

Capítulo 2

Espacios vectoriales

2.1. Definición y ejemplos

Imagina que en uno de esos documentales de ínfima audiencia que todos decimos que vemos mencionan los espacios vectoriales y te entra curiosidad acerca de qué es aquello. Si preguntas a un matemático corres el riesgo de que te diga que son realmente, su definición rigurosa, y te quedes como antes y además con pocas ganas de estudiar matemáticas. Lo mismo se aplica a otras estructuras algebraicas¹. Estas estructuras, entre las que se cuenta la de espacio vectorial, son una abstracción *a posteriori* motivada por ciertos ejemplos destacados o al menos por situaciones matemáticas naturales. Aunque en los textos aparezcan en los primeros capítulos por la sana o insana obsesión de los matemáticos por guardar el orden deductivo, históricamente son culminaciones. Por ello causan tanta perplejidad fuera del círculo de los devotos e iniciados. Hay que recordar que no ya la estructura de espacio vectorial sino la propia álgebra lineal tuvo detractores en sus orígenes².

Antes de asustar a nadie, la idea básica es que un *espacio vectorial* V es cualquier conjunto de cosas, que denominaremos *vectores*, tales que se pueden multiplicar por números y sumar entre sí. En breve, tal que dados $\vec{u}, \vec{v} \in V$ se tenga que $\lambda\vec{u} + \mu\vec{v}$ también pertenece a V , que también sea un vector.

En la práctica esto es lo que tienes que tener en mente y al final bajo ciertas hipótesis recobramos de aquí una definición rigurosa. La pregunta es por qué si buscamos la definición en la Wikipedia, en un libro o en la lección de un matemático recibimos una lista rara y larga de propiedades. La respuesta es que lo anterior es demasiado impreciso como para considerarlo una definición de verdad porque no

¹En mi opinión, los físicos son mucho más agradecidos. Si les preguntas qué es un electrón o qué es el modelo estándar te dirán una mentirijilla metafórica aunque en su investigación estén pensando en espinores o en representaciones de grupos, que son temas avanzados de estructuras algebraicas.

²Se considera a H. Grassman el padre del álgebra lineal junto con el *álgebra exterior*. Un crítico fue su contemporáneo W.R. Hamilton quien formuló parte de la física de su tiempo con sus *cuaterniones*, unos vectores de cuatro coordenadas que también se podían multiplicar. Aunque el álgebra lineal y el álgebra exterior ganaron la partida, la notación de Hamilton ha perdurado en esas letras **i**, **j**, **k** que todavía se emplean al calcular productos vectoriales.

especifica las propiedades básicas que deben tener una suma de vectores y una multiplicación por números que en contextos más o menos exóticos no tienen nada que ver con las que nos son familiares. Por mencionar algo más concreto, un punto importante es precisar qué significa “número”, por ejemplo permitiremos \mathbb{R} y \mathbb{C} e incluso en ciertas aplicaciones se utilizan números que desconoces pero prohibiremos \mathbb{Z} . Una definición matemática no puede permitirse cabos sueltos, en nuestro caso debe ser una lista completa de lo que se puede y no se puede hacer para que algo merezca llamarse espacio vectorial.

A pesar de no tener todavía la definición que no subleva a los matemáticos, veamos algunos ejemplos. Más adelante los comprobaremos con más cuidado.

El espacio de vectores de toda la vida \mathbb{R}^n es un espacio vectorial con la suma coordenada a coordenada y la multiplicación simultánea de todas las coordenadas por un número real. De la misma forma \mathbb{C}^n es un espacio vectorial escogiendo como números \mathbb{C} . De hecho si decretamos que solo está permitida la multiplicación de números reales, todavía \mathbb{C}^n sería un espacio vectorial, que a veces se llama $\mathbb{C}_{\mathbb{R}}^n$ para distinguirlo del habitual, ya que $\vec{u}, \vec{v} \in \mathbb{C}^n$ y $\lambda, \mu \in \mathbb{R}$ implican $\lambda\vec{v} + \mu\vec{u} \in \mathbb{C}^n$.

También son espacios vectoriales $\mathcal{M}_{m \times n}(\mathbb{R})$ y $\mathcal{M}_{m \times n}(\mathbb{C})$ pues a efectos de hacer sumas o multiplicaciones por números siempre podemos recolocar los elementos en una columna para tener vectores de \mathbb{R}^{mn} o \mathbb{C}^{mn} .

