ÁLGEBRA II / ÁLGEBRA - NOTAS DEL TEÓRICO

SILVINA RIVEROS, ALEJANDRO TIRABOSCHI Y AGUSTÍN GARCÍA IGLESIAS

Año 2021
FAMAF - UNC

LEER

ÍNDICE GENERAL

I VECTORES Y SISTEMAS LINEALES EN $\mathbb{R}^n$

II ÁLGEBRA LINEAL

III APÉNDICES

A NÚMEROS COMPLEJOS 9
A.1 Cuerpos . . . 9
A.1.1 Un cuerpo finito . . . 10
A.2 Números complejos . . . 11
B FUNCIONES POLINÓMICAS 17
B.1 Definición de funciones polinómicas . . . 17
B.2 División de polinomios . . . 20
C MULTIPLICACIÓN DE POLINOMIOS POR FFT 23
C.1 Representación de polinomios por valores . . . 23
C.2 Transformada de Fourier discreta . . . 24
C.3 Transformada rápida de Fourier . . . 29
C.4 La antitransformada de Fourier . . . 32
D DETERMINANTE 33
D.1 Determinantes . . . 33
D.2 Regla de Cramer . . . 41

IV ÍNDICE

Indice alfabético 47

ÍNDICE DE FIGURAS

Figura 1 Representación gráfica de los números complejos. . . 13
Figura 2 Ejemplos de la representación gráfica de los números complejos. 13

PREFACIO

Las siguientes notas se han utilizado para el dictado del curso "Álgebra II / Álgebra / Álgebra Lineal" del primer año de las licenciaturas y profesorados de FAMAF. Han sido las notas principales en el dictado del año 2018 y 2020, y se limitan casi exclusivamente al contenido dictado en el curso. Las partes señaladas con (*) y los apéndices son optativos.

Estas notas están basadas principalmente en *Apuntes de Álgebra II - Año 2005* de Silvina Riveros y han sido revisadas, modificadas y ampliadas por Alejandro Tiraboschi y Agustín García Iglesias.

También hemos utilizado como bibliografía de apoyo los siguientes:

- Serge Lang: Álgebra Lineal, Fondo Educativo Interamericano (1976). Puede descargarse de:

  https://archive.org/details/IntroduccionAlAlgebraLinealSergeLang

- *Álgebra Lineal.* Autores: Gabriela Jerónimo, Juan Sabia y Susana Tesauri. Año 2008. Puede descargarse del Departamento de Matemática de la UBA, en la dirección

  http://mate.dm.uba.ar/~jeronimo/algebra_lineal/AlgebraLineal.pdf

- *Linear Algebra.* Autores: Jim Hefferon. Se descarga en

  http://joshua.smcvt.edu/linearalgebra/

- *Álgebra Lineal.* Autores: Kenneth Hoffman y Ray Kunze. Año: 1973. Editorial: Prentice Hall.

  Ficha en biblioteca de FAMAF: http://bit.ly/2tn3eRc

Contenidos mínimos

Resolución de ecuaciones lineales. Matrices. Operaciones elementales. Matriz inversa. Espacios vectoriales sobre $\mathbb{R}$ y $\mathbb{C}$. Subespacios. Independencia lineal. Bases y dimensión Rectas y planos en $\mathbb{R}^n$. Transformaciones lineales y matrices. Isomorfismos. Cambio de bases. Núcleo e imagen de transformaciones lineales. Rango fila y columna. Determinante de una matriz. Cálculo y propiedades básicas. Espacios con producto interno. Desigualdad de Cauchy-Schwartz. Desigualdad triangular. Teorema de Pitágoras. Ortonormalización de Gram-Schmidt. Ecuaciones de rectas y planos en $\mathbb{R}^n$. Distancias. Introducción a vectores y valores propios. Aplicaciones. Diagonalización de matrices simétricas.

Parte I

VECTORES Y SISTEMAS LINEALES EN $\mathbb{R}^n$

Parte II

ÁLGEBRA LINEAL

Parte III

APÉNDICES

A

NÚMEROS COMPLEJOS

A.1 CUERPOS

En el cuatrimestre pasado se ha visto el concepto de cuerpo, del cual haremos un repaso.

(Ver también https://es.wikipedia.org/wiki/Cuerpo_(matemáticas)).

Definición A.1.1. Un conjunto $\mathbb{K}$ es un *cuerpo* si es un anillo de división conmutativo, es decir, un anillo conmutativo con unidad en el que todo elemento distinto de cero es invertible respecto del producto. Por tanto, un cuerpo es un conjunto $\mathbb{K}$ en el que se han definido dos operaciones, '$+$' y '$\cdot$', llamadas *adición* y *multiplicación* respectivamente, que cumplen las propiedades **I1**,. . ., **I7** que se listan más abajo.

Sean a, b, c elementos arbitrarios de $\mathbb{K}$, y 0 y 1 dos elementos especiales de $\mathbb{K}$. Entonces se satisfacen:

I1. $a + b$ y $a \cdot b$ pertenecen a $\mathbb{K}$.

I2. *Conmutatividad.* $a + b = b + a$; $ab = ba$.

I3. *Asociatividad.* $(a + b) + c = a + (b + c)$; $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

I4. *Existencia de elemento neutro.* Existen números $0, 1 \in \mathbb{K}$ con $0 \neq 1$ tal que $a + 0 = a$; $a \cdot 1 = a$.

I5. *Distributividad.* $a \cdot (b + c) = a \cdot b + a \cdot c$.

I6. *Existencia del inverso aditivo.* Por cada a en $\mathbb{K}$ existe un único $-a$ en $\mathbb{K}$ tal que $a + (-a) = 0$.

I7. *Existencia de inverso multiplicativo.* Si a es distinto de o, existe un único elemento $a^{-1} \in \mathbb{K}$ tal que $a \cdot a^{-1} = 1$.

Muchas veces denotaremos el producto yuxtaponiendo los elementos, es decir $ab := a \cdot b$, para $a, b \in \mathbb{K}$. Debido a la ley de asociatividad para la suma (axioma **I3**) $(a + b) + c$ es igual a $a + (b + c)$ y por lo tanto podemos eliminar los paréntesis sin ambigüedad. Es decir, denotamos

$$a + b + c := (a + b) + c = a + (b + c).$$

De forma análoga, usaremos la notación

$$abc = (ab)c = a(bc).$$

Debido a la ley de conmutatividad (axioma **I2**), es claro que del axioma **I4** se deduce que $0 + a = a + 0 = a$ y $1a = a1 = a$. Análogamente, por **I2** e **I6** obtenemos que $-a + a = a + (-a) = 0$, y por **I6** que $aa^{-1} = a^{-1}a = 1$.

Todos los axiomas corresponden a propiedades familiares de los cuerpos que ya conocemos, como ser el cuerpo de los números reales, denotado $\mathbb{R}$ y el cuerpo de los números racionales (fracciones), denotado $\mathbb{Q}$. De ellas pueden deducirse la mayoría de las reglas comunes a los cuerpos. Por ejemplo, podemos *definir* la operación de sustracción diciendo que $a - b$ es lo mismo que $a + (-b)$; y deducir las reglas elementales por ejemplo,

$$a - (-b) = a + b, \qquad -(-a) = a.$$

También podemos deducir

$$(ab)^{-1} = a^{-1}b^{-1}$$

con tal que a y b sean diferentes de cero. Otras reglas útiles incluyen

$$-a = (-1)a$$

y más generalmente

$$-(ab) = (-a)b = a(-b),$$

y también

$$ab = (-a)(-b),$$

así como

$$a \cdot 0 = 0,$$

todas reglas familiares de la aritmética elemental.

A.1.1 *Un cuerpo finito*

A modo de ejemplo, y para entrenar la intuición de que un cuerpo no necesariamente tiene un número infinito de elementos, consideremos el conjunto con dos elementos $\mathbb{F}_2 = \{0, 1\}$. Definimos la suma $+ \colon \mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{F}_2$ mediante la regla

$$0 + 0 = 0, \qquad 0 + 1 = 1, \qquad 1 + 0 = 1, \qquad 1 + 1 = 0$$

y el producto $\cdot \colon \mathbb{F}_2 \times \mathbb{F}_2 \to \mathbb{F}_2$ como

$$0 \cdot 0 = 0, \qquad 0 \cdot 1 = 0, \qquad 1 \cdot 0 = 0, \qquad 1 \cdot 1 = 1.$$

Dejamos como ejercicio para el lector comprobar que estas operaciones así definidas satisfacen los axiomas **I1** a **I7** y por lo tanto $\mathbb{F}_2$ es un cuerpo, con dos elementos.

Observación. El lector suspicaz reconocerá en estas operaciones a la suma y el producto definidos en el conjunto $\mathbb{Z}_2 = \{0, 1\}$ de congruencias módulo 2 definido en Álgebra I / Matemática Discreta I. En efecto, resultados desarrollados en ese curso permiten demostrar que los conjuntos $\mathbb{Z}_p$, con p primo, son ejemplos de cuerpos, en este caso con p elementos.

Ejemplo. Sea p un número primo y

$$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$$

el conjunto de restos de dividir por p. Definimos suma y producto en $\mathbb{Z}_p$ de la siguiente manera: sean $a, b \in \mathbb{Z}_p$, entonces

$$\begin{aligned} a + b &= c \quad \text{si} \quad a + b \equiv c \pmod{p} \ \wedge \ 0 \leqslant c \leqslant p-1, \\ a \cdot b &= d \quad \text{si} \quad a \cdot b \equiv d \pmod{p} \ \wedge \ 0 \leqslant d \leqslant p-1. \end{aligned}$$

No es complicado, usando lo que conocemos de congruencia, probar que $\mathbb{Z}_p$ es un cuerpo. La única propiedad cuya prueba no es obvia es **I7**, la existencia de inverso. Esta propiedad se deduce del teorema que enuncia la existencia de soluciones de la ecuación lineal de congruencia.

A.2 NÚMEROS COMPLEJOS

La ecuación polinómica $x^2 + 1 = 0$ (¿cuál es el número que elevado al cuadrado y adicionado 1 da 0?) no tiene solución dentro del cuerpo de los números reales, pues todos sabemos que $x^2 \geqslant 0$ para todo $x \in \mathbb{R}$ y por lo tanto $x^2 + 1 > 0 \ \forall \ x \in \mathbb{R}$. Sin embargo, podemos extender $\mathbb{R}$ a otro cuerpo, de tal forma que *toda* ecuación polinómica con coeficientes en $\mathbb{R}$ tenga solución.

Definición A.2.1. Los *números complejos* es el conjunto $\mathbb{C}$ de los pares ordendados (a, b), denotados $a + ib$, con a, b en $\mathbb{R}$, con las operaciones '$+$' y '$\cdot$', definidas

$$(a + ib) + (c + id) := (a + c) + i(c + d), \tag{A.2.1}$$

$$(a + ib) \cdot (c + id) := (ac - bd) + i(ad + bc). \tag{A.2.2}$$

Al número complejo $i = 0 + i \cdot 1$ lo llamamos el *imaginario puro*. Si $z = a + ib$ es un número complejo, diremos que a es la *parte real* de z y la denotamos $a = \operatorname{Re} z$. Por otro lado, b es la *parte imaginaria* de z que es denotada $b = \operatorname{Im} z$.

