de transición es

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

En la tercera generación, $T^2(\frac{6}{10},\frac{2}{10},\frac{2}{10})^t = (\frac{7}{20},\frac{5}{10},\frac{3}{20})$. Si $w = (w_1,w_2,w_3)$ es vector de probabilidad y autovector de autovalor 1, entonces $\lim\limits_{n\to\infty} T^n(\frac{6}{10},\frac{2}{10},\frac{2}{10})^t = w^t$. $\mathrm{Ker}(T-\mathrm{Id}) = \langle(1,2,1)\rangle$, luego $w = (\frac{1}{4},\frac{1}{2},\frac{1}{4})$.

P39. a) De AA y AA sale con probabilidad 1 AA. De Aa y Aa sale AA, Aa, aA, aa, es decir, con probabilidad $\frac{1}{4}$ AA, con probabilidad $\frac{1}{2}$ Aa y con probabilidad $\frac{1}{4}$ aa. De aa y aa sale con probabilidad 1 aa. La primera matriz es la matriz de transición.

b) $c_A(x) = -(x-1)^2(x-\frac{1}{2})$. $\mathrm{Ker}(A-\mathrm{Id}) = \langle(1,0,0),(0,0,1)\rangle$ y $\mathrm{Ker}(A-\frac{1}{2}\mathrm{Id}) = \langle(1,-2,1)\rangle$.
$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}$ y $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$.

c) $\lim_{n\to\infty} D^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Por tanto,

$$\lim_{n\to\infty} g_n^t = \lim_{n\to\infty} A^n g_0^t = \lim_{n\to\infty} PD^nP^{-1}g_0^t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}^{-1} g_0^t$$
$$= \frac{1}{2}\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \frac{2x_0+y_0}{2} \\ 0 \\ \frac{y_0+2z_0}{2} \end{pmatrix}.$$

P40. **a)** Si el ratón pasa por la puerta A, al día siguiente el ratón pasará con probabilidad $0'5$ por la puerta A y con probabilidad $0'5$ por la puerta B. Si el ratón pasa por la puerta B, al día siguiente el ratón pasará con probabilidad $0'2$ por la puerta A y con probabilidad $0'8$ por la puerta B. Supongamos que 100 ratones han pasado por A y 200 por B, entonces al día siguiente $100\cdot 0'5+200\cdot 0'2$ pasarán por A y $100\cdot 0'5+200\cdot 0'8$ pasarán por B. Si dividimos por el número de ratones $100+200=300$, diremos que si la probabilidad que un ractón pase por A es $\frac{100}{300}=\frac{1}{3}$ y la que pase por B es $\frac{200}{300}=\frac{2}{3}$ (y esta información la codificamos en el vector $(\frac{1}{3},\frac{2}{3})$), entonces al día siguiente la probabilidad de que pase por A es $\frac{100}{300}\cdot 0'5+\frac{200}{300}\cdot 0'2$ y de que pase por B es $\frac{100}{300}\cdot 0'5+\frac{200}{300}\cdot 0'8$, (y esta información la codificamos en $(\frac{1}{3}\cdot 0'5+\frac{2}{3}\cdot 0'2, \frac{1}{3}\cdot 0'5+\frac{2}{3}\cdot 0'8)$). Observemos que

$$\begin{pmatrix} \frac{1}{3}\cdot 0'5+\frac{2}{3}\cdot 0'2 \\ \frac{1}{3}\cdot 0'5+\frac{2}{3}\cdot 0'8 \end{pmatrix} = \begin{pmatrix} 0'5 & 0'2 \\ 0'5 & 0'8 \end{pmatrix}\cdot\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}.$$

La matriz de transición es la segunda.

b) $F^2(0'5,0'5)^t = (0'305,0'695)^t$, $F^3(0'5,0'5)^t = (0'2915,0'7085)$.

c) $c_F(x) = (x-1)(x-0'3)$, luego los autovalores son $1, 0'3$. $\mathrm{Ker}(F-\mathrm{Id}) = \langle(2,5)\rangle$ y $\mathrm{Ker}(F-0'3\,\mathrm{Id}) = \langle(1,-1)\rangle$. $P = \begin{pmatrix} 2 & 1 \\ 5 & -1 \end{pmatrix}$ y $D = \begin{pmatrix} 1 & 0 \\ 0 & 0'3 \end{pmatrix}$.

d) $w = (\frac{2}{7},\frac{5}{7})$ es el vector de probabilidad de autovalor 1. Entonces,

$$\lim_{n\to\infty} F^n(0'5,0'5)^t = w^t = (\frac{2}{7},\frac{5}{7})^t.$$

P41. Sea S el conjunto de todas las sucesiones de números complejos y consideremos el operador lineal $\tilde{\nabla}\colon S^2 \to S^2$, $\tilde{\nabla}((p_n),(q_n)) = ((p_{n+1}),(q_{n+1}))$. Tenemos que calcular los $((p_n),(1-p_n)) \in S^2$ tales que

$$(\tilde{\nabla}^2 - A\cdot\tilde{\nabla} - B)((p_n),(1-p_n)) = 0.$$

Es decir, $(\nabla^2 - \frac{0'9}{4}\cdot\nabla - 0'1\cdot\mathrm{Id})(p_n) = \frac{0'7}{4}$. Una solución particular de esta ecuación en diferencias es $(p_n) = (\frac{7}{27})$, y todas las soluciones son

$$\begin{aligned}&(\frac{7}{27}) + \mathrm{Ker}(\nabla^2 - \frac{0'9}{4}\cdot\nabla - 0'1\cdot\mathrm{Id}) = (\frac{7}{23}) + \mathrm{Ker}(\nabla + 0'22\cdot\mathrm{Id}) + \mathrm{Ker}(\nabla - 0'44\cdot\mathrm{Id})\\ &= \{(\frac{7}{27} + a\cdot(-0'22)^n + b\cdot 0'44^n)\}.\end{aligned}$$

Luego, $\lim\limits_{n\to\infty} p_n = \frac{7}{27}$.

P42. Escribamos el sistema de ecuaciones de modo matricial y “diagonalicemos”:

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 6\\ 4 & 2 & 1 & 5\end{array}\right) \overset{C_1 x C_3}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 6\\ 1 & 2 & 4 & 5\end{array}\right) \overset{F_2 - F_1}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 6\\ 0 & -1 & 2 & -1\end{array}\right) \underset{C_3-2C_1}{\overset{C_2-3C_1}{\rightsquigarrow}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 6\\ 0 & -1 & 2 & -1\end{array}\right)$$
$$\overset{C_3+2\cdot C_2}{\rightsquigarrow} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 6\\ 0 & -1 & 0 & -1\end{array}\right).$$

Las soluciones del sistema

$$\begin{aligned} x' &= 6\\ -y' &= -1\end{aligned}$$

son $x' = 6$, $y' = 1$ y $z' = z'$. Sea $C = C_1 \times C_3 \circ (C_2 - 3C_1) \circ (C_3 - 2C_1) \circ (C_3 + 2\cdot C_2)$ la composición de las transformaciones elementales por columnas efectuadas. Las soluciones de nuestro sistema son $(x,y,z)^t = C\cdot(6,1,z')^t$, es decir,

$$\begin{pmatrix} 6\\ 1\\ z'\end{pmatrix} \overset{F_2+2F_3}{\rightsquigarrow} \begin{pmatrix} 6\\ 1+2z'\\ z'\end{pmatrix} \overset{F_1-2F_3}{\rightsquigarrow} \begin{pmatrix} 6-2z'\\ 1+2z'\\ z'\end{pmatrix} \overset{F_1-3F_2}{\rightsquigarrow} \begin{pmatrix} 3-8z'\\ 1+2z'\\ z'\end{pmatrix} \overset{F_1\times F_3}{\rightsquigarrow} \begin{pmatrix} z'\\ 1+2z'\\ 3-8z'\end{pmatrix} = \begin{pmatrix} x\\ y\\ z\end{pmatrix}.$$

P43. Hagamos $t_n = 1 - p_n - s_n$, para todo n. Entonces, obtenemos

$$\begin{pmatrix} p_{n+2}\\ s_{n+2}\end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0'5 & 0'1\\ 0'1 & 0'4\end{pmatrix}\cdot\begin{pmatrix} p_{n+1}\\ s_{n+1}\end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0'3 & 0'1\\ 0'2 & 0'4\end{pmatrix}\cdot\begin{pmatrix} p_n\\ s_n\end{pmatrix} + \begin{pmatrix} 0'1\\ 0'1\end{pmatrix}.$$

Que escribimos

$$\begin{pmatrix} \nabla^2 - \frac{0'5}{2}\nabla - \frac{0'3}{2} & -\frac{0'1}{2}\nabla - \frac{0'1}{2}\\ -\frac{0'1}{2}\nabla - 0'1 & \nabla^2 - 0'2\nabla - 0'2\end{pmatrix}\begin{pmatrix} p_n\\ s_n\end{pmatrix} = \begin{pmatrix} 0'1\\ 0'1\end{pmatrix}.$$

Que vamos a resolver por transformaciones elementales

$$\left(\begin{array}{cc|c} 20\nabla^2-5\nabla-3 & -\nabla-1 & 2 \\ -\nabla-2 & 20\nabla^2-4\nabla-4 & 2 \end{array}\right) \underset{F_1+(20\nabla-45)F_2}{\rightsquigarrow} \left(\begin{array}{cc|c} 87 & 179+99\nabla-980\nabla^2+400\nabla^3 & -48 \\ -\nabla-2 & 20\nabla^2-4\nabla-4 & 2 \end{array}\right)$$

$$\underset{F_2+\frac{\nabla+2}{87}F_1}{\rightsquigarrow} \left(\begin{array}{cc|c} 87 & 179+99\nabla-980\nabla^2+400\nabla^3 & -48 \\ 0 & \frac{10+29\nabla-121\nabla^2-180\nabla^3+400\nabla^4}{87} & \frac{10}{29} \end{array}\right) \underset{C_2-\frac{179+99\nabla-980\nabla^2+400\nabla^3}{87}C_1}{\rightsquigarrow}$$

$$\left(\begin{array}{cc|c} 87 & 0 & -48 \\ 0 & \frac{10+29\nabla-121\nabla^2-180\nabla^3+400\nabla^4}{87} & \frac{10}{29} \end{array}\right)$$

Las soluciones de este último sistema

$$\begin{array}{rcl} 87\cdot p'_n & = & -48 \\ \frac{10+29\nabla-121\nabla^2-180\nabla^3+400\nabla^4}{87}(s'_n) & = & \frac{10}{29} \end{array}$$

son $p'_n=\frac{-16}{29}$, $s'_n=\frac{5}{23}+\lambda\cdot(-0'39)^n+\mu\cdot(-0'24)^n+\alpha\cdot(0'42)^n+\beta\cdot(0'65)^n$. Escribamos $P(\nabla)=\frac{179+99\nabla-980\nabla^2+400\nabla^3}{87}$. Entonces,

$$\begin{pmatrix} p'_n \\ s'_n \end{pmatrix} \underset{F_1-P(\nabla)F_2}{\rightsquigarrow} \begin{pmatrix} p'_n-P(\nabla)s'_n \\ s'_n \end{pmatrix} = \begin{pmatrix} p_n \\ s_n \end{pmatrix}.$$

$P(\nabla)(\gamma^n)=(P(\gamma)\cdot\gamma^n)$, por tanto,

$$\begin{array}{rcl} p_n & = & \frac{14}{69}+\lambda\cdot 0'32\cdot(-0'39)^n-\mu\cdot 1'01\cdot(-0'24)^n-\alpha\cdot 0'90\cdot(0'42)^n+\beta\cdot 0'73\cdot(0'65)^n \\ s_n & = & \frac{5}{23}+\lambda\cdot(-0'39)^n+\mu\cdot(-0'24)^n+\alpha\cdot(0'42)^n+\beta\cdot(0'65)^n \end{array}$$

Luego, $\lim\limits_{n\to\infty}\begin{pmatrix} p_n \\ s_n \end{pmatrix}=\begin{pmatrix} \frac{14}{69} \\ \frac{5}{23} \end{pmatrix}$ y $\lim\limits_{n\to\infty} t_n = 1-\frac{14}{69}-\frac{5}{23}=\frac{40}{69}$. Observemos que el vector $w=(\frac{14}{69},\frac{5}{23},\frac{40}{69})$ cumple que

$$w^t=\frac{1}{2}\begin{pmatrix} 0'6 & 0'2 & 0'1 \\ 0'2 & 0'5 & 0'1 \\ 0'2 & 0'3 & 0'8 \end{pmatrix}\cdot w^t+\frac{1}{2}\begin{pmatrix} 0'4 & 0'2 & 0'1 \\ 0'3 & 0'5 & 0'1 \\ 0'3 & 0'3 & 0'8 \end{pmatrix}\cdot w^t$$

es decir, es un autovector de autovalor 1 de $\frac{1}{2}\begin{pmatrix} 0'6 & 0'2 & 0'1 \\ 0'2 & 0'5 & 0'1 \\ 0'2 & 0'3 & 0'8 \end{pmatrix}+\frac{1}{2}\begin{pmatrix} 0'4 & 0'2 & 0'1 \\ 0'3 & 0'5 & 0'1 \\ 0'3 & 0'3 & 0'8 \end{pmatrix}$.

P44. De B a A la corriente es de intensidad I_1+I_2. La suma de diferencia de potenciales, en todo circuito cerrado, es nula. Luego,

$$\begin{array}{rcl} 4I_1'+4(I_1'+I_2')+4I_1 & = & 0 \\ 4I_2'+2I_2''+4(I_1'+I_2')+2I_2'+I_2 & = & \cos t \end{array}$$

Dividamos la primera ecuación por 4 y escribamos las ecuaciones matricialmente,

$$\begin{pmatrix} 2D+1 & D \\ 4D & 2D^2+10D+1 \end{pmatrix}\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \cos t \end{pmatrix}$$

Resolvamos el sistema de ecuaciones mediante transformaciones elementales:

$$\left.\begin{array}{cc} 2D+1 & D \\ 4D & 2D^2+10D+1 \end{array}\right| \begin{array}{c} 0 \\ \cos t \end{array} \overset{F_1-\frac{1}{2}F_2}{\rightsquigarrow} \left.\begin{array}{cc} 1 & -D^2-4D-\frac{1}{2} \\ 4D & 2D^2+10D+1 \end{array}\right| \begin{array}{c} -\frac{1}{2}\cos t \\ \cos t \end{array} \overset{F_2-4DF_1}{\rightsquigarrow}$$

$$\left.\begin{array}{cc} 1 & -D^2-4D-\frac{1}{2} \\ 0 & 4D^3+18D^2+12D+1 \end{array}\right| \begin{array}{c} -\frac{1}{2}\cos t \\ \cos t-2\operatorname{sen} t \end{array} \overset{C_2+(D^2+4D+\frac{1}{2})C_1}{\rightsquigarrow}$$

$$\left.\begin{array}{cc} 1 & 0 \\ 0 & 4D^3+18D^2+12D+1 \end{array}\right| \begin{array}{c} -\frac{1}{2}\cos t \\ \cos t-2\operatorname{sen} t \end{array}$$

Tenemos que resolver el sistema de ecuaciones

$$\begin{array}{rcl} J_1 & = & -\frac{1}{2}\cos t \\ (4D^3+18D^2+12D+1)J_2 & = & \cos t-2\operatorname{sen} t \end{array}$$

Sus soluciones son

$$\begin{array}{rcl} J_1 & = & -\frac{\cos t}{2} \\ J_2 & = & \frac{-\cos t}{353}+\frac{42\operatorname{sen} t}{353}+\lambda e^{-0'1t}+\mu e^{-0'7t}+\gamma e^{-3'7t} \end{array}$$

Luego,

$$\begin{pmatrix} J_1 \\ J_2 \end{pmatrix} = \begin{pmatrix} -\frac{\cos t}{2} \\ \frac{-\cos t}{353}+\frac{42\operatorname{sen} t}{353}+\lambda e^{-0'1t}+\mu e^{-0'7t}+\gamma e^{-3'7t} \end{pmatrix} \overset{F_1+(D^2+4D+\frac{1}{2})F_2}{\rightsquigarrow}$$

$$\begin{pmatrix} \frac{-8\cos t-17\operatorname{sen} t}{353}+0'12\lambda e^{-0'1t}-1'7\mu e^{-0'7t}-0'58\gamma e^{-3'7t} \\ \frac{-\cos t+42\operatorname{sen} t}{353}+\lambda e^{-0'1t}+\mu e^{-0'7t}+\gamma e^{-3'7t} \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}.$$

Si $t>>0$ entonces $(I_1,I_2)\simeq(\frac{-8\cos t-17\operatorname{sen} t}{353},\frac{-\cos t+42\operatorname{sen} t}{353})$.

Capítulo 4

Geometría euclídea

4.1. Introducción

Tenemos la intuición clara de lo que es la distancia entre dos puntos de $\mathbb{R}^3$ y el ángulo entre dos vectores. No nos es tan inmediato cómo calcularlos o definirlos. Veremos que con la definición de producto escalar en $\mathbb{R}^3$ (y en otros espacios vectoriales) podemos definir con rigor todos estos conceptos y calcularlos.

La operación fundamental para cálculos y demostraciones será la proyección ortogonal de un vector sobre un subespacio vectorial. Con ella calcularemos bases ortonormales mediante el proceso de ortonormalización de Gram-Schmidt y calcularemos la solución aproximada óptima de un sistema de ecuaciones lineales.

Toda la teoría que primero desarrollamos en $\mathbb{R}$-espacios vectoriales reales puede desarrollarse de modo totalmente análogo en $\mathbb{C}$-espacios vectoriales. Lo cual es particularmente necesario en Mecánica Cuántica. En este contexto será sencillo probar que todo operador hermítico diagonaliza en una base ortonormal y obtendremos la clasificación euclídea de las cuádricas.

4.2. Espacio vectorial euclídeo

Comencemos con la definición de producto escalar y la definición de espacio vectorial euclídeo.

1. Definiciones: Dar un producto escalar en un $\mathbb{R}$-espacio vectorial E es asignar a cada par de vectores $e, v \in E$ un numero real, que denotaremos $e \cdot v$ ó $\langle e|v\rangle$, de modo que

1. Es simétrico: $e \cdot v = v \cdot e$.

2. Es lineal a la derecha y a la izquierda:

$$e\cdot(\lambda v+\mu v') = \lambda(e\cdot v)+\mu(e\cdot v'), \quad (\lambda v+\mu v')\cdot e = \lambda(v\cdot e)+\mu(v'\cdot e).$$

3. Es definido-positivo: $e\cdot e \geq 0$, y solo se da la igualdad cuando $e=0$.

Un $\mathbb{R}$-espacio vectorial E con un producto escalar $\cdot$ se dice que es un **espacio vectorial euclídeo**, también diremos "$(E,\cdot)$ es un espacio vectorial euclídeo".

2. Ejemplos: 1. El producto escalar usual en $\mathbb{R}^2$ es

$$(x_1,x_2)\cdot(y_1,y_2) := x_1y_1+x_2y_2.$$

2. El producto escalar usual en $\mathbb{R}^3$ es

$$(x_1,x_2,x_3)\cdot(y_1,y_2,y_3) := x_1y_1+x_2y_2+x_3y_3.$$

3. El producto escalar usual en $\mathbb{R}^n$ es

$$(x_1,\dots,x_n)\cdot(y_1,\dots,y_n) := x_1y_1+\ldots+x_ny_n.$$

4. Un producto escalar en $F :=$ {Aplicaciones $f\colon[0,1]\to\mathbb{R}$ continuas} $=\{e^x+\pi,\cos(x), x^2+1,$ etc.$\}$ es

$$\langle f|g\rangle := \int_0^1 f(x)g(x)\,\mathrm{d}x.$$

3. Ejercicio: Calcula en el ejemplo segundo anterior $(1,2,3)\cdot(1,-1,-2)$. Calcula en el ejemplo cuarto anterior $\langle x|x^2+1\rangle$.

Si $(E,\cdot)$ es un espacio vectorial euclídeo y $W\subseteq E$ es un subespacio vectorial, entonces en particular tenemos en W un producto escalar, luego W es de modo natural un espacio vectorial euclídeo.

4. Definición: El **módulo** de un vector $e\in E$ es el número real $\|e\| := +\sqrt{e\cdot e}$.

Observemos que $\|e\|^2 = e\cdot e$.

5. Proposición: *Se cumple que* $\|\lambda\cdot e\| = |\lambda|\cdot\|e\|$.

Demostración. En efecto,

$$\|\lambda\cdot e\| = \sqrt{(\lambda e)\cdot(\lambda e)} = \sqrt{\lambda^2(e\cdot e)} = |\lambda|\cdot\|e\|.$$

□

Por tanto, $\frac{e}{\|e\|}$ (con $e \neq 0$) es un vector de módulo 1.

6. Definición: Diremos que dos vectores $e, v \in E$ son ortogonales cuando $e \cdot v = 0$.

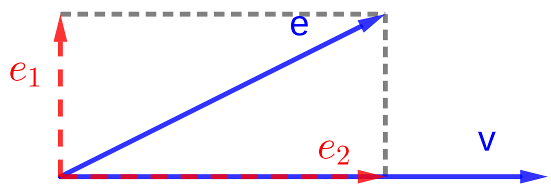
7. Ejercicio: Calcula un vector de $\mathbb{R}^2$ ortogonal a $(1,2)$.

8. Teorema de Pitágoras: *Sea $(E, \cdot)$ un espacio euclídeo y $e, v \in E$. Entonces:*

$$\text{“Los vectores } e,v \text{ son ortogonales} \iff \|e+v\|^2 = \|e\|^2 + \|v\|^2.\text{”}$$

Demostración. Como $\|e+v\|^2 = (e+v)\cdot(e+v) = e\cdot e + v\cdot v + 2(e\cdot v) = \|e\|^2 + \|v\|^2 + 2(e\cdot v)$, es fácil concluir el teorema. □

9. Proyección ortogonal de un vector e sobre $\langle v\rangle$: Suponemos que $v \neq 0$ y queremos expresar $e = e_1 + e_2$ de modo que e_1 sea ortogonal a v y $e_2 = \lambda \cdot v$, para cierto $\lambda \in \mathbb{R}$. Diremos que e_2 es la proyección ortogonal de e sobre la recta $\langle v\rangle$.



Calculemos e_2 o equivalentemente λ:

$e\cdot v = (e_1 + \lambda\cdot v)\cdot v = e_1\cdot v + \lambda(v\cdot v) = \lambda(v\cdot v)$.

Por tanto, $\lambda = \frac{e\cdot v}{v\cdot v}$ y $e_2 = \frac{e\cdot v}{v\cdot v}\cdot v$. Si definimos $e_1 := e - \frac{e\cdot v}{v\cdot v}\cdot v$, tenemos que $e_1\cdot v = 0$ y

$$e = e_1 + e_2.$$

10. Desigualdad de Cauchy-Schwarz: $|e\cdot v| \leq \|e\|\cdot\|v\|$.

Demostración. Sea $e_2 = \lambda\cdot v$ la proyección ortogonal de e sobre $\langle v\rangle$. Por el teorema de Pitágoras, $||e_2|| \leq ||e||$. Además,

$$|e\cdot v| = |e_2\cdot v| = |\lambda\cdot v\cdot v| = |\lambda|\cdot||v||^2 = ||e_2||\cdot||v|| \leq ||e||\cdot||v||.$$

□

11. Desigualdad triangular: $\|e+v\| \leq \|e\| + \|v\|$.

Demostración. Basta probar que $\|e+v\|^2 \leq (\|e\|+\|v\|)^2$. Tenemos

$$\begin{aligned}\|e+v\|^2 &= (e+v)\cdot(e+v) = e\cdot e + v\cdot v + 2\cdot e\cdot v = \|e\|^2 + \|v\|^2 + 2\cdot e\cdot v \\ &\underset{4.2.10}{\leq} \|e\|^2 + \|v\|^2 + 2\|e\|\cdot\|v\| = (\|e\|+\|v\|)^2.\end{aligned}$$

□

12. Definición: La **distancia** entre dos puntos $p,q \in E$ es $d(p,q) := \|q-p\|$.

13. Ejemplo: Calculemos la distancia entre $(1,2,3)$ y $(1,1,1)$:

$$d((1,2,3),(1,1,1)) = \|(1,1,1)-(1,2,3)\| = \|(0,-1,-2)\| = \sqrt{1+4} = \sqrt{5}.$$

Observemos que $d(p,q) = \|q-p\| = \|p-q\| = d(q,p)$.
Por la desigualdad triangular tenemos que

$$d(p,q)+d(q,r) = \|q-p\| + \|r-q\| \geq \|r-p\| = d(p,r).$$

14. Ejercicio: La configuración espacial del ion hexaacuocobre $Cu(H_2O)_6$ es la de un octaedro regular de vértices las moléculas de agua y centro el átomo de cobre. Si la longitud de los enlaces $Cu-OH_2$ es 167 pm (pm= $10^{-12}m$), calcula la distancia entre los átomos de oxígeno de dos moléculas de agua adyacentes.