Un ejemplo general que contiene a muchos de los que se te puedan ocurrir son las funciones $f : X \rightarrow \mathbb{R}$ o $f : X \rightarrow \mathbb{C}$ con X un conjunto fijado. Si X es el conjunto de seres humanos, un vector de este espacio sería asignarles a cada uno su número de pasaporte o su altura o a unos π y a otros e . El caso es que aunque no sepamos sumar personas para que den otra persona o multiplicarlas por números, sí podemos hacerlo con los números que les asignamos y eso es suficiente.

Si $X = \mathbb{R}$ un ingeniero puede considerar que este es un espacio vectorial de señales que para cada tiempo da por ejemplo el valor en ese instante de la onda de radio entrante en un receptor. Si se ha procedido a un muestreo de la señal examinándola en tiempos discretos, como es natural en el procesamiento digital, entonces quizá prefiera considerar $X = \mathbb{Z}$ y así el espacio vectorial está formado por tiras infinitas a izquierda y derecha $(x_n)_{n=-\infty}^{\infty}$.

Vamos por fin con la temida definición rigurosa de espacio vectorial. Es raro que conozcas todos los términos que aparecen en ella. La iremos desentrañando poco a poco. La historia del concepto está descrita en [4].

Sean V un conjunto, K un cuerpo y dos operaciones $+$: $V \times V \rightarrow V$, \cdot : $K \times V \rightarrow V$. Se dice que V es un *espacio vectorial* sobre K con estas operaciones si $(V, +)$ es un grupo abeliano y además para $\vec{u}, \vec{v} \in V$, $a, b \in K$ cualesquiera se cumplen las propiedades:

$$\begin{aligned} 1 \cdot \vec{u} &= \vec{u}, & (a + b) \cdot \vec{u} &= a \cdot \vec{u} + b \cdot \vec{u}, \\ a \cdot (b \cdot \vec{u}) &= (ab) \cdot \vec{u}, & a \cdot (\vec{u} + \vec{v}) &= a \cdot \vec{u} + a \cdot \vec{v}. \end{aligned}$$

A los elementos de K se les suele llamar *escalares* para diferenciarlos de los de V que son *vectores*. Aparte de esto, un par de observaciones con respecto a la notación:

incluso si el símbolo “.” no representara el producto de toda la vida, se suele omitir y en la primera propiedad 1 es el elemento unidad de K , el que cumple $1a = a$ para $a \in K$.

Seguramente tu primera duda es qué es un *cuerpo*. Es un conjunto donde hay definidas una suma, una resta, una multiplicación y una división (salvo por cero) con las propiedades habituales. Si esto te resulta vago, en rigor $(K, +)$ y $(K - \{0\}, \times)$ deben ser grupos abelianos (sigue leyendo) y además se deben cumplir las *propiedades distributivas*, las reglas de operar paréntesis. En este curso $K = \mathbb{R}$ o \mathbb{C} y es difícil que se te ocurran más ejemplos aparte de \mathbb{Q} aunque hay otros bien distintos e importantes³. Conservaremos no obstante en la teoría K como un cuerpo genérico.

Así, no permitimos $K = \mathbb{N}$ ni $K = \mathbb{Z}$ porque, en general, en el primero no se puede restar sin salirse de K y en el segundo no se puede dividir.

El otro término que te puede resultar extraño en la definición es lo de *grupo abeliano*. Un conjunto con una operación, en nuestro caso V con $+$, se dice que es un grupo abeliano si satisface la *propiedad conmutativa*, la *propiedad asociativa*, tiene elemento neutro y tiene elemento inverso⁴. Estos dos últimos se suelen indicar en espacios vectoriales con $\vec{0}$ y $-\vec{u}$.

Las cuatro propiedades al final de la definición podrán parecer arbitrarias pero se dejan leer, no involucran términos nuevos.