Es claro que $z = a + ib$ es igual a $w = c + id$ si coinciden su parte real e imaginaria, es decir

$$a + bi = c + di \quad \Leftrightarrow \quad a = c \ \wedge \ b = d.$$

Podemos ver a $\mathbb{R}$ contenido en $\mathbb{C}$, con la correspondencia $a \to a + i \cdot 0$ y observamos que si nos restringimos a $\mathbb{R}$, tenemos las reglas de adición y multiplicación usuales.

La definición de la suma de dos números complejos no debería sorprendernos, pues es la suma "coordenada a coordenada". La definición del producto se basa en que deseamos que $i^2 = -1$, es decir que i sea la solución de la ecuación polinómica $x^2 + 1 = 0$, y que el producto sea distributivo.

Primero, comprobemos que $i^2 = -1$. Esto es debido a que

$$i^2 = (0 + i \cdot 1)(0 + i \cdot 1) = (0 \cdot 0 - 1 \cdot 1) + i(0 \cdot 1 + 1 \cdot 0) = -1,$$

y por lo tanto $i^2 + 1 = -1 + 1 = 0$.

Sean $0 = 0 + i \cdot 0, 1 = 1 + i \cdot 0 \in \mathbb{C}$, es fácil comprobar que son los elementos neutros de la suma y el producto, respectivamente. Por otro lado, si $z = a + ib$, entonces $-z = -a - ib$ es el opuesto aditivo de z. El inverso multiplicativo es un poco más complicado. Primero observemos que dado $a + ib \in \mathbb{C}$,

$$(a + ib)(a - ib) = aa - b(-b) = a^2 + b^2 \in \mathbb{R}.$$

Supongamos que $a + ib \neq 0$, encontremos a partir de las reglas de adición y multiplicación la inversa de z. Sea $c + id$ tal que $(a + ib)(c + id) = 1$, luego

$$\begin{aligned} c + id = \frac{1}{a + ib} = \frac{1}{a + ib}\frac{a - ib}{a - ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2} \\ = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2} \end{aligned}$$

(observar que como $a + ib \neq 0$, entonces $a^2 + b^2 > 0$.)

Usando lo anterior, y un poco más de trabajo, obtenemos

Proposición A.2.2. *Sean $0 = 0 + i \cdot 0, 1 = 1 + i \cdot 0 \in \mathbb{C}$. Entonces, $\mathbb{C}$ con las operaciones '$+$' y '$\cdot$', definidas en (A.2.1) y (A.2.2), respectivamente, es un cuerpo con elementos neutros 0 y 1, y*

$$\begin{aligned} -(a + ib) &= -a - ib \\ (a + ib)^{-1} &= \frac{a - ib}{a^2 + b^2}, \qquad \text{para } a + ib \neq 0. \end{aligned}$$

Demostración. Ejercicio. □

Hemos definido los números complejos como pares ordenados y como tales es posible representarlos en el plano $\mathbb{R} \times \mathbb{R}$:

Por el teorema de Pitágoras, la distancia del número complejo $a + ib$ al 0 es $\sqrt{a^2 + b^2}$.

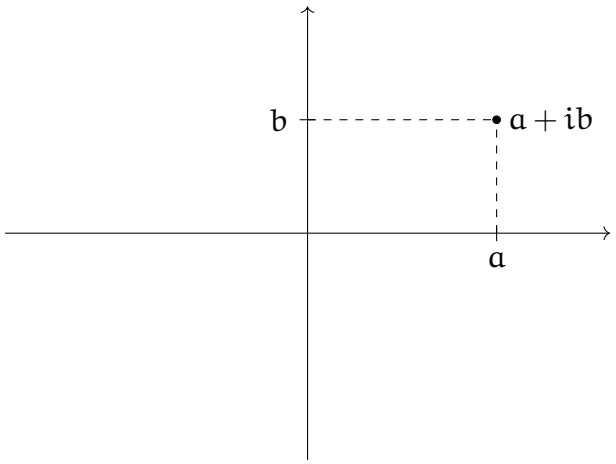

Figura 1: Representación gráfica de los números complejos.

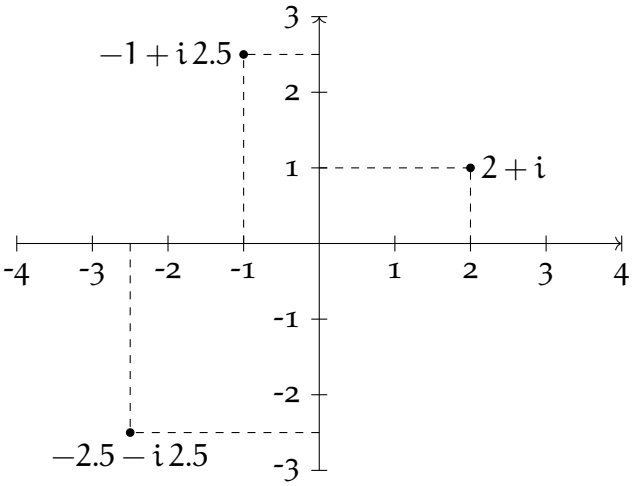

Figura 2: Ejemplos de la representación gráfica de los números complejos.

Definición A.2.3. Sea $z = a + ib \in \mathbb{C}$. El *módulo* de z es

$$|z| = \sqrt{a^2 + b^2}.$$

El *conjugado* de z es

$$\bar{z} = a - ib.$$

Ejemplo. $|4 + 3i| = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$, $\overline{4 + 3i} = 4 - 3i$.

Proposición A.2.4. *Sean z y w números complejos.*

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

(2) *Si* $z \neq 0$, $z^{-1} = \dfrac{\overline{z}}{|z|^2}$.

(3) $\overline{z + w} = \overline{z} + \overline{w}$.

(4) $\overline{zw} = \overline{z}\,\overline{w}$.

Demostración. Son comprobaciones rutinarias. Para ejemplificar, hagamos la demostración de (4).

Si $z = a + bi$ y $w = c + di$, entonces $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$. Por lo tanto,

$$\overline{zw} = (ac - bd) - (ad + bc)i.$$

Como $\overline{z} = a - bi$ y $\overline{w} = c - di$,

$$\overline{z}\,\overline{w} = (ac - (-b)(-d)) + (a(-d) + b(-c))i = (ac - bd) - (ad + bc)i.$$

Por lo tanto $\overline{zw} = \overline{z}\,\overline{w}$. □

Ejercicio. Determinar el número complejo $2 - 3i + \dfrac{i}{1-i}$.

Solución. El ejercicio nos pide que escribamos el número en el formato $a + bi$, con $a, b \in \mathbb{R}$. En general, para eliminar un cociente donde el divisor tiene parte imaginaria no nula, multiplicamos arriba y abajo por el conjugado del divisor, como $z\overline{z} \in \mathbb{R}$, obtenemos un divisor real. En el ejemplo:

$$\begin{aligned} 2 + 3i + \frac{i}{1-i} &= 2 + 3i + \frac{i}{1-i} \cdot \frac{1+i}{1+i} \\ &= 2 + 3i + \frac{i(1+i)}{(1-i)(1+i)} \\ &= 2 + 3i + \frac{i-1}{2} \\ &= 2 + 3i + \frac{i}{2} - \frac{1}{2} \\ &= \frac{3}{2} + i\frac{7}{2} \end{aligned}$$

□

Un poco de trigonometría. Recordemos que dado un punto $p = (x, y)$ en el plano, la recta que une el origen con p determina un ángulo θ con el eje x y entonces

$$x = r\,\mathrm{sen}(\theta), \qquad y = r\cos(\theta)$$

donde r es la longitud del segmento determinado por $(0, 0)$ y (x, y). En el lenguaje de los números complejos, si $z = a + bi$ y θ el ángulo determinado por z y el eje horizontal, entonces

$$a = |z|\,\mathrm{sen}(\theta), \qquad b = |z|\cos(\theta),$$

es decir

$$z = |z|(\cos(\theta) + i\,\mathrm{sen}(\theta)). \tag{A.2.3}$$

Si $z \in \mathbb{C}$, la fórmula (A.2.3) e llamada la *forma polar* de z y θ es llamado el *argumento de* z.

Notación exponencial. Otra notación para representar a los números complejos es la *notación exponencial*, en la cual se denota

$$e^{i\theta} = \cos(\theta) + i\,\mathrm{sen}(\theta). \tag{A.2.4}$$

Por lo tanto si $z \in \mathbb{C}$ y θ es el argumento de z,

$$z = re^{i\theta}$$

donde $r = |z|$. No perder de vista, que la notación exponencial no es más que una notación (por ahora).

Proposición A.2.5. *Sean* $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$, *entonces*

$$z_1 z_2 = r_1 r_2\, e^{i(\theta_1+\theta_2)}.$$

Demostración. $z_1 = r_1(\cos(\theta_1) + i\,\mathrm{sen}(\theta_1))$, $z_2 = r_2(\cos(\theta_2) + i\,\mathrm{sen}(\theta_2))$, luego

$$\begin{aligned} z_1 z_2 &= r_1 r_2(\cos(\theta_1) + i\,\mathrm{sen}(\theta_1))(\cos(\theta_2) + i\,\mathrm{sen}(\theta_2)) \\ &= r_1 r_2(\cos(\theta_1)\cos(\theta_2) + i\cos(\theta_1)\,\mathrm{sen}(\theta_2) + i\,\mathrm{sen}(\theta_1)\cos(\theta_2) \\ &\qquad + i^2\,\mathrm{sen}(\theta_1)\,\mathrm{sen}(\theta_2)) \\ &= r_1 r_2((\cos(\theta_1)\cos(\theta_2) - \mathrm{sen}(\theta_1)\,\mathrm{sen}(\theta_2)) + i(\,\mathrm{sen}(\theta_1)\cos(\theta_2) \\ &\qquad + \cos(\theta_1)\,\mathrm{sen}(\theta_2))) \\ &\overset{(*)}{=} r_1 r_2(\cos(\theta_1+\theta_2) + i\,\mathrm{sen}(\theta_1+\theta_2)) = r_1 r_2 e^{i(\theta_1+\theta_2)}. \end{aligned}$$

La igualdad $(*)$ se debe a las tradicionales fórmulas trigonométrica del coseno y seno de la suma de ángulos. □

Observación (Justificación de la notación exponencial). Los alumnos que conozcan las series de Taylor reconocerán inmediatamente las fórmulas

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \tag{*}$$

donde x es un número real y

$$\begin{aligned} \cos(\theta) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\theta^{2k} \\ \mathrm{sen}(\theta) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\theta^{2k+1}, \end{aligned}$$

donde $0 \leqslant \theta < 2\pi$. Ahora bien, remplacemos x por $i\theta$ en la fórmula (*) y obtenemos

$$\begin{aligned} e^{i\theta} &= \sum_{n=0}^{\infty} \frac{1}{n!}(i\theta)^n \\ &= \sum_{k=0}^{\infty} \frac{1}{(2k)!}(i\theta)^{2k} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}(i\theta)^{2k+1}. \end{aligned} \tag{**}$$

No es difícil ver que $i^{2k} = (-1)^k$ y por lo tanto $i^{2k+1} = i^{2k} \cdot i = (-1)^k i$. Luego, por (**),

$$\begin{aligned} e^{i\theta} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\theta^{2k} + i\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\theta^{2k+1} \\ &= \cos(\theta) + i\,\mathrm{sen}(\theta), \end{aligned}$$

recuperando así la fórmula (A.2.4), llamada *fórmula de Euler.*

Observación (Identidad de Euler). Observemos que especializando la fórmula de Euler en π obtenemos

$$e^{i\pi} = \cos(\pi) + i\,\mathrm{sen}(\pi) = -1.$$

Escrito de otra forma

$$e^{i\pi} - 1 = 0. \tag{A.2.5}$$

Esta última expresión es denominada la *identidad de Euler* y es considerada una de las fórmulas más relevantes de la matemática, pues comprende las cinco constantes matemáticas más importantes:

(1) El número **0**.