15. Definición: Cuando $e,v \neq 0$ tenemos que $-1 \leq \frac{e\cdot v}{\|e\|\cdot\|v\|} \leq 1$, y diremos que el **coseno** del ángulo α que forman e y v es

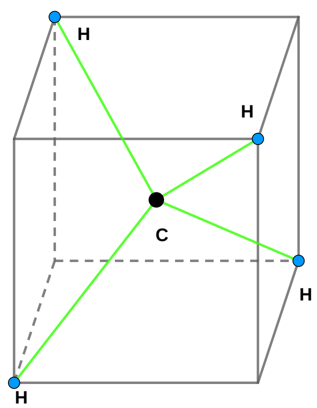
$$\cos\alpha := \frac{e\cdot v}{\|e\|\cdot\|v\|},$$

de modo que $\alpha = \arccos(\frac{e\cdot v}{\|e\|\cdot\|v\|})$, que es un número real entre 0 y π (medido en radianes). Observemos que $\alpha = \frac{\pi}{2}$ justo cuando $e\cdot v = 0$. Denotaremos $\alpha =: \widehat{ev}$.

El ángulo que forman 3 puntos distintos ordenados a,b,c se define como el ángulo que forman los vectores $\vec{ba} := a-b$ y $\vec{bc} := c-b$.

16. Ejercicio: Sean e y v dos vectores de módulo 1 y α el ángulo que forman. Prueba que la proyección ortogonal de e en v es igual a $(\cos\alpha)\cdot v$.

17. Ejercicio: La molécula de metano CH_4 tiene la siguiente configuración espacial: El átomo de carbono está situado en el centro de un cubo y los átomos de hidrógeno están localizados alternativamente en los vértices del cubo. Puede pensarse también, que los átomos de hidrógeno son los vértices de un tetraedro de centro el átomo de carbono. Calcula el ángulo que forman H,C,H.



4.3. Bases ortonormales

1. Definición: Diremos que una base $B = \{e_1, \dots, e_n\}$ de un espacio vectorial euclídeo $(E, \cdot)$ es ortogonal si $e_i \cdot e_j = 0$ para todo $i \neq j$; diremos que es una base ortonormal si $e_i \cdot e_j = 0$ para todo $i \neq j$ y $e_i \cdot e_i = 1$, para todo i.

2. Ejemplo: Si $(\mathbb{R}^3, \cdot)$ es el espacio vectorial euclídeo usual, entonces la base usual $B = \{(1,0,0),(0,1,0),(0,0,1)\}$ es una base ortonormal.

3. Sea $B = \{e_1, \dots, e_n\}$ una base ortonormal de E, $v_B = (x_1, \dots, x_n)$ y $v'_B = (x'_1, \dots, x'_n)$, entonces

$$v \cdot v' = (x_1e_1 + x_2e_2 + \cdots + x_ne_n) \cdot (x'_1e_1 + x'_2e_2 + \cdots + x'_ne_n) = \textstyle\sum_{i=1}^n x_ix'_i = v_B \cdot v'_B.$$
$$\|v\| = \sqrt{v \cdot v} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \|v_B\|.$$
$$d(v, v') = \|v - v'\| = \|v_B - v'_B\| = d(v_B, v'_B).$$

Observemos que $v \cdot e_i = (x_1, \dots, x_n) \cdot (0, \overset{i}{\dots}, 1, \dots, 0) = x_i$, luego $v_B = (v \cdot e_1, \dots, v \cdot e_n)$.

4. Ejercicio: Si $\{e_1, \dots, e_n\}$ es una base ortogonal de un espacio vectorial euclídeo E, probad que $\frac{e_1}{\|e_1\|}, \dots, \frac{e_n}{\|e_n\|}$ es una base ortonormal de E.

5. Ejercicio: Sea $(E, \cdot)$ un espacio vectorial euclídeo y $e_1, \dots, e_r$ vectores que cumplen que $e_i \cdot e_j = 0$ si $i \neq j$. Probad que $e_1, \dots, e_r$ son linealmente independientes.

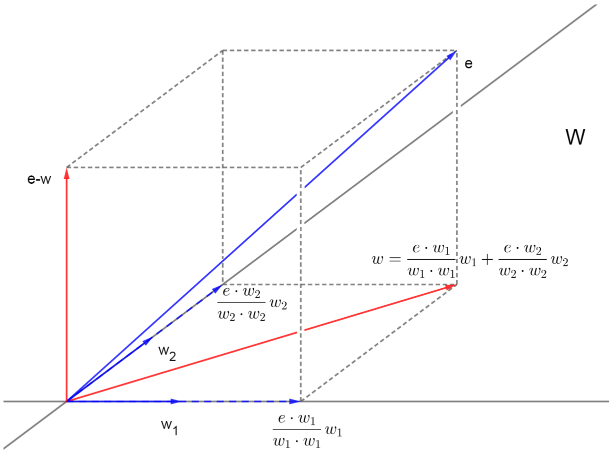
6. Definición: Sea $(E, \cdot)$ un espacio vectorial euclídeo, $e \in E$ y $W \subseteq E$ un subespacio. Diremos que e es ortogonal a W, si $e \cdot w = 0$ para todo $w \in W$.

7. Ejercicio: Calcula un vector de $\mathbb{R}^3$ perpendicular al plano $x_1 + 2x_2 + 5x_3 = 0$ (el plano $x_1 + 2x_2 + 5x_3 = 0$ es el subespacio vectorial de $\mathbb{R}^3$ formado por los $(x_1, x_2, x_3) \in \mathbb{R}^3$ tales que $x_1 + 2x_2 + 5x_3 = 0$.)

8. Proyección ortogonal sobre un subespacio: Sea $(E, \cdot)$ un espacio vectorial euclídeo y $W \subseteq E$ un subespacio. Sea $\{w_1, \dots, w_n\}$ una base ortogonal de W y $e \in E$ un vector. Diremos que

$$w = \frac{e \cdot w_1}{w_1 \cdot w_1} w_1 + \cdots + \frac{e \cdot w_n}{w_n \cdot w_n} w_n.$$

es la proyección ortogonal de e sobre W.



Observemos que $(e-w)\cdot w_i = 0$, para todo i, luego $e-w$ es ortogonal a W. Por último, w es el vector de W más cercano a e: Dado $w' \in W$, sea $w'' = w' - w \in W$. Entonces, $w' = w + w''$ y

$$d(e,w')^2 = \|e-w'\|^2 = \|(e-w)+(-w'')\|^2$$
$$= \|e-w\|^2 + \|w''\|^2 = d(e,w)^2 + \|w''\|^2.$$

Luego, $d(e,w') \geq d(e,w)$ y $d(e,w') = d(e,w)$ si y solo si $w'' = 0$, es decir, $w' = w$.

9. Proceso de ortogonalización de Gram-Schmidt: Sea $\{e_1,\dots,e_n\}$ una base de un espacio vectorial euclídeo $(E,\cdot)$. Vamos a definir una base ortogonal $\{v_1,\dots,v_n\}$ de E: Sea

$$v_1 = e_1.$$

Se cumple que $\langle e_1\rangle = \langle v_1\rangle$. Sea

$$\begin{aligned} v_2 &= e_2 - \text{la proyección ortogonal de } e_2 \text{ sobre } \langle e_1\rangle \\ &= e_2 - \text{la proyección ortogonal de } e_2 \text{ sobre } \langle v_1\rangle \\ &= e_2 - \tfrac{e_2\cdot v_1}{v_1\cdot v_1}\cdot v_1. \end{aligned}$$

Se cumple que v_2, v_1 son ortogonales y además $\langle e_1,e_2\rangle = \langle v_1,e_2\rangle = \langle v_1,v_2\rangle$. Sea

$$\begin{aligned} v_3 &= e_3 - \text{ la proyección ortogonal de } e_3 \text{ sobre } \langle e_1,e_2\rangle \\ &= e_3 - \text{ la proyección ortogonal de } e_3 \text{ sobre } \langle v_1,v_2\rangle \\ &= e_3 - \tfrac{e_3\cdot v_1}{v_1\cdot v_1}\cdot v_1 - \tfrac{e_3\cdot v_2}{v_2\cdot v_2}\cdot v_2. \end{aligned}$$

Se cumple que v_3, v_2, v_1 son ortogonales y además $\langle e_1,e_2,e_3\rangle = \langle v_1,v_2,e_3\rangle = \langle v_1,v_2,v_3\rangle$. Así sucesivamente vamos construyendo una base ortogonal $\{v_1,\dots,v_n\}$ de E. Por lo tanto, $\{\frac{v_1}{\|v_1\|},\dots,\frac{v_n}{\|v_n\|}\}$ es una base ortonormal de E.

10. Ejemplo: Mediante el proceso de ortogonalización de Gram-Schmidt, obtengamos una base ortogonal a partir de la base $B = \{(1,2,2),(1,1,0),(0,1,1)\}$ de $\mathbb{R}^3$:

$$v_1 = (1,2,2)$$
$$v_2 = (1,1,0) - \tfrac{(1,1,0)\cdot v_1}{v_1\cdot v_1}\cdot v_1 = (1,1,0) - \tfrac{3}{9}(1,2,2) = (2/3,1/3,-2/3)$$
$$v_3 = (0,1,1) - \tfrac{(0,1,1)\cdot v_1}{v_1\cdot v_1} v_1 - \tfrac{(0,1,1)\cdot v_2}{v_2\cdot v_2} v_2 = (0,1,1) - \tfrac{4}{9}(1,2,2) + \tfrac{1}{3}(2/3,1/3,-2/3)$$
$$= (-2/9,2/9,-1/9).$$

11. Teorema: *En todos los espacios vectoriales existen bases, luego en todos los espacios vectoriales euclídeos existen bases ortogonales, luego existen bases ortonormales.*

12. Definición: Sea $(E,\cdot)$ un espacio vectorial euclídeo. Dado un subespacio vectorial $L \subseteq E$ se define el ortogonal de L, que denotaremos $L^\perp$, como

$$L^\perp := \{e \in E : e \cdot l = 0, \forall l \in L\}.$$

$L^\perp$ es un subespacio vectorial de E: En efecto, dados $v_1, v_2 \in L^\perp$ y $\lambda \in \mathbb{R}$, entonces $(v_1 + \lambda \cdot v_2) \cdot l = v_1 \cdot l + \lambda \cdot (v_2 \cdot l) = 0$, para todo $l \in L$, luego $v_1 + \lambda \cdot v_2 \in L^\perp$.

13. Propiedades del ortogonal: *Sea $(E,\cdot)$ un espacio vectorial euclídeo y $L.L' \subseteq E$ subespacios vectoriales. Entonces,*

1. $E^\perp = \{0\}$ *y* $\{0\}^\perp = E$.
2. *Si* $L \subseteq L'$ *entonces* $L^\perp \supseteq L'^\perp$.
3. $(L^\perp)^\perp = L$.
4. $E = L \oplus L^\perp$ *(*$E = L + L^\perp$ *y* $L \cap L^\perp = \{0\}$*).*
5. $\dim E = \dim L + \dim L^\perp$.
6. $(L + L')^\perp = L^\perp \cap L'^\perp$ *y* $(L \cap L')^\perp = L^\perp + L'^\perp$.

Demostración. 1. Si $e \in E^\perp$ entonces $e \cdot e = 0$, luego $e = 0$ y $E^\perp = 0$. Obviamente, $\{0\}^\perp = E$.

2. Obvio.

3. 4. y 5. Sea $\langle e_1, \dots, e_n \rangle$ una base ortonormal de L. Ampliando esta base, sea $\{e_1, \dots, e_n, e'_1, \dots, e'_m\}$ una base de E. Por el proceso de ortonormalización de Gram-Schmidt, obtenemos una base ortonormal

$$\{e_1, \dots, e_n, e''_1, \dots, e''_m\}$$

de E. Ahora ya es claro que $L^\perp = \langle e''_1, \dots, e''_m \rangle$ y $(L^\perp)^\perp = \langle e_1, \dots, e_n \rangle = L$.

Además, $E = L \oplus L^\perp$ y tomando dimensiones

$$\dim E = \dim L + \dim L^\perp$$

6. Obviamente, $(L+L')^\perp = L^\perp \cap L'^\perp$. Veamos la otra igualdad. Sea $N = L^\perp$ y $N' = L'^\perp$. Entonces sabemos que $(N + N')^\perp = N^\perp \cap N'^\perp$. Tomando ortogonales, por 3. obtenemos

$$N + N' = (N^\perp \cap N'^\perp)^\perp = (L \cap L')^\perp.$$

Luego, $L^\perp + L'^\perp = (L \cap L')^\perp$.

□

14. Ejemplo : Consideremos el plano Π de $\mathbb{R}^3$ definido por la ecuación

$$3x_1 + x_2 - 5x_3 = 0.$$

Es decir, $\Pi = \{(x_1, x_2, x_3) \in \mathbb{R}^3$ tales que $3x_1 + x_2 - 5x_3 = 0\}$. Sea $v := (3, 1, -5)$, entonces

$$\begin{aligned}\langle v\rangle^{\perp} &= \{(x_1,x_2,x_3) \in \mathbb{R}^3 \text{ tales que } v \cdot (x_1,x_2,x_3) = 0\}\\ &= \{(x_1,x_2,x_3) \in \mathbb{R}^3 \text{ tales que } 3x_1 + x_2 - 5x_3 = 0\} = \Pi.\end{aligned}$$

Por tanto, $\Pi^{\perp} = (\langle v\rangle^{\perp})^{\perp} = \langle v\rangle = \langle(3,1,-5)\rangle$.

Calculemos la distancia de $p = (1,1,-1) \in \mathbb{R}^3$ a Π, $d(p,\Pi)$. Se define $d(p,\Pi)$ como la distancia de p al punto de Π más cercano a p, es decir, a la proyección ortogonal de p sobre Π. Consideremos la proyección de p sobre $\langle v\rangle$,

$$p_1 = \frac{p \cdot v}{v \cdot v} \cdot v = \frac{9}{35}(3,1,-5).$$

Entonces, la proyección ortogonal de p sobre Π es $p_2 = p - p_1$ y

$$d(p,\Pi_1) = d(p,p_2) = \|p - p_2\| = \|p_1\| = \|\frac{9}{35}(3,1,-5)\| = \frac{9}{35}\sqrt{35} = \frac{9}{\sqrt{35}}.$$

15. Ejemplo : Consideremos ahora la recta R de $\mathbb{R}^3$ definida por las ecuaciones

$$\begin{aligned}3x_1 + x_2 - 5x_3 &= 0\\ x_1 + x_2 + x_3 &= 0\end{aligned}$$

Es decir, $R = \Pi_1 \cap \Pi_2$, donde Π_1 es el plano $3x_1 + x_2 - 5x_3 = 0$ y Π_2 es el plano $x_1 + x_2 + x_3 = 0$. Por tanto,

$$R^{\perp} = (\Pi_1 \cap \Pi_2)^{\perp} = \Pi_1^{\perp} + \Pi_2^{\perp} = \langle(3,1,-5),(1,1,1)\rangle.$$

16. Ejemplo : Dados dos vectores $(3,1,-5),(1,1,1) \in \mathbb{R}^3$, calculemos $\langle(3,1,-5),(1,1,1)\rangle^{\perp}$. Observemos que $\langle(3,1,-5),(1,1,1)\rangle$ es el plano de ecuación

$$\begin{vmatrix} x_1 & x_2 & x_3\\ 3 & 1 & -5\\ 1 & 1 & 1\end{vmatrix} = 0.$$

Es decir, $\langle(3,1,-5),(1,1,1)\rangle$ es el plano de ecuación $6x_1 - 8x_2 + 2x_3 = 0$. Por lo tanto, $\langle(3,1,-5),(1,1,1)\rangle^{\perp} = \langle(6,-8,2)\rangle$.

4.3.1. Solución aproximada óptima de un sistema de ecuaciones lineales

Consideremos un sistema de ecuaciones lineales

$$\begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & b_1 \\ \cdots & = & \cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n & = & b_m \end{array}$$

que escribimos de modo conciso $A\cdot x^t = b^t$.

17. Definición: Diremos que $\tilde{x}\in\mathbb{R}^n$ es una solución aproximada del sistema de ecuaciones lineales si $\|A\tilde{x}^t - b^t\| \le \|Ax^t - b^t\|$, para todo $x\in\mathbb{R}^n$.

18. Proposición: *Sea $b'\in \operatorname{Im} A$ la proyección ortogonal de b sobre $\operatorname{Im} A$. Entonces, $\tilde{x}\in\mathbb{R}^n$ es una solución aproximada del sistema de ecuaciones lineales si y solo si $A\tilde{x}^t = b'$.*

Demostración. Recordemos que b' es el vector de $\operatorname{Im} A$ más cercano a b. Por tanto, $\|Ax^t - b^t\| > \|b'-b\|$, para todo $x\in\mathbb{R}^n$ tal que $Ax^t \neq b'$.

□

19. Corolario: *Sea $\tilde{x}$ una solución aproximada del sistema de ecuaciones lineales. Entonces, x' es una solución aproximada del sistema si y solo si $x'-\tilde{x}\in \operatorname{Ker} A$.*

20. Proposición: *Sea $A\in M_{m\times n}(\mathbb{R})$. Entonces,*

$$(\operatorname{Im} A)^{\perp} = \operatorname{Ker} A^t.$$

Demostración. $b\in(\operatorname{Im} A)^{\perp} \iff b\cdot Ax^t = 0$, para todo $x\in\mathbb{R}^n \iff b\cdot A = 0 \iff A^t b^t = 0 \iff b\in \operatorname{Ker} A^t$. □

21. Teorema: *$\tilde{x}\in\mathbb{R}^n$ es una solución aproximada del sistema de ecuaciones lineales $Ax^t = b^t \iff \tilde{x}$ es solución del sistema de ecuaciones lineales $A^tAx^t = A^tb^t$.*

Demostración. Observemos que $b'\in \operatorname{Im} A$ es la proyección ortogonal de b en $\operatorname{Im} A$ si y solo si $b-b'\in(\operatorname{Im} A)^{\perp} = \operatorname{Ker} A^t$. Es decir, $b'\in \operatorname{Im} A$ es la proyección ortogonal de b en $\operatorname{Im} A$ si y solo si $A^t(b-b')^t = 0$, o equivalentemente $A^tb'^t = A^tb^t$. Sea b' la la proyección ortogonal de b en $\operatorname{Im} A$, entonces $Ax^t = b'^t$ si y solo si $A^tAx^t = A^tb^t$.

□

22. Ejemplo : Calculemos la función $f(x) = ax + b$ cuya gráfica "mejor" se aproxime a los puntos $(1,1),(3,3),(-1,5)$ y $(2,-2)$. Con mayor precisión, queremos una solución aproximada del sistema de ecuaciones lineales

$$\begin{array}{rcr} a+b & \simeq & 1 \\ 3a+b & \simeq & 3 \\ -a+b & \simeq & 5 \\ 2a+b & \simeq & -2 \end{array}$$

Sea $A^t = \begin{pmatrix} 1 & 3 & -1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$. Buscamos una solución del sistema de ecuaciones $A^tAx^t = A^t(1,3,5,-2)^t$, es decir, $\begin{pmatrix} 15 & 5 \\ 5 & 4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$. Luego, $a = \frac{-31}{35}$ y $b = \frac{20}{7}$.

23. Definición : Diremos que una solución aproximada $\tilde{x}$ es óptima si $\|\tilde{x}\| \leq \|x'\|$ para toda solución aproximada x' del sistema de ecuaciones.

24. Proposición : *Sea $(E,\cdot)$ un espacio vectorial euclídeo y $T\colon E \to E'$ una aplicación lineal. La aplicación lineal $T'\colon (\operatorname{Ker} T)^\perp \to \operatorname{Im} T$, $T'(e) := T(e)$ es un isomorfismo lineal.*

Demostración. T' es biyectiva: Sea $T(e) \in \operatorname{Im} T$ (con $e \in E$). Existen $e_1 \in \operatorname{Ker} T$ y $e_2 \in (\operatorname{Ker} T)^\perp$ (únicos) de modo que $e = e_1 + e_2$. Entonces,

$$T(e) = T(e_1 + e_2) = T(e_1) + T(e_2) = T(e_2) = T'(e_2).$$

Si $T'(e'_2) = T'(e_2)$ ($e'_2 \in (\operatorname{Ker} T)^\perp$) entonces $T(e'_2) = T(e_2)$, que equivale a $T(e'_2 - e_2) = 0$. Por tanto, $e'_2 - e_2 \in \operatorname{Ker} T \cap (\operatorname{Ker} T)^\perp = \{0\}$, luego $e'_2 - e_2 = 0$ y $e'_2 = e_2$. □

25. Proposición : *Sea $\tilde{x}$ una solución aproximada del sistema de ecuaciones lineales $Ax^t = b^t$. Entonces, $\tilde{x}$ es la solución aproximada óptima del sistema de ecuaciones lineales $Ax^t = b^t$ si y solo si $\tilde{x} \in (\operatorname{Ker} A)^\perp$.*

Demostración. Sea $b' \in \operatorname{Im} A$ la proyección ortogonal de b sobre $\operatorname{Im} A$ e $\tilde{y} \in (\operatorname{Ker} A)^\perp$ tal que $A\tilde{y}^t = b'^t$, que sabemos que existe por la proposición 4.3.24. Por el teorema 4.3.21, $\tilde{y}$ es una solución aproximada, luego $z = \tilde{x} - \tilde{y} \in \operatorname{Ker} A$. Además,

$$\|\tilde{x}\| = \|z + \tilde{y}\| = \sqrt{\|z\|^2 + \|\tilde{y}\|^2} \geq \sqrt{\|\tilde{y}\|^2} = \|\tilde{y}\|.$$

Por tanto, $\tilde{y}$ es la solución aproximada óptima y $\tilde{x}$ es óptima si y solo si $z = 0$, es decir, $\tilde{x} = \tilde{y}$. □

26. Proposición: *Sea $A \in M_{m\times n}(\mathbb{R})$. Entonces,*

$$(\operatorname{Ker} A)^{\perp} = \operatorname{Im} A^t.$$

Demostración. Por la proposición 4.3.20, $(\operatorname{Ker} A)^{\perp} = ((\operatorname{Im} A^t)^{\perp})^{\perp} = \operatorname{Im} A^t$. □

27. Teorema: *Si $\tilde{y}$ es una solución del sistema de ecuaciones lineales $A^tAA^ty^t = A^tb^t$, entonces $A^t\tilde{y}^t$ es la solución aproximada óptima del sistema de ecuaciones lineales $Ax^t = b^t$.*

Demostración. Por el teorema 4.3.21, $A^t\tilde{y}^t$ es solución aproximada del sistema $Ax^t = b^t$. Por la proposición 4.3.25, $A^t\tilde{y}^t \in \operatorname{Im} A^t = (\operatorname{Ker} A)^{\perp}$ es la solución aproximada óptima de $Ax^t = b^t$.

□

28. Ejemplo: Calculemos la solución aproximada óptima del sistema de ecuaciones lineales

$$\begin{array}{rcl} x+y+2z &=& 1 \\ x-y+z &=& 1 \\ 3x+y+5z &=& 1 \end{array}$$

Sea $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \\ 3 & 1 & 5 \end{pmatrix}$. Entonces, $A^tAA^t = \begin{pmatrix} 50 & 26 & 126 \\ 18 & 6 & 42 \\ 84 & 42 & 210 \end{pmatrix}$. Una solución del sistema $A^tAA^ty^t = A^t(1,1,1)^t$ es $y = (-1/42, 5/21, 0)$. Entonces la solución aproximada óptima es $A^ty^t = (3/14, -11/42, 4/21)^t$.

4.4. Isometrías. Matrices ortogonales

1. Notación: Supondremos en esta sección que E y E' son espacios vectoriales euclídeos y que $T: E \to E'$ un isomorfismo $\mathbb{R}$-lineal.

2. Definición: Diremos que T es una isometría si "conserva distancias", es decir,

$$d(T(e), T(v)) = d(e, v), \text{ para todo } e, v \in E.$$

3. Ejemplo: Sea $B = \{e_1, \ldots, e_n\}$ una base ortonormal de E. La aplicación $\mathbb{R}$-lineal $E \to \mathbb{R}^n$, $e \mapsto e_B$ es una isometría.

4. Proposición: *Las siguientes afirmaciones son equivalentes:*

1. *T es una isometría.*

2. *$\|T(e)\| = \|e\|$, para todo $e \in E$.*

3. *"T conserva el producto escalar": $T(e)\cdot T(v) = e\cdot v$, para todo $e, v \in E$.*

Demostración. Observemos que $d(T(e),T(v)) = \|T(e)-T(v)\| = \|T(e-v)\|$ y $d(e,v) = \|e-v\|$. Por tanto, T conserva distancias si y solo si $\|T(w)\| = \|w\|$, para todo $w \in E$, que equivale a decir que $T(w)\cdot T(w) = w\cdot w$.