Ahora si queremos probar de verdad que fijado X el conjunto de funciones $V = \{f : X \rightarrow \mathbb{R}\}$ con la suma y producto habituales es un espacio vectorial sobre \mathbb{R} , tenemos que comprobar en principio un montón de cosas. El truco para evitarlo es apelar a lo que ya conocemos o supuestamente conocemos. Primero, todos sabemos sumar, restar, multiplicar y dividir números reales, por tanto \mathbb{R} es un cuerpo. ¿Es $(V, +)$ un grupo abeliano? Sí y resulta que probarlo no entraña nada nuevo porque sumar funciones es sumar sus valores por tanto se sigue de que $(\mathbb{R}, +)$ es grupo abeliano por ser K cuerpo. De la misma forma, las cuatro propiedades al final de la definición de espacio vectorial se seguirían de que \mathbb{R} es un cuerpo si $\vec{u}, \vec{v} \in \mathbb{R}$ y eso es todo lo que necesitamos porque, como acabamos de decir, las funciones se operan operando sus valores por separado.

Si $X = \{1, 2, \dots, n\}$ podemos identificar $f : X \rightarrow \mathbb{R}$ con un vector de toda la vida de \mathbb{R}^n , donde la i -ésima coordenada dada por el valor de $f(i)$. De lo anterior se deduce que \mathbb{R}^n es un espacio vectorial sobre \mathbb{R} . De la misma forma se tiene que $\mathcal{M}_{m \times n}(\mathbb{R})$ es un espacio vectorial sobre \mathbb{R} tomando como X todas las posibles posiciones de los elementos, $X = \{(1, 1), (1, 2), \dots, (m, n)\}$ e identificando $f(i, j)$ con a_{ij} . Por supuesto lo mismo se aplica a \mathbb{C}^n y a $\mathcal{M}_{m \times n}(\mathbb{C})$.

³Uno poco útil en ingeniería es $K = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$, la división se puede hacer sin salirse de K racionalizando. Un ejemplo todavía más raro y que, lo creas o no, es parte de una familia útil en ingeniería, es $K = \{0, 1, \heartsuit, \clubsuit\}$ donde el 0 y 1 tienen las propiedades de siempre ($0a = 0$, $1a = a$, $0 + a = a$) mientras que \heartsuit y \clubsuit son dos “números” nuevos sujetos a las reglas $\heartsuit + \clubsuit = 1$, $1 + 1 = \heartsuit + \heartsuit = \clubsuit + \clubsuit = 0$, $\heartsuit^2 = \clubsuit$, $\clubsuit^2 = \heartsuit$ y $\heartsuit\clubsuit = 1$. Por ejemplo, si queremos calcular \heartsuit/\clubsuit la penúltima igualdad dice que $\clubsuit\clubsuit = \heartsuit$ así que $\clubsuit = \heartsuit/\clubsuit$.

⁴Si no conoces estas propiedades de memoria, en fórmulas son $\vec{u} + \vec{v} = \vec{v} + \vec{u}$, $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$, $\vec{u} + \vec{0} = \vec{u}$ y $\vec{u} + (-\vec{u}) = \vec{0}$.

Ahora vamos a dar un último pase mágico por el que la definición no rigurosa pasará a serlo en los ejemplos del curso y uno podrá olvidarse ya de este lío de propiedades.

Dado V un espacio vectorial, se dice que W es un *subespacio vectorial* de V si $W \subset V$ y es un espacio vectorial sobre el mismo cuerpo con las operaciones definidas en V .

Fácil, ¿verdad? Un subespacio vectorial es solo un espacio vectorial dentro de otro. Pues bien, para los subespacios no es necesario comprobar toda la lista de propiedades porque ya se cumplen en un sitio mayor. Lo único que hay que comprobar es ¡la definición intuitiva!

Proposición 2.1.1. *Sea V un espacio vectorial sobre K . Un subconjunto $W \subset V$ define un subespacio vectorial de V (y en particular es un espacio vectorial) si y solo si para todo $\lambda, \mu \in K$ y $\vec{u}, \vec{v} \in W$ se tiene $\lambda\vec{u} + \mu\vec{v} \in W$.*

Demostración. Primero hay que saber que en todo espacio vectorial se cumple $0\vec{u} = \vec{0}$ y $\vec{u} + (-1)\vec{u} = \vec{0}$. Aunque te parezca escandaloso, esto necesita una pequeña demostración pues no está en la lista de propiedades. Dándola por hecha⁵, de $\lambda\vec{u} + \mu\vec{v} \in W$ se tiene tomando $\lambda = \mu = 0$ y $\lambda = -1, \mu = 0$ que los elementos neutro e inverso están en W . Por otro lado, con $\lambda = \mu = 1$ y $\lambda = 1, \mu = 0$ se deduce que las operaciones $+$ y \cdot están bien definidas en W , no se salen de W . La conmutativa y asociativa de $(W, +)$ y las cuatro propiedades de la definición para W se deducen de las de V porque $W \subset V$.