(2) El numero **1**.

(3) El número $\boldsymbol{\pi}$, número irracional que es la relación entre la circunferencia de un círculo y su diámetro. Es aproximadamente **3.14159 . . .**.

(4) El número $\boldsymbol{e}$, también un número irracional. Es la base de los logaritmos naturales y surge naturalmente a través del estudio del interés compuesto y el cálculo. El número $\boldsymbol{e}$ está presente en una gran cantidad de ecuaciones importantes. Es aproximadamente **2.71828 . . .**.

(5) El número $\boldsymbol{i}$, el más fundamental de los números imaginarios.

B

FUNCIONES POLINÓMICAS

En este apéndice se definirán las funciones polinómicas y se mostrarán algunas de sus propiedades fundamentales. Trabajaremos sobre $\mathbb{K}$ cuerpo con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

B.1 DEFINICIÓN DE FUNCIONES POLINÓMICAS

Definición B.1.1. Una función $f : \mathbb{K} \to \mathbb{K}$ es *polinomial* o *polinómica* o directamente decimos que f es un *polinomio*, si existen $a_0, a_1, \ldots, a_n \in \mathbb{K}$ tal que

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \tag{B.1.1}$$

para todo $x \in \mathbb{K}$. En este caso diremos que f *tiene grado* $\leqslant n$. Si $a_n \neq 0$ diremos que f *tiene grado* n y se denota $\operatorname{gr}(f) = n$.

En el caso del polinomio 0, el grado no está definido y se usa la convención $\operatorname{gr}(0) = -\infty$.

Diremos también que $a_0, \ldots, a_n$ son los *coeficientes* de f, a_0 es el *término constante* de f y a_n el *coeficiente principal.*

Observación B.1.2. Para la definición formal de función polinómica o polinomio deberíamos ser más cuidadosos, pues en realidad no sabemos a priori si la escritura de una función polinómica es única. Es decir, existe la posibilidad de f se escriba de otra forma y, en particular, el coeficiente más significativo sea diferente. No es muy complicado demostrar que esto no puede ocurrir, pero no lo haremos en este apunte.

Sea f un polinomio. Si c es un número tal que $f(c) = 0$, entonces llamamos a c *una raíz de* f. Veremos en un momento que un polinomio distinto de cero puede tener solo un número finito de raíces, y daremos un límite para la cantidad de estas raíces.

Ejemplo. Sea $f(x) = x^2 - 3x + 2$. Entonces $f(1) = 0$ y por lo tanto, 1 es una raíz de f. Además, $f(2) = 0$. Por lo tanto, 2 es también una raíz de f.

Ejemplo. Sean $a, b, c \in \mathbb{R}$ y $f(x) = ax^2 + bx + c$, un polinomio en $\mathbb{R}$. Si $b^2 - 4ac = 0$, entonces el polinomio tiene una raíz real, que es

$$-\frac{b}{2a}.$$

Si $b^2 - 4ac > 0$, entonces el polinomio tiene dos raíces reales distintas que son

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{y} \quad \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

En el caso que $b^2 - 4ac < 0$ el polinomio no tiene raíces reales.

Teorema B.1.3. *Sea* f *un polinomio de grado* $\leqslant n$ *y sea* c *una raíz. Entonces existe un polinomio* g *de grado* $\leqslant n-1$ *tal que para todo* x *se cumple*

$$f(x) = (x-c)g(x).$$

Demostración. Escribamos $f(x)$ en función de las potencias de x:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Veremos a continuación que f puede también escribirse en potencias de $x-c$: escribamos

$$x = (x-c) + c,$$

luego

$$f(x) = a_n((x-c)+c)^n + a_{n-1}((x-c)+c)^{n-1} + \cdots + a_1((x-c)+c) + a_0.$$

Expandiendo las potencias de los binomios $((x-c)+c)^k$ $(1 \leqslant k \leqslant n)$, obtenemos

$$f(x) = b_n(x-c)^n + b_{n-1}(x-c)^{n-1} + \cdots + b_1(x-c) + b_0,$$

para ciertos $b_0, b_1, \ldots, b_n \in \mathbb{K}$. Como $f(c) = 0$, entonces $0 = f(c) = b_0$, luego

$$\begin{aligned} f(x) &= b_n(x-c)^n + b_{n-1}(x-c)^{n-1} + \cdots + b_1(x-c) \\ &= (x-c)(b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1) \\ &= (x-c)g(x), \end{aligned}$$

con $g(x) = b_n(x-c)^{n-1} + b_{n-1}(x-c)^{n-2} + \cdots + b_1$, que es una función polinómica de grado $\leqslant n-1$, y vemos que nuestro teorema está probado. □

El polinomio f es el *polinomio nulo* si $f(x) = 0$ para toda $x \in \mathbb{K}$. Si f es el polinomio nulo, denotamos $f = 0$.

Teorema B.1.4. *Sea* f *un polinomio de grado* $n \geqslant 0$*, entonces* f *tiene a lo más* n *raíces.*

Demostración. Sea

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

con $a_n \neq 0$.

Probaremos el resultado haciendo inducción sobre n.

Si $n = 0$, $a_0 \neq 0$, es decir $f(x) = a_0 \neq 0$, que es lo que teníamos que probar (f no tiene raíces).

Sea $n > 0$. Sea c raíz de f. Por el teorema B.1.3,

$$f(x) = (x - c)g(x),$$

con

$$g(x) = b_{n-1}x^{n-1} + \cdots + b_1x + b_0.$$

Es claro que $b_{n-1} = a_n \neq 0$ y por lo tanto, por hipótesis inductiva, $g(x)$ tiene a lo más $n-1$ raíces. Ahora bien

$$0 = f(x) = (x - c)g(x) \quad \Leftrightarrow \quad x - c = 0 \text{ o } g(x) = 0.$$

Es decir x es raíz de f si y solo si $x = c$ o x es raíz de g. Como g tiene a lo más $n-1$ raíces, f tiene a lo más n raíces. □

Observemos que si f y g son polinomios con

$$f(x) = a_nx^n + \cdots + a_1x + a_0 \quad \text{y} \quad g(x) = b_nx^n + \cdots + b_1x + b_0,$$

entonces como $ax^i + bx^i = (a+b)x^i$, tenemos que $f + g$ es un polinomio definido por

$$(f + g)(x) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

Por otro lado, debido a que $(ax^i)(bx^j) = abx^{i+j}$, el producto de dos polinomios también es un polinomio y el cálculo de los coeficientes de fg se hace aplicando la propiedad distributiva. Más precisamente,

$$(fg)(x) = a_nb_mx^{n+m} + (a_{n-1}b_m + a_nb_{m-1})X^{m+n-1} + \cdots .$$

Proposición B.1.5. *Sean f y g polinomios de grado n y m, respectivamente. Entonces fg es un polinomio de grado $n + m$*

Demostración. Sean

$$f(x) = a_nx^n + \cdots + a_1x + a_0 \quad \text{y} \quad g(x) = b_mx^m + \cdots + b_1x + b_0,$$

con $a_n, b_m \neq 0$. Entonces,

$$(fg)(x) = a_nb_mx^{n+m} + h(x), \tag{B.1.2}$$

con $h(x)$ un polinomio de grado menor a $n + m$. Por lo tanto, el coeficiente principal de fg es $a_nb_m \neq 0$ y, en consecuencia fg tiene grado $n + m$. □

Ejemplo. Sean $f(x) = 4x^3 - 3x^2 + x + 2$ y $g(x) = x^2 + 1$. Entonces,

$$\begin{aligned}(f + g)(x) &= (4 + 0)x^3 + (-3 + 1)x^2 + (1 + 0)x + (2 + 1)\\ &= 4x^3 - 2x^2 + x + 3,\end{aligned}$$

y

$$\begin{aligned}(fg)(x) &= (4x^3 - 3x^2 + x + 2)(x^2 + 1)\\ &= (4x^3 - 3x^2 + x + 2)x^2 + (4x^3 - 3x^2 + x + 2)1\\ &= 4x^5 - 3x^4 + x^3 + 2x^2 + 4x^3 - 3x^2 + x + 2\\ &= 4x^5 - 3x^4 + 5x^3 - x^2 + x + 2\end{aligned}$$

B.2 DIVISIÓN DE POLINOMIOS

Si f y g son polinomios, entonces no necesariamente la función f/g está bien definida en todo punto y puede que tampoco sea un polinomio. Cuando trabajamos con enteros, en cursos anteriores, probamos la existencia del algoritmo de división, más precisamente.

Sean n, d enteros positivos. Entonces existe un entero r tal que $0 \leqslant r < d$ un entero $q \geqslant 0$ tal que

$$n = qd + r.$$

Ahora describiremos un procedimiento análogo para polinomios.

Algoritmo de División. *Sean f y g polinomios distintos de cero. Entonces existen polinomios q, r tales que $\mathrm{gr}(r) < \mathrm{gr}(g)$ y tales que*

$$f(x) = q(x)g(x) + r(x).$$

A $q(x)$ lo llamamos el *cociente* de la *división polinomial* y a $r(x)$ lo llamamos el *resto* de la división polinomial.

No veremos aquí la demostración del algoritmo de división, basta decir que es muy similar a la demostración del algoritmo de división para números enteros. En los siguientes ejemplos se verá como se calculan el cociente y resto de la división polinomial.

Ejemplo. Sean $f(x) = 4x^3 - 3x^2 + x + 2$ y $g(x) = x^2 + 1$. Para encontrar la división polinomial, debemos multiplicar por un monomio ax^k a $g(x)$ de tal forma que el coeficiente principal de $ax^k g(x)$ sea igual al coeficiente principal de $f(x)$. En este caso, multiplicamos a $g(x)$ por $4x$ y nos queda

$$f(x) = 4xg(x) + r_1(x) = (4x^3 + 4x) + (-3x^2 - 3x + 2)$$

Ahora, con $r_1(x) = -3x^2 - 3x + 2$ hacemos el mismo procedimiento, es decir multiplicamos por -3 a $g(x)$ y vemos que es lo que "falta":

$$r_1(x) = (-3)g(x) + r(x) = (-3x^2 - 3) + (-3x + 5).$$

Como $r(x) = -3x + 5$ tiene grado menor que 2, tenemos que

$$\begin{aligned} f(x) &= 4xg(x) + r_1(x) \\ &= 4xg(x) + (-3)g(x) + r(x) \\ &= (4x - 3)g(x) + r(x). \end{aligned}$$

Es decir,

$$f(x) = q(x)g(x) + r(x),$$

con $q(x) = 4x - 3$ y $r(x) = -3x + 5$.