Si T conserva distancias, como

$$T(e+v)\cdot T(e+v) = (T(e)+T(v))\cdot(T(e)+T(v)) = T(e)\cdot T(e)+T(v)\cdot T(v)+2T(e)\cdot T(v)$$
$$T(e+v)\cdot T(e+v) = (e+v)\cdot(e+v) = e\cdot e+v\cdot v+2e\cdot v.$$

obtenemos que $T(e)\cdot T(v) = e\cdot v$. Recíprocamente, si $T(e)\cdot T(v) = e\cdot v$, para todo $e, v \in E$, entonces $T(w)\cdot T(w) = w\cdot w$, para todo $w \in E$ y T conserva distancias.

□

5. Proposición : *Sea $\{e_1,\dots,e_n\}$ una base ortonormal de E. T es una isometría si y solo si $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal.*

Demostración. $\Rightarrow$) $T(e_i)\cdot T(e_i) = e_i\cdot e_i = 1$, para todo i y $T(e_i)\cdot T(e_j) = e_i\cdot e_j = 0$, para todo $i \neq j$, luego $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal.

$\Leftarrow$) Dado $e = x_1e_1+\cdots+x_ne_n$, tenemos que $T(e) = x_1T(e_1)+\cdots+x_nT(e_n)$ y

$$||e|| = \sqrt{x_1^2+\cdots+x_n^2} = ||T(e)||.$$

□

6. Corolario : *Sean $B = \{e_1,\dots,e_n\}$ y $B' = \{e'_1,\dots,e'_n\}$ bases ortonormales de E y E', repectivamente. Sea A la matriz asociada a T en dichas bases. Entonces, la aplicación lineal T es una isometría $\iff A^tA = \mathrm{Id}$.*

Demostración. La columna i de A es el vector $T(e_i)^{B'}$. $A^tA = \mathrm{Id}$ si y solo si el producto escalar de la fila i de A^t con la columna j de A^t es 0 si $i \neq j$ y es 1 si $i = j$. Es decir, $A^tA = \mathrm{Id}$ si y solo si $T(e_i)_{B'}\cdot T(e_j)_{B'} = 0$ si $i \neq j$ y $T(e_i)_{B'}\cdot T(e_j)_{B'} = 1$ si $i = j$. Luego, $A^tA = \mathrm{Id}$ si y solo si $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal. Por la proposición anterior $A^tA = \mathrm{Id}$ si y solo si T es una isometría.

□

7. Nota: Las matrices $A \in M_n(\mathbb{R})$ que cumplen que $A^t \cdot A = \mathrm{Id}$ se las llama matrices ortogonales. Si A es ortogonal, entonces $\det(A)^2 = \det(A^tA) = \det(\mathrm{Id}) = 1$, luego $\det(A) = \pm 1$. Por tanto, si $T\colon E \to E$ es una isometría entonces $\det(T) = \pm 1$.

8. Ejemplo: Los giros del plano alrededor del origen son isometrías: La matriz de un giro $g_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ de ángulo α, en la base usual, es

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

Como

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}^t \cdot \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

g_α es una isometría. Observemos que $\det(g_\alpha) = 1$.

9. Ejemplo: Sea $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ la simetría respecto del eje OX. La matriz de S en la base usual es $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Como $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^t \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ entonces S es una isometría. Observemos que $\det(S) = -1$.

10. Ejemplo: Dado un hiperplano $H \subset \mathbb{R}^n$ (que pasa por el origen) sea e un vector perpendicular a H.La aplicación lineal S definida por $S(e) = -e$ y $S(h) = h$ para todo $h \in H$, es una isometría, denominada simetría respecto del hiperplano plano H. Observemos que $S = S^{-1}$. Una isometría que sea la identidad sobre un hieprplano H, pero que no sea igual a $\mathrm{Id}\colon \mathbb{R}^n \to \mathbb{R}^n$, es justamente la simetría respecto del hiperplano H.

11. Proposición: *Toda isometría $T\colon E \to E$ es igual a la composición de n o menos simetrías.*

Demostración. Si $\dim E = 1$ entonces $T = \mathrm{Id}$ o la simetría respecto del origen. Supongamos que hemos demostrado el teorema cuando $\dim E = 1, \dots, n-1$. Basta demostrarlo para el caso $\dim E = n$. Sea $\{e_1, \dots, e_n\}$ una base ortonormal de E. Probemos que componiendo T con cierta simetría S_1 (o con $S_1 = \mathrm{Id}$) obtenemos que $(S_1 \circ T)(e_1) = e_1$. Si $T(e_1) = e_1$ tomamos $S_1 = \mathrm{Id}$. Si $T(e_1) = -e_1$, entonces S_1 es la simetría respecto del hiperplano $H = \langle e_1 \rangle^\perp$. En los demás casos, definimos S_1 como la simetría respecto del hiperplano $H = \langle e_1 + T(e_1) \rangle \oplus \langle e_1, T(e_1) \rangle^\perp$. Entonces, $(S_1 \circ T)(\langle e_1 \rangle^\perp) = \langle e_1 \rangle^\perp$. Como $\dim \langle e_1 \rangle^\perp = n-1$, sabemos que existen simetrías $S'_2, \dots, S'_n$ de $\langle e_1 \rangle^\perp$ tales que $S'_2 \circ \cdots \circ S'_n = (S_1 \circ T)_{\langle e_1 \rangle^\perp}$. Sea $S_i\colon E \to E$ la aplicación lineal tal que $S_i(e_1) = e_1$ y $S_{i,|\langle e_1 \rangle^\perp} = S'_i$, que es una simetría. Entonces, $S_2 \circ \cdots \circ S_n = S_1 \circ T$ y $T = S_1 \circ \cdots \circ S_n$. $\square$

12. Isometrías del plano: Toda isometría $S:\mathbb{R}^2\to\mathbb{R}^2$ es la identidad o una simetría o una composición de dos simetrías. Sea S_1 la simetría respecto un eje $\langle e_1\rangle$ y S_2 es la simetría respecto un eje $\langle e_2\rangle$. Podemos suponer $e_1\cdot e_2>0$ (si no tomaríamos $-e_1$ en vez de e_1). Entonces $S_2\circ S_1$ es el giro G (en sentido antihorario) de ángulo $2\cdot\widehat{e_1e_2}$. En efecto, puede comprobarse que $G(e_1)=S_2(e_1)=(S_2\circ S_1)(e_1)$ y que $\det(G)=1=\det(S_2\circ S_1)$, entonces como $\det(G^{-1}\circ S_2\circ S_1)=1$ y $(G^{-1}\circ S_2\circ S_1)(e_1)=e_1$ se tiene que $G^{-1}\circ S_2\circ S_1=\mathrm{Id}$. En conclusión las isometrías del plano son las simetrías y los giros.

En general, la composición de dos simetrías de $\mathbb{R}^n$ respecto de dos hiperplanos distintos H y H', es un giro alrededor del "eje" $H\cap H'$.

4.5. Volumen de un paralelepípedo

1. Proposición: *Sean $B=\{e_1,\dots,e_n\}$ y $B'=\{e'_1,\dots,e'_n\}$ dos bases ortonormales de un espacio vectorial euclídeo E. Sean $v_1,\dots,v_n\in E$ y escribamos $v_{iB}=(a_{i1},\dots,a_{in})$ y $v_{iB'}=(a'_{i1},\dots,a'_{in})$. Entonces,*

$$|\det(a_{ij})|=|\det(a'_{ij})|.$$

Demostración. Consideremos la isometría $\mathrm{Id}\colon E\to E$, luego $\mathrm{Id}_B^{B'}$ es una matriz ortogonal. Además, $v_i^{B'}=\mathrm{Id}_B^{B'}\cdot v_i^B$, para todo i. Luego, $(a'_{ij})^t=\mathrm{Id}_B^{B'}\cdot(a_{ij})^t$ y

$$|\det(a'_{ij})|=|\det(\mathrm{Id}_B^{B'})|\cdot|\det(a_{ij})|=|\det(a_{ij})|.$$

□

2. Definición: Sea $B=\{e_1,\dots,e_n\}$ una base ortonormal de E y $v_1,\dots,v_n\in E$ vectores de coordenadas $v_{iB}=(a_{i1},\dots,a_{in})$. Llamaremos volumen del paralelepípedo[1] definido por los vectores $v_1,\dots,v_n$, al número real no negativo

$$Vol(v_1,\dots,v_n):=\left|\begin{vmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{vmatrix}\right|.$$

(si $n=2$, denotamos $Area(v_1,v_2)=Vol(v_1,v_2)$, si $n=1$ denotamos $Long(v_1)=Vol(v_1)$).

Si $n=1$, entonces $v_1=a_{11}e_1$. Obviamente $|a_{11}|=\|v_1\|$ y $Long(v_1)=\|v_1\|$.

[1]Esta definición no depende de la base ortonormal escogida por la proposición 4.5.1.

3. Ejemplo: Calculemos el volumen del paralelepípedo definido por los vectores $e_1 = (1,1,2), e_2 = (-1,1,3), e_3 = (1,2,2) \in \mathbb{R}^3$:

$$Vol(e_1,e_2,e_3) = |\begin{vmatrix} 1 & 1 & 2 \\ -1 & 1 & 3 \\ 1 & 2 & 2 \end{vmatrix}| = |\begin{vmatrix} 1 & 1 & 2 \\ 0 & 2 & 5 \\ 0 & 1 & 0 \end{vmatrix}| = |-5| = 5$$

4. Propiedades del volumen:

1. *$Vol(v_1,\dots,v_n) = 0$ si los vectores $v_1,\dots,v_n$ son linealmente dependientes.*
2. *$Vol(v_1,\overset{i}{\dots},v_i+\lambda v_j,\dots,v_n) = Vol(v_1,\overset{i}{\dots},v_i,\dots,v_n)$, para $i \neq j$.*
3. *$Vol(v_1,\dots,\lambda v_i,\dots,v_n) = |\lambda| \cdot Vol(v_1,\dots,v_i,\dots,v_n)$.*
4. *$Vol(v_1,\dots,v_i,\dots,v_j,\dots,v_n) = Vol(v_1,\dots,v_j,\dots,v_i,\dots,v_n)$.*

Demostración. Son consecuencia de las propiedades del determinante de una matriz. □

5. Observación: Si $v_1,\dots,v_r \in E$ son vectores linealmente independientes entonces $E' = \langle v_1,\dots,v_r\rangle$ es un espacio vectorial euclídeo de dimensión r y podemos definir $Vol(v_1,\dots,v_r)$.

6. Proposición: *Sea $(E,\cdot)$ un espacio vectorial euclídeo y $v_1\dots,v_n \in E$ vectores linealmente independientes. Entonces,*

$$Vol(v_1,\dots,v_n) = \sqrt{\det((v_i\cdot v_j)_{ij})}.$$

Demostración. Sea $\{e_1,\dots,e_n\}$ una base ortonormal de $\langle v_1,\dots,v_n\rangle$ y escribamos $v_i = \sum_j a_{ij}e_j$. Entonces, $(v_i\cdot v_j) = (a_{ij})\cdot(a_{ij})^t$ y

$$Vol(v_1,\dots,v_n) = |\det(a_{ij}))| = \sqrt{\det((v_i\cdot v_j))}.$$

□

7. Proposición: *Sean $v_1,\dots,v_n \in E$ vectores linealmente independientes, v'_n la proyección ortogonal de v_n sobre $\langle v_1,\dots,v_{n-1}\rangle$ y $v''_n = v_n - v'_n$. Entonces,*

$$Vol(v_1,\dots,v_n) = Vol(v_1,\dots,v_{n-1})\cdot \|v''_n\|.$$

Si $n=3$, estamos diciendo que "el volumen de un paralelepípedo es igual al área de la base por la altura".

Demostración. En efecto

$$Vol(v_1,\dots,v_{n-1},v_n) = Vol(v_1,\dots,v_{n-1},v'_n+v''_n) \overset{4.5.4}{=} Vol(v_1,\dots,v_{n-1},v''_n)$$
$$= \begin{vmatrix} v_1\cdot v_1 & \cdots & v_1\cdot v_{n-1} & 0 \\ \vdots & & \vdots & \vdots \\ v_{n-1}\cdot v_1 & \cdots & v_{n-1}\cdot v_{n-1} & 0 \\ 0 & \cdots & 0 & v''_n\cdot v''_n \end{vmatrix}^{\frac{1}{2}} = Vol(v_1,\dots,v_{n-1})\cdot \|v''_n\|.$$

□

8. Ejemplo: Calculemos el área del paralelogramo definido por los vectores de $\mathbb{R}^3$, $v_1=(1,1,1), v_2=(1,-1,1)$. Entonces,

$$Area(v_1,v_2) = \sqrt{\begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix}} = \sqrt{8}.$$

De otro modo: sea

$$v'_2 = \frac{v_2\cdot v_1}{v_1\cdot v_1} v_1 = \frac{(1,-1,1)\cdot(1,1,1)}{(1,1,1)\cdot(1,1,1)}(1,1,1) = \frac{1}{3}\cdot(1,1,1)$$

la proyección ortogonal de $v_2=(1,-1,1)$ sobre $v_1=(1,1,1)$. Sea

$$v''_2 = v_2 - v'_2 = (2/3,-4/3,2/3).$$

Entonces,

$$Area(v_1,v_2) = Long(v_1)\cdot\|v''_2\| = \|(1,1,1)\|\cdot\|(2/3,-4/3,2/3)\| = \sqrt{8}.$$

4.5.1. Producto vectorial

A partir de ahora, en esta subsección, todos los vectores serán vectores de $\mathbb{R}^3$.

9. Definición: Sea $\{\vec{i}=(1,0,0), \vec{j}=(0,1,0), \vec{k}=(0,0,1)\}$ la base ortonormal usual de $\mathbb{R}^3$. Dados dos vectores $e=(a,b,c)$ y (a',b',c') definimos

$$e\times v = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & c \\ a' & b' & c' \end{vmatrix} := (bc'-b'c)\cdot\vec{i} - (ac'-a'c)\cdot\vec{j} + (ab'-a'b)\cdot\vec{k}.$$

y decimos que $e\times v$ es el producto vectorial de los vectores e y v.

Es fácil comprobar que

$$(e+\lambda e')\times v = e\times v+\lambda\cdot(e'\times v),\quad e\times(v+\lambda v') = e\times v+\lambda\cdot(e\times v') \text{ y } e\times v = -v\times e.$$

10. Proposición: $(e\times v)\cdot w = \det(e,v,w)$.

Demostración. Escribamos $e=(a,b,c)$, $v=(a',b',c')$ y $w=(a'',b'',c'')$. Entonces,

$$(e\times v)\cdot w = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a & b & c \\ a' & b' & c' \end{vmatrix}\cdot w = ((bc'-b'c)\cdot\vec{i}-(ac'-a'c)\cdot\vec{j}+(ab'-a'b)\cdot\vec{k})\cdot(a'',b'',c'')$$

$$=(bc'-b'c)\cdot a''-(ac'-a'c)\cdot b''+(ab'-a'b)\cdot c'' = \begin{vmatrix} a'' & b'' & c'' \\ a & b & c \\ a' & b' & c' \end{vmatrix}.$$

□

Recordemos que $|(e\times v)\cdot w| = |\det(e,v,w)| = Vol(e,v,w)$.
La siguiente proposición determina geométricamente qué vector es $e\times w$.

11. Proposición: *Se cumple que*

1. $e\times w$ *es ortogonal a* e *y* w.

2. $\|e\times v\|$ *es igual al área del paralelogramo definido por* e *y* v.

3. $\det(e,v,e\times v)\geq 0$.

Demostración. 1. $(e\times w)\cdot e = \det(e,w,e)=0$ y $(e\times w)\cdot w = \det(e,w,w)=0$.

2. $e\times v$ es ortogonal a e y w. Entonces, $\|e\times v\|^2 = (e\times v)\cdot(e\times v) = Vol(e,v,e\times v) = Vol(e,v)\cdot\|e\times v\|$. Luego, $\|e\times v\| = Vol(e,v)$.

3. $\det(e,v,e\times v) = (e\times v)\cdot(e\times v)\geq 0$. □

12. Cuestiones: ¿Sabrías definir el producto vectorial de dos vectores de un espacio vectorial euclídeo de dimensión 3? ¿Sabrías definir el producto vectorial de tres vectores de $\mathbb{R}^4$?

4.6. Espacio vectorial euclídeo complejo

Hemos definido los $\mathbb{R}$-espacios vectoriales euclídeos. De modo totalmente análogo podemos definir los $\mathbb{C}$-espacios vectoriales euclídeos complejos. Tenemos también resultados análogos con demostraciones totalmente análogas.

Dado un número complejo $z = a + bi \in \mathbb{C}$ denotamos $\bar{z} = a - bi$ su conjugado.

1. Definiciones : Dar un producto escalar en un $\mathbb{C}$-espacio vectorial E es asignar a cada par de vectores $e, v \in E$ un número complejo, que denotaremos $e * v$ ó $\langle e|v\rangle$, de modo que

1. $e * v = \overline{v * e}$.
2. Es lineal a la izquierda y antilineal a la derecha[2]:

$$(\lambda v + \mu v') * e = \lambda(v * e) + \mu(v' * e), \quad e * (\lambda v + \mu v') = \bar{\lambda}(e * v) + \bar{\mu}(e * v').$$

3. Es definido-positivo: $e * e \geq 0$, y solo se da la igualdad cuando $e = 0$.

Un $\mathbb{C}$-espacio vectorial E con un producto escalar $*$ se dice que es un $\mathbb{C}$-**espacio vectorial euclídeo complejo**, también se dice que "$(E, *)$ es un espacio vectorial hermítico".

2. Ejemplos :

1. El producto escalar usual en $\mathbb{C}^2$ es

$$(x_1, x_2) * (y_1, y_2) := x_1\overline{y_1} + x_2\overline{y_2}.$$

2. El producto escalar usual en $\mathbb{C}^3$ es

$$(x_1, x_2, x_3) * (y_1, y_2, y_3) := x_1\overline{y_1} + x_2\overline{y_2} + x_3\overline{y_3}.$$

3. El producto escalar usual en $\mathbb{C}^n$ es

$$(x_1, \dots, x_n) * (y_1, \dots, y_n) := x_1\overline{y_1} + \cdots + x_n\overline{y_n}.$$

4. Un producto escalar en $F := \{\text{Aplicaciones } f : [0,1] \to \mathbb{C} \text{ continuas}\} = \{e^{xi}, 1 + ix^3, \text{sen}(x), \text{etc.}\}$ es

$$\langle f|g\rangle := \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x.$$

[2] En los textos de Física y Química suele seguirse la convención de que sea lineal a la derecha y antilineal a la izquierda.

3. Ejercicio : Calcula en el ejemplo segundo anterior $(1,i,3)*(1,-1,-2)$.

4. Ejercicio : Calcula en el ejemplo cuarto anterior $\langle e^{ix}|e^{2ix}\rangle$.

Si $(E,*)$ es un $\mathbb{C}$-espacio vectorial hermítico y $W\subseteq E$ es un $\mathbb{C}$-subespacio vectorial, entonces en particular tenemos en W un producto escalar, luego W es de modo natural un $\mathbb{C}$-espacio vectorial hermítico.

5. Definición : El **módulo** de un vector $e\in E$ es el número real $\|e\| := +\sqrt{e*e}$. Dados $e,v\in E$, diremos que la **distancia** de e a v es el número real $d(e,v) := \|v-e\|$,

Observemos que $\|e\|^2 = e*e$.

6. Proposición: *Se cumple que* $\|\lambda\cdot e\| = |\lambda|\cdot\|e\|$.

Demostración. En efecto, $\|\lambda\cdot e\| = \sqrt{(\lambda e)*(\lambda e)} = \sqrt{\lambda\cdot\bar{\lambda}\cdot(e*e)} = |\lambda|\cdot\|e\|$. □

Por tanto, $\frac{e}{\|e\|}$ (con $e\neq 0$) es un vector de módulo 1.

7. Teorema de Pitágoras: *Sea* $(E,*)$ *un* $\mathbb{C}$*-espacio vectorial hermítico y* $e,v\in E$*. Si* $e*v=0$*, entonces* $\|e+v\|^2 = \|e\|^2+\|v\|^2$.

Demostración. $\|e+v\|^2 = (e+v)*(e+v) = e*e+v*v+e*v+v*e = \|e\|^2+\|v\|^2$. □

8. Definición : Diremos que dos vectores $e,v\in E$ son ortogonales cuando $e*v=0$.

4.6.1. Bases ortonormales

9. Definición : Diremos que una base $B=\{e_1,\dots,e_n\}$ de un $\mathbb{C}$-espacio vectorial hermítico $(E,*)$ es ortogonal si $e_i*e_j=0$ para todo $i\neq j$; diremos que es una base ortonormal si $e_i*e_j=0$ para todo $i\neq j$ y $e_i*e_i=1$, para todo i.

10. Ejemplo : Si $(\mathbb{C}^3,\cdot)$ es el $\mathbb{C}$-espacio vectorial hermítico usual, entonces la base usual $B=\{(1,0,0),(0,1,0),(0,0,1)\}$ es una base ortonormal.

11. Sea $B=\{e_1,\dots,e_n\}$ una base ortonormal de E, $v_B=(x_1,\dots,x_n)$ y $v'_B=(x'_1,\dots,x'_n)$, entonces

$$\begin{aligned}
&v*v' = (x_1e_1+x_2e_2+\cdots+x_ne_n)*(x'_1e_1+x'_2e_2+\cdots+x'_ne_n) = \textstyle\sum_{i=1}^n x_i\overline{x'_i} = v_B*v'_B\\
&\|v\| = \sqrt{v*v} = \sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2} = \|v_B\|\\
&d(v,v') = \|v-v'\| = \|v_B-v'_B\| = d(v_B,v'_B).
\end{aligned}$$

Observemos que $v*e_i = (x_1,\dots,x_n)*(0,\dots,\overset{i}{1},\dots,0) = x_i$, luego $v_B = (v*e_1,\dots,v*e_n)$.

12. Definición: Sea $(E,*)$ un espacio vectorial hermítico y $W \subseteq V$ un subespacio. Diremos que $e \in E$ es ortogonal a W, si $e * w = 0$ para todo $w \in W$.

13. Proyección ortogonal sobre un subespacio: Sea $(E,*)$ un espacio vectorial hermítico y $W \subseteq E$ un subespacio vectorial. Sea $\{w_1,\dots,w_n\}$ una base ortogonal de W y $e \in E$ un vector. Diremos que

$$w = \frac{e*w_1}{w_1*w_1}w_1 + \cdots + \frac{e*w_n}{w_n*w_n}w_n$$

es la proyección ortogonal de e sobre W. Observemos que $w \in W$ y $(e-w)*w_i = 0$, para todo i, luego $e-w$ es ortogonal a W.

14. Ejemplo: Consideremos el $\mathbb{C}$-espacio vectorial $F = \{f\colon [0,1] \to \mathbb{C}, \text{ continuas}\}$, con el producto escalar

$$\langle f|g\rangle = \int_0^1 f\bar{g}dx.$$

Probemos que las funciones $1, e^{2\pi ix}, e^{4\pi ix}, e^{6\pi ix}, \dots$ son ortogonales entre sí: Supongamos $n \neq m$, entonces

$$\begin{aligned}\langle e^{n2\pi ix}|e^{m2\pi ix}\rangle &= \int_0^1 e^{n2\pi ix}\cdot e^{-m2\pi ix}dx = \int_0^1 e^{(n-m)2\pi ix}dx = \frac{e^{(n-m)2\pi ix}}{(n-m)2\pi i}\Big]_0^1 \\ &= \frac{1-1}{(n-m)2\pi i} = 0.\end{aligned}$$

Si $n = m$, entonces $\langle e^{n2\pi ix}|e^{n2\pi ix}\rangle = \int_0^1 e^{n2\pi ix}\cdot e^{-n2\pi ix}dx = \int_0^1 dx = x]_0^1 = 1$. Calculemos la proyección ortogonal w de $x \in F$ sobre el subespacio vectorial $W = \langle 1, e^{2\pi ix}, e^{4\pi ix}, e^{6\pi ix}\rangle$:

$$\begin{aligned}w &= \frac{\langle x|1\rangle}{\langle 1|1\rangle}1 + \frac{\langle x|e^{2\pi ix}\rangle}{\langle e^{2\pi ix}|e^{2\pi ix}\rangle}e^{2\pi ix} + \frac{\langle x|e^{4\pi ix}\rangle}{\langle e^{4\pi ix}|e^{4\pi ix}\rangle}e^{4\pi ix} + \frac{\langle x|e^{6\pi ix}\rangle}{\langle e^{6\pi ix}|e^{6\pi ix}\rangle}e^{6\pi ix} \\ &= \frac{1}{2} - \frac{e^{2\pi ix}}{2\pi i} - \frac{e^{4\pi ix}}{4\pi i} - \frac{e^{6\pi ix}}{6\pi i}.\end{aligned}$$

(observemos que $\int_0^1 xe^{ax}dx = \frac{xe^{ax}}{a}\Big]_0^1 - \int_0^1 \frac{e^{ax}}{a}dx = \frac{e^a}{a} - (\frac{e^a}{a^2} - \frac{1}{a^2}) = \frac{e^a}{a} - \frac{e^a-1}{a^2}$).