La otra implicación es obvia, si W es un espacio vectorial el producto por escalares y la suma de vectores están bien definidos en W y por tanto $\lambda\vec{u} + \mu\vec{v} \in W$. \square

Terminemos con un ramillete de ejemplos y contraejemplos.

Fijada $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, el conjunto $V = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$ es un espacio vectorial (sobre \mathbb{R}). Claramente $V \subset \mathbb{R}^n$ y basta ver que es subespacio usando la Proposición 2.1.1. Si $\vec{u}, \vec{v} \in V$ se cumple $A\vec{u} = A\vec{v} = \vec{0}$, por tanto $A(\lambda\vec{u} + \mu\vec{v}) = \vec{0}$, así pues $\lambda\vec{u} + \mu\vec{v} \in V$.

Dados $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ y $\vec{b} \neq \vec{0}$, el conjunto $V = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{b}\}$ no es un espacio vectorial. Si $\vec{u} \in V$ entonces $2\vec{u} \notin V$ ya que $A(2\vec{u}) = 2A\vec{u} = 2\vec{b} \neq \vec{b}$. Esto significa que multiplicar por dos no está bien definido en V . No tenemos una operación $\cdot : K \times V \rightarrow V$.

Los polinomios con coeficientes reales $\mathbb{R}[x]$ forman un espacio vectorial sobre \mathbb{R} . Son un subconjunto de todas las funciones $\mathbb{R} \rightarrow \mathbb{R}$ así que podemos aplicar la Proposición 2.1.1. Si $P, Q \in \mathbb{R}[x]$ y $\lambda, \mu \in \mathbb{R}$ entonces es obvio que $\lambda P + \mu Q \in \mathbb{R}[x]$.

De la misma forma los polinomios de $\mathbb{R}[x]$ con grado menor o igual que 3, denotados a veces con $\mathbb{R}_3[x]$, también forman un espacio vectorial. Sin embargo los de grado exactamente 3 no, ya que por ejemplo $P = 2x^3 + x$ y $Q = 1 - 3x^3$ tienen grado 3 pero $3P + 2Q$ no lo tiene.

⁵Si no tienes gana de pensarlo, lo primero se sigue de $\vec{0} = \vec{u} + (-\vec{u}) = (0+1)\vec{u} + (-\vec{u}) = 0\vec{u} + \vec{0} = 0\vec{u}$ y lo segundo, una vez sabido lo primero, de $\vec{0} = 0\vec{u} = (1-1)\vec{u} = \vec{u} + (-1)\vec{u}$.

Las matrices cuadradas simétricas reales $V = \{A \in \mathcal{M}_n(\mathbb{R}) : A = A^t\}$ forman un espacio vectorial sobre \mathbb{R} porque $V \subset \mathcal{M}_n(\mathbb{R})$ y si $A, B \in V$ entonces $(\lambda A + \mu B)^t = \lambda A^t + \mu B^t = \lambda A + \mu B$.

Sin embargo el análogo natural complejo $V = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^\dagger\}$ no es un espacio vectorial sobre \mathbb{C} , solo sobre \mathbb{R} .

Dados $a, b > 0$ consideremos el conjunto V formado por todas las funciones reales $y = y(x)$ con dos derivadas tales que satisfacen la ecuación del oscilador armónico amortiguado $y'' + ay' + by = 0$. Se cumple que V es un espacio vectorial sobre \mathbb{R} porque es un subconjunto del espacio vectorial de todas las funciones $\mathbb{R} \rightarrow \mathbb{R}$ y si $y, z \in V$, $\lambda, \mu \in \mathbb{R}$ se tiene

$$(\lambda y + \mu z)'' + a(\lambda y + \mu z)' + b(\lambda y + \mu z) = \lambda(y'' + ay' + by) + \mu(z'' + az' + bz) = 0,$$

ya que cada uno de los paréntesis de la última expresión es nulo.