Observemos que se puede hacer un esquema parecido a la división de números enteros, el cual nos facilita el cálculo:

$$\begin{array}{rrrl} 4x^3 - 3x^2 & +x+2 & = (x^2+1)(4x-3) - 3x + 5 \\ \underline{-4x^3 \qquad} & \underline{-4x} & \\ -3x^2 - 3x & +2 & \\ \underline{3x^2 \qquad} & \underline{+3} & \\ -3x & +5 & \end{array}$$

Ejemplo. Sean

$$f(x) = 2x^4 - 3x^2 + 1 \quad \text{y} \quad g(x) = x^2 - x + 3.$$

Deseamos encontrar $q(x)$ y $r(x)$ como en el algoritmo de Euclides. Haciendo la división como en el ejercicio anterior:

$$\begin{array}{rrrrl} 2x^4 & -3x^2 & & +1 & = (x^2-x+3)(2x^2+2x-7) - 13x + 22 \\ \underline{-2x^4 + 2x^3} & \underline{-6x^2} & & & \\ 2x^3 & -9x^2 & & & \\ \underline{-2x^3 + 2x^2} & \underline{-6x} & & & \\ -7x^2 & -6x & +1 & & \\ \underline{7x^2} & \underline{-7x} & \underline{+21} & & \\ -13x & +22 & & & \end{array}$$

Es decir $q(x) = 2x^2 + 2x - 7$ y $r(x) = -13x + 22$.

Observemos que el algoritmo de división nos dice que si dividimos un polinomio por uno de grado 1, entonces el resto es una constante (que puede ser 0). Más aún:

Teorema B.2.1 (Teorema del resto)**.** *Sea f polinomio y $c \in \mathbb{K}$. Entonces, el resto de dividir f por $x - c$ es $f(c)$.*

Demostración. Por el algoritmo de Euclides

$$f(x) = q(x)(x-c) + r,$$

con r de grado < 1, es decir $r \in \mathbb{K}$. Ahora bien

$$f(c) = q(c)(c-c) + r = r,$$

luego $f(c)$ es el resto de dividir f por $x - c$. □

Observar que esto nos da otra prueba del teorema B.1.4: $f(c) = 0$, luego por teorema del resto $f(x) = q(x)(x-c)$.

C

MULTIPLICACIÓN DE POLINOMIOS POR FFT

Como vimos en B.1 el producto de polinomios se calcula usando que $x^i x^j = x^{i+j}$ y la propiedad distributiva. Si un polinomio tiene grado n y el otro tiene grado m, entonces son necesarias nm multiplicaciones de coeficientes ("todos contra todos").

También puede plantearse de esta forma: si necesitamos multiplicar polinomios de grado n entonces la multiplicación de dos polinomios requiere n^2 multiplicaciones. Como la multiplicación es la operación más costosa del procedimiento, podemos decir que multiplicar dos polinomios de grado n requiere *alrededor de* n^2 *operaciones.*

Este nivel de complejidad (n^2) parece ser razonable a nivel computacional, pero si los polinomios a multiplicar tiene grados muy altos puede ser necesario contar con métodos más rápidos, o que requieran menos operaciones. En este apéndice mostraremos la multiplicación de polinomios usando la trasformada de Fourier discreta implementándola con la transformada rápida de Fourier (FFT) y mostraremos que usando este método se puede multiplicar dos polinomios de grado n en *alrededor de* $n \log_2(n)$ *operaciones.*

C.1 REPRESENTACIÓN DE POLINOMIOS POR VALORES

La primero observación importante es que todo polinomio de grado $< n$ está determinado por n valores que toma.

Proposición C.1.1. *Sa* f *un polinomio de grado menor que* n *y* $x_0, \ldots, x_{n-1} \in \mathbb{R}$ *todos distintos entre sí. Sea* $y_i = f(x_i)$, $0 \leqslant i < n$. *Si* g *polinomio de grado menor que* n *tal que* $g(x_i) = y_i$ *con* $0 \leqslant i < n$, *entonces se cumple que* $g = f$.

Demostración. Sea $h = f - g$, es claro que $\operatorname{gr}(h) < n$. Si $h \neq 0$, por la proposición B.1.4, h tiene a lo más $n-1$ raíces. Sin embargo $h(x_i) = f(x_i) - g(x_i) = y_i - y_i = 0$, es decir h tiene al menos n raíces. Esto provoca un absurdo que vino de suponer que $h \neq 0$. Por lo tanto, $h = 0$ y en consecuencia $f = g$. □

Definición C.1.2. Sea $n \in \mathbb{N}$ y $X = [x_0, \ldots, x_{n-1}]$ un conjunto ordenado de n puntos distintos. Si

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

un polinomio de grado menor que n, diremos que $[a_0, a_1, \ldots, a_{n-1}]$ es la *representación por coeficientes* de f y que $[f(x_0), f(x_1), \ldots, f(x_{n-1})]$ es la *representación por valores* de f (respecto a X).

Debido a la proposición C.1.1 es claro que una representación por valores de un polinomio lo determina unívocamente. La transformada de Fourier rápida es un método eficiente para calcular la representación por valores de un polinomio a partir de la representación por coeficientes. El mismo método pero con una pequeña modificación nos devuelve la representación por coeficientes a partir de una representación por valores. Ahora bien ¿para qué nos sirve esto para multiplicar polinomios? La respuesta la da la proposición siguiente. Este resultado se basa en la sencilla idea que si x_0 es un número, entonces $(fg)(x_0) = f(x_0)g(x_0)$, es decir calcular el producto de dos polinomios representados por valores conlleva un número de operaciones similares a la cantidad de valores evaluados.

Proposición C.1.3. *Sea $n \in \mathbb{N}$ y $X = [x_0, \ldots, x_n]$ un conjunto ordenado de n puntos distintos y sean f, g polinomios de grado menor que $n/2$ con representación por valores $[y_0, y_1, \ldots, y_n]$ y $[z_0, z_1, \ldots, z_n]$, respectivamente. Entonces la representación por valores de fg es $[y_0z_0, y_1z_1, \ldots, y_nz_n]$*

Demostración. Como $y_i = f(x_i)$, $z_i = g(x_i)$, es claro que $(fg)(x_i) = f(x_i)g(x_i) = y_iz_i$. Como $gr(f), gr(g) < n/2$, entonces $gr(fg) < n$ y por lo tanto $[y_0z_0, y_1z_1, \ldots, y_nz_n]$ determina unívocamente fg. □

La idea entonces para multiplicar polinomios usando la transformada rápida de Fourier es: sean f, g polinomios de grado $< n$,

(1) Calcular FFT(f) y FFT(g) (del orden de $2n \log_2(2n)$ operaciones). Esto nos devuelve una representación por valor de f y g.

(2) Calcular la representación por valor de fg haciendo el producto coordenada a coordenada de las representaciones por valor de f y g (del orden de $2n$ operaciones) .

(3) Calcular IFFT(fg), la inversa de FFT, que devuelve la representación por coeficientes de fg (del orden de $2n \log_2(2n)$ operaciones).

Implementando lo anterior, la cantidad de operaciones para multiplicar dos polinomios de grado $< n$ es $n \log_2(n)$ (salvo suma y multiplicación por constantes), que en la práctica y para n grande es *mucho* menor que n^2, el número de operaciones requeridas si se hiciera la multiplicación de la forma usual.

C.2 TRANSFORMADA DE FOURIER DISCRETA

La series de Fourier permiten representar una función periódica y continua a trozos como una combinación de funciones armónicas puras. Son usadas en muchas ramas de la ingeniería, además de ser una herramienta sumamente útil en la matemática abstracta. Sus áreas de aplicación incluyen análisis vibratorio, acústica, óptica, procesamiento de imágenes y señales, y compresión de datos.

Teorema C.2.1. *Sea* $f : \mathbb{R} \to \mathbb{R}$ *una función de período* 1 *y continua a trozos, entonces podemos escribir de una única forma*

$$f(x) = \sum_{j=-\infty}^{\infty} c_j e^{2\pi i j x}, \tag{C.2.1}$$

con $c_j \in \mathbb{C}$ *para* $j \in \mathbb{Z}$.

La demostración del teorema anterior se basa en una generalización a espacios de dimensión infinita de los conceptos de bases ortonormales en un espacio vectorial. Diremos que la serie de la (C.2.1) es la *serie de Fourier de* f.

Ahora bien, en el mundo de la computación no es posible trabajar con funciones continuas y series y nos debemos restringir a valores de una función y sumas finitas, respectivamente.

La discretización de teoremas análogos al teorema C.2.1 que nos permitan trabajar con computadoras ha llevado a los matemáticos a definir la transformada de Fourier discreta.

Definición C.2.2. La *transformada de Fourier discreta* transforma una secuencia de n números complejos $f_0, f_1, \ldots, f_{n-1}$ en otra secuencia de n números complejos:

$$c_k = \sum_{j=0}^{n-1} f_j e^{-2\pi i j k/n} \qquad (0 \leqslant k \leqslant n-1).$$

Ejemplo C.2.3. Sea f un polinomios de grado $n-1$,

$$f = f_0 + f_1 x + \cdots + f_{n-1} x^{n-1}.$$

Podemos representar f como una n-upla con sus coeficientes:

$$f = (f_0, f_1, f_2, \ldots, f_{n-1}).$$

Observemos que la transformada de Fourier discreta de f

$$c = (c_0, c_1, c_2, \ldots, c_{n-1})$$

no es otra cosa que

$$c = (f(1), f(e^{-2\pi i/n}), f(e^{-2\pi i 2/n}), \ldots, f(e^{-2\pi i(n-1)/n})).$$

Es decir la transformada de Fourier discreta de un polinomio f es la representación por valores de f respecto a $X = \{1, e^{-2\pi i/n}, e^{-2\pi i 2/n}, \ldots, e^{-2\pi i(n-1)/n}\}$ (ver la definición C.1.2).

Veremos ahora esta definición desde el punto de vista del álgebra lineal.

Definición C.2.4. Dado $n \in \mathbb{N}$, se llama *raíz n-ésima de la unidad* a cualquiera de los números complejos que satisfacen la ecuación

$$z^n = 1.$$

Para cada n, las n diferentes raíces n-ésimas de la unidad son:

$$e^{2\pi ik/n} \text{ donde } k = 0, 1, 2, \ldots, n-1.$$

Si denotamos $w = e^{2\pi i/n}$, entonces

$$w^0, w^1, \ldots, w^{n-1}$$

son las n raíces n-ésimas de la unidad. Observar que si z es una raíz de la unidad, entonces $\bar{z}$ también lo es y $z\bar{z} = 1$.

Sea

$$F = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \vdots & & & & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{bmatrix}.$$

Es decir

$$[F]_{jk} = w^{(j-1)(k-1)}.$$

Teorema C.2.5. *Para cada $n \in \mathbb{N}$ la matrices F y $\overline{F}$, la matriz conjugada de F, satisfacen*

$$F\overline{F} = \overline{F}F = nI_n \quad \textit{o, equivalentemente,} \quad F^{-1} = \frac{1}{n}\overline{F}.$$

Además, la transformada de Fourier discreta de la secuencia de números complejos $f = (f_0, \ldots, f_{n-1})$ es $c = \frac{1}{n}\overline{F}f$.