15. Ejercicio: Prueba las desigualdades de Cauchy-Schwartz y triangular en espacios vectoriales hermíticos.

16. Proceso de ortogonalización de Gram-Schmidt: Sea $B = \{e_1,\dots,e_n\}$ una base de un espacio vectorial hermítico $(E,*)$. Obtengamos una base ortogonal $\{v_1,\dots,v_n\}$ de E a partir de B: Sea

$$v_1 = e_1.$$

Se cumple que $\langle e_1\rangle = \langle v_1\rangle$. Sea

$$\begin{aligned} v_2 &= e_2 - \text{la proyección ortogonal de } e_2 \text{ sobre } \langle e_1\rangle \\ &= e_2 - \text{la proyección ortogonal de } e_2 \text{ sobre } \langle v_1\rangle \\ &= e_2 - \frac{e_2 * v_1}{v_1 * v_1}\cdot v_1. \end{aligned}$$

Se cumple que v_2, v_1 son ortogonales y además $\langle e_1, e_2\rangle = \langle v_1, e_2\rangle = \langle v_1, v_2\rangle$. Sea

$$\begin{aligned} v_3 &= e_3 - \text{ la proyección ortogonal de } e_3 \text{ sobre } \langle e_1, e_2\rangle \\ &= e_3 - \text{ la proyección ortogonal de } e_3 \text{ sobre } \langle v_1, v_2\rangle \\ &= e_3 - \frac{e_3 * v_1}{v_1 * v_1}\cdot v_1 - \frac{e_3 * v_2}{v_2 * v_2}\cdot v_2. \end{aligned}$$

Se cumple que v_3, v_2, v_1 son ortogonales y además $\langle e_1, e_2, e_3\rangle = \langle v_1, v_2, e_3\rangle = \langle v_1, v_2, v_3\rangle$. Así sucesivamente vamos construyendo una base ortogonal $\{v_1, \dots, v_n\}$ de E. Por lo tanto, $\{\frac{v_1}{\|v_1\|}, \dots, \frac{v_n}{\|v_n\|}\}$ es una base ortonormal de E.

17. Teorema : *En todos los espacios vectoriales sobre los números complejos existen bases, luego en todos los $\mathbb{C}$-espacios vectoriales euclídeos existen bases ortonormales.*

18. Definición : Sea $(E, *)$ un espacio vectorial hermítico. Dado un subespacio vectorial $L \subseteq E$ se define el ortogonal de L, que denotaremos $L^\perp$ como

$$L^\perp := \{e \in E : e * l = 0, \forall l \in L\}.$$

$L^\perp$ es un subespacio vectorial de E: En efecto, dados dos vectores $v_1, v_2 \in L^\perp$, entonces $(v_1 + v_2) * l = v_1 * l + v_2 * l = 0$, para todo $l \in L$, luego $v_1 + v_2 \in L^\perp$. Dados $\lambda \in \mathbb{C}$ y $v \in L^\perp$, entonces $(\lambda \cdot v) * l = \lambda \cdot (v * l) = 0$ para todo $l \in L$, luego $\lambda \cdot v \in L^\perp$.

19. Proposición : *Se cumplen las siguientes propiedades:*

1. $E^\perp = \{0\}$ *y* $\{0\}^\perp = E$.
2. *Si* $L \subseteq L'$ *entonces* $L^\perp \supseteq L'^\perp$.
3. $(L^\perp)^\perp = L$.
4. $E = L \oplus L^\perp$ *(*$E = L + L^\perp$ *y* $L \cap L^\perp = \{0\}$*).*
5. $\dim_{\mathbb{C}} E = \dim_{\mathbb{C}} L + \dim_{\mathbb{C}} L^\perp$.
6. $(L + L')^\perp = L^\perp \cap L'^\perp$ *y* $(L \cap L')^\perp = L^\perp + L'^\perp$.

Demostración. 1. Si $e \in E^{\perp}$ entonces $e * e = 0$, luego $e = 0$ y $E^{\perp} = 0$. Obviamente, $\{0\}^{\perp} = E$.

2. Obvio.

3. 4. y 5. Sea $\langle e_1, \dots, e_n \rangle$ una base ortonormal de L. Ampliando esta base, sea $\{e_1, \dots, e_n, e'_1, \dots, e'_m\}$ una base de E. Por el proceso de ortonormalización de Gram-Schmidt, obtenemos una base ortonormal

$$\{e_1, \dots, e_n, e''_1, \dots, e''_m\}$$

de E. Ahora ya es claro que $L^{\perp} = \langle e''_1, \dots, e''_m \rangle$ y $(L^{\perp})^{\perp} = \langle e_1, \dots, e_n \rangle = L$.

Además, $E = L \oplus L^{\perp}$ y tomando dimensiones

$$\dim_{\mathbb{C}} E = \dim_{\mathbb{C}} L + \dim_{\mathbb{C}} L^{\perp}.$$

6. Obviamente, $(L+L')^{\perp} = L^{\perp} \cap L'^{\perp}$. Veamos la otra igualdad. Sea $N = L^{\perp}$ y $N' = L'^{\perp}$. Entonces sabemos que $(N+N')^{\perp} = N^{\perp} \cap N'^{\perp}$. Tomando ortogonales, por 3. obtenemos

$$N + N' = (N^{\perp} \cap N'^{\perp})^{\perp} = (L \cap L')^{\perp}.$$

Luego, $L^{\perp} + L'^{\perp} = (L \cap L')^{\perp}$.

□

4.6.2. Isomorfismos lineales unitarios

20. Notación: Supondremos en esta subsección que E y E' son $\mathbb{C}$-espacios vectoriales euclídeos (o hermíticos) y que $T \colon E \to E'$ un isomorfismo $\mathbb{C}$-lineal.

21. Definición: Diremos que T es unitario si "conserva distancias", es decir,

$$d(T(e), T(v)) = d(e, v), \text{ para todo } e, v \in E.$$

22. Ejemplo: Sea B una base ortonormal de E. El isomorfismo lineal $E \to \mathbb{C}^n$, $e \mapsto e_B$ es unitario.

23. Proposición: *Las siguientes afirmaciones son equivalentes*

1. *T es unitario.*
2. *$\|T(e)\| = \|e\|$, para todo $e \in E$.*
3. *"T conserva el producto escalar": $T(e) * T(v) = e * v$, para todo $e, v \in E$.*

Demostración. Observemos que $d(T(e),T(v)) = \|T(e)-T(v)\| = \|T(e-v)\|$ y $d(e,v) = \|e-v\|$. Por tanto, T conserva distancias si y solo si $\|T(w)\| = \|w\|$, para todo $w \in E$, que equivale a decir que $T(w)*T(w) = w*w$.

Si T conserva distancias, entonces como

$$\begin{aligned} &T(e+v)*T(e+v) = T(e)*T(e)+T(v)*T(v)+T(e)*T(v)+T(v)*T(e) \\ &(e+v)*(e+v) = e*e+v*v+e*v+v*e \end{aligned}$$

entonces, $T(e)*T(v)+T(v)*T(e) \overset{*}{=} e*v+v*e$. Cambiando, e por ie, obtenemos

$$i(T(e)*T(v)) - i(T(v)*T(e)) = i(e*v) - i(v*e).$$

Luego, $T(e)*T(v)-T(v)*T(e) = e*v-v*e$. Por $\overset{*}{=}$, $T(e)*T(v) = e*v$.

Recíprocamente, si $T(e)*T(v) = e*v$, para todo $e,v \in E$, entonces $T(w)*T(w) = w*w$, para todo $w \in E$ y T conserva distancias.

□

24. Proposición: *Sea $\{e_1,\dots,e_n\}$ una base ortonormal de E. T es unitario si y solo si $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal.*

Demostración. $\Rightarrow$) $T(e_i)*T(e_i) = e_i*e_i = 1$, para todo i y $T(e_i)*T(e_j) = e_i*e_j = 0$, para todo $i \neq j$, luego $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal.

$\Leftarrow$) Dado $e = x_1e_1+\cdots+x_ne_n$, tenemos que $T(e) = x_1T(e_1)+\cdots+x_nT(e_n)$ y

$$||e|| = \sqrt{x_1^2+\cdots+x_n^2} = ||T(e)||.$$

□

25. Definición: Dada una matriz $(a_{ij}) \in M_{m\times n}(\mathbb{C})$ denotaremos $(a_{ij})^* := (\overline{a_{ij}})^t$. Diremos que una matriz cuadrada $(a_{ij}) \in M_n(\mathbb{C})$ es unitaria si $(a_{ij})^* \cdot (a_{ij}) = \mathrm{Id}$.

26. Corolario: *Sean $B = \{e_1,\dots,e_n\}$ y $B' = \{e'_1,\dots,e'_n\}$ bases ortonormales de E y E' respectivamente. Sea A la matriz asociada a T en la bases B y B'. Entonces, T es unitario $\iff A^*\cdot A = \mathrm{Id}$.*

Demostración. Las columnas de A son los vectores $T(e_i)^{B'}$. $A^*\cdot A = \mathrm{Id}$ si y solo si el producto de la fila i de A^* con la columna j de A es 0 si $i \neq j$ y es 1 si $i = j$. Es decir, conjugando, $T(e_j)_{B'} * T(e_i)_{B'} = 0$ si $i \neq j$ y $T(e_j)_{B'} * T(e_i)_{B'} = 1$ si $i = j$. Es decir, si y solo si $T(e_j)*T(e_i) = 0$ si $i \neq j$ y $T(e_j)*T(e_i) = 1$ si $i = j$. Luego, $A^*A = \mathrm{Id}$ si y solo si $\{T(e_1),\dots,T(e_n)\}$ es una base ortonormal. Por la proposición anterior hemos terminado.

□

4.6.3. Operadores hermíticos

27. Definición: Sea $(E,*)$ un espacio vectorial hermítico. Diremos que un operador lineal $T\colon E\to E$ es hermítico si $T(e)*v=e*T(v)$, para todo $e,v\in E$.

28. Proposición: *Sea $(E,*)$ un espacio vectorial hermítico y $B=\{e_1,\dots,e_n\}$ una base ortonormal de E. Un operador $\mathbb{C}$-lineal $T\colon E\to E$ es hermítico si y solo si la matriz A asociada a T en la base B es hermítica, es decir, $A^*=A$.*

Demostración. Sea $e_B=(x_1,\dots,x_n)$ e $v_B=(y_1,\dots,y_n)$. Entonces, $T(e)*v=e*T(v)$ si y solo si

$$(x_1,\dots,x_n)A^t(\overline{y_1},\dots,\overline{y_n})^t=(x_1,\dots,x_n)\bar{A}(\overline{y_1},\dots,\overline{y_n})^t$$

es decir, $A^t=\bar{A}$, o equivalentemente $A^*=A$. □

Observemos que si $A\in M_n(\mathbb{R})$, entonces A es hermítica si y solo si es simétrica.

29. Teorema de descomposición espectral : *Sea $(E,*)$ un $\mathbb{C}$-espacio vectorial hermítico de dimensión finita y sea $T\colon E\to E$ un operador hermítico. Entonces, todos los autovalores de T son reales y existe una base ortonormal de E donde T diagonaliza.*

Demostración. Sea $\alpha\in\mathbb{C}$ un autovalor de T y e un autovector de autovalor α. Entonces,

$$\alpha\cdot(e*e)=T(e)*e=e*T(e)=e*(\alpha e)=\bar{\alpha}\cdot(e*e),$$

luego $\alpha=\bar{\alpha}$, es decir, $\alpha\in\mathbb{R}$.

Veamos que existe una base ortonormal donde T diagonaliza. Si $\dim_{\mathbb{C}}E=1$ la afirmación es obvia. Supongamos que la hemos probado para todo operador hermítico cuando $\dim_{\mathbb{C}}E=1,\dots,n-1$. Tenemos que probarla cuando $\dim_{\mathbb{C}}E=n$.

Sea $e_1\in E$ un autovector, de módulo 1, de autovalor α. Sea $E'=\langle e_1\rangle^\perp$. Observemos que para todo $e'\in E'$ se cumple que $T(e')\in E'$, ya que

$$T(e')*e_1=e'*T(e_1)=e'*(\alpha e_1)=\bar{\alpha}(e'*e_1)=0.$$

El operador $T'\colon E'\to E'$, $T'(e'):=T(e')$ es obviamente hermítico. Como $\dim_{\mathbb{C}}E'=n-1$ entonces existe una base ortonormal $\{e_2,\dots,e_n\}$ de E' en la que T' diagonaliza. Obviamente, $\{e_1,e_2,\dots,e_n\}$ es una base ortonormal de E y T diagonaliza en ella.

□

30. Corolario : *Sea $A\in M_n(\mathbb{R})$ una matriz cuadrada simétrica (luego hermítica). Existe una matriz ortogonal $P\in M_n(\mathbb{R})$ ($P^t=P^{-1}$) y una matriz diagonal $D\in M_n(\mathbb{R})$, tales que*

$$A=PDP^t.$$

Las columnas de P son una base ortonormal de autovectores de A y los coeficientes de la diagonal de D son los autovalores de A.

Demostración. A es una matriz hermítica, luego todos sus autovalores son reales. La demostración del teorema anterior, muestra que si $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es el operador lineal de matriz A en la base usual U de $\mathbb{R}^n$, existe una base $B = \{e_1, \dots, e_n\}$ de autovectores de T ortonormal. Sean $P = \mathrm{Id}_B^U$ y $D = T_B^B$, entonces P es ortogonal (luego, $P^{-1} = P^t$), D diagonal y

$$A = T_U^U = \mathrm{Id}_B^U \cdot T_B^B \cdot \mathrm{Id}_U^B = P \cdot D \cdot P^t.$$

□

31. Ejemplo : Calculemos una base ortonormal donde el operador lineal hermítico $T\colon \mathbb{C}^2 \to \mathbb{C}^2$ de matriz $\begin{pmatrix} 3 & 1-2i \\ 1+2i & 7 \end{pmatrix}$ en la base usual, diagonaliza. Calculemos $c_T(x) = (x-2)(x-8)$, luego los autovalores son $2, 8$. Calculemos una base de autovectores. $\mathrm{Ker}(T - 2\,\mathrm{Id}) = \langle(-1+2i, 1)\rangle$ y $\mathrm{Ker}(T - 8\,\mathrm{Id}) = \langle(1, 1+2i)\rangle$. Entonces,

$$\{(\frac{-1+2i}{\sqrt{6}}, \frac{1}{\sqrt{6}}), (\frac{1}{\sqrt{6}}, \frac{1+2i}{\sqrt{6}})\}$$

es una base ortonormal donde T diagonaliza.

32. Ejemplo : Sea $X = (a_1, b_1) \times (a_2, b_2) \times (a_3, b_3) \subset \mathbb{R}^3$ y B las seis caras del borde del cubo X. Consideremos el espacio vectorial euclídeo complejo

$$F := \{\text{Funciones } f : X \to \mathbb{C} \text{ infinit. derivables, tales que } \lim_{(x,y,z)\to b} f(x,y,z) = 0, \forall b \in B\},$$

con el producto escalar

$$\langle f|g\rangle := \int_X f(x,y,z)\overline{g(x,y,z)}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_{a_3}^{b_3}\int_{a_2}^{b_3}\int_{a_1}^{b_1} f(x,y,z)\overline{g(x,y,z)}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

Por tanto, $\|f\|^2 = \int_X |f|^2\,\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

Cada función continua compleja $\psi(x,y,z) \in F$ describe el estado de una partícula cuántica (confinada en X), entendiendo que la probabilidad de encontrar a la partícula en cierta región $\Omega \subseteq X$ es

$$(*) \qquad \frac{\displaystyle\int_\Omega |\psi(x,y,z)|^2\,\mathrm{d}x\mathrm{d}y\mathrm{d}z}{\displaystyle\int_X |\psi(x,y,z)|^2\,\mathrm{d}x\mathrm{d}y\mathrm{d}z}.$$

Las funciones $\psi(x,y,z)$ se denominan funciones de onda o de estado (del sistema cuántico). Una función $\psi(x,y,z)$ y $\lambda \cdot \psi(x,y,z)$, $0 \neq \lambda \in \mathbb{C}$, describen el mismo comportamiento, como puede comprobarse en $(*)$. Las funciones de onda $\psi(x,y,z)$ van evolucionando a lo largo del tiempo t, escribamos la evolución $\psi_t(x,y,z) = \Psi(x,y,z,t)$. La evolución viene determinada por la ecuación de Schrödinger:

$$\frac{\hbar}{i}\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}) + V\cdot\Psi,$$

donde V es la energía potencial de la partícula, $\hbar = \frac{h}{2\pi}$ y h es la constante de Planck, m la masa de la partícula. De la ecuación de Schrödinger se deduce que existe una familia uniparamétrica de operadores unitarios $U_t\colon F \to F$, con $t \in \mathbb{R}$, de modo que $U_{t+s} = U_t \circ U_s$ y $U_t(\Psi(x,y,z,s)) = \Psi(x,y,z,t+s)$.

Toda magnitud medible de una función de onda está descrita por un operador hermítico $T\colon F \to F$. Los posibles valores de la medición de esa magnitud son los autovalores α_n de T, que siempre son números reales. Si $\psi(x,y,z)$ es un autovector v de autovalor α, entonces el valor de la medición es α con toda seguridad; pero en general si $\psi = v_1 + \ldots + v_r$, donde los v_i son autovectores de autovalores distintos $\alpha_1,\ldots,\alpha_r$, entonces los posibles valores de la medición son $\alpha_1,\ldots,\alpha_r$, y la probabilidad de que sea α_n es

$$\mathscr{P}(\alpha_n) = \frac{\|v_n\|^2}{\|\psi\|^2} = \frac{\langle v_n|v_n\rangle}{\langle\psi|\psi\rangle}$$

y, por el teorema de Pitágoras, $\|\psi\|^2 = \|v_1\|^2 + \ldots + \|v_r\|^2$, $\mathscr{P}(\alpha_1) + \ldots + \mathscr{P}(\alpha_r) = 1$. En particular, el valor medio de las observaciones en ese estado es

$$\sum_n \alpha_n \mathscr{P}(\alpha_n) = \frac{\langle \sum_n \alpha_n v_n | \sum_n v_n\rangle}{\langle\psi|\psi\rangle} = \frac{\langle T(\psi)|\psi\rangle}{\langle\psi|\psi\rangle}.$$

Si el resultado de la medición es α_n, ψ pasa a ser $v_n = P_n(\psi)$, donde $P_n\colon F \to F$ es la proyección ortogonal sobre el subespacio $\langle v_n\rangle$.

Ejemplos: el momento en la dirección del eje x viene descrito por el operador hermítico $\frac{\hbar}{i}\frac{\partial}{\partial x}\colon F \to F$, y la energía cinética por $-\frac{\hbar^2}{2m}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})\colon F \to F$. Se dice que una función de onda de energía potencial V es estacionaria, si es un autovector del operador hermítico

$$-\frac{\hbar^2}{2m}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) + V\cdot\mathrm{Id}$$

y el autovalor del autovector se dice que es la energía total E de la función de onda. Estamos describiendo la denominada ecuación de Schrödinger independiente del tiempo para funciones de onda estacionarias

$$-\frac{\hbar^2}{2m}(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}) + V\psi = E\psi.$$

4.7. Métricas

1. Definición: Sea E un $\mathbb{R}$-espacio vectorial. Dar un producto bilineal $\star$ en E es asignar a cada par de vectores $e, v \in E$ un número real $e \star v \in \mathbb{R}$ de modo que

$$(e_1 + \lambda e_2) \star v = e_1 \star v + \lambda \cdot (e_2 \star v) \quad \text{y} \quad e \star (v_1 + \mu v_2) = e \star v_1 + \mu \cdot (e \star v_2)$$

(para todo $e, e_1, e_2, v, v_1, v_2 \in E$ y todo $\lambda, \mu \in \mathbb{R}$).

2. Ejemplo: Los productos escalares son productos bilineales.

3. Matriz asociada a un producto bilineal: Sea $B = \{e_1, \ldots, e_n\}$ una base de E. Diremos que $A = (a_{ij}) \in M_n(\mathbb{R})$, donde $a_{ij} := e_i \star e_j$, es la matriz asociada a $\star$ (en la base B). Observemos que si $e_B = (x_1, \ldots, x_n)$ y $v_B = (y_1, \ldots, y_n)$ entonces

$$\begin{aligned} e \star v &= (x_1 e_1 + \cdots + x_n e_n) \star (y_1 e_1 + \cdots + y_n e_n) = \sum_{ij} x_i y_j e_i \star e_j = \sum_{ij} x_i y_j a_{ij} \\ &= (\sum_i x_i a_{i1}, \ldots, \sum_i x_i a_{in}) \cdot (y_1, \ldots, y_n)^t) = (x_1, \ldots, x_n) \cdot A \cdot (y_1, \ldots, y_n)^t. \end{aligned}$$

Recíprocamente, dada una matriz $A \in M_n(\mathbb{R})$ podemos definir el siguiente "producto bilineal" $\star$ en E: dados $e_B = (x_1, \ldots, x_n)$ y $e'_B = (x'_1, \ldots, x'_n)$ definimos

$$e \star e' := (x_1, \ldots, x_n) \cdot A \cdot (x_1, \ldots, x_n)^t.$$

que es bilineal.

Sea $(E, \cdot)$ un espacio vectorial euclídeo. $B = \{e_1, \ldots, e_n\}$ es una base ortonormal de E si y solo si la matriz asociada a $\cdot$ en la base B es la matriz identidad.

4. Cambio de base: Sea $B' = \{e'_1, \ldots, e'_n\}$ otra base de E, $e'_i = \sum_j p_{ji} \cdot e_j$ y $P = (p_{ij})$ la matriz de cambio de base de B' a B. Calculemos la matriz A' de $\star$ en la base B': Dados $e_{B'} = (x'_1, \ldots, x'_n)$ y $v_{B'} = (y'_1, \ldots, y'_n)$ tenemos que $e_B = x = x' \cdot P^t$ y $v_B = y = y' \cdot P^t$, luego

$$e \star v = x \cdot A \cdot y^t = (x' \cdot P^t) \cdot A \cdot (P \cdot y'^t).$$

Luego, $A' = P^t A P$.

5. Definición: Diremos que un producto bilineal $\star$ es una métrica cuando $e \star e' = e' \star e$ para todo $e, e' \in E$.

Si $A = (a_{ij})$ es la matriz asociada a un producto bilineal $\star$ en una base $\{e_i\}$, y $\star$ es una métrica, entonces

$$a_{ij} = e_i \star e_j = e_j \star e_i = a_{ji}$$

es decir, la matriz (a_{ij}) es simétrica. Recíprocamente, si (a_{ij}) es simétrica entonces $\star$ es una métrica.

6. Teorema de Sylvester: *Sea E un $\mathbb{R}$-espacio vectorial y $\star$ una métrica en E. Entonces existe una base en la que la matriz asociada a $\star$ es igual a*

$$\begin{pmatrix} \pm1 & & & & & \\ & \ddots & & & & \\ & & \pm1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Demostración. Basta demostrar que existe una base ortogonal $\{e_i\}$, porque si $e_i \star e_i \neq 0$, tomaríamos $\frac{e_i}{\sqrt{|e_i \star e_i|}}$ en vez de e_i.

Si $e \star e = 0$ para todo $e \in E$, entonces $0 = (e+e') \star (e+e') = 2 \cdot e \star e'$ y $e \star e' = 0$ para todo $e, e' \in E$, luego la matriz asociada e $\star$ en cualquier base es la matriz cero.