Exprimiendo el silicio [opcional]. Esta sección ha tenido demasiado de atracán teórico como para que podamos hacer muchas cosas con el ordenador. Solo vamos a usarlo para intrigar al lector con uno de los espacios vectoriales más pequeños añadiendo la promesa de que la teoría subyacente, que excede este curso, es importante en ciertas aplicaciones.

El cuerpo más pequeño que se puede construir es el que solo contiene el elemento neutro de la suma y de la multiplicación, $K = \{0, 1\}$. Para que la suma no se salga de K la única posibilidad es decretar $1 + 1 = 0$. En la práctica K se usa para representar un *bit* pero si el decreto anterior te parece extraño en términos informáticos, puedes pensar que 0 significa “ser par” y 1 significa “ser impar” lo que cuadra con todas las operaciones en K .

Las líneas 1–8 del siguiente código `sagemath` definen el espacio vectorial $V = K^2$, formado por pares de elementos de K y muestra todos los vectores posibles. Por supuesto son $(0,0)$, $(1,0)$, $(0,1)$ y $(1,1)$.

```

1 # K es el cuerpo con dos elementos {0,1}
2 K = GF(2)
3 # V es el espacio vectorial de pares de bits
4 V = VectorSpace(K,2)
5 print('Lista de vectores:')
6 for v in V:
7     print(v)
8 print('-----')
9
10 cero = vector( K, [0,0] )
11 uno = vector( K, [1,0] )
12 corazon = vector( K, [0,1] )
13 trebol = vector( K, [1,1] )
14
15 print('Tabla de sumas:')
16 print('C+T=', corazon + trebol)
17 print('1+1=', uno + uno)
18 print('C+C=', corazon + corazon)
19 print('T+T=', trebol + trebol)

```

En las líneas 10–13 se dan respectivamente los curiosos nombres `cero`, `uno`, `corazon` y `trebol` a los vectores de V . El resto del programa comprueba que con estos nombres

la tabla de sumas es exactamente igual que la del extraño cuerpo $\{0, 1, \heartsuit, \clubsuit\}$ de cuatro elementos mencionado en una nota a pie de página. Esto implica que hay una forma no solo de sumar y restar vectores de V sino también de multiplicarlos y dividirlos (salvo por cero). Pues bien, esto es un hecho general. Un teorema asegura que para cualquier cuerpo finito K y cualquier n hay una manera de multiplicar y dividir vectores en K^n que convierte este espacio vectorial en un cuerpo⁶.

Antes de que pienses que los matemáticos se preocupan por cosas muy raras, debes saber que en el caso de $K = \{0, 1\}$ esta estructura adicional sobre tiras de bits que permite sumarlas, restarlas, multiplicarlas y dividir las conservando su longitud, es crucial en la corrección de errores al reproducir soportes digitales como CD, DVD o Blu-ray. Tal corrección no es un lujo, simplemente sin ella no funcionarían porque incluso recién salidos de fábrica, antes de que los rayes, contienen muchos errores de grabación. Si consideras que estos soportes son antediluvianos, debes saber que la corrección de errores también se utilizan en las memorias USB y que alarga su vida útil. Quizá mientras lees estos apuntes se haya averiado alguno de los bits en el *pendrive* donde los almacenas pero no sufrirás de pánico informático hasta que el número de fallos colapse el algoritmo de corrección de errores.

2.2. Bases y dimensión

Pensemos en \mathbb{R}^2 que identificamos con todos los puntos o direcciones del plano. Tenemos claro que su dimensión es dos porque necesitamos especificar una x y una y . Alguien con educación matemática elemental nos podría decir que hay cuatro puntos cardinales y nosotros convencerle de que usando números negativos hay redundancia en ellos, 3 pasos al sur son -3 pasos al norte. Todavía hay más redundancia si introducimos las direcciones intermedias de la rosa de los vientos, como el noreste. Aun así, siempre que elijamos dos direcciones que no tengan nada que ver, la redundancia desaparece y obtenemos un sistema válido para hacer mediciones. Por ejemplo, un paso al este y dos al norte son $\sqrt{2}$ pasos al noreste y uno al norte, ya que un un paso de longitud 1 al noreste te lleva a $(1/\sqrt{2}, 1/\sqrt{2})$.