Demostración. Probemos primero que $F\overline{F} = nI_n$. Observemos que el producto de la fila $j+1$ con la fila de F con la columna $k+1$ de $\overline{F}$ $(0 \leqslant j, k < n)$ es

$$1 \cdot 1 + w^j\overline{w}^k + w^{2j}\overline{w}^{2k} + \cdots + w^{(n-1)j}\overline{w}^{(n-1)k}. \tag{C.2.2}$$

Si $j = k$ entonces todos los términos de la suma son 1 y por lo tanto la expresión C.2.2 es igual a n. Si $j \neq k$, denotemos $r = w^j\overline{w}^k$, entonces la expresión C.2.2 es igual a la serie geométrica

$$1 + r + r^2 + \cdots + r^{n-1} = \frac{r^n - 1}{r - 1}.$$

Ahora bien, como $1 = w^n = \overline{w}^n$, es claro que $r^n = 1$ y por lo tanto $r^n - 1/r - 1$ es 0.

Para probar la segunda afirmación del teorema multiplicamos la matriz $\overline{F}$ por el vector f:

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \overline{w} & \overline{w}^2 & \cdots & \overline{w}^{n-1} \\ 1 & \overline{w}^2 & \overline{w}^4 & \cdots & \overline{w}^{2(n-1)} \\ \vdots & & & & \vdots \\ 1 & \overline{w}^{n-1} & \overline{w}^{2(n-1)} & \cdots & \overline{w}^{(n-1)^2} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} \sum_{j=0}^{n-1} f_j \\ \sum_{j=0}^{n-1} \overline{w}^j f_j \\ \sum_{j=0}^{n-1} \overline{w}^{2j} f_j \\ \vdots \\ \sum_{j=0}^{n-1} \overline{w}^{(n-1)j} f_j \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix}.$$

Reacomodando cada sumatoria y considerando que $\overline{w} = e^{-2\pi i/n}$ obtenemos

$$c_k = \sum_{j=0}^{n-1} f_j e^{-2\pi ijk/n},$$

para $0 \leqslant k \leqslant n-1$, que es lo que queríamos probar. □

La matriz F se la llama la *matriz de Fourier* y, por lo visto en el teorema anterior, la transformada de Fourier discreta de f es $c = F^{-1}f = \overline{F}f/n$.

En base al teorema podemos dar una definición equivalente de la transformada de Fourier discreta y definir la antitransformada.

Definición C.2.6. Sea F la matriz de Fourier $n \times n$. Sea $f = (f_0, \ldots, f_{n-1}), \in \mathbb{C}^n$, entonces la *transformada de Fourier discreta (DFT)* de f es $c = F^{-1}f$. La *transformada inversa de Fourier discreta (IDFT)* de $c = (c_0, \ldots, c_{n-1}) \in \mathbb{C}^n$ es Fc.

Convolución discreta

Ahora bien, ¿por qué es importante la matriz de Fourier F? Una posible explicación es la siguiente: hay un cálculo que aparece constantemente en las aplicaciones y hay dos formas de hacerlo. El método directo se incluye en la definición. El método indirecto usa F y F^{-1} y, aunque es más complicado, se puede implementar para que sea mucho más rápido que el cálculo directo. El cálculo se llama convolución, y la regla que permite realizarlo mediante F y F^{-1} es la regla de convolución.

Definición C.2.7. Sean $f = (f_0, \ldots, f_{n-1})$ y $g = (g_0, \ldots, g_{n-1})$ dos vectores en $\mathbb{C}^n$, entonces la *convolución* de f y g es

$$f * g = \left(\sum_{j+k \equiv 0\,(n)} f_j q_k, \sum_{j+k \equiv 1\,(n)} f_j q_k, \ldots, \sum_{j+k \equiv n-1\,(n)} f_j q_k \right).$$

Ejemplo C.2.8. La convolución entre $(1,2,3)$ y $(4,5,6)$ Es

$$\begin{aligned} (1,2,3) * (4,5,6) &= (1\cdot 4 + 2\cdot 6 + 3\cdot 5, 1\cdot 5 + 2\cdot 4 + 3\cdot 6, 1\cdot 6 + 2\cdot 5 + 3\cdot 4) \\ &= (31, 31, 28). \end{aligned}$$

Ejemplo C.2.9. El ejemplo más notable en el contexto que estamos estudiando es la multiplicación de polinomios, que puede ser vista como una convolución. Veamos un caso especial de dos polinomios de grado 2.

Multiplicar $f_0 + f_1x + f_2x^2$ por $g_0 + g_1x + g_2x^2$ es exactamente como hacer la convolución, con una diferencia esencial: el producto es un polinomio de grado 4. y por lo tanto tiene cinco coeficientes, mientras f y g tienen tres. Dado que la convolución produce una secuencia de salida de la misma longitud que las entradas, agregamos dos ceros a la entrada:

$$f = (f_0, f_1, f_2, 0, 0) \quad \text{y} \quad g = (g_0, g_1, g_2, 0, 0)$$

La convolución de f y g (con $n = 5$) es

$$f * g = (f_0g_0, f_0g_1 + f_1g_0, f_0g_2 + f_1g_1 + f_2g_0, f_1g_2 + f_2g_1, f_2g_2).$$

Es claro que con la convolución obtenemos entonces los coeficientes del producto fg.

Esto se puede generalizar a cualquier grado: para multiplicar dos polinomios f y g de grado $< n/2$ completamos los coeficientes de cada polinomio con 0 hasta grado n y hacemos la convolución. De esa forma obtenemos los coeficientes del polinomio fg.

Teorema C.2.10. *Sean* $f, g \in \mathbb{C}^n$ *y sean* $c = F^{-1}f$, $d = F^{-1}g$, *las transformadas de Fourier discretas de* f *y* g *respectivamente, entonces*

$$f * g = n\,F(cd), \tag{C.2.3}$$

donde cd *indica el producto coordenada a coordenada de* c *por* d.

Demostración. Se deja a cargo del lector. □

La expresión dada por la ecuación (C.2.3) se llama la *regla de convolución.*

Ejemplo. Los ejemplos C.2.3 y C.2.9 y el teorema C.2.10 nos muestran como podemos obtener la multiplicación de polinomios utilizando la transformada de Fourier discreta. Aquí haremos un repaso de como hacerlo.

Sea $n \in \mathbb{N}$ y f, g dos polinomios complejos de grado $< n/2$. Sean

$$f = (f_0, f_1, \ldots, f_{n-1}) \quad \text{y} \quad g = (g_0, g_1, \ldots, g_{n-1})$$

las representaciones por coeficientes de f y g respectivamente (donde, evidentemente, los últimos coeficientes van a ser 0). Sean

$$c = (c_0, c_1, \ldots, c_{n-1}) \quad \text{y} \quad d = (d_0, d_1, \ldots, d_{n-1})$$

las representaciones de f y g, respectivamente, por valores respecto al conjunto $X = \{1, e^{-2\pi i/n}, e^{-2\pi i2/n}, \ldots, e^{-2\pi i(n-1)/n}\}$. Es decir, $c = F^{-1}f$ y $d = F^{-1}g$, luego el producto f y g como funciones polinómicas tiene coeficientes

$$f * g = n\,F(cd).$$

Observar que el hecho de que F sea invertible y que al aplicar transformada discreta de Fourier a un polinomio se obtiene una representación por valores del mismo, junto al teorema C.2.10, hacen que no sea necesario utilizar la proposición C.1.3 para calcular el producto de dos polinomios utilizando la representación por valores.

C.3 TRANSFORMADA RÁPIDA DE FOURIER

La transformada de Fourier rápida, o FFT por sus siglas en inglés, no es nada más que un método eficiente para calcular la transformada de Fourier discreta. La FFT se calcula en forma recursiva y su implementación se basa en ideas ingeniosas que describiremos a lo largo de esta sección.

Nosotros la aplicaremos a polinomios de grado arbitrario, pero es mucho más fácil de explicar cuando el grado de los polinomios es $2^k - 1$ para algún $k \in \mathbb{N}$. Como todo $j \in \mathbb{N}$ cumple que para algún k, $2^{k-1} \leqslant j < 2^k$ es claro que podemos extender el método a cualquier polinomio. Por lo tanto, de ahora en más consideraremos polinomios de grado menor que n donde $n = 2^k$.

Ejemplo. Ejemplificaremos el caso $n = 4$. Es decir, calcularemos la transformada de Fourier discreta para polinomios de grado menor o igual a 3. Sea $f(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3$, calcular $F^{-1} f$ directamente conllevaría 12 multiplicaciones: multiplicar la fila 2 de F^{-1} por f, la fila3 por f y y fila 4 por f (multiplicar por la fila 1 no implica agregar multiplicaciones). Veremos que con el método de la transformada rápida de Fourier podremos obtener la representación de f por valores con muchas menos multiplicaciones.

Como $n = 4$ las raíces cuartas de la unidad son 1, i, -1, $-i$ y la matriz F es

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$$

y

$$4F^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$$

La transformada de Fourier discreta aplicada a f se hace calculando $F^{-1} f = (\overline{F}/n) f$ y devuelve $f(1)$, $f(-i)$, $f(-1)$, $f(i)$, que son los valores que interesan.

Ahora procedemos a escribir f como suma de una función par f_+, y una impar f_-. Es decir:

$$f = f_+ + f_-,$$

con

$$f_+(x) = a_0 + a_2x^2 \quad \text{y} \quad f_-(x) = a_1 + a_3x^3.$$

Si definimos

$$\tilde{f}_+(x) = a_0 + a_2x, \quad \text{y} \quad \tilde{f}_-(x) = a_1 + a_3x,$$

obtenemos entonces que

$$f(x) = \tilde{f}_+(x^2) + x\tilde{f}_-(x^2).$$

Luego,

$$\begin{array}{lcl} f(1) & = & \tilde{f}_+(1) + \tilde{f}_-(1), \\ f(-i) & = & \tilde{f}_+(-1) - i\,\tilde{f}_-(-1) \\ f(-1) & = & \tilde{f}_+(1) - \tilde{f}_-(1), \\ f(i) & = & \tilde{f}_+(-1) + i\,\tilde{f}_-(-1), \\ . & & \end{array} \qquad (*)$$

Las funciones $\tilde{f}_+$ y $\tilde{f}_-$ son de grado 1 y requieren solo una multiplicación para ser calculadas. Por (*), para calcular la representación por valores de f, solo debemos calcular $\tilde{f}_\pm(\pm 1)$, es decir esto nos lleva 4 multiplicaciones. Finalmente, debemos calcular $i \cdot \tilde{f}_-(-1)$ que es una multiplicación más.

Concluyendo: con 5 multiplicaciones, en vez de 12, pudimos calcular la representación de f por valores o, lo que es lo mismo, la transformada de Fourier discreta.

¿Como generalizamos el ejemplo anterior? Una de las claves del ejemplo anterior es que comenzamos trabajando en $n = 4$ valores y redujimos el cálculo a $n/2 = 2$ valores. Veamos como hacemos esto en general.