Podemos suponer que existe $e_1 \in E$ tal que $e_1 \star e_1 \neq 0$. La dimensión de $\langle e_1 \rangle^\perp := \{e' \in E \colon e' \star e_1 = 0\}$ es $\dim E - 1$, porque $\langle e_1 \rangle^\perp$ es igual al núcleo de la aplicación lineal $E \to \mathbb{R}$, $e' \mapsto e' \star e_1$ (cuya imagen es $\mathbb{R}$). Entonces, $E = \langle e_1 \rangle \oplus \langle e_1 \rangle^\perp$. Si $e \star e = 0$ para todo $e \in \langle e_1 \rangle^\perp$, entonces de nuevo $e \star e' = 0$ para todo $e, e' \in \langle e_1 \rangle^\perp$. Si $\{e_2, \dots, e_n\}$ es una base de $\langle e_1 \rangle^\perp$, entonces $\{e_1, \dots, e_n\}$ es una base ortogonal de E. Podemos suponer que existe $e_2 \in \langle e_1 \rangle^\perp$ tal que $e_2 \star e_2 \neq 0$. De nuevo, la dimensión de $\langle e_1, e_2 \rangle^\perp = \{e' \in \langle e_1 \rangle^\perp \colon e' \star e_2 = 0\}$ es $\dim \langle e_1, e_2 \rangle^\perp - 1$ y

$$E = \langle e_1 \rangle \oplus \langle e_1 \rangle^\perp = \langle e_1 \rangle \oplus \langle e_2 \rangle \oplus \langle e_1, e_2 \rangle^\perp.$$

Procediendo del mismo modo sucesivamente, obtendremos una base ortogonal de E para la métrica $\star$.

□

Reordenando la base, podemos suponer en el teorema anterior que

$$\star \equiv \begin{pmatrix} 1 & & & & & & & & \\ & \overset{p}{\ddots} & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \overset{q}{\ddots} & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

Veamos que el número p no depende de la base escogida: para todo $e \in E' = \langle e_1, \dots, e_p \rangle$, no nulo, se tiene que $e \star e > 0$; del mismo modo para todo $e \in E'' = \langle e_{p+1}, \dots e_n \rangle$ no nulo se cumple que $e \star e \leq 0$. Supongamos que existe un subespacio vectorial V tal que para todo $v \in V$ no nulo $v \star v > 0$. Obviamente $V \cap E'' = 0$, luego $\dim V \leq \dim E - \dim E'' = \dim E' = p$. En conclusión, p es la dimensión máxima de los subespacios $E' \subset E$ que son euclídeos con $\star$. Si definimos $e \star' v = -e \star v$, tendremos que q no depende tampoco de la base escogida. Se dice que $\star$ tiene signatura (p, q).

Sea A una matriz cuadrada. Llamaremos menor principal de orden r de la matriz A al determinante de la matriz cuadrada formada por los coeficientes de la matriz A que están en las r primeras filas y en las r primeras columnas.

7. Proposición : *Sea $\star$ una métrica y A la matriz asociada a $\star$ en una base $B = \{e_1, \dots, e_n\}$. Entonces, $\star$ es un producto escalar $\iff$ todos los menores principales de A son positivos.*

Demostración. $\Rightarrow$) Sabemos que existe una base ortonormal B' para $\star$. Si P es la matriz de cambio de base de B a B', entonces

$$A = P^t \cdot \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \cdot P$$

y $|A| = |P|^2 > 0$. Por tanto, el determinante de la matriz asociada a cualquier producto escalar es positivo.

El producto escalar $\star$ restringido a $\langle e_1, \dots, e_i \rangle$ es un producto escalar y su matriz asociada es la submatriz de $A = (a_{rs})_{r,s \leq i}$. Por tanto, $|(a_{rs})_{r,s \leq i}| > 0$. Es decir, todos los menores principales de A son positivos.

$\Leftarrow$) Procedemos por inducción sobre n. Si $n = 1$ entonces $A = (\lambda)$ y $\lambda = |A| > 0$, luego $(\mu \cdot e_1) \star (\mu \cdot e_1) = \mu^2 \cdot \lambda > 0$, para todo $\mu \neq 0$ y $\star$ es un producto escalar..

Supongamos que el teorema es cierto para matrices de orden $\leq n - 1$.

La matriz asociada a la restricción de $\star$ a $\langle e_1, \dots, e_{n-1} \rangle$ es $(a_{rs})_{r,s \leq n-1}$. Todos los menores principales de $(a_{rs})_{r,s \leq n-1}$ son positivos. Por hipótesis de inducción la restricción de $\star$ a $\langle e_1, \dots, e_{n-1} \rangle$ es un producto escalar. Por tanto, existe una base ortonormal $\{e'_1, \dots, e'_{n-1}\}$ de la restricción de $\star$ a $\langle e_1, \dots, e_{n-1} \rangle$. Sea $e'_n = e_n - \sum_{i=1}^{n-1} (e_n \star e'_i) \cdot e'_i$. La matriz de $\star$ en la base $B' = \{e'_1, \dots, e'_n\}$ es

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & e'_n \star e'_n \end{pmatrix}$$

Si Q es la matriz de cambio de base de B a B' tenemos que

$$Q^t \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & e'_n \star e'_n \end{pmatrix} Q = A$$

y $0 < |A| = e'_n \star e'_n \cdot |Q|^2$. Luego, $e'_n \star e'_n > 0$ y $\star$ es un producto escalar.

□

8. Sea $(E,\cdot)$ un $\mathbb{R}$-espacio vectorial euclídeo. Dado un operador lineal autoadjunto $T\colon E \to E$ (es decir, $T(e)\cdot v = e\cdot T(v)$) entonces la métrica $\star$ definida por $e \star v := e \cdot T(v)$ es una métrica: la matriz (a_{ij}) asociada a $\star$ en una base ortonormal B de E coincide con T_B porque

$$e_B \cdot (a_{ij}) \cdot v^B = e \star v = e \cdot T(v) = e_B \cdot T_B \cdot v^B.$$

Recíprocamente, dada una métrica $\star$ si $T\colon E \to E$ es el operador lineal determinado por $e \star v = e \cdot T(v)$ (la matriz asociada a T en una base ortonormal de E es igual a la matriz asociada a $\star$ en dicha base), entonces

$$T(e)\cdot v = v \cdot T(e) = v \star e = e \star v = e \cdot T(v)$$

y T es autoadjunto.

Reformulemos el teorema de representación espectral para cuando tengamos una pareja de métricas (o al menos una de ellas sea un producto escalar).

9. Teorema: *Sea $(E,\cdot)$ un espacio vectorial euclídeo y $\star$ una métrica. Existe una base ortonormal para $\cdot$ que es ortogonal para $\star$.*

Demostración. Sea $T\colon E \to E$ la aplicación lineal que cumpla que $e \star v = e \cdot T(v)$. T es un operador lineal autoadjunto. Por el teorema de descomposición espectral existe una base ortonormal $\{e'_1,\dots,e'_n\}$ de E formada por autovectores de T. Observemos que $e'_i \star e'_j = e'_i \cdot T(e'_j) = e'_i \cdot \lambda_j e'_j = 0$, para todo $i \neq j$, es decir, esta base es ortogonal para $\star$.

□

10. Corolario: *Sea $S\colon \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal. Existen bases ortonormales $B = \{e_1,\dots,e_n\}$ de $\mathbb{R}^n$ y $B' = \{v_1,\dots,v_m\}$ de $\mathbb{R}^m$ de modo que $S(e_1) = \lambda_1 \cdot v_1$, …, $S(e_r) = \lambda_r v_r$ y $S(e_{r+1}) = \cdots = S(e_n) = 0$.*[3]

[3]Si U y U' son las bases usuales de $\mathbb{R}^n$ y $\mathbb{R}^m$, tenemos que $S_U^{U'} = \mathrm{Id}_{B'}^{U'} \cdot S_B^{B'} \cdot \mathrm{Id}_U^B$, donde Id_U^B y $\mathrm{Id}_{B'}^{U'}$ son matrices ortogonales y $S_B^{B'}$ es una matriz "diagonal". A este producto de matrices se le denomina la descomposición en valores singulares de $S_U^{U'}$.

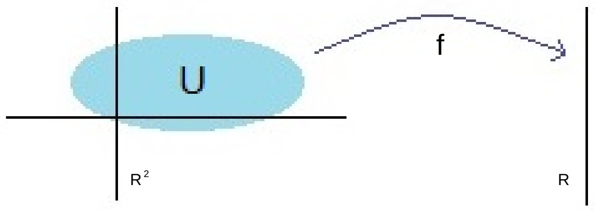
Demostración. Sea $\star$ la métrica en $\mathbb{R}^n$ definida por $e \star v = S(e) \cdot S(v)$. Existe una base ortonormal $\{e_1, \dots, e_n\}$ de $(\mathbb{R}^n, \cdot)$ ortogonal para $\star$. Entonces, $S(e_i) \cdot S(e_j) = e_i \star e_j = 0$, para $i \neq j$. Reordenando la base podemos suponer $\{S(e_1), \dots, S(e_r)\}$ es una base de $\mathrm{Im} S$. Para $i > r$, $S(e_i) = 0$ porque es ortogonal a $\mathrm{Im} S$ y pertenece a $\mathrm{Im} S$. Sea $\{S(e_1), \dots, S(e_r), v'_{r+1}, \dots, v'_m\}$ una base ortogonal de $(\mathbb{R}^m, \cdot)$. Dividiendo por los módulos, obtenemos la base ortonormal $\{v_1, \dots, v_m\}$ buscada de $\mathbb{R}^m$.

□

4.8. Aplicaciones

4.8.1. Máximos y mínimos de funciones diferenciables

Dado $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ y $r \in \mathbb{R}$, llamaremos bola abierta de radio r centrada en a, que denotaremos $B_r(a)$, a $B_r(a) := \{x \in \mathbb{R}^n : \|x - a\| \leq r\}$. Diremos que un subconjunto $U \subset \mathbb{R}^n$ es un abierto si es unión de bolas abiertas.



Sea $U \subset \mathbb{R}^n$ un abierto. Diremos que una función $f : U \to \mathbb{R}$ es infinitamente derivable si podemos derivarla parcialmente todas las veces que queramos.
Sea $f : U \to \mathbb{R}$ infinitamente derivable. Un resultado fundamental del Análisis Funcional nos dice que si $f(a) = 0$ entonces existen funciones infinitamente derivables $h_i : U \to \mathbb{R}$ tales que

$$f = h_1 \cdot (x_1 - a_1) + \cdots + h_n \cdot (x_n - a_n)$$

Por tanto, si $g : U \to \mathbb{R}$ es infinitamente derivable tenemos que $f = g - g(a)$ cumple que $f(a) = 0$, luego $f = h_1 \cdot (x_1 - a_1) + \cdots + h_n \cdot (x_n - a_m)$ y

$$g = g(a) + h_1 \cdot (x_1 - a_1) + \cdots + h_n \cdot (x_n - a_n).$$

Observemos que $\frac{\partial g}{\partial x_i}(a) = h_i(a)$. Si $h_i(a) = \frac{\partial g}{\partial x_i}(a) > 0$, como h_i es continua, existe una bola abierta "pequeña" centrada en a, $B_r(a)$ de modo que para todo $x \in B_r(a)$ se cumple que $h_i(x) > 0$. Por tanto, dado $x' = (a_1, \dots, a_i + \epsilon. \dots, a_n)$, con $|\epsilon| < r$ tendremos que $g(x') = g(a) + h_i(x') \cdot \epsilon$, que es mayor que $g(a)$ si $\epsilon > 0$ y menor que $g(a)$ si $\epsilon < 0$. De modo equivalente razonaremos si $h_i(a) < 0$. En conclusión: $g(a)$ no es el máximo ni el mínimo que alcanza g en $B_r(a)$ (para cualquier r muy pequeño) si $\frac{\partial g}{\partial x_i}(a) \neq 0$, algún i.

1. Proposición: *Una condición necesaria para que $g\colon U \to \mathbb{R}$ alcance un máximo o un mínimo local en $a \in U$ es que $\frac{\partial g}{\partial x_1}(a) = \cdots = \frac{\partial g}{\partial x_n}(a) = 0$.*

Hemos probado que $g = g(a) + h_1(x)\cdot(x_1 - a_1) + \cdots + h_n(x)\cdot(x_n - a_n)$. Igualmente, $h_i(x) = h_i(a) + \sum_j h_{ij}(x)(x_j - a_j)$. Por tanto,

$$g(x) = g(a) + \sum_i \frac{\partial g}{\partial x_i}(a)(x_i - a_i) + \sum_{ij} h_{ij}(x)(x_i - a_i)(x_j - a_j).$$

Sea $H_{ij}(x) = h_{ij}(x) + h_{ji}(x)$, entonces

$$g(x) = g(a) + \left(\frac{\partial g}{\partial x_i}(a)\right)\cdot(x-a)^t + \frac{1}{2}(x-a)\cdot\big(H_{ij}(x)\big)\cdot(x-a)^t.$$

Observemos que $\frac{\partial^2 g}{\partial x_i \partial x_j}(a) = H_{ij}(a)$. Si $\left(\frac{\partial^2 g}{\partial x_i \partial x_j}(a)\right)$ es una matriz simétrica definida positiva, entonces $\big(H_{ij}(x)\big)$ es una matriz simétrica definida positiva para todo $x \in B_r(a)$, con r suficientemente pequeño. No es difícil probar el siguiente enunciado.

2. Proposición: *Sean $U \subset \mathbb{R}^n$ un abierto, $a \in U$ y $g\colon U \to \mathbb{R}$ una función infinitamente derivable. Supongamos que $\frac{\partial g}{\partial x_1}(a) = \cdots = \frac{\partial g}{\partial x_n}(a) = 0$ y que $\det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}(a)\right) \neq 0$. Entonces,*

1. *a es un mínimo local de g si y solo si $\left(\frac{\partial^2 g}{\partial x_i \partial x_j}(a)\right)$ es definida positiva.*
2. *a es un máximo local de g si y solo si $\left(\frac{\partial^2 g}{\partial x_i \partial x_j}(a)\right)$ es definida negativa.*

Dado $a \in \mathbb{R}$, diremos que $V(a) = 0$. Dados $a, b \in \mathbb{R}$ no nulos diremos que $V(a,b) = 0$ si $\operatorname{sign}(a) = \operatorname{sign}(b)$ y $V(a,b) = 1$ si $\operatorname{sign}(a) = -\operatorname{sign}(b)$. Sean $a_1, \ldots, a_n \in \mathbb{R}$, no nulos definimos

$$V(a_1,\ldots,a_n) = V(a_1,a_2) + V(a_2,a_3) + \cdots + V(a_{n-1},a_n),$$

En general, dados $a_1,\ldots,a_n \in \mathbb{R}$, $V(a_1,\ldots,a_n)$ se define como lo hemos hecho anteriormente pero quitando de la lista $a_1,\ldots,a_n$ los que sean nulos. Por ejemplo,

$$V(1,-1,2,2,0,-1) = V(1,-1,2,2,-1) = 3.$$

3. Regla de Descartes: *Sea $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$ un polinomio de grado n, tal que el número de raíces reales contando multiplicidades es n. Entonces, el número de raíces de $p(x)$ estrictamente mayores que cero, contando multiplicidades, es $V(a_0, a_1, \ldots, a_n)$.*

Demostración. Procedemos por inducción sobre n. Para $n = 1$ es claro. Suponemos que la regla de Descartes es válida para polinomios de grado $n-1$. Si $a_0 = 0$, entonces $p(x) = x\cdot(a_1+\cdots+a_nx^{n-1}) = x\cdot q(x)$ y las raíces reales positivas de $p(x)$ son las de $q(x)$ y por hipótesis de inducción terminamos. Podemos suponer que $a_0 > 0$ multiplicando $p(x)$ por -1 si es necesario. Sean $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_n$ las raíces de $p(x)$, y supongamos que $\alpha_i < 0 < \alpha_{i+1}$, es decir que el número de raíces estrictamente positivas de $p(x)$ es $n-i$. Por el teorema de Rolle, las raíces de la derivada $p(x)'$ son $\{\beta_i\}_{1\le i\le n-1}$ con $\alpha_i \le \beta_i \le \alpha_{i+1}$.

Si $a_1 > 0$, entonces $p(x)$ es creciente a la derecha de 0, como $p(0) > 0$ y $p(\alpha_{i+1}) = 0$ entonces $0 < \beta_i < \alpha_{i+1}$. El número de raíces estrictamente positivas (contando multiplicidades) de $p(x)'$ es $n-i$ y por hipótesis de inducción $n-i = V(a_1, 2a_2,\dots,na_n) = V(a_1,a_2,\dots,a_n) = V(a_0,a_1,a_2,\dots,a_n)$.

Si $a_1 < 0$, entonces $p(x)$ es decreciente a la izquierda de 0, como $p(0) > 0$ y $p(\alpha_i) = 0$ entonces $\alpha_i < \beta_i < 0$. El número de raíces estrictamente positivas (contando multiplicidades) de $p(x)'$ son $n-i-1$ y por hipótesis de inducción $n-i-1 = V(a_1,2a_2,\dots,na_n)$ y $n-i = V(a_0,a_1,a_2,\dots,a_n)$.

Si $a_1 = 0$, entonces $\beta_i = 0$. Si $a_2 > 0$ entonces $p(0) > 0$ sería un mínimo local y como $p(0) > 0$ y $p(\alpha_i) = 0 = p(\alpha_{i+1})$, por el teorema de Weierstrass $p(x)$ tomaría valores máximos en $(\alpha_i, 0)$ y $(0, \alpha_{i+1})$, luego habría otras dos raíces de $p(x)'$ entre α_i y α_{i+1}. Si $a_2 = 0$ entonces 0 es una raíz múltiple de $p(x)'$. Luego, $a_2 < 0$. El número de raíces estrictamente positivas (contando multiplicidades) de $p(x)'$ son $n-i-1$ y por hipótesis de inducción $n-i-1 = V(a_1,2a_2,\dots,na_n)$ y $n-i = V(a_0,a_1,a_2,\dots,a_n)$.

□

4. Proposición: *Sea $\star$ una métrica, $A \in M_n(\mathbb{R})$ la matriz asociada a $\star$ en una base y $c_A(x) = \pm(x^n + a_{n-1}x^{n-1}+\cdots+a_0)$ el polinomio característico de A. Entonces, $\star$ es un producto escalar si y solo si $V(a_0,a_1,\dots,a_{n-1},1) = n$.*

Demostración. Por el teorema de descomposición espectral, existe una matriz ortogonal P tal que $P^{-1}AP = P^tAP = D$ es diagonal. Entonces, $\star$ es un producto escalar si y solo si los coeficientes de la diagonal de D son positivos, es decir, todas las raíces del polinomio característico de A, que coincide con el de D, son positivas. □

4.8.2. Formas cuadráticas

5. Definición: Una forma cuadrática en $\mathbb{R}^n$ es un polinomio homogéneo de grado dos en n-variables,

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{q} & \mathbb{R} \\ (x_1,\dots,x_n) & \mapsto & q(x_1,\dots,x_n) = \sum_{i\le j} a_{ij}x_ix_j\,. \end{array}$$

6. Ejemplo: La forma cuadrática en $\mathbb{R}^3$, $q(x,y,z)=x^2+2y^2-z^2+4xy+5yz$.

Dada una matriz simétrica $A=(a_{ij})$ tenemos la forma cuadrática

$$q(x_1,\dots,x_n):=(x_1,\dots,x_n)\cdot A\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\sum_{ij}a_{ij}x_ix_j=\sum_i a_{ii}x_i^2+\sum_{i<j}2a_{ij}x_ix_j.$$

Recíprocamente, dada una forma cuadrática, $q(x_1,\dots,x_n)=\sum_{i\leq j}b_{ij}x_ix_j$ definamos la matriz simétrica $A=(a_{ij})$, donde $a_{ij}=a_{ji}:=b_{ij}/2$ si $i<j$ y $a_{ii}:=b_i$. Es fácil comprobar que

$$(x_1,\dots,x_n)\cdot A\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\sum_{i\leq j}b_{ij}x_ix_j.$$

Sea $q(x_1,\dots,x_n)=(x_1,\dots,x_n)\cdot A\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ una forma cuadrática. Por el corolario 4.6.30 sabemos que existe una matriz ortogonal B (las columnas de B es una base ortonormal formada por autovectores de A, de autovalores $b_1,\dots,b_n$) tal que

$$A=B\cdot D\cdot B^{-1}=B\cdot D\cdot B^t,\quad \text{siendo } D=\begin{pmatrix}b_1&\cdots&0\\ \vdots&\ddots&\vdots\\ 0&\cdots&b_n\end{pmatrix}.$$

Si consideramos la transformación ortogonal de coordenadas

$$\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=B^t\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\qquad \text{(luego } \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=B\cdot\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\text{)},$$

entonces,

$$q(x_1,\dots,x_n)=(x_1,\dots,x_n)\cdot A\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=(x_1,\dots,x_n)\cdot B\cdot D\cdot B^t\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=(y_1,\dots,y_n)\cdot D\cdot\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$
$$=\sum_i b_i {y_i}^2.$$

4.8.3. Clasificación afín euclídea de cuádricas

Consideremos un polinomio de grado dos en las variables $x_1,\dots,x_n$:

$$p(x_1,\dots,x_n)=\sum_{i\le j} a_{ij}x_i\cdot x_j+\sum_i a_ix_i+a.$$

El conjunto de las soluciones de la ecuación

$$\sum_{i\le j} a_{ij}x_i\cdot x_j+\sum_i a_ix_i+a=0$$

se dice que es una cuádrica de $\mathbb{R}^n$. Queremos encontrar un sistema de coordenadas en el que la ecuación de la cuádrica sea lo más sencilla posible, con otras palabras queremos calcular la "ecuación reducida" de la cuádrica. En los cambios de coordenadas que realicemos, sólo admitiremos translaciones y transformaciones lineales que sean isometrías (conservan las distancias).

En la subsección anterior, hemos probado que por una transformación ortogonal de coordenadas (b_{ij}),

$$\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix}=(b_{ij})\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}$$

podemos conseguir que $\sum_{i\le j} a_{ij}x_ix_j=b_1y_1^2+\cdots+b_ry_r^2$ (con $b_i\neq 0$). Por tanto,

$$p(x_1,\dots,x_n)=b_1y_1^2+\cdots+b_ry_r^2+\sum_i b'_iy_i+b'.$$

Por la traslación $y_i=z_i-\frac{b'_i}{2\cdot b_i}$, para todo $i\le r$, $y_j=z_j$ para todo $j>r$, tenemos que

$$p(x_1,\dots,x_n)=b_1z_1^2+\cdots+b_rz_r^2+b_{r+1}z_{r+1}+\cdots+b_nz_n+d.$$

- Si $b_{r+1}=b_{r+2}=\cdots=b_n=0$, entonces

$$\boxed{p(x_1,\dots,x_n)=b_1z_1^2+\cdots+b_rz_r^2+d}$$

(aquí hay que distinguir dos casos $d=0$ y $d\neq 0$).

- Si $b_{r+i}\neq 0$, para algún $i>0$, por otra transformación ortogonal de coordenadas

$$u_1=z_1,\dots,u_r=z_r,u_{r+1}=\frac{b_{r+1}z_{r+1}+\cdots+b_nz_n}{\sqrt{\sum_i b_{r+i}^2}},\text{ etc.}$$

obtenemos

$$p(x_1,\dots,x_n)=b_1u_1^2+\cdots+b_ru_r^2+c\cdot u_{r+1}+d$$

(con $c\neq 0$). Con la traslación $u_{r+1}=v_{r+1}-\frac{d}{c}$, $u_i=v_i$ para todo $i\neq r+1$ tenemos que

$$\boxed{p(x_1,\dots,x_n)=b_1v_1^2+\cdots+b_rv_r^2+c\cdot v_{r+1}}$$

Vamos a considerar $p(x_1,\dots,x_n)$ equivalente a $\lambda\cdot p(x_1,\dots,x_n)$, para todo $\lambda\in\mathbb{R}$, porque queremos calcular las "ecuaciones reducidas de la cuádrica" $p(x_1,\dots,x_n)=0$, después de una transformación afín ortogonal de coordenadas. Entonces podemos suponer que c o d son iguales a -1. Reordenando las r primeras variables, si es necesario podemos suponer que $b_i>0$, para $i\leq s$ y $b_j<0$ para $s<j\leq r$. En conclusión, por una transformación afín ortogonal de coordenadas

$$\{p(x_1,\dots,x_n)=0\}=\left\{\begin{array}{l} 1.\ \ b_1v_1^2+\cdots+b_sv_s^2-b_{s+1}\cdot v_{s+1}^2-\cdots-b_r\cdot v_r^2-1=0.\\ \text{ó}\\ 2.\ \ b_1v_1^2+\cdots+b_sv_s^2-b_{s+1}\cdot v_{s+1}^2-\cdots-b_r\cdot v_r^2=0.\\ \text{ó}\\ 3.\ \ b_1v_1^2+\cdots+b_sv_s^2-b_{s+1}\cdot v_{s+1}^2-\cdots-b_r\cdot v_r^2-v_{r+1}=0.\end{array}\right.$$

(con $b_1,\dots,b_r>0$).