En álgebra lineal los diferentes sistemas de medida constituyen lo que se llaman *bases*. En términos prácticos uno puede preguntarse por qué no fijar uno de esos sistemas convencionalmente, igual que hay un acuerdo en el meridiano de Greenwich. La respuesta es que la conveniencia de uno u otro depende de la situación. En astronomía las elipses de los planetas son muy perfectas cuando se miden desde el centro de masas del Sistema Solar, desde allí la física se vuelve más fácil pero a la hora de hacer observaciones la mayor parte de los telescopios están en la Tierra y es más natural medir en ese sistema rotatorio en el que las órbitas de los planetas son curvas complicadas.

Un ejemplo en el ámbito de la ingeniería es que emplear un sistema u otro al

⁶Un pionero en la introducción de estos cuerpos finitos fue É. Galois, un matemático de vida azarosa en tiempos de la restauración monárquica francesa que murió en un duelo con tan solo 20 años. La notación \mathbf{GF} que emplea `sagemath` son las iniciales de *Galois field*.

describir una señal de audio sirve para separar algunas de sus características como el ruido o los graves o los agudos, que permitan procesarla adecuadamente.

Lo primero que haremos es dar un nombre a la situación en la que unos vectores no ofrezcan información redundante, esto es, que ninguno se despeje en función de los otros.

En un espacio vectorial V sobre K se dice que los vectores $\vec{v}_1, \dots, \vec{v}_n$ son *linealmente independientes* si la única solución de

$$(2.1) \quad \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n = \vec{0}$$

con $\lambda_1, \dots, \lambda_n \in K$ es $\lambda_1 = \dots = \lambda_n = 0$ (aunque es irrelevante en este curso, para un conjunto infinito de vectores, la independencia lineal se define pidiendo que ocurra para cualquiera de sus subconjuntos finitos).

En un alarde de originalidad, cuando hay más soluciones se dice que los vectores son *linealmente dependientes*. En este caso algún $\lambda_i \neq 0$ y se despeja \vec{v}_i en términos del resto de los vectores por medio de $\vec{v}_i = -\sum_{j \neq i} \lambda_j / \lambda_i \vec{v}_j$.

Seguro que no asombra que el estudio de la dependencia o independencia lineal se reduzca habitualmente a discutir un sistema homogéneo.

Por ejemplo, estudiemos si los vectores de \mathbb{R}^3

$$\vec{v}_1 = (2, 1, -1)^t, \quad \vec{v}_2 = (3, 0, 1)^t, \quad \text{y} \quad \vec{v}_3 = (-1, 4, -7)^t$$

son linealmente independientes. La ecuación (2.1) se puede escribir como

$$A\vec{\lambda} = \vec{0} \quad \text{con} \quad A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 4 \\ -1 & 1 & -7 \end{pmatrix}, \quad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

En general en \mathbb{R}^n o \mathbb{C}^n el sistema será $A\vec{\lambda} = \vec{0}$ con A la matriz formada por los vectores en columna. Es fácil ver que $\text{rg}(A) = 2$, concretamente aplicando eliminación de Gauss intercambiando las dos primeras filas:

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 3 & -1 \\ -1 & 1 & -7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 3 & -9 \\ 0 & 1 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 3 & -9 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por tanto el sistema tiene infinitas soluciones y los vectores son linealmente dependientes.

La dependencia lineal de dos vectores se reduce a comprobar que uno es múltiplo del otro. En \mathbb{R}^n esto normalmente se decide a ojo mientras que en \mathbb{C}^n quizá nuestro ojo tenga que trabajar más y pedir ayuda a nuestra mano. Por ejemplo, los vectores $\vec{v}_1 = (1 + i, 1 + 2i, i - 1)^t$ y $\vec{v}_2 = (1 + 2i, 3i - 1, i - 2)^t$ de \mathbb{C}^3 son linealmente dependientes si y solo si $\vec{v}_2 = \frac{1+2i}{1+i} \vec{v}_1$, para que cuadren las primeras coordenadas, y un cálculo muestra que esta relación se cumple.