Sea f es una función, entonces se puede obtener como la suma de una función par y una función impar:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2},$$

luego si

$$f_+ := \frac{f(x) + f(-x)}{2}, \quad \text{y} \quad f_-(x) := \frac{f(x) - f(-x)}{2},$$

tenemos que $f = f_+ + f_-$ donde f_+ es una función par ($f_+(-x) = f_+(x), \forall x$) y f_- es una función impar ($f_-(-x) = -f_-(x), \forall x$).

En el caso que $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ sea un polinomio tenemos:

$$f_+(x) = a_0 + a_2x^2 + a_4x^4 + \cdots +, \qquad f_-(x) = a_1x + a_3x^3 + a_5x^5 + \cdots$$

Luego si definimos

$$\tilde{f}_+(x) = a_0 + a_2x^1 + a_4x^2 + \cdots = \sum_{i<n/2} a_{2i}x^i$$

$$\tilde{f}_-(x) = a_1 + a_3x^1 + a_5x^2 + \cdots = \sum_{i<n/2} a_{2i+1}x^i,$$

obtenemos que

$$f(x) = \tilde{f}_+(x^2) + x\tilde{f}_-(x^2). \tag{C.3.1}$$

Ahora bien, hemos reducido el cálculo de $f(x)$ de grado $< n$ al cálculo de dos funciones de grado $< n/2$, pero veremos a continuación que la ganancia en el tiempo del cálculo se obtiene debido a que los valores donde calculamos f son raíces de la unidad.

Observación C.3.1. Si n par y $w = e^{2\pi i/n}$, entonces $1, w, w^2, \ldots, w^{n-1}$ las raíces n-ésimas de la unidad, y

$$\{(w^2)^k : 0 \leqslant k < n/2\}$$

es el conjunto de las $n/2$-ésimas raíces de la unidad. Más aún, si denotamos $u = \overline{w} = e^{-2\pi i/n}$, entonces $1, u, u^2, \ldots, u^{n-1}$ es también el conjunto de raíces n-ésimas de la unidad y $1, u^2, (u^2)^2, \ldots, (u^2)^{n-1}$ es el conjunto de las $n/2$-ésimas raíces de la unidad.

Por la fórmula (C.3.1),

$$f(u^k) = \tilde{f}_+((u^2)^k) + u^k\tilde{f}_-((u^2)^k), \quad \text{para } 0 \leqslant k < n,$$

Sea n par, entonces para $0 \leqslant k < n/2$ y observemos que

$$\begin{aligned} (u^2)^k &= e^{-\frac{4k\pi i}{n}} \\ (u^2)^{k+n/2} &= e^{-\frac{4(k+n/2)\pi i}{n}} = e^{-\frac{4k\pi i}{n}} e^{-2\pi i} = (u^2)^k. \end{aligned}$$

Entonces

$$\begin{aligned} f(u^k) &= \tilde{f}_+((u^2)^k) + u^k\tilde{f}_-((u^2)^k), \quad \text{para } 0 \leqslant k < n/2, \\ f(u^{k+n/2}) &= \tilde{f}_+((u^2)^k) - u^k\tilde{f}_-((u^2)^k), \quad \text{para } 0 \leqslant k < n/2. \end{aligned} \tag{*}$$

En la segunda formula utilizamos que $u^{k+n/2} = -u^k$. Entonces, calcular $f(u^k)$ para $0 \leqslant k < n$ se reduce a calcular

$$\tilde{f}_+((u^2)^k) \quad \text{y} \quad \tilde{f}_-((u^2)^k), \quad \text{para } 0 \leqslant k < n/2,$$

y luego aplicar las fórmulas (*).

Por lo tanto, hemos reducido de la transformada de Fourier discreta de f cálculo de la transformada de Fourier discreta de $\tilde{f}_+$ y $\tilde{f}_-$.

Repitiendo el razonamiento que hicimos para f a $\tilde{f}_+$ y $\tilde{f}_-$ podemos calcular la transformada de Fourier discreta de f en forma recursiva ($n = 2^m$ y observación C.3.1).

El ahorro de operaciones que se obtiene, como ya dijimos, al calcular de esta forma la representación de f por n valores se debe a que reducimos ese cálculo al cálculo de la representación de dos funciones por $n/2$ valores. Se puede probar entonces que el cálculo de la transformada de Fourier discreta de f conlleva alrededor $n \log_2(n)$ operaciones.

El algoritmo es sencillo de programar. La siguiente sería una implementación en Python, con algo de pseudocódigo.

TRANSFORMADA RÁPIDA DE FOURIER

```
def FFT(f):
    # pre: f = [f_0, f_1, ..., f_(n-1)], n = 2**k (k >= 0)
    # post: devuelve c = [f(u**0), f(u**1), ..., f(u**(n-1))]
    #       donde u = e**(-2*pi*i/n)
    n = len(f)
    if n == 1:
        c = 1
    else:
        u = e**(-2*pi*1j/n)
        f_p, f_i = f[::2], f[1::2] # coeficientes pares e impares
        c_p, c_i = FFT(f_p), FFT(f_i)
        c = [0] * n # lista de longitud n con 0's
        for j in range(n // 2):
            c[j] = c_p[j] + u**j * c_i[j]
            c[j + n // 2] = c_p[j] - u**j * c_i[j]
    return c
```

Observación. Observar que el algoritmo se aplica a cualquier secuencia $f = (f_0, f_1, \ldots, f_{n-1})$ donde no necesariamente los f_j deben ser los coeficientes de un polinomio. En todos los casos se obtiene $c = F^{-1}f$, la transformada discreta de Fourier.

C.4 LA ANTITRANSFORMADA DE FOURIER

D

DETERMINANTE

En el apéndice se harán las demostraciones de los resultados correspondientes a la sección de determinantes (sección **??**).

D.1 DETERMINANTES

Lo primero que veremos será la demostración del teorema **??**. Los tres resultados de ese teorema los demostraremos en forma separada: serán los teoremas D.1.1, D.1.3 y D.1.4.

Teorema D.1.1. *Sea* $A \in M_n(\mathbb{K})$ *y sea* $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *multiplicando la fila* r *por* c*, es decir* $A \xrightarrow{cF_r} B$*, entonces* $\det B = c \det A$.

Demostración. Si multiplicamos la fila r por c obtenemos

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ ca_{r1} & ca_{r2} & \cdots & ca_{rn} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{bmatrix}.$$

Observemos que al hacer el desarrollo por la primera columna obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} C_{1i}^B + ca_{r1} C_{r1}^B + \sum_{i=r+1}^{n} a_{i1} C_{1i}^B.$$

Ahora bien, si $i \neq r$, la matriz $B(i|1)$ es la matriz $A(i|1)$ con una fila multiplicada por c, luego $|B(i|1)| = c|A(i|1)|$ y, en consecuencia $C_{i1}^B = c\, C_{i1}^A$. Además, $B(r|1) = A(r|1)$, luego $C_{r1}^B = C_{r1}^A$. Por lo tanto, reemplazando en la ecuación anterior C_{i1}^B por $c\, C_{i1}^A$ si $i \neq r$ y C_{r1}^B por C_{r1}^A, obtenemos

$$|B| = \sum_{i=1}^{r-1} a_{i1} c\, C_{1i}^A + ca_{r1} C_{r1}^A + \sum_{i=r+1}^{n} a_{i1} c C_{1i}^A = c|A|.$$

□

Lema D.1.2. *Sean* A, B, C *matrices* $n \times n$ *tal que*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

y

$$C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & \cdots & a_{rn} + b_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Es decir B *es igual a* A *pero con la fila* r *cambiada y* C *es como* A *y* B *excepto en la fila* r *donde cada coeficiente el la suma del de* A *y* B *correspondiente. Entonces* $\det(C) = \det(A) + \det(B)$.

Demostración. Se hará por inducción en n. Para $n = 1$, del resultado se reduce a probar que $\det[a + b] = \det[a] + \det[b]$, lo cual es trivial, pues el determinante en matrices 1×1 es la identidad.

Primero consideremos el caso $r = 1$. En este caso tenemos que $A(1|1) = B(1|1) = C(1|1)$, pues en la única fila que difieren las matrices es en la primera. Además, si $i > 1$, $A(i|1)$, $B(i|1)$ y $C(i|1)$ son iguales, excepto que difieren en la primera fila donde los coeficientes de $C(i|1)$ son la suma de los de $A(i|1)$ y $B(i|1)$, entonces, por hipótesis inductiva, $\det C(i|1) = \det A(i|1) + \det B(i|1)$. Concluyendo, tenemos que

$$\begin{aligned} \det A(1|1) &= \det B(1|1) = \det C(1|1), \\ \det C(i|1) &= \det A(i|1) + \det B(i|1), \qquad i > 1 \end{aligned} \quad ,$$

lo cual implica que

$$\begin{aligned} C^C_{11} &= C^A_{11} = C^B_{11}, \\ C^C_{i1} &= C^A_{i1} + C^B_{i1}, \qquad i > 1 \end{aligned} \quad .$$

Luego

$$\begin{aligned} \det C &= (a_{11} + b_{11})C^C_{11} + \sum_{i=2}^{n} a_{i1} C^C_{i1} \\ &= a_{11} C^C_{11} + b_{11} C^C_{11} + \sum_{i=2}^{n} a_{i1}(C^A_{i1} + C^B_{i1}) \\ &= a_{11} C^A_{11} + b_{11} C^B_{11} + \sum_{i=2}^{n} a_{i1}(C^A_{i1} + C^B_{i1}) \\ &= a_{11} C^A_{11} + \sum_{i=2}^{n} a_{i1} C^A_{i1} + b_{11} C^B_{11} + \sum_{i=2}^{n} a_{i1} C^B_{i1} \\ &= \det A + \det B. \end{aligned}$$

El caso $r > 1$ se demuestra de manera similar o, si se prefiere, puede usarse el teorema D.1.4, observando que la permutación entre la fila 1 y la fila r cambia el signo del determinante. □

Teorema D.1.3. *Sea* $A \in M_n(\mathbb{K})$. *Sea* $c \in \mathbb{K}$ *y* B *la matriz que se obtiene de* A *sumando a la fila* r *la fila* s *multiplicada por* c, *es decir* $A \xrightarrow{F_r + cF_s} B$, *entonces* $\det B = \det A$.

Demostración. A y B difieren solo en la fila r, donde los coeficientes de B son los los de A más c por los de la fila s. Luego si

$$A = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r \\ \vdots \\ F_n \end{bmatrix}, \qquad B = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_r + cF_s \\ \vdots \\ F_n \end{bmatrix}, \qquad A' = \begin{bmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ cF_s \\ \vdots \\ F_n \end{bmatrix},$$

el lema anterior nos dice que

$$\det B = \det A + \det A'. \tag{D.1.1}$$

Ahora bien, por teorema D.1.1,

$$\det A' = c \begin{vmatrix} F_1 \\ \vdots \\ F_s \\ \vdots \\ F_s \\ \vdots \\ F_n \end{vmatrix},$$

y este último determinante es cero, debido a que la matriz tiene dos filas iguales. Luego, $\det B = \det A$. □

Teorema D.1.4. *Sea* $A \in M_n(\mathbb{K})$ *y sean* $1 \leqslant r, s \leqslant n$. *Sea* B *la matriz que se obtiene de* A *permutando la fila* r *con la fila* s, *es decir* $A \xrightarrow{F_r \leftrightarrow F_s} B$, *entonces* $\det B = -\det A$.