7. Observación: El caso 3. $p(x_1,\dots,x_n)=b_1v_1^2+\cdots+b_sv_s^2-b_{s+1}\cdot v_{s+1}^2-\cdots-b_r\cdot v_r^2-v_{r+1}$, $b_i>0$ es equivalente (multiplicando $p(x)$ por -1, haciendo el cambiando v_{r+1} por $-v_{r+1}$ y reordenando las r primeras coordenadas) a $b_1v_1^2+\cdots+b_{r-s}v_{r-s}^2-b_{r-s+1}\cdot v_{r-s+1}^2-\cdots-b_r\cdot v_r^2-v_{r+1}$, $b_i>0$ (con $b_i=b_{s+i}$, para $i\leq r-s$ y $b_{r-s+i}=b_i$, para $i\leq s$).

Igualmente, el caso 2. $p(x_1,\dots,x_n)=b_1v_1^2+\cdots+b_sv_s^2-b_{s+1}\cdot v_{s+1}^2-\cdots-b_r\cdot v_r^2$ es equivalente a $b_1v_1^2+\cdots+b_{r-s}v_{r-s}^2-b_{r-s+1}\cdot v_{r-s+1}^2-\cdots-b_r\cdot v_r^2$.

En conclusión, en los casos 2. y 3., podemos suponer que s es mayor o igual que $r/2$ (pues s ó $r-s$ es mayor o igual que $r/2$).

8. Clasificación afín euclídea de las cónicas: Demos la clasificación afín euclídea de los polinomios en dos variables de grado dos (salvo un factor multiplicativo). Sigamos notaciones anteriores ($a,b>0$),

1. $r=2,s=2$: $\left\{\begin{array}{l} ax^2+by^2-1=0 \ \ \text{(elipse).}\\ ax^2+by^2=0 \ \ \text{(dos rectas imaginarias no paralelas).}\end{array}\right.$

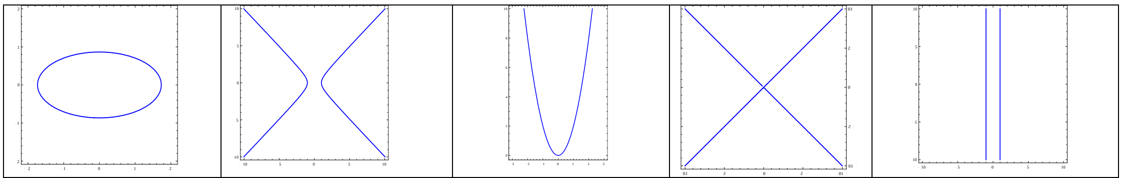
2. $r=2,s=1$: $\left\{\begin{array}{l} ax^2-by^2-1=0 \ \ \text{(hipérbola).}\\ ax^2-by^2=0 \ \ \text{(par de rectas, no paralelas).}\end{array}\right.$

3. $r=2, s=0$: $-ax^2-by^2-1=0$ (elipse imaginaria).

4. $r=1, s=1$: $\begin{cases} ax^2-1=0 \ \text{(dos rectas paralelas).} \\ ax^2=0 \ \text{(recta doble).} \\ ax^2-y=0 \ \text{(parábola).} \end{cases}$

5. $r=1, s=0$: $-ax^2-1=0$ (dos rectas imaginarias paralelas).

6. $r=0, s=0$: $\begin{cases} -1=0 \ \text{(el vacío).} \\ 0=0 \ \text{(todo el plano).} \\ -x=0 \ \text{(una recta).} \end{cases}$



Elipse	Hipérbola	Parábola	2 Rectas	Rect.Paral

En rigor, los casos de 6. no son casos a tener en cuenta porque los polinomios considerados son de grado menor que 2.

9. Clasificación euclídea afín de las cuádricas de $\mathbb{R}^3$: Demos la clasificación afín euclídea de los polinomios en tres variables de grado dos (salvo un factor multiplicativo). Sigamos notaciones anteriores ($a,b,c>0$),

1. $r=3, s=3$: $\begin{cases} ax^2+by^2+cz^2-1=0 \ \text{(elipsoide).} \\ ax^2+by^2+cz^2=0 \ \text{(cono imaginario).} \end{cases}$

2. $r=3, s=2$: $\begin{cases} ax^2+by^2-cz^2-1=0 \ \text{(hiperboloide reglado),} \\ ax^2+by^2-cz^2=0 \ \text{(cono).} \end{cases}$

3. $r=3, s=1$: $ax^2-by^2-cz^2-1=0$ (hiperboloide no reglado).

4. $r=3, s=0$: $-ax^2-by^2-cz^2-1=0$ (elipsoide imaginario).

5. $r=2, s=2$: $\begin{cases} ax^2+by^2-1=0 \ \text{(cilindro elíptico),} \\ ax^2+by^2=0 \ \text{(dos hiperplanos imaginarios no paralelos),} \\ ax^2+by^2-z=0 \ \text{(paraboloide).} \end{cases}$

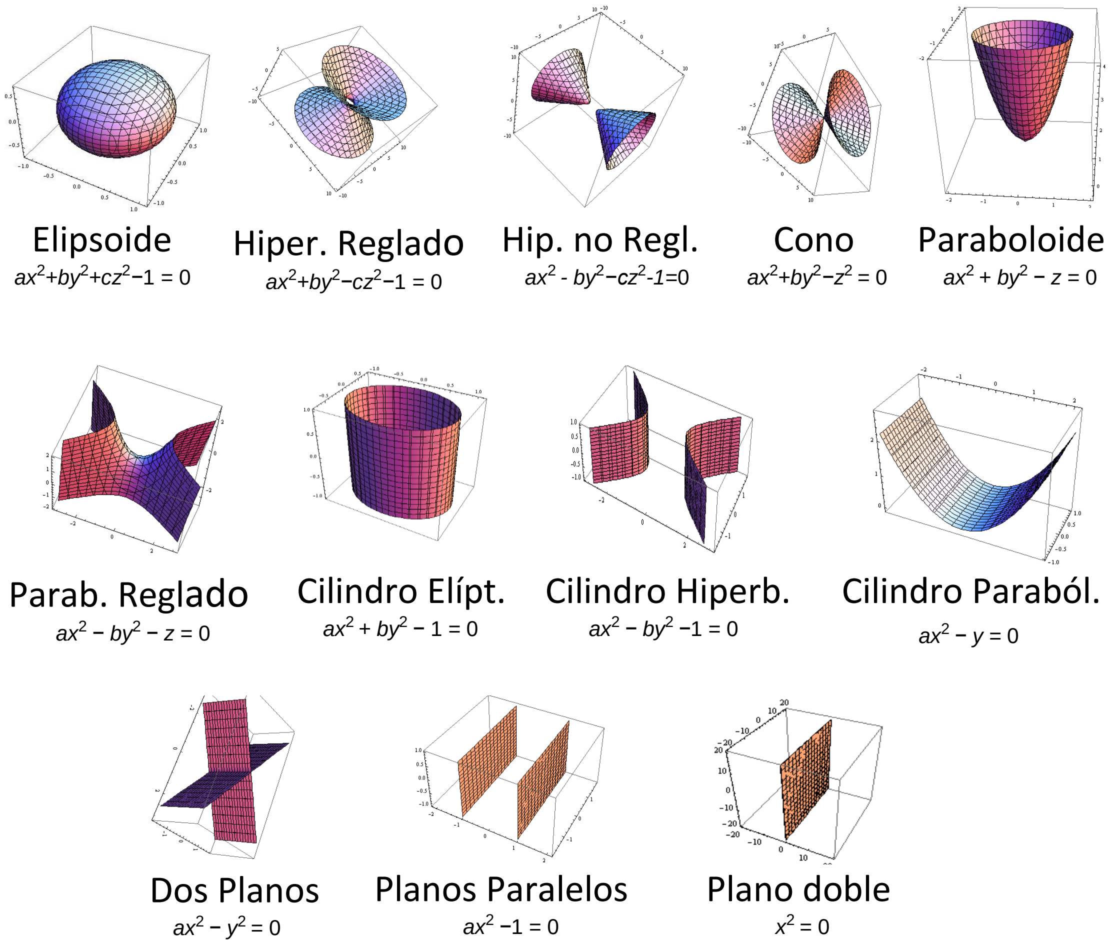
6. $r=2, s=1$: $\begin{cases} ax^2-by^2-1=0 \ \text{(cilindro hiperbólico).} \\ ax^2-by^2=0 \ \text{(dos hiperplanos no paralelos).} \\ ax^2-by^2-z=0 \ \text{(paraboloide reglado).} \end{cases}$

7. $r=2, s=0$: $-ax^2-by^2-1=0$ (cilindro imaginario).

8. $r=1, s=1$: $\begin{cases} ax^2-1=0 \text{ (par de planos paralelos).} \\ ax^2=0 \text{ (plano doble).} \\ ax^2-y=0 \text{ (cilindro parabólico).} \end{cases}$

9. $r=1, s=0$: $-ax^2-1=0$ (par de planos imaginarios paralelos).

10. $r=0, s=0$: $\begin{cases} -1=0 \text{ (vacío).} \\ 0=0 \text{ (todo } \mathbb{R}^3\text{).} \\ -x=0 \text{ (un plano).} \end{cases}$



10. Ejemplo: Calculemos las ecuaciones reducidas de la cónica

$$2x^2+2y^2+2xy+x+y-3=0.$$

Consideremos la forma cuadrática $2x^2+2y^2+2xy$. Si

$$A=\begin{pmatrix}2&1\\1&2\end{pmatrix}$$

entonces $2x^2+2y^2+2xy=(x,y)\cdot A\cdot(x,y)^t$. Los autovalores de A son las raíces de $c_A(x)=(x-3)\cdot(x-1)$. Como $\operatorname{Ker}(A-\operatorname{Id})=\langle(1,-1)\rangle$ y $\operatorname{Ker}(A-3\operatorname{Id})=\langle(1,1)\rangle$, entonces una base ortonormal de autovectores es $\{\frac{1}{\sqrt{2}}\cdot(1,-1),\frac{1}{\sqrt{2}}\cdot(1,1)\}$. Si hacemos el cambio de coordenadas

$$\begin{pmatrix}x\\y\end{pmatrix}=\frac{1}{\sqrt{2}}\cdot\begin{pmatrix}1&1\\-1&1\end{pmatrix}\begin{pmatrix}x'\\y'\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}x'+y'\\-x'+y'\end{pmatrix}$$

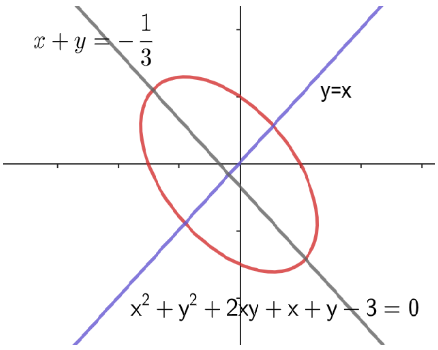
Obtenemos

$$2x^2+2y^2+2xy+x+y-3=x'^2+3y'^2+\sqrt{2}\cdot y'-3$$

Si hacemos la traslación $x'=x''$, $y'=y''-\frac{\sqrt{2}}{6}$ obtenemos

$$x'^2+3y'^2+\sqrt{2}\cdot y'-3=x''^2+3y''^2-\frac{19}{6}$$

Entonces, la ecuación reducida de $2x^2+2y^2+2xy+x+y-3=0$ es



$$\frac{6}{19}x''^2+\frac{18}{19}y''^2-1=0$$

(donde $x=\frac{1}{\sqrt{2}}(x''+y''-\frac{\sqrt{2}}{6})$ e $y=\frac{1}{\sqrt{2}}(-x''+y''-\frac{\sqrt{2}}{6})$), que es la ecuación de una elipse. Cuyos ejes son $x''=0$ y $y''=0$, es decir, la recta $x=y$ y la recta $x+y=\frac{-1}{3}$. El centro es $x''=0$, $y''=0$, es decir, el punto $(\frac{-1}{6},\frac{-1}{6})$.

11. Ejemplo: Calculemos las ecuaciones reducidas de la cuádrica

$$7x^2+4y^2+7z^2+4xy-10xz+4yz+4z-1=0$$

Consideremos la forma cuadrática $7x^2+4y^2+7z^2+4xy-10xz+4yz$ y $A=\begin{pmatrix}7&2&-5\\2&4&2\\-5&2&7\end{pmatrix}$.

Entonces,

$$7x^2+4y^2+7z^2+4xy-10xz+4yz=(x,y,z)\cdot A\cdot(x,y,z)^t$$

Los autovectores de A son $\langle(1,-1,1)\rangle$, $\langle(1,2,1)\rangle$ y $\langle(-1,0,1)\rangle$ de autovalores $0,6,12$ respectivamente. Si hacemos el cambio de coordenadas

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{-1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}}x' + \frac{1}{\sqrt{6}}y' - \frac{1}{\sqrt{2}}z' \\ \frac{-1}{\sqrt{3}}x' + \frac{2}{\sqrt{6}}y' \\ \frac{1}{\sqrt{3}}x' + \frac{1}{\sqrt{6}}y' + \frac{1}{\sqrt{2}}z' \end{pmatrix}$$

Obtenemos

$$7x^2+4y^2+7z^2+4xy-10xz+4yz+4z-1 = 6y'^2+12z'^2+\frac{4}{\sqrt{3}}x'+\frac{2\sqrt{2}}{\sqrt{3}}y'+2\sqrt{2}z'-1$$

Después de la traslación $x'=u, y'=v-\frac{1}{3\sqrt{6}}, z'=w-\frac{1}{6\sqrt{2}}$ obtenemos

$$6y'^2+12z'^2+\frac{4}{\sqrt{3}}x'+\frac{2\sqrt{2}}{\sqrt{3}}y'+2\sqrt{2}z'-1 = 6v^2+12w^2+\frac{4}{\sqrt{3}}u-\frac{23}{18}.$$

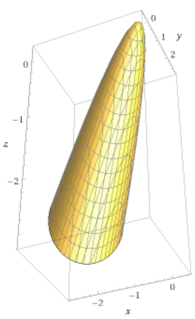
Después de la traslación $u = s-\frac{23\sqrt{3}}{72}$ obtenemos

$$6v^2+12w^2+\frac{4}{\sqrt{3}}u-\frac{23}{18} = 6v^2+12w^2+\frac{4}{\sqrt{3}}s.$$

Después de la simetría $s=-t$, obtenemos

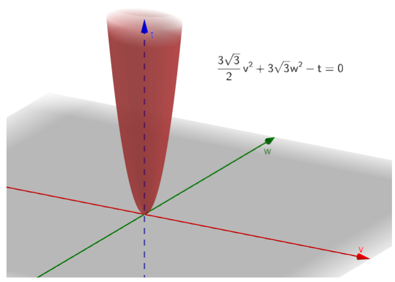
$$6v^2+12w^2+\frac{4}{\sqrt{3}}s = 6v^2+12w^2-\frac{4}{\sqrt{3}}t.$$

La ecuación reducida de $7x^2+4y^2+7z^2+4xy-10xz+4yz+4z-1=0$ es



$$\frac{3\sqrt{3}}{2}\cdot v^2+3\sqrt{3}\cdot w^2-t=0,$$

que es un paraboloide.



4.8.4. Matriz de covarianza

12. Proposición: *Sea $(E,\cdot)$ un espacio vectorial euclídeo y $\star$ una métrica en E. Sea $T\colon E\to E$ el operador lineal autoadjunto asociado ($e\star v = e\cdot T(v)$).*

1. *Consideremos la función $G\colon E\backslash\{0\}\to\mathbb{R}$, $G(e):=\frac{e\star e}{e\cdot e}$. Entonces, $G(e)$ es un máximo de G si y solo si e es un autovector de T de autovalor máximo.*

2. *En el conjunto de los subespacios vectoriales no nulos de E consideremos la función $G(H):=\operatorname{tr}(T_H)$, donde T_H es el operador lineal autoadjunto asociado a $\cdot$ y $\star$ restringidos a H, y tr es la traza de la matriz asociada a T_H (es decir, la suma de los coeficientes de la diagonal). Si $\{e_1,\dots,e_r\}$ es una base ortonormal de H, entonces $G(H)=e_1\star e_1+\cdots+e_r\star e_r$. Si $G(H)$ es un máximo de G, en el conjunto de los subespacios vectoriales r-dimensionales de E, entonces H está generado por una base ortonormal de autovectores de T.*

Demostración. 1. Sea $B=\{e_1,\dots,e_n\}$ una base ortonormal de E y ortogonal para $\star$, luego $T(e_i)=\lambda_i\cdot e_i$ y $e_i\star e_i=\lambda_i$. Reordenando los vectores podemos suponer que $\lambda_1=\cdots=\lambda_r>\cdots\geq\lambda_n$. Dividiendo e por su módulo podemos suponer que $e\cdot e=1$. Escribamos $e_B=(\mu_1,\dots,\mu_n)$. Entonces,

$$G(e)=\frac{e\star e}{e\cdot e}=\lambda_1\cdot\mu_1^2+\cdots+\lambda_n\cdot\mu_n^2\leq\lambda_1\cdot(\mu_1^2+\cdots+\mu_n^2)=\lambda_1=G(e_1)$$

y se tiene la igualdad si y solo si $\mu_{r+1}=\cdots=\mu_n=0$, es decir, e es un autovector de T de autovalor máximo.

2. La matriz asociada a T_H en la base ortonormal $\{e_1,\dots,e_r\}$ coincide con la matriz asociada a $\star$ en esta base, luego $G(H)=\operatorname{tr}(T_H)=e_1\star e_1+\cdots+e_r\star e_r$.

Si H_1 es ortogonal a H_2, $G(H_1\perp H_2)=G(H_1)+G(H_2)$: Sean $e_1,\dots,e_s$ y $e_{s+1},\dots,e_{s+t}$ bases ortonormales de H_1 y H_2. Entonces,

$$G(H_1\perp H_2)=\sum_{i=1}^{s+t}e_i\star e_i=\sum_{i=1}^{s}e_i\star e_i+\sum_{i=1}^{t}e_{s+i}\star e_{s+i}=G(H_1)+G(H_2).$$

Procedamos por inducción sobre r. Si $r=1$, hemos acabado por el apartado 1. Si H no contiene ningún autovector de T, sea $e\in E$ un autovector de autovalor máximo, $R=\langle e\rangle$, $H'=R^\perp\cap H$ y R' tal que $H=R'\perp H'$. Entonces,

$$G(H)=G(R')+G(H')<G(R)+G(H')=G(R\perp H').$$

Contradicción. Podemos afirmar que H contiene un autovector e, sea $R=\langle e\rangle$ y $H=R\perp H'$. Sea H'' entre todos los subespacios vectoriales $r-1$-dimensionales contenidos en $R^\perp$, tal que $G(H'')$ sea máximo. Sabemos que

$$G(H)=G(R)+G(H')\leq G(R)+G(H'')=G(R\perp H'')$$

luego $G(H')=G(H'')$ y H' contiene una base ortonormal de autovectores de T, luego H también. □

13. Teorema: *Sea C un conjunto de m puntos (con repeticiones o no) de $\mathbb{R}^n$.*

1. *El punto más cercano a los puntos de C es $\frac{\sum_{a\in C} a}{m}$.*
2. *Las rectas que mejor se aproximan a los puntos de C son las rectas que pasan por $\mu := \frac{\sum_{a\in C} a}{m}$ y de vector director un autovector de autovalor máximo de la matriz de covarianza*[4] *$A := \sum_{a\in C}(a-\mu)^t\cdot(a-\mu)$.*
3. *Las subvariedades lineales r-dimensionales que mejor se aproximan a los puntos de C son las subvariedades que pasan por μ y de espacio vectorial director un espacio vectorial de dimensión r generado por una base ortonormal de autovectores de A cuya suma de autovalores sea la máxima posible.*

Demostración. 1. Calculemos el punto $x\in\mathbb{R}^n$ tal que $F(x)=\sum_{a\in C} d(a,x)^2$ sea mínima. Tenemos que $F(x)=\sum_{a\in C}((x-a)\cdot(x-a))=\sum_{a\in C}(x\cdot x-2a\cdot x+a\cdot a)$. Derivando parcialmente en cada variable ha de salir cero. Luego, $0=2\cdot(mx_i-\sum_{a\in C}a_i)$. Luego, $x=\frac{\sum_{a\in C} a}{m}$.

2. y 3. Consideremos la subvariedad lineal que pasa por $\gamma\in\mathbb{R}^n$ y de espacio vectorial director H. Cada $v\in\mathbb{R}^n$ es igual a $v=v^{H^\perp}+v^H$, siendo $v^{H^\perp}$ la proyección ortogonal de v sobre $H^\perp$ y v^H la proyección ortogonal de v sobre H. La distancia (al cuadrado) de $a\in\mathbb{R}^n$ a esta subvariedad es $(a-\gamma)\cdot(a-\gamma)-(a-\gamma)^H\cdot(a-\gamma)^H$. Tenemos que calcular γ y H en los que la función

$$F(\gamma,H)=\sum_{a\in C}((a-\gamma)\cdot(a-\gamma)-(a-\gamma)^H\cdot(a-\gamma)^H)$$

sea mínima. La subvariedad de espacio vectorial director H que pasa por γ, es la misma que la subvariedad de espacio vectorial director H que pasa por $\gamma-\gamma^H=\gamma^{H^\perp}$. Entonces,

$$F(\gamma,H)=F(\gamma^{H^\perp},H)=\sum_{a\in C}((a-\gamma^{H^\perp})\cdot(a-\gamma^{H^\perp})-a^H\cdot a^H)$$

Fijado H, $\gamma^{H^\perp}$ tal que $\sum_{a\in C}((a^{H^\perp}-\gamma^{H^\perp})\cdot(a^{H^\perp}-\gamma^{H^\perp}))$ es mínimo es $\gamma^{H^\perp}=\frac{\sum_{a\in C} a^{H^\perp}}{m}$ y podemos tomar $\gamma=\frac{\sum_{a\in C} a^{H^\perp}}{m}+\frac{\sum_{a\in C} a^{H}}{m}=\frac{\sum_{a\in C} a}{m}=\mu$, que no depende de H.

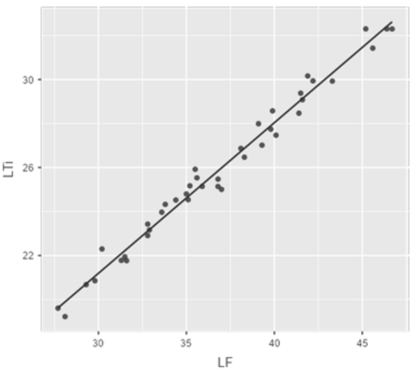
Nos falta calcular H tal que $G(H):=\sum_{a\in C}(a-\mu)^H\cdot(a-\mu)^H$ sea máximo. Sea $\star$ la métrica de $\mathbb{R}^n$ definida por $x\star y=x\cdot A\cdot y^t$, luego la matriz de $\star$ en la base usual es A y la matriz del operador autoadjunto T asociado a $\cdot$ y $\star$ es A. Si $\{e_1,\dots,e_r\}$ es una base ortonormal de H, entonces

$$\begin{aligned}G(H)&=\sum_{a\in C}(a-\mu)^H\cdot(a-\mu)^H=\sum_{a\in C}(a-\mu)^{\langle e_1\rangle}\cdot(a-\mu)^{\langle e_1\rangle}+\cdots+\sum_{a\in C}(a-\mu)^{\langle e_r\rangle}\cdot(a-\mu)^{\langle e_r\rangle}\\&=\sum_{a\in C}(e_1\cdot(a-\mu))\cdot(e_1\cdot(a-\mu))+\cdots+\sum_{a\in C}(e_r\cdot(a-\mu))\cdot(e_r\cdot(a-\mu))\\&=e_1\cdot A\cdot e_1^t+\cdots+e_r\cdot A\cdot e_r^t=e_1\star e_1+\cdots+e_r\star e_r.\end{aligned}$$

[4] En la bibliografía la matriz de covarianza es igual a $\frac{1}{m}\cdot A$.

Concluimos por la proposición anterior.

□



14. Ejemplo: En un hospital se quiere estudiar la relación entre la longitud del fémur y la tibia de los fetos de 26 semanas. Se ha tomado una muestra de 40 fetos y se ha obtenido los resultados que se muestra en la imagen. De la muestra se obtiene que la media es $\mu = (36'82, 25'86)$ y que la matriz de covarianza es $A = \begin{pmatrix} 26'26 & 17'97 \\ 17'97 & 12'53 \end{pmatrix}$. Calculemos la recta que mejor se aproxima a estos datos.

Los autovalores de A son $\lambda_1 = 38'64$ y $\lambda_2 = 0'16$ y los autovectores son $v_1 = (1'45, 1)$ y $v_2 = (-0'68, 1)$. Luego la recta que mejor se aproxima es $\frac{x-36'82}{1'45} = y - 25'86$. Un número que nos indicaría lo bien ajustada que está la recta a los datos de la muestra es $\frac{\lambda_1}{\operatorname{tr}(A)} = \frac{38'64}{38'64+0'16} = 0'99$, que al estar muy próximo a 1 (y no a $\frac{1}{n} = \frac{1}{2}$) indica que es una muy buena aproximación.