Consideremos ahora los polinomios $P_1 = 1 + x + x^2$, $P_2 = 3 + 5x + x^2$, $P_3 = 7 - x^2 + x^3$ del espacio vectorial de polinomios reales en x . La ecuación relevante es

$$\lambda_1(1 + x + x^2) + \lambda_2(3 + 5x + x^2) + \lambda_3(7 - x^2 + x^3) = 0 \quad \text{con } \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}.$$

Agrupando los términos del mismo grado,

$$(\lambda_1 + 3\lambda_2 + 7\lambda_3) + (\lambda_1 + 5\lambda_2)x + (\lambda_1 + \lambda_2 - \lambda_3)x^2 + \lambda_3x^3 = 0.$$

El polinomio nulo es el que tiene todos sus coeficientes nulos, por tanto llegamos al sistema homogéneo

$$A\vec{\lambda} = \vec{0} \quad \text{con } A = \begin{pmatrix} 1 & 3 & 7 \\ 1 & 5 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

Está claro que $\text{rg}(A) = 3$ simplemente considerando las tres últimas filas (el rango no puede ser mayor que tres). Entonces el sistema es compatible determinado, solo tiene la solución trivial, y los polinomios P_1, P_2, P_3 resultan ser linealmente independientes.

Las expresiones como las del primer miembro de (2.1) se dice que son *combinaciones lineales* de los vectores $\vec{v}_1, \dots, \vec{v}_n$. El conjunto de todas las combinaciones lineales lo denotaremos con $\mathcal{L}(\{\vec{v}_1, \dots, \vec{v}_n\})$. Está claro tras la Proposición 2.1.1 que forman un subespacio W de V . Se dice que $\{\vec{v}_1, \dots, \vec{v}_n\}$ genera W o que es un *sistema de generadores* de W . Si C es un conjunto infinito se define $\mathcal{L}(C)$ como la unión de los $\mathcal{L}(C')$ para todos los subconjuntos finitos $C' \subset C$. De esta forma se eliminan problemas de convergencia que pertenecen al negociado de los cursos de análisis.

Con ello ya estamos preparados para dar una definición matemática de base que concrete la idea de un sistema de medida, una elección de puntos cardinales, en el que no hay redundancia y que permite representar cualquier elemento, cualquier observación, en nuestro espacio.

Dado un espacio vectorial V , se dice que un conjunto $\mathcal{B} \subset V$ es una *base* de V si los vectores de \mathcal{B} son linealmente independientes y $\mathcal{L}(\mathcal{B}) = V$.

Fijada una base finita⁷ $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ de un espacio vectorial V cada vector $\vec{u} \in V$ se escribe de forma única como $\vec{u} = \lambda_1\vec{v}_1 + \dots + \lambda_n\vec{v}_n$ porque si hubiera dos combinaciones lineales que dieran \vec{u} , al restarlas llegaríamos a algo del tipo (2.1) que contradiría la independencia lineal. Estos λ_j se dice que son las *coordenadas* de \vec{u} en la base \mathcal{B} e incluso a veces se escribe, abusando un poco de la notación, $\vec{u} = (\lambda_1, \dots, \lambda_n)^t$ especificando la base en la que trabajamos.

⁷En este curso solo nos ocuparemos de hacer cálculos con bases finitas, sin embargo hay espacios de gran interés en ingeniería que se salen de este esquema y aparecerán en otras asignaturas. Desde el punto de vista matemático son tema de estudio del *análisis funcional*, a medio camino entre el álgebra y el análisis. Su desarrollo estuvo inicialmente a la formalización de la física cuántica.

En \mathbb{R}^2 la base más sencilla es $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ con $\vec{e}_1 = (1, 0)^t$ y $\vec{e}_2 = (0, 1)^t$, llamada la *base canónica* que se extiende a \mathbb{R}^n y \mathbb{C}^n (y en general a K^n) de la manera obvia tomando \vec{e}_i con un 1 en el lugar i y ceros en el resto. Volviendo al ejemplo del principio de la sección, el vector $\vec{u} = (1, 2)^t \in \mathbb{R}^2$ que corresponde a un paso al este y dos al norte es $\vec{u} = \vec{e}_1 + 2\vec{e}_2$. Sin embargo en la base $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})^t$ y $\vec{v}_2 = (0, 1)^t$, que corresponde a usar el noreste y el norte, se tiene $\vec{u} = \sqrt{2}\vec{v}_1 + \vec{v}_2$. Identificando \vec{u} con sus coordenadas, \vec{u} es $(1, 2)^t$ cuando empleamos la base canónica y es $(\sqrt{2}, 1)^t$ cuando usamos la base \mathcal{B}' .