Demostración. Primero probaremos el teorema bajo el supuesto de que la fila 1 es permutada con la fila k, para $k > 1$. Esto será suficiente para probar el teorema, puesto que intercambiar las filas k y k_0 es equivalente a realizar tres permutaciones de filas: primero intercambiamos las filas 1 y k, luego las filas 1 y k_0, y finalmente intercambiando las filas 1 y k. Cada permutación cambia el signo del determinante y al ser tres permutaciones, el intercambio de la fila k con la fila k_0 cambia el signo.

La prueba es por inducción en n. El caso base $n = 1$ es completamente trivial. (O, si lo prefiere, puede tomar $n = 2$ como el caso base, y el teorema

es fácilmente probado usando la fórmula para el determinante de una matriz 2×2). Las definiciones de los determinantes de A y B son:

$$\det(A) = \sum_{i=1}^{n} a_{i1} C_{i1}^{A} \quad \text{y} \quad \det(B) = \sum_{i=1}^{n} b_{i1} C_{i1}^{B}.$$

Supongamos primero que $i \neq 1, k$. En este caso, está claro que $A(i|1)$ y $B(i|1)$ son iguales, excepto que dos filas se intercambian. Por lo tanto, por hipótesis inductiva $C_{i1}^{A} = -C_{i1}^{B}$. Ya que también $a_{i1} = b_{i1}$, tenemos entonces que

$$a_{i1} C_{i1}^{A} = -b_{i1} C_{i1}^{B}, \quad \text{para } i \neq 1, k. \tag{D.1.2}$$

Queda por considerar los términos $i = 1$ y $i = k$. Nosotros afirmamos que

$$-a_{k1} C_{k1}^{A} = b_{11} C_{11}^{B} \quad \text{y} \quad -a_{11} C_{11}^{A} = b_{k1} C_{k1}^{B}. \tag{D.1.3}$$

Si probamos esto, entonces

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} a_{i1} C_{i1}^{A} \\
&= a_{11} C_{11}^{A} + \sum_{i=2}^{k-1} a_{i1} C_{i1}^{A} + a_{k1} C_{k1}^{A} + \sum_{i=k+1}^{n} a_{i1} C_{i1}^{A} \qquad \text{(D.1.2) y (D.1.3)} \\
&= -b_{k1} C_{k1}^{B} - \sum_{i=2}^{k-1} b_{i1} C_{i1}^{B} - b_{11} C_{11}^{B} - \sum_{i=k+1}^{n} b_{i1} C_{i1}^{B} \\
&= -\sum_{i=1}^{n} b_{i1} C_{i1}^{B} = -\det(B).
\end{aligned}$$

Luego el teorema está probado. Por lo tanto debemos probar (D.1.3). Por simetría, basta probar la primera identidad de (D.1.3), es decir que $a_{k1} C_{k1}^{A} = -b_{11} C_{11}^{B}$.

Para esto, primero debemos observar que $a_{k1} = b_{11}$, por lo tanto sólo hace falta probar que $-C_{k1}^{A} = C_{11}^{B}$. En segundo lugar, debemos tener en cuenta que $B(1|1)$ se obtiene de $A(k|1)$ reordenando las filas $1, 2, \ldots, k-1$ de $A(k|1)$ en el orden $2, 3, \ldots, k-1, 1$. Este reordenamiento puede hacerse permutando la fila 1 con la fila 2, luego permutando esa fila con la fila 3, etc., terminando con una permutación con la fila $k-1$. Esto es un total de $k-2$ permutaciones de fila. Asi que, por hipótesis inductiva,

$$\begin{aligned}
\det(B(1|1)) &= (-1)^{k-2} \det(A(k|1)) = (-1)^{k} \det(A(k|1)) \\
&= -(-1)^{k+1} \det(A(k|1)),
\end{aligned}$$

es decir $C_{11}^{B} = -C_{k1}^{A}$. Esto completa la demostración del teorema. □

Observación. Del resultado anterior se deduce fácilmente que si una matriz tiene dos filas iguales entonces su determinante es 0. Esto se debe a que, intercambiando las dos filas iguales obtenemos la misma matriz, pero calculando el determinante con el teorema anterior vemos que cambia de signo y el único número en $\mathbb{K}$ que es igual a su opuesto es el 0.

Corolario D.1.5. *Consideremos matrices elementales en* $\mathbb{K}^{n\times n}$.

(1) Sea E *la matriz elemental que se obtiene multiplicando por* $c \neq 0$ *la matriz* Id_n. *Entonces* $\det(E) = c$.

(2) Sea E *la matriz elemental que se obtiene a partir de* Id_n *sumando* c *veces* F_r *a* F_s *(*$r \neq s$*). Entonces* $\det(E) = 1$.

(3) Sea E *la matriz elemental que se obtiene a partir de* Id_n *de permutando la* F_r *con* F_s *(*$r \neq s$*). Entonces* $\det(E) = -1$.

Demostración. Se demuestra trivialmente considerando que en todos los casos $E = e(\mathrm{Id}_n)$ donde e es una operación elemental por fila, considerando que $\det(\mathrm{Id}_n) = 1$ y aplicando los teoremas D.1.1, D.1.3 y D.1.4, según corresponda. □

A continuación veremos que el determinante del producto de matrices es el producto de los determinantes de las matrices.

Teorema D.1.6. *Sea* $A \in M_n(\mathbb{K})$ *y* E *una matriz elemental* $n \times n$. *Entonces*

$$\det(EA) = \det E \det A. \tag{D.1.4}$$

Demostración. En todos los casos $EA = e(A)$ donde e es una operación elemental por fila (teorema **??**).

(1) Si $c \neq 0$, y E es la matriz elemental que se obtiene de multiplicar por c la fila r de Id_n, luego

$$\det(EA) = \det(e(A)) \overset{\text{Teor. D.1.1}}{=} c \cdot \det(A) \overset{\text{Cor. D.1.5.(1)}}{=} \det(E)\det(A).$$

(2) Si E es la matriz elemental que se obtiene de sumar a la fila r de Id_n la fila s multiplicada por c, entonces $\det E = 1$. Por otro lado $\det(EA) = \det(A)$, por lo tanto $\det(EA) = \det(E)\det(A)$.

(3) Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de Id_n, entonces $\det E = -1$. Por otro lado $\det(EA) = -\det(A)$, por lo tanto $\det(EA) = \det(E)\det(A)$. □

Corolario D.1.7. *Sea* $A \in M_n(\mathbb{K})$ *y* $E_1, \ldots, E_k$ *matrices elementales* $n \times n$. *Entonces*

$$\det(E_k E_{k-1} \ldots E_1 A) = \det(E_k)\det(E_{k-1}) \ldots \det(E_1)\det(A).$$

Demostración. Por la aplicación reiterada del teorema D.1.6 tenemos,

$$\begin{aligned}\det(E_k E_{k-1} \dots E_1 A) &= \det(E_k)\det(E_{k-1}\dots E_1 A)\\ &= \det(E_k)\det(E_{k-1})\det(E_{k-2}\dots E_1 A)\\ &\vdots\\ &= \det(E_k)\det(E_{k-1})\det(E_{k-2})\dots\det(E_1)\det(A).\end{aligned}$$

□

Teorema D.1.8. *$A \in \mathbb{K}^{n\times n}$ es invertible si y solo si* $\det(A) \neq 0$.

Demostración. ($\Rightarrow$) A invertible, luego por el teorema **??**, A es producto de matrices elementales, es decir $A = E_1 E_2 \cdots E_k$ donde $E_1, E_2, \dots, E_k$ son matrices elementales.

Por el corolario anterior, $\det(A) = \det(E_1)\det(E_2)\dots\det(E_k)$. Como el determinante de matrices elementales es distinto de cero,

$$\det(A) = \det(E_1)\det(E_2)\dots\det(E_k) \neq 0.$$

($\Leftarrow$) Sean $E_1, E_2, \dots, E_k$ matrices elementales tales que $R = E_1 E_2 \cdots E_k A$ y R es MERF. Luego,

$$\det(R) = \det(E_1)\det(E_2)\cdots\det(E_k)\det(A).$$

Como los determinantes de matrices elementales son no nulos

$$\frac{\det(R)}{\det(E_1)\det(E_2)\cdots\det(E_k)} = \det(A). \qquad (*)$$

Supongamos que R no es la identidad. Entonces, por el corolario **??**, $\det(R) = 0$, por lo tanto, $\det(A) = 0$, lo cual contradice la hipótesis y llegamos a un absurdo.

Esto implica que $R = \mathrm{Id}_n$ y en consecuencia A es equivalente por filas a Id_n y por lo tanto invertible. □

Teorema D.1.9. *Sean $A, B \in M_n(\mathbb{K})$, entonces*

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

Demostración. Separemos la prueba en dos casos A es invertible y A no es invertible.

A invertible. Entonces $A = E_1 \cdots E_k$ producto de matrices elementales. Por lo tanto $AB = E_1 \cdots E_k B$, luego por el corolario D.1.7 $\det(AB) = \det(E_1)\cdots\det(E_k)\det(B) = \det(A)\det(B)$.

A no invertible. Entonces A es equivalente por filas a una MERF R con la última fila nula. Es decir $R = E_1 \cdots E_k A$ y R tiene la última fila nula, por lo tanto $A = E_k^{-1} E_{k-1}^{-1} \dots E_1^{-1} R$.

Como R tiene la última fila nula, no es difícil ver que RB tiene tiene también la última fila nula y por lo tanto $\det(RB) = 0$. Luego

$$\det(AB) = \det(E_k^{-1}) \dots \det(E_1^{-1}) \det(RB) = 0.$$

Como $\det(A) = 0$, tenemos también qure

$$\det(A) \det(B) = 0.$$

□

Haremos ahora la demostración del teorema **??**.

Teorema D.1.10. *Sea* E *matriz elemental, entonces* E^t *es matriz elemental del mismo tipo y* $\det(E) = \det(E^t)$.

Demostración. Si $c \neq 0$ y E es la matriz elemental que se obtiene de multiplicar por c la fila r de Id_n, es claro que $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$.

Si E es la matriz elemental que se obtiene de sumar a la fila r de Id_n la fila s multiplicada por $c \in \mathbb{K}$, entonces E^t es la matriz elemental que se obtiene de sumar a la fila s de Id_n la fila r multiplicada por c. Luego, $\det(E) = \det(E^t) = 1$.

Finalmente, si E es la matriz elemental que se obtiene de intercambiar la fila r por la fila s de Id_n,entonces $E^t = E$ y por lo tanto $\det(E) = \det(E^t)$. □

Teorema D.1.11. *Sea* $A \in M_n(\mathbb{K})$, *entonces* $\det(A) = \det(A^t)$

Demostración. Si A es invertible, entonces $A = E_k E_{k-1} \dots E_1$ con E_i elemental, por lo tanto $\det(A) = \det(E_k) \det(E_{k-1}) \dots \det(E_1)$. Luego,

$$\begin{aligned} \det(A^t) &= \det(E_1^t \dots E_k^t) = \det(E_1^t) \dots \det(E_k^t) = \det(E_1) \dots \det(E_k) \\ &= \det(A). \end{aligned}$$

Si A no es invertible, entonces A^t no es invertible y en ese caso $\det(A) = \det(A^t) = 0$. □

Finalmente, demostremos el teorema **??**.