4.8.5. Relatividad Especial

Hipótesis del espacio-tiempo newtoniano

$\mathbb{R}^4$ es un ejemplo de espacio-tiempo newtoniano. Dado $x_t = (t, x_1, x_2, x_3)$ decimos que t es la coordenada temporal de x_t y que (x_1, x_2, x_3) son las coordenadas espaciales de x_t. Dados dos puntos x_t y $x'_{t'}$ decimos que el tiempo transcurrido entre x_t y $x'_{t'}$ es $t' - t$. Decimos que x_t y $x'_{t'}$ son simultáneos si $t = t'$ y decimos que la distancia (espacial) entre x_t y x'_t es $\sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + (x'_3 - x_3)^2}$. Consideremos en $\mathbb{R}^4$ la métrica $\star$ definida por $x_t \star x'_{t'} := tt'$, que se denomina métrica del tiempo y es una métrica de signatura $(1,0)$ y $\operatorname{rad}\star := \{x'_{t'} \in \mathbb{R}^4 \colon\ x'_{t'} \star x_t = 0,\ \forall x_t \in \mathbb{R}^4\} = \{0\} \times \mathbb{R}^3$. Consideremos en $\operatorname{rad}\star = \{0\} \times \mathbb{R}^3$ la métrica euclídea $\cdot$ usual. Entonces, la distancia temporal de x_t a $x'_{t'}$ con la métrica métrica del tiempo $\star$ es igual a

$$d(x'_{t'}, x_t) = \sqrt{(x'_{t'} - x_t) \star (x'_{t'} - x_t)} = \sqrt{(t'-t)^2} = |t' - t|.$$

La distancia espacial entre dos puntos simultáneos x_t y x'_t, con la métrica del espacio, es igual a

$$\sqrt{(x'_t - x_t) \cdot (x'_t - x_t)} = \sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + (x'_3 - x_3)^2}.$$

Decimos que $\sigma\colon \mathbb{R} \to \mathbb{R}^4$, $\sigma(t) = (t, x_1(t), x_2(t), x_3(t))$ es un observador o un móvil de $\mathbb{R}^4$.

Los conceptos de tiempo y espacio son absolutos. No dependen del observador. Ésta es la intución que tenemos del universo en el que vivimos. Pero en $\mathbb{R}^4$ tenemos unas rectas privilegiadas: los ejes coordenados y en particular un observador privilegiado $\sigma(t) = (t, 0, 0, 0)$, inexistentes en el universo en el que creemos vivir. Para evitar estas críticas damos la siguiente definición.

15. Definición : Un espacio-tiempo newtoniano es un $\mathbb{R}$-espacio vectorial E de dimensión 4, junto con una métrica $\star$ de signatura $(1,0)$ en E, denominada métrica del tiempo, y junto con un producto escalar $\cdot$ en el $\mathbb{R}$-subespacio vectorial de E de dimensión 3, $\operatorname{rad}\star := \{e' \in E : e' \star e = 0,\ \forall e \in E\}$.

La longitud (temporal) de un segmento qp $(p, q \in E)$, $d(q,p) := \sqrt{(p-q)\star(p-q)}$, se dice que es el tiempo en valor absoluto transcurrido a lo largo de este segmento. El tiempo en el espacio-tiempo newtoniano está orientado, se habla de pasado y futuro. Para orientar el tiempo fijemos un vector $v \in E$ tal que $v \star v = 1$. Diremos que $e \in E$ está positivamente orientado si $e \star v > 0$. Observemos que $E = \langle v \rangle \oplus \operatorname{rad}\star$. Si $p - q$ está positivamente orientado, entonces $p - q = \lambda \cdot v + e'$, con $\lambda > 0$ y $e' \in \operatorname{rad}\star$, luego

$$\sqrt{(p-q)\star(p-q)} = \lambda = (p-q)\star v,$$

Si $p-q$ no está positivamente orientado, entonces $(p-q)\star v = -\sqrt{(p-q)\star(p-q)}$. Si w está positivamente orientado y $w \star w = 1$, entonces $w = v + e''$ con $e'' \in \operatorname{rad}\star$ y para todo $e \in E$ se cumple que $e \star v = e \star w$. Luego, para orientar el tiempo y medirlo podemos tomar w en vez de v.

Se dice que el tiempo transcurrido a lo largo de una curva, digamos por sencillez que está formada por segmentos, es la suma de los tiempos (con signos, orientados) transcurrido a lo largo de cada segmento. El tiempo trancurrido a lo largo de $qp \cup pr$ es igual al tiempo transcurrido a lo largo de qr:

$$(p-q)\star v + (r-p)\star v = (r-q)\star v.$$

Por tanto, el tiempo transcurrido a lo largo de una curva entre dos de sus puntos a y b es igual a $(b-a)\star v$. Dado $p \in E$, se dice que $p + \operatorname{rad}\star$ es el espacio simultáneo a p, porque

$$(q-p)\star v = 0 \iff q - p \in \operatorname{rad}\star \iff q \in p + \operatorname{rad}\star.$$

Obsérvese que el tiempo y el espacio $\operatorname{rad}\star$ son datos absolutos del espacio-tiempo newtoniano.[5]

[5]El físico riguroso cuestionará con razón que haya una unidad de tiempo absoluta. No habría que

Un sistema de referencia inercial es una base e_0, e_1, e_2, e_3 de E tal que $e_0 \star e_0 = 1$, e_0 está positivamente orientado, y e_1, e_2, e_3 es una base ortonormal de $(\operatorname{rad}\star, \cdot)$. Si (s, x_1, x_2, x_3) son las coordenadas de $e \in E$ en el sistema de referencia inercial, se dice que s es la coordenada temporal y que (x_1, x_2, x_3) son las coordenadas espaciales de e en el sistema de referencia inercial.

Diremos que un observador (o móvil) es una aplicación diferenciable $\sigma\colon \mathbb{R} \to E$, tal que $\sigma'(t) \star \sigma'(t) = 1$ y $\sigma'(t)$ está positivamente orientado y $\sigma(0) \in \operatorname{rad}\star$. En coordenadas, $\sigma(t) = (t, x_1(t), x_2(t), x_3(t))$. Diremos que las rectas $\sigma(t) = t \cdot w$ (con $w \star w = 1$ y w positivamente orientado) son observadores inerciales (en el origen). Dado éste observador y una base ortonormal e_1, e_2, e_3 de $\operatorname{rad}\star$, consideremos el sistema de referencia inercial w, e_1, e_2, e_3, entonces en coordenadas $\sigma(t) = (t, 0, 0, 0)$. Dado un móvil $\gamma\colon \mathbb{R} \to E$, $\gamma(t) = (t, x_1(t), x_2(t), x_3(t))$, se dice que $\gamma(t) - \sigma(t) \in \operatorname{rad}\star$ es la posición espacial del móvil relativa al observador inercial. En coordenadas $\gamma(t) - \sigma(t) = (x_1(t), x_2(t), x_3(t))$ y se dice que $(x_1'(t), x_2'(t), x_3'(t))$ es el vector (espacial) de velocidad de γ medido por el observador inercial, cuyo módulo respecto de la métrica espacial es $\sqrt{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2}$.

Hipótesis del espacio-tiempo de la relatividad especial

16. Definición: El espacio-tiempo de la relatividad especial es un $\mathbb{R}$-espacio vectorial E de dimensión 4 junto con una métrica simétrica no singular $\star$ de signatura $(1,3)$, métrica que denominaremos métrica del tiempo.

Por sencillez, como estamos en Álgebra Lineal, las curvas estarán formadas por segmentos. Diremos que la longitud del arco de una curva en E entre dos puntos, medida con la métrica del tiempo, es el tiempo absoluto transcurrido a lo largo del arco de curva.

Cada v, tal que $v \star v = 1$ define un espacio-tiempo newtoniano particular: Sea $\star_v$ la métrica definida por $e \star_v e' := (e \star v) \cdot (e' \star v)$. Observemos que

$$\operatorname{rad}\star_v = \{e' \in E : e' \star_v e = 0, \forall e \in E\} = \langle v \rangle^\perp,$$

luego $\star_v$ es una métrica de signatura $(1,0)$. La métrica $-c^2\star$ restringida a $\operatorname{rad}\star_v = \langle v \rangle^\perp$ es un producto escalar, porque si v_1, v_2, v_3 son una base ortogonal de $\langle v \rangle^\perp$, con $|v_i \star v_i| = 1$ (para $i = 1,2,3$), ha de cumplirse que $v_i \star v_i = -1$ (para $i = 1,2,3$), porque la signatura de $\star$ es $(1,3)$. Por tanto, $(E, \star_v)$ junto con $(\operatorname{rad}\star_v, -c^2\star)$ es un espacio-tiempo newtoniano. Se dice que $(\langle v \rangle^\perp, -c^2\star)$ es el espacio relativo a v. Dado $e \in E$, se escribe

dar una métrica $\star$, sino ésta y sus proporcionales y fijar la unidad de tiempo sería fijar una de éstas métricas. Lo mismo podríamos decir, después, del producto escalar y la unidad de longitud espacial. También cuestionará que exista un punto tan singular como el $0 \in E$. Esto obliga a los matemáticos a hablar de espacios afines en vez de espacios vectoriales.

de modo único como $e = \lambda\cdot v + e'$, con $\lambda \in \mathbb{R}$ y $e' \in \langle v\rangle^{\perp}$. Se dice que $\lambda = e\star v = e\star_v v$ es el tiempo de e relativo a v y que $e' = e-(e\star v)\cdot v$ es la posición espacial de e relativa a v.

En relatividad no hay una noción absoluta de espacio.

Nos falta orientar el tiempo.

Definición: Diremos que $e\in E$ es un vector tipo tiempo si $e\star e > 0$.

Lema: Sean $v_1, v_2, v_3 \in E$ vectores tipo tiempo y supongamos que $v_1\star v_2 > 0$. Entonces, $v_1\star v_3 > 0 \iff v_2\star v_3 > 0$.

Demostración. Por simetría, solo tenemos que probar la implicación directa. Dividiendo v_i por $\sqrt{v_i\star v_i}$, podemos suponer que $v_i\star v_i = 1$, para todo i. Sea $E' = \langle e_1\rangle^{\perp}$. Entonces, E' con la métrica $\cdot$ igual a $-\star$ es un espacio euclídeo. Escribamos $v_2 = \lambda v_1 + v_2'$ y $v_3 = \mu v_1 + v_3'$ (donde $\lambda, \mu > 0$ y $v_2', v_3' \in E'$). Entonces, $1 = v_2\star v_2 = \lambda^2 - v_2'\cdot v_2'$ y $1 = v_3\star v_3 = \mu^2 - v_3'\cdot v_3'$. Luego

$$v_2\star v_3 = \lambda\mu - v_2'\cdot v_3' > \sqrt{\lambda^2-1}\cdot\sqrt{\mu^2-1} - v_2'\cdot v_3' = \sqrt{v_2'\cdot v_2'}\cdot\sqrt{v_3'\cdot v_3'} - v_2'\cdot v_3' \geq 0.$$

□

Definición: Fijemos un vector tipo tiempo v (o cualquier otro v' tal que $v\star v' > 0$). Diremos que un vector de tipo tiempo w está positivamente orientado si $v\star w > 0$.

Un observador o un móvil es una aplicación diferenciable $\sigma\colon [a,b]\to E$ tal que $\sigma'(t)\star\sigma'(t) > 0$ y $\sigma'(t)$ está positivamente orientado. Reparametrizando podemos suponer que $\sigma'(t)\star\sigma'(t) = 1$. Las rectas $e+\mathbb{R}\cdot v'$ (con $v'\star v' = 1$ y $v\star v' > 0$) diremos que son observadores inerciales. Consideremos un punto $p = \sigma(t_0)$ en la curva σ y sea $v_0 = \sigma'(t_0)$. El observador inercial $p+\mathbb{R}\cdot v_0$ tiene un modo particular de medir el tiempo

$$t_{v_0}\colon E\to\mathbb{R},\ t_{v_0}(e') := (e'-p)\star v_0.$$

Diremos que $p+\langle v_0\rangle^{\perp}$ es el espacio simultáneo a p, para el observador inercial y para el observador σ (pues consideramos que cerca de p las observaciones de σ y el observador inercial son muy parecidas, luego en p ambos observan el mismo espacio simultáneo).

Sea $q = \sigma(t_1)$ otro punto de la curva σ y $v_1 = \sigma'(t_1)$. Si $p' \in p+\langle v_0\rangle^{\perp}$ y $q'\in q+\langle v_1\rangle^{\perp}$, entonces el observador σ dirá que la distancia temporal entre p' y q' es igual a la longitud del arco de curva de σ entre p y q. Observemos que hay un tiempo absoluto: $\star$. Cada observador mide el tiempo de un modo particular o relativo, si bien el tiempo con el que se mide a sí mismo coincide con el tiempo absoluto.

17. Sistema de referencia inercial: Llamaremos sistema de referencia inercial (en el origen) a una base ortogonal e_0, e_1, e_2, e_3 tal que $e_0 \star e_0 = 1$, $e_0 \star v > 0$ y $e_i \star e_i = \frac{-1}{c^2}$, para todo $i > 0$. Es decir, es dar un observador inercial $\{t \cdot e_0\}$ (con $e_0 \star e_0 = 1$ y $v \star e_0 > 0$) y una base ortonormal en $\langle e_0 \rangle^\perp$ (que es euclídeo con la métrica $-c^2\star$). Si (t, x_1, x_2, x_3) son las coordenadas de un vector e en el sistema de referencia inercial, entonces se dice que $t = e \star e_0$ y $(0, x_1, x_2, x_3) = e - (e \star e_0) \cdot e_0 \in \langle e_0 \rangle^\perp$ (o abreviadamente (x_1, x_2, x_3)) son el tiempo y posición de e medidos por el observador inercial.

Si (t, x_1, x_2, x_3) son las coordenadas de un vector e en un sistema de referencia inercial entonces $e \star e = t^2 - \frac{x_1^2 + x_2^2 + x_3^2}{c^2}$. Diremos que las trayectorias de los rayos de luz son las rectas $p + \mathbb{R} \cdot e$, con e isótropo ($e \star e = 0$). El conjunto de los rayos de luz que salen del origen es el cono $t^2 - \frac{x_1^2 + x_2^2 + x_3^2}{c^2} = 0$.

Dada una curva $\gamma\colon [a, b] \to E$ (parametrizada por el observador inercial de modo que $\gamma'(t) \star e_0 = 1$, para todo t), es decir, $\gamma(t) = (t, x_1(t), x_2(t), x_3(t))$, diremos que $(x_1'(t), x_2'(t), x_3'(t))$ es el vector (espacial) de velocidad de γ medido por el observador inercial, cuyo módulo respecto de la métrica espacial es $\sqrt{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2}$.

18. Velocidad máxima: Si γ es un observador entonces

$$\begin{aligned} 0 < \gamma'(t) \star \gamma'(t) &= (1, x_1'(t), x_2'(t), x_3'(t)) \star (1, x_1'(t), x_2'(t), x_3'(t)) \\ &= 1 - \frac{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2}{c^2}, \end{aligned}$$

luego el módulo de la velocidad de γ medido por el observador inercial es menor que c.

19. Velocidad de la luz: Si $\gamma(t)$ es un rayo de luz entonces

$$\begin{aligned} 0 = \gamma'(t) \star \gamma'(t) &= (1, x_1'(t), x_2'(t), x_3'(t)) \star (1, x_1'(t), x_2'(t), x_3'(t)) \\ &= 1 - \frac{x_1'(t)^2 + x_2'(t)^2 + x_3'(t)^2}{c^2}, \end{aligned}$$

luego el módulo de la velocidad del rayo de luz medido por el observador inercial es igual a c.

20. Paradoja de los gemelos: Sean $v_1, v_2, v_3 \in E$ vectores tiempo positivamente orientados, tales que $v_1 = v_2 + v_3$. Entonces,

$$\sqrt{v_1 \star v_1} \geq \sqrt{v_2 \star v_2} + \sqrt{v_3 \star v_3}$$

y se da la igualdad si y solo si $v_2 \in \langle v_1 \rangle$. En efecto, $E = \langle v_1 \rangle \perp E'$, donde $e' \star e' < 0$, para todo $e' \in E'$ no nulo. Entonces, $v_2 = \lambda v_1 + e'$ y $v_2 = \mu v_1 - e'$, donde $e' \in E'$ no nulo, $\lambda \cdot \mu > 0$

y $\lambda+\mu=1$ (luego, $|\lambda|+|\mu|=1$). Entonces,

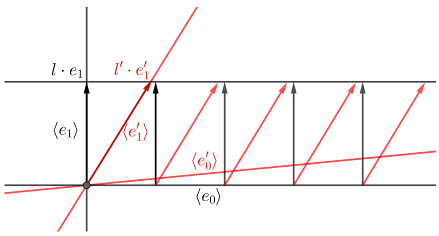
$$\begin{aligned}\sqrt{v_2\star v_2}+\sqrt{v_3\star v_3}&\leq \sqrt{\lambda v_1\star \lambda v_1}+\sqrt{\mu v_1\star v_1}\\ &=(|\lambda|+|\mu|)\cdot\sqrt{v_1\star v_1}=\sqrt{v_1\star v_1}.\end{aligned}$$

y se da la igualdad si y solo si $v_2\in\langle v_1\rangle$. Sean $v_1,\dots,v_n\in E$ vectores tiempo positivamente orientados, tales que $v_1=v_2+\cdots+v_n$. Puede probarse inductivamente que

$$\sqrt{v_1\star v_1}\geq\sum_{i=2}^n\sqrt{v_i\star v_i}$$

y se da la igualdad si y solo si $v_i\in\langle v_1\rangle$, para todo i.

Como consecuencia el camino (poligonal) más largo que une dos puntos p y q (con $(p-q)\star(p-q)>0$), recorrido por un observador, es la línea recta. Es decir, un observador que coincide en p y q con un observador inercial, dirá que el tiempo que ha transcurrido (en su viaje) es menor que el tiempo que dice el observador inercial que ha trancurrido (en su viaje). Además un observador no puede viajar al pasado, porque tendríamos un camino en E que une un punto consigo mismo, de longitud (temporal) menor que cero.



21. Contracción de longitudes: Consideremos una varilla de longitud l, en reposo en un sistema de referencia inercial e_0,e_1,e_2,e_3, y calculemos su longitud l' para un observador inercial que se mueve con velocidad v en la dirección de la varilla (digamos e_1). Consideremos el sistema de referencia inercial

$$e_0'=\frac{1}{\sqrt{1-(\frac{v}{c})^2}}\cdot(e_0+v\cdot e_1),\ e_1'=\frac{1}{\sqrt{1-(\frac{v}{c})^2}}\cdot(\frac{v}{c^2}\cdot e_0+e_1),\ e_2'=e_2,\ e_3'=e_3.$$

El extremo (superior) de la varilla sigue la recta $le_1+\langle e_0\rangle$ (y el extremo inferior la recta $\langle e_0\rangle$). El punto de corte de $le_1+\langle e_0\rangle$ con $\langle e_0'\rangle^\perp$ es $l'e_1'$. Entonces, $l'e_1'=le_1+\lambda e_0$, donde

$$0=(le_1+\lambda e_0)\star(e_0+ve_1)=\lambda-\frac{lv}{c^2}.$$

Luego, $\lambda=\frac{lv}{c^2}$ y

$$\begin{aligned}l'&=-c^2(l'e_1'\star e_1')=-c^2(le_1+\lambda e_0)\star(\frac{1}{\sqrt{1-(\frac{v}{c})^2}}\cdot(\frac{v}{c^2}\cdot e_0+e_1))\\ &=\frac{l}{\sqrt{1-(\frac{v}{c})^2}}-\frac{\lambda v}{\sqrt{1-(\frac{v}{c})^2}}=l\sqrt{1-(\frac{v}{c})^2}.\end{aligned}$$

4.9. Biografía de Charles Hermite

HERMITE BIOGRAPHY

Charles Hermite was the sixth of his parents seven children. Education was not a high priority for Charles's parents but despite not taking too much personal interest in their children's education, nevertheless they did provide them with good schooling. Charles was something of a worry to his parents for he had a defect in his right foot which meant that he moved around only with difficulty. It was clear that this would present him with problems in finding a career. However he had a happy disposition and bore his disability with a cheerful smile.

Charles attended the Collège de Nancy, then went to Paris where he attended the Collège Henri. In 1840-41 he studied at the Collège Louis-le-Grand where some fifteen years earlier Galois had studied. In fact he was taught mathematics there by Louis Richard who had taught Galois. In some ways Hermite was similar to Galois for he preferred to read papers by Euler, Gauss and Lagrange rather than work for his formal examinations.

If Hermite neglected the studies that he should have concentrated on, he was showing remarkable research ability publishing two papers while at Louis-le-Grand. Also like Galois he was attracted by the problem of solving algebraic equations and one of the two papers attempted to show that the quintic cannot be solved in radicals. That he was unfamiliar with Galois's contributions, despite being at the same school, is not at all surprising since the mathematical community were completely unaware of them at this time. However he might reasonably have known of the contributions of Ruffini and Abel to this question, but apparently he did not.

Again like Galois, Hermite wanted to study at the École Polytechnique and he took a year preparing for the examinations. He was tutored by Catalan in 1841-42 and certainly Hermite fared better than Galois had done for he passed. However it was not a glorious pass for he only attained sixty-eighth place in the ordered list. After one year at the École Polytechnique Hermite was refused the right to continue his studies because of his disability. Clearly this was an unfair decision and some important people were prepared to take up his case and fight for him to have the right to continue as a student at the École Polytechnique. The decision was reversed so that he could continue his studies but strict conditions were imposed. Hermite did not find these conditions acceptable and decided that he would not graduate from the École Polytechnique.

Hermite made friends with important mathematicians at this time and frequently visited Joseph Bertrand. On a personal note this was highly significant for he would marry Joseph Bertrand's sister. More significantly from a mathematical point of view he began corresponding with Jacobi and, despite not shining in his formal education, he was already producing research which was ranking as a leading world-class mathematician. The letters he exchanged with Jacobi show that Hermite had discovered some differential equations satisfied by theta-functions and he was using Fourier series to study them. He had found general solutions to the equations in terms of theta-functions. Hermite may have still been an undergraduate but it is likely that his ideas from around 1843 helped Liouville to his important 1844 results which include the result now known as Liouville's theorem.

After spending five years working towards his degree he took and passed the examinations for the baccalauréat and licence which he was awarded in 1847. In the following year he was appointed to the École Polytechnique, the institution which had tried to prevent him continuing his studies some four years earlier; he was appointed répétiteur and admissions examiner.

Hermite made important contributions to number theory and algebra, orthogonal polynomials, and elliptic functions. He discovered his most significant mathematical results over the ten years following his appointment to the École Polytechnique. In 1848 he proved that doubly periodic functions can be represented as quotients of periodic entire functions. In 1849 Hermite submitted a memoir to the Académie des Sciences which applied Cauchy's residue techniques to doubly periodic functions. Sturm and Cauchy gave a good report on this memoir in 1851 but a priority dispute with Liouville seems to have prevented its publication.

Another topic on which Hermite worked and made important contributions was the theory of quadratic forms. This led him to study invariant theory and he found a reciprocity law relating to binary forms. With his understanding of quadratic forms and invariant theory he created a theory of transformations in 1855. His results on this topic provided connections between number theory, theta functions, and the transformations of abelian functions.

On 14 July 1856 Hermite was elected to the Académie des Sciences. However, despite this achievement, 1856 was a bad year for Hermite for he contracted smallpox. It was Cauchy who, with his strong religious conviction, helped Hermite through the crisis. This had a profound effect on Hermite who, under Cauchy's influence, turned to the Roman Catholic religion.

The next mathematical result by Hermite which we must mention is one for which he is rightly famous. Although an algebraic equation of the fifth degree cannot be solved in radicals, a result which was proved by Ruffini and Abel, Hermite showed in 1858 that an algebraic equation of the fifth degree could be solved using elliptic functions. He applied these results to number theory, in particular to class number relations of quadratic forms.

The year 1869 saw him become a professor when he succeeded Duhamel as professor of analysis both at the École Polytechnique and at the Sorbonne. Hermite resigned his chair at the École Polytechnique in 1876 but continued to hold the chair at the Sorbonne until he retired in 1897.

The 1870s saw Hermite return to problems which had interested him earlier in his career such as problems concerning approximation and interpolation. In 1873 Hermite published the first proof that e is a transcendental number. This is another result for which he is rightly famous. Using method's similar to those of Hermite, Lindemann established in 1882 that π was also transcendental. Hermite is now best known for a number of mathematical entities that bear his name: Hermite polynomials, Hermite's differential equation, Hermite's formula of interpolation and Hermitian matrices.