Si tenemos una base con n elementos ($\neq \infty$) lo que hagamos con las combinaciones lineales de vectores potencialmente raros (polinomios, señales, etc.) se traduce fielmente en lo que hagamos con sus vectores de coordenadas que están en K^n , que en el curso serán los familiares \mathbb{R}^n o \mathbb{C}^n . Por si no está claro, aquí va como enunciado teórico:

Lema 2.2.1. *Sea V un espacio vectorial sobre K con una base finita de n elementos y sean $\vec{c}_1, \dots, \vec{c}_m \in K^n$ los vectores formados por las coordenadas de $\vec{u}_1, \dots, \vec{u}_m \in V$. Entonces $\sum_j \lambda_j \vec{u}_j = \vec{0}$ con $\lambda_j \in K$ si y solo si $\sum_j \lambda_j \vec{c}_j = \vec{0}$.*

El número de grados de libertad de un sistema físico es independiente de la manera en que lo describamos. El análogo en el contexto de los espacios vectoriales es el siguiente resultado-definición:

Proposición 2.2.2. *Todas las bases de un espacio vectorial V tienen el mismo número de elementos. Tal número se llama dimensión de V y se indica con $\dim V$.*

Inciendo en un comentario anterior, podría ocurrir $\dim V = \infty$ pero no vamos a hallar bases en este curso cuando ocurra esto. Si te empeñas mucho en tener un ejemplo, $\{1, x, x^2, x^3, \dots\}$ es una base del espacio de polinomios.

Si ahora leemos con esta notación y con ojos de ver el Teorema 1.2.2, nos percatamos de que nos asegura lo siguiente:

Corolario 2.2.3. *Sea $V = \{\vec{x} \in K^n : A\vec{x} = \vec{0}\}$ con $A \in \mathcal{M}_{m \times n}(K)$. Se cumple $\dim V = n - \text{rg}(A)$.*

En este curso $K = \mathbb{R}, \mathbb{C}$ aunque el resultado tiene validez en cualquier cuerpo ya que sumas, restas, multiplicaciones y divisiones es todo lo que necesitamos para hacer y deshacer las transformaciones elementales de la eliminación de Gauss. Una base de V es $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_{n-r}\}$ con \vec{v}_j como en el Teorema 1.2.2. Obviamente $\mathcal{L}(\mathcal{B}) = V$ y los \vec{v}_j son linealmente independientes porque en otro caso habría diferentes posibles elecciones de λ_j dando $\vec{x} = \vec{0}$.

En la práctica, a menudo sabemos la dimensión o bien por este corolario o porque existe una base fácil. Con ello no es necesario comprobar que un candidato a base genera.

Proposición 2.2.4. *Si V es un espacio vectorial con $\dim V = n < \infty$, cualesquiera n vectores linealmente independientes forman una base.*

Hay una interpretación del rango en términos de dimensiones distinta de la del Corolario 2.2.3 que implica que el rango de una matriz y de su traspuesta coinciden, lo cual suena muy sorprendente porque los cálculos de la eliminación de Gauss sobre A y A^t no guardan ninguna relación. Entender el milagro requiere leer la demostración. Lo hagas o no, conviene que recuerdes a partir de ejemplos anteriores que para cualquier matriz A el producto $A\vec{\lambda}$ es la combinación lineal $\lambda_1\vec{c}_1 + \dots + \lambda_n\vec{c}_n$ con \vec{c}_j las columnas de A .

Proposición 2.2.5. *Sea $A \in \mathcal{M}_{m \times n}(K)$ entonces $\text{rg}(A)$ es igual a la dimensión del espacio generado por sus columnas y también a la del espacio generado por sus filas. En particular $\text{rg}(A) = \text{rg}(A^t)$.*

Las pruebas de las proposiciones anteriores son sencillas e ilustrativas pero las pospondremos en favor de presentar cuanto antes algunos ejemplos.

Como primer ejemplo, calculemos la dimensión y una base del subespacio