Teorema D.1.12. *El determinante de una matriz* A *de orden* $n \times n$ *puede ser calculado por la expansión de los cofactores en cualquier columna o cualquier fila. Más específicamente,*

(1) si usamos la expansión por la j*-ésima columna,* $1 \leqslant j \leqslant n$, *tenemos*

$$\begin{aligned} \det A &= \sum_{i=1}^{n} a_{ij} C_{ij} \\ &= a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + a_{nj} C_{nj}. \end{aligned}$$

(2) si usamos la expansión por la i-ésima fila, $1 \leqslant i \leqslant n$, tenemos

$$\begin{aligned}\det A &= \sum_{j=1}^{n} a_{ij} C_{ij} \\ &= a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in};\end{aligned}$$

Demostración. (1) Primero hagamos la demostración para $j = 2$, es decir para el desarrollo por la segunda columna. Escribamos A en función de sus columnas, es decir

$$A = \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_n \end{bmatrix},$$

donde C_k es la columna k de A. Sea $B = [b_{ij}]$ la matriz definida por

$$B = \begin{bmatrix} C_2 & C_1 & C_3 & \cdots & C_n \end{bmatrix}.$$

Entonces, $\det(B) = -\det(A)$. Por otro lado, por la definición de determinante,

$$\begin{aligned}\det(B) &= \sum_{i=1}^{n} b_{i1} C_{i1}^{B} \\ &= \sum_{i=1}^{n} b_{i1} (-1)^{i+1} B(i|1) \\ &= \sum_{i=1}^{n} a_{i2} (-1)^{i+1} B(i|1).\end{aligned}$$

Ahora bien, es claro que $B(i|1) = A(i|2)$, por lo tanto

$$\det(B) = \sum_{i=1}^{n} a_{i2} (-1)^{i+1} A(i|2) = -\sum_{i=1}^{n} a_{i2} C_{i2}.$$

Es decir, $\det(A) = -\det(B) = \sum_{i=1}^{n} a_{i2} C_{i2}$.

El caso $j > 2$ se demuestra de forma similar: si B es la matriz

$$B = \begin{bmatrix} C_j & C_1 & C_2 & \cdots & C_{j-1} & C_{j+1} & \cdots & C_n \end{bmatrix}.$$

entonces $\det(B) = (-1)^{j-1} \det(A)$, pues son necesarios $j - 1$ permutaciones para recuperar la matriz A (es decir, llevar la columna j a su lugar). Como $B(i|1) = A(i|j)$, desarrollando por la primera columna el determinante de B obtenemos el resultado.

(2) Observemos primero que $A^t(j|i) = A(i|j)^t$, por lo tanto, si calculamos $\det(A^t)$ por desarrollo por columna i, obtenemos

$$\begin{aligned}\det A = \det(A^t) &= \sum_{j=1}^{n} [A^t]_{ji}(-1)^{i+j} \det(A^t(j|i)) \\ &= \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det(A(i|j)^t) \\ &= \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det(A(i|j)).\end{aligned}$$

□

D.2 REGLA DE CRAMER

Veremos ahora que la inversa de una matriz invertible se puede escribir en términos de determinantes de algunas matrices relacionadas y esto, junto a otros resultados, nos permitirá resolver ecuaciones lineales con n-variables y n-incógnitas cuya matriz asociada es invertible.

Teorema D.2.1. *Sea A matriz $n \times n$, entonces*

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} A = \begin{bmatrix} 0 & \cdots & 0 & \det A & 0 & \cdots & 0 \\ & & & \uparrow i & & & \end{bmatrix}.$$

Es decir, la matriz fila formada por los cofactores correspondientes a la columna i multiplicada por la matriz A es igual a la matriz fila con valor $\det A$ en la posición i y 0 en las otras posiciones.

Demostración. Si C_j denota la matriz formada por la columna j de A debemos probar que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^{n} a_{kj} C_{ki} = \begin{cases} \det(A) & \text{si } j = i \\ 0 & \text{si } j \neq i. \end{cases}$$

Ahora bien,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_i = \sum_{j=1}^{n} C_{ji} a_{ji},$$

y esto último no es más que el cálculo del determinante por desarrollo de la columna i, es decir, es igual a $\det(A)$.

Para ver el caso $i \neq j$, primero observemos que si

$$B = \begin{bmatrix} C_1 & C_2 & \cdots & C_j & \cdots & C_j & \cdots & C_{n-1} & C_n \end{bmatrix}, \\ \uparrow i \qquad \uparrow j$$

es decir, B es la matriz A donde reemplazamos la columna i por la columna j, entonces como B tiene dos columnas iguales, $\det(B) = 0$. Por lo tanto, si calculamos el determinante de B por el desarrollo en la columna i, obtenemos

$$0 = \det(B) = \sum_{k=1}^{n} a_{kj} C_{ki}. \tag{D.2.1}$$

Por otro lado,

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = \sum_{k=1}^{n} C_{ki} a_{kj},$$

luego, por la ecuación (D.2.1) tenemos que

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} C_j = 0$$

si $i \neq j$.

□

Definición D.2.2. Sea A matriz $n \times n$, la *matriz de cofactores* es la matriz cuyo coeficiente ij vale C_{ij}. La matriz de cofactores de A se denota $\operatorname{cof}(A)$. La *matriz adjunta de A* es $\operatorname{adj}(A) = \operatorname{cof}(A)^t$.

Teorema D.2.3. *Sea A matriz $n \times n$, entonces*

$$\operatorname{adj}(A) \cdot A = \det(A) \operatorname{Id}_n .$$

Demostración. Observar que la fila i de $\operatorname{adj}(A)$ es $\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix}$. Por lo tanto, la fila i de $\operatorname{adj}(A) \cdot A$ es

$$\begin{bmatrix} C_{1i} & C_{2i} & \cdots & C_{ni} \end{bmatrix} A,$$

que por el teorema D.2.1 es una matriz fila con el valor $\det A$ en la posición i y todos los demás coeficientes iguales a o. Luego

$$\begin{aligned} \operatorname{adj}(A) \cdot A &= \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \cdot A \\ &= \begin{bmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det A \end{bmatrix} \\ &= \det(A) \operatorname{Id}_n \end{aligned}$$

□

Corolario D.2.4. *Si A es invertible, entonces*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

Demostración.

$$\frac{1}{\det A} \operatorname{adj} A \cdot A = \frac{1}{\det A} \det A \operatorname{Id}_n = \operatorname{Id}_n .$$

□

Teorema D.2.5 (Regla de Cramer)**.** *Sea $AX = Y$ un sistema de ecuaciones tal que $A \in M_n(\mathbb{K})$ es invertible. Entonces, el sistema tiene una única solución $(x_1, \ldots, , x_n)$ con*

$$x_j = \frac{\det A_j}{\det A}, \qquad j = 1, \ldots, n,$$

donde A_j es la matriz $n \times n$ que se obtiene de A remplazando la columna j de A por Y.

Demostración. Haremos la demostración para matrices 3×3. La demostración en el caso general es completamente análoga.

Como A es invertible, existe A^{-1} y multiplicamos la ecuación a izquierda por A^{-1} y obtenemos que $A^{-1}AX = A^{-1}Y$, es decir $X = A^{-1}Y$ y esta es la única solución. Luego

$$\begin{aligned} A^{-1}Y &= \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ &= \frac{1}{\det A} \begin{bmatrix} y_1 C_{11} + y_2 C_{21} + y_3 C_{31} \\ y_1 C_{12} + y_2 C_{22} + y_3 C_{32} \\ y_1 C_{13} + y_2 C_{23} + y_3 C_{33} \end{bmatrix} \qquad (*) \end{aligned}$$

Ahora bien, $y_1 C_{11} + y_2 C_{21} + y_3 C_{31}$ es el cálculo de determinante por desarrollo de la primera columna de la matriz

$$\begin{bmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_1 & a_{32} & a_{33} \end{bmatrix},$$

y, de forma análoga, el segundo y tercer coeficiente de la matriz $(*)$ son el determinante de las matrices 3×3 que se obtienen de A remplazando la columna 2 y 3, respectivamente, de A por Y. Es decir

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^{-1}Y = \frac{1}{\det A} \begin{bmatrix} \det A_1 \\ \det A_2 \\ \det A_3 \end{bmatrix} = \begin{bmatrix} \frac{\det A_1}{\det A} \\ \frac{\det A_2}{\det A} \\ \frac{\det A_3}{\det A} \end{bmatrix},$$

luego $x_j = \dfrac{\det A_j}{\det A}$ para $j = 1, 2, 3$. □

Ejemplo. Resolvamos usando la regla de Cramer el siguiente sistema:

$$\begin{aligned} x_1 + x_2 - x_3 &= 6 \\ 3x_1 - 2x_2 + x_3 &= -5 \\ x_1 + 3x_2 - 2x_3 &= 14. \end{aligned}$$

La matriz asociada al sistema es

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 3 & -2 & 1 \\ 1 & 3 & -2 \end{bmatrix}.$$

Luego

$$A_1 = \begin{bmatrix} 6 & 1 & -1 \\ -5 & -2 & 1 \\ 14 & 3 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 6 & -1 \\ 3 & -5 & 1 \\ 1 & 14 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 & 6 \\ 3 & -2 & -5 \\ 1 & 3 & 14 \end{bmatrix},$$

y

$$\det A = -3, \qquad \det A_1 = -3, \qquad \det A_2 = -9, \qquad \det A_3 = 6.$$

Por lo tanto,

$$\begin{aligned} x_1 &= \frac{\det A_1}{\det A} = \frac{-3}{-3} = 1 \\ x_2 &= \frac{\det A_2}{\det A} = \frac{-9}{-3} = 3 \\ x_3 &= \frac{\det A_3}{\det A} = \frac{6}{-3} = -2. \end{aligned}$$

Observación. La regla de Cramer implementada de una manera ingenua es ineficiente computacionalmente para sistemas de más de dos o tres ecuaciones. En el caso de n ecuaciones con n incógnitas, requiere el cálculo de $n+1$ determinantes, mientras que el método de eliminación de Gauss o eliminación gaussiana produce el resultado con la misma complejidad computacional que el cálculo de un solo determinante. Sin embargo, recientemente se ha demostrado que la regla de Cramer se puede implementar en el tiempo $O(n^3)$, que es comparable a los métodos más utilizados para la obtención de soluciones de sistemas de ecuaciones lineales, como ser la eliminación gaussiana (ver https://en.wikipedia.org/wiki/Cramer's_rule y https://es.wikipedia.org/wiki/Eficiencia_Algorítmica).

Sin embargo, la regla de Cramer tiene propiedades numéricas muy pobres, por lo que no es adecuada para resolver incluso sistemas pequeños de forma fiable, a menos que las operaciones se realicen en aritmética racional con precisión ilimitada.

Parte IV

ÍNDICE

ÍNDICE ALFABÉTICO

cuerpo, 9

forma polar, 14
fórmula de Euler, 16

identidad de Euler, 16

matriz
    de cofactores, 42

notación exponencial, 15
números complejos, 11

polinomio, 17

regla de Cramer, 43

serie de Fourier, 25