Article by: J.J. O'Connor and E.F. Robertson: http://mathshistory.st-andrews.ac.uk/Biographies/Hermite.html.

4.10. Cálculos con ordenador

1. Escribe $\{1,2,1\}.\{1,1,1\}$ para calcular el producto escalar de $(1,2,1)$ y $(1,1,1)$. Escribe $Norm[(1,2,1)]$ para calcular $\|(1,2,1)\|$.

2. Escribe $Projection[\{1,2,3\},\{1,1,1\}]$ para calcular la proyección ortogonal de $(1,2,3)$ sobre $\langle(1,1,1)\rangle$.

3. Escribe $Orthogonalize[\{\{1,1,0\},\{1,2,,1\},\{0,1,-1\}\}]$ para calcular una base ortonormal por el método de Gram-Schmidt a partir de $\{(1,1,0),(1,2,,1),(0,1,-1)\}$.

4. Escribe $Cross[\{1,2,3\},\{1,1,1\}]$ para calcular $(1,2,3)\times(1,1,1)$.

5. Escribe $PseudoInverse[\{\{1,1,2\},\{1,-1,1\},\{3,1,5\}\}].\{\{1\},\{1\},\{1\}\}$ para calcular la solución aproximada óptima del sistema de ecuaciones lineales del ejemplo 4.3.28.

4.11. Problemas

1. Demuestra que si e y v son dos vectores que tienen el mismo módulo entonces $e+v$ y $e-v$ son ortogonales. Demuestra que las diagonales de un rombo (paralelogramo cuyos lados son todos de igual longitud) son perpendiculares.

2. Demuestra que la distancia de un punto (x_0,y_0,z_0) al plano de ecuación
$$ax+by+cz-d=0$$
es igual a $\frac{|ax_0+by_0+cz_0-d|}{\sqrt{a^2+b^2+c^2}}$.

3. Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ la simetría especular respecto del plano de ecuación $x+y+z=0$. Calcula la matriz de T en la base usual y prueba que T es una isometría.

4. Si $A \in M_n(\mathbb{R})$ es una matriz ortogonal, prueba que $\det(A) = \pm 1$.

5. Sea $V = \langle (1,1,1),(1,-1,1) \rangle \subset \mathbb{R}^3$. Calcula $\dim_{\mathbb{R}} V^{\perp}$.

6. Sea $V \subset \mathbb{R}^4$ un subespacio vectorial de dimensión 2 ¿Qué dimensión tiene $V^{\perp}$?

7. Sea $Ax^t = b^t$ un sistema de ecuaciones lineales y supongamos que $\operatorname{Ker} A = 0$. Prueba que $(A^tA)^{-1}A^tb^t$ es la solución aproximada óptima del sistema de ecuaciones lineales $Ax^t = b^t$.

8. Sea $(E,\cdot)$ un espacio vectorial euclídeo y $L \subseteq E$ un subespacio vectorial. Consideremos la proyección ortogonal de E sobre L

$$\begin{array}{ccc} E = L \oplus L^{\perp} & \xrightarrow{\pi} & L \\ e = e_1 + e_2 & \mapsto & e_1, \end{array}$$

es decir, $\pi(e) = e_1$ es la proyección ortogonal de e en L. Sea $B = \{e_1,\dots,e_n\}$ una base ortonormal de E y $B' = \{e'_1,\dots,e'_m\}$ una base de L y $Q \in M_{n\times m}(\mathbb{R})$ la matriz de la inclusión $L \overset{i}{\hookrightarrow} E$ en las bases B' y B. Calcula la matriz P de la aplicación lineal $\pi: E \to L$ en las bases B y B'.

9. Calcula el área del paralelogramo definido por los vectores $(1,1,1),(2,1,2) \in \mathbb{R}^3$.

10. Calcula el volumen del paralelepípedo definido por los vectores $(1,1,-1),(-1,1,1)$ y $(0,0,1) \in \mathbb{R}^3$. Calcula $(1,1,-1)\times(-1,1,1)$ y $((1,1,-1)\times(-1,1,1))\cdot(0,0,1)$.

11. Fórmula de Herón (10-70): Sea T un triángulo de lados de longitud a,b,c y área A. Prueba que

$$A = \sqrt{\frac{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}{16}} = \frac{1}{4}\sqrt{\begin{vmatrix} 0 & a^2 & b^2 & 1 \\ a^2 & 0 & c^2 & 1 \\ b^2 & c^2 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{vmatrix}}.$$

12. Fórmula de Tartaglia (1499-1557): Prueba que el volumen V del tetraedro de vértices $p_1, p_2, p_3, p_4 \in \mathbb{R}^3$ es igual a

$$V = \sqrt{\frac{1}{288}\begin{vmatrix} d_{11} & d_{12} & d_{13} & d_{14} & 1 \\ d_{21} & d_{22} & d_{23} & d_{24} & 1 \\ d_{31} & d_{32} & d_{33} & d_{34} & 1 \\ d_{41} & d_{42} & d_{43} & d_{44} & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix}}$$

donde $d_{ij} = d(p_i, p_j)^2$.

13. Calcula el área del cuadrilatero de vértices $(1,1)$, $(2,2)$, $(3,2)$ y $(5,4)$.

14. Calcula la recta perpendicular al plano $x+y+z=3$ que pasa por el punto $(1,0,1)$.

15. Calcula el paralelepípedo rectangular de volumen v de menor superficie.

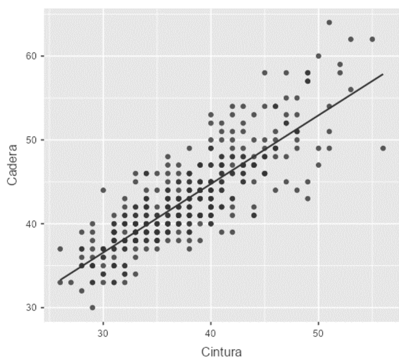
16. Demuestra el teorema 4.3.21 usando la proposición 4.8.1.

17. Descomposición en valores singulares: Dada $A \in M_{n\times m}(\mathbb{R})$, prueba que existen matrices ortogonales $P \in M_{n\times n}(\mathbb{R})$ y $Q \in M_{m\times m}(\mathbb{R})$ y una matriz "diagonal" $D \in M_{n\times m}(\mathbb{R})$ tales que
$$A = P \cdot D \cdot Q.$$

18. Sea $\star$ una métrica de $\mathbb{R}^3$ de matriz asociada en la base usual de $\mathbb{R}^3$, $\begin{pmatrix} -3 & -1 & 0 \\ -1 & 5 & 4 \\ 0 & 4 & 3 \end{pmatrix}$.
Calcula una base ortogonal $\{e_1, e_2, e_3\}$ tal que $e_i \star e_i = 0$ ó 1 ó, -1, para todo i.

19. Calcula las ecuaciones reducidas de la cónica $2xy + x + y + 1 = 0$.

20. En la ciudad de Buckingham se ha hecho un estudio sobre la relación entre el perímetro medido en pulgadas de cinturas y caderas en una muestra de 401 individuos (cuyos datos se muestran en la imagen). De la muestra se obtiene que la media es $\mu = (37'90, 43'04)$ y que la matriz de covarianza es $A = \begin{pmatrix} 32'82 & 27'12 \\ 27'12 & 31'99 \end{pmatrix}$. Calcula la recta que mejor se aproxima a los datos.



4.12. Solución de los problemas

P1. Sabemos que $\|e\| = \|v\|$, luego $e \cdot e = v \cdot v$. Entonces,

$$(e+v)\cdot(e-v) = e\cdot e + e\cdot v - v\cdot e - v\cdot v = e\cdot e - v\cdot v = 0.$$

Si e y v son dos lados del paralelogramo, entonces las diagonales son $e+v$ y $e-v$.

P2. Un vector perpendicular al plano es (a,b,c). Sea p un punto del plano. La distancia de (x_0,y_0,z_0) al plano es el módulo de la proyección ortogonal de $(x_0,y_0,z_0)-p$ sobre $\langle(a,b,c)\rangle$, que es

$$|((x_0,y_0,z_0)-p)\cdot\frac{(a,b,c)}{\|(a,b,c)\|}| = \frac{|ax_0+by_0+cz_0-d|}{\sqrt{a_0^2+b_0^2+c_0^2}}.$$

P3. Consideremos la base $B' = \{(1,1,1),(1,-1,0),(0,1,-1)\}$. Entonces,

$$T(1,1,1) = -(1,1,1),\ T(1,-1,0) = (1,-1,0) \text{ y } T(0,1,-1) = (0,1,-1).$$

Luego $T_{B'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, luego la matriz de T en la base usual es

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}\cdot\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}^{-1} = \frac{1}{3}\cdot\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}.$$

P4. Sabemos que $A^tA = \mathrm{Id}$, luego $|A|^2 = |A^tA| = |\mathrm{Id}| = 1$ y $|A| = \pm 1$.

P5. Como $\dim V = 2$, $\dim V^\perp = 3 - \dim V = 1$.

P6. Tenemos que $\dim V^\perp = \dim \mathbb{R}^4 - \dim V = 2$.

P7. Observemos que $\mathrm{Ker}(A^tA) = \mathrm{Ker}\, A$: Si $A^tAx^t = 0$ entonces $xA^tAx^t = 0$, luego $Ax^t = 0$, y recíprocamente si $Ax^t = 0$ entonces $A^tAx^t = 0$. Por lo tanto, A^tA es invertible. Por el teorema 4.3.27, $A^t\tilde{y}^t$ es la solución aproximada óptima del sistema, donde $A^t\tilde{y}^t$ cumple que $A^tAA^t\tilde{y}^t = A^tb^t$. Entonces, $A^t\tilde{y} = (A^tA)^{-1}A^tb^t$ es la solución aproximada óptima del sistema.

P8. Como consecuencia del problema 7, puede probarse que $P = (Q^tQ)^{-1}Q^t$. Demos otra demostración. Dado $v \in E$ sabemos que $v - \pi(v) \in L^\perp$, o dicho de otro modo,

$$v\cdot l = \pi(v)\cdot l,\ \forall l \in L.$$

Sea $v_B = (x_1,\dots,x_n)$ y $l_{B'} = (y_1,\dots,y_m)$. Entonces, $l^B = Q\cdot(y_1,\dots,y_m)^t$ y

$$v\cdot l = (x_1,\dots,x_n)\cdot Q\begin{pmatrix} y_1\\ \vdots \\ y_m\end{pmatrix}.$$

Por otra parte, $\pi(v)^{B'} = P\cdot(x_1,\dots,x_n)^t$, luego $\pi(v)^B = Q\cdot P\cdot(x_1,\dots,x_n)^t$. Por tanto,

$$\pi(v)\cdot l = (x_1,\dots,x_n)P^tQ^tQ\begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix}.$$

Entonces, $Q = P^tQ^tQ$, luego $Q^t = (Q^tQ)P$ y $\boxed{P = (Q^tQ)^{-1}Q^t}$.

La matriz de la composición $E \xrightarrow{\pi} L \xrightarrow{i} E$ en la base B es $QP = Q(Q^tQ)^{-1}Q^t$.

P9. La proyección ortogonal de $(2,1,2)$ sobre $(1,1,1)$ es $\frac{(2,1,2)\cdot(1,1,1)}{(1,1,1)\cdot(1,1,1)}(1,1,1) = \frac{5}{3}\cdot(1,1,1)$. Luego el área es igual a la longitud de la base por altura:

$$\|(1,1,1)\|\cdot\|(2,1,2)-\frac{5}{3}\cdot(1,1,1)\| = \sqrt{3}\cdot\|(\frac{1}{3},\frac{-2}{3},\frac{1}{3})\| = \sqrt{3}\cdot\frac{\sqrt{2}}{\sqrt{3}} = \sqrt{2}.$$

De otro modo: el área es $\sqrt{\begin{vmatrix}(2,1,2)\cdot(2,1,2) & (2,1,2)\cdot(1,1,1)\\ (1,1,1)\cdot(2,1,2) & (1,1,1)\cdot(1,1,1)\end{vmatrix}} = \sqrt{2}.$

P10. El volumen es igual a $|\begin{vmatrix}1 & 1 & -1\\ -1 & 1 & 1\\ 0 & 0 & 1\end{vmatrix}| = 2$, $(1,1,-1)\times(-1,1,1) = \begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ 1 & 1 & -1\\ -1 & 1 & 1\end{vmatrix} =$ $2\vec{i}+2\vec{k} = (2,0,2)$ y $((1,1,-1)\times(-1,1,1))\cdot(0,0,1) = 2$.

P11. Podemos suponer que un vértice es el origen y los otros dos e y v. Entonces, $a^2 = e\cdot e$, $b^2 = v\cdot v$ y $c^2 = (e-v)\cdot(e-v) = e\cdot e + v\cdot v - 2e\cdot v$. Entonces,

$$\begin{aligned} A &= \frac{1}{2}\sqrt{\begin{vmatrix} e\cdot e & e\cdot v\\ e\cdot v & v\cdot v\end{vmatrix}} \\ &= \sqrt{\frac{1}{4}\begin{vmatrix} a^2 & \frac{a^2+b^2-c^2}{2}\\ \frac{a^2+b^2-c^2}{2} & b^2\end{vmatrix}} \overset{6}{=} \cdots = \sqrt{\frac{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}{16}}.\end{aligned}$$

$$A=\frac{1}{2}\sqrt{\begin{vmatrix} e\cdot e & e\cdot v\\ e\cdot v & v\cdot v\end{vmatrix}}=\frac{1}{4}\sqrt{\begin{vmatrix} -2e\cdot e & -2e\cdot v\\ -2e\cdot v & -2v\cdot v\end{vmatrix}}=\frac{1}{4}\sqrt{\begin{vmatrix} 0 & e\cdot e & v\cdot v & 1\\ e\cdot e & -2e\cdot e & -2e\cdot v & 0\\ v\cdot v & -2e\cdot v & -2v\cdot v & 0\\ -1 & 0 & 0 & 0\end{vmatrix}}\overset{F_2+F_1}{\underset{F_3+F_1}{=}}\cdots$$

$$\overset{C_2+C_1}{\underset{C_3+C_1}{=}}\frac{1}{4}\sqrt{\begin{vmatrix} 0 & e\cdot e & v\cdot v & 1\\ e\cdot e & 0 & e\cdot e+v\cdot v-2e\cdot v & 1\\ v\cdot v & e\cdot e+v\cdot v-2e\cdot v & 0 & 1\\ -1 & -1 & -1 & 0\end{vmatrix}}=\frac{1}{4}\sqrt{\begin{vmatrix} 0 & a^2 & b^2 & 1\\ a^2 & 0 & c^2 & 1\\ b^2 & c^2 & 0 & 1\\ -1 & -1 & -1 & 0\end{vmatrix}}.$$

P12. Podemos suponer que $p_1=0$. V es un sexto del volumen del paralelepípedo de lados los vectores p_2,p_3,p_4 y sus trasladados (es consecuencia de que el volumen de una pirámide es igual al área de las base por altura dividido por tres). Observemos que $d_{ij}=(p_i-p_j)\cdot(p_i-p_j)=p_i\cdot p_i+p_j\cdot p_j-2\cdot p_i\cdot p_j=d_{i1}+d_{1j}-2\cdot p_i\cdot p_j$. Luego,

$$\sqrt{\begin{vmatrix} d_{11} & d_{12} & d_{13} & d_{14} & 1\\ d_{21} & d_{22} & d_{23} & d_{24} & 1\\ d_{31} & d_{32} & d_{33} & d_{34} & 1\\ d_{41} & d_{42} & d_{43} & d_{44} & 1\\ 1 & 1 & 1 & 1 & 0\end{vmatrix}}\overset{\substack{F_2-F_1\\ F_3-F_1}}{\underset{F_4-F_1}{=}}\cdots\overset{\substack{C_2-C_1\\ C_3-C_1}}{\underset{C_4-C_1}{=}}\sqrt{\begin{vmatrix} 0 & p_2\cdot p_2 & p_3\cdot p_3 & p_4\cdot p_4 & 1\\ p_2\cdot p_2 & -2p_2\cdot p_2 & -2p_2\cdot p_3 & -2p_2\cdot p_4 & 0\\ p_3\cdot p_3 & -2p_3\cdot p_2 & -2p_3\cdot p_3 & -2p_3\cdot p_4 & 0\\ p_4\cdot p_4 & -2p_4\cdot p_2 & -2p_4\cdot p_3 & -2p_4\cdot p_4 & 0\\ 1 & 0 & 0 & 0 & 0\end{vmatrix}}$$

$$=\sqrt{-\det(-2p_i\cdot p_j)_{i,j>1}}=\sqrt{8}\cdot\sqrt{\det(p_i\cdot p_j)_{i,j>1}}=\sqrt{8}\cdot 6\cdot V=\sqrt{288}\cdot V.$$

P13. Es la suma del área del triángulo de vértices $(1,1)$, $(2,2)$ y $(3,2)$ y del triángulo de vértices $(3,2)$, $(5,4)$ y $(2,2)$.

P14. Un vector perpendicular al plano es $(1,1,1)$. Luego la recta pedida es de ecuaciones implícitas. $\frac{x-1}{1}=\frac{y}{1}=\frac{z-1}{1}$.

P15. Si P es un paralelepípedo de lados de longitud a,b,c de volumen v, entonces $v=abc$. La superficie de P es $S=2\cdot(ab+ac+bc)=2\cdot(\frac{v}{bc}\cdot b+\frac{v}{bc}\cdot c+bc)=2\cdot(\frac{v}{c}+\frac{v}{b}+bc)$. Para b pequeño o c pequeño S es grande, si b y c no son pequeños para b grande

[6] Es una sencilla comprobación que los radicandos coinciden. Pero puede usarse el siguiente argumento: El primer radicando es un polinomios en a,b,c de grado 4, invariante al multiplicar cada variable por ±1. Es divisible por $-a+b+c$, porque el resto dividir por él es un polinomio en b y c, nulo para todo valor positivo de b y c; ya que cuando $a=b+c$ el área del triángulo es nula. Por tanto, es divisible por los cuatro factores que aparecen en el segundo radicando. Luego los dos radicandos son proporcionales. Entonces solo hay que comprobar la fórmula para un triángulo concreto, por ejemplo el triángulo rectángulo de catetos de longitud 1.

o c grande S es grande. Entonces, S alcanza un mínimo en un mínimo local. Se ha de cumplir que

$$0=\frac{1}{2}\frac{\partial S}{\partial b}=-\frac{v}{b^2}+c,\ 0=\frac{1}{2}\frac{\partial S}{\partial c}=-\frac{v}{c^2}+b.$$

Luego, $b=c=\sqrt[3]{v}$ y $a=\sqrt[3]{v}$. Es decir, el paralelepípedo es un cubo.

P16. Tenemos que calcular un mínimo de la función $g(x)=(xA^t-b)\cdot(Ax^t-b^t)$. Por la proposición 4.8.1, $0=\frac{\partial g}{\partial x_i}=(0,\dots,\overset{i}{1},\dots,0)\cdot A^t\cdot(Ax^t-b^t)+(xA^t-b)\cdot A\cdot(0,\dots,\overset{i}{1},\dots,0)^t$, para todo i. Luego, $(0,\dots,\overset{i}{1},\dots,0)\cdot A^t\cdot Ax^t=(0,\dots,\overset{i}{1},\dots,0)\cdot A^t\cdot b^t$, para todo i y

$$A^t\cdot Ax^t=A^t\cdot b^t.$$

P17. Consideremos la aplicación lineal $T\colon\mathbb{R}^m\to\mathbb{R}^n$ de matriz A en las bases usuales U y U' de $\mathbb{R}^m$ y $\mathbb{R}^n$ respectivamente. Por la proposición 4.7.10, existen bases ortonormales $B=\{e_1,\dots,e_m\}$ y $B'=\{e'_1,\dots,e'_n\}$ en $\mathbb{R}^m$ y $\mathbb{R}^n$ respectivamente de modo que $T_B^{B'}=D$ es diagonal. Entonces, $T_U^{U'}=\mathrm{Id}_B^{U'}\cdot T_B^{B'}\cdot\mathrm{Id}_U^{B'}$.

P18. Construyamos una base ortogonal. Sea $e_1=(1,0,0)$, luego $e_1\star e_1=-3$. Calculemos $\langle e_1\rangle^\perp=\{(x,y,z)\in\mathbb{R}^3\colon (x,y,z)\star(1,0,0)=0\}=\{(x,y,z)\in\mathbb{R}^3\colon -3x-y=0\}=\langle(1,-3,0),(0,0,1)\rangle$. Sea $e_2=(0,0,1)$, luego $e_2\star e_2=3$. Calculemos

$$\begin{aligned}\langle e_1,e_2\rangle^\perp&=\{v=\lambda(1,-3,0)+\mu(0,0,1)\colon v\star e_2=0\}\\&=\{\lambda(1,-3,0)+\mu(0,0,1)\colon -12\lambda+3\mu=0=0\}=\langle(1,-3,4)\rangle.\end{aligned}$$

Observemos que $(1,-3,4)\star(1,-3,4)=0$. Dividamos por los módulos y obtenemos que la base ortogonal buscada es $(\frac{1}{\sqrt3},0,0),(0,0,\frac{1}{\sqrt3}),(1,-3,4)\}$.

P19. La matriz asociada a la forma cuadrática $2xy$ es $\begin{pmatrix}0&1\\1&0\end{pmatrix}$, cuyos autovectores son $(1,1)$ y $(1,-1)$ de autovalores 1 y -1 respectivamente. Consideremos la base ortonormal $\{\frac{1}{\sqrt2}(1,1),\frac{1}{\sqrt2}(1,-1)$ y el cambio de coordenadas

$$\begin{pmatrix}x\\y\end{pmatrix}=\frac{1}{\sqrt2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}x'\\y'\end{pmatrix}=\frac{1}{\sqrt2}\begin{pmatrix}x'+y'\\x'-y'\end{pmatrix}.$$

Entonces,

$$2xy+x+y+1=x'^2-y'^2+\frac{x'+y'}{\sqrt2}+\frac{x'-y'}{\sqrt2}+1=x'^2-y'^2+\sqrt2x'+1.$$

Mediante la traslación $x' = u - \frac{1}{\sqrt{2}}$, $y' = v$, tenemos que

$$x'^2 - y'^2 + \sqrt{2}x' + 1 = u^2 - v^2 + \frac{1}{2}.$$

Por tanto, la cónica $2xy + x + y + 1 = 0$ es igual a la cónica $2v^2 - 2u^2 - 1 = 0$, que es una hipérbola (de asíntotas $u + v = 0$ y $u - v = 0$, es decir, las rectas $x = \frac{-1}{2}$ y $y = \frac{-1}{2}$).

P20. Los autovalores de A son $\lambda_1 = 59'52$ y $\lambda_2 = 5'28$; y los autovectores asociados son $v_1 = (1'01, 1)$ y $v_2 = (-0{,}98, 1)$ respectivamente. Por tanto, la recta que mejor se aproxima a los datos es $\frac{x-37'90}{1'01} = y - 43'04$. Por otra parte $\frac{\lambda_1}{\mathrm{tr}(A)} = \frac{59'52}{32'82+31'99} = 0'91$, número cercano a 1 (y lejano a $\frac{1}{2}$) que indica que la recta es una buena aproximación a los datos de la muestra.

Bibliografía

[1] BOLOS, J.V., CAYETANO, J., REQUEJO, B.: *Álgebra Lineal y Geometría*, Colección Manuales Uex **50**, 2007.

[2] DEL VALLE, J.C., *Álgebra Lineal para estudiantes de Ingenierías y Ciencias*, McGraw-Hill/Interamericana de México, 2011.

[3] HERNÁNDEZ, D.: *Álgebra Lineal*, Ediciones de la Universidad de Salamanca, 1994.

[4] LAMBERSON, R.H., MCKELVEY, R., NOON, B.R.: *A Dynamic Analysis of the Viability of the Northen Spotted Owl in a Fragmented Forest Enviroment,* Conservation Biology **6** 505-512, 1992.

[5] LAY, D.C.: *Álgebra Lineal y sus Aplicaciones*, Ed. Pearson, 2007.

[6] LEVINE, I.N.: *Química Cuántica*, Ed. Prentice Hall, 2001.

[7] NAVARRO, J.A.: *Álgebra Conmutativa Básica*, Manuales Uex **19**, 1996.

[8] MONK, P., MUNRO L.J.: *Maths for Chemistry*, Oxford University Press, 2010.

[9] MULERO, M.A., OJEDA, I.: *Matemáticas para primero de Ciencias,* Colección Manuales Uex **54**, 2008.

[10] STEINER, E.: *Matemáticas para las Ciencias Aplicadas*, Ed Reverté, 2005.

Índice alfabético

colec

UNIVERSIDAD DE EXTREMADURA

man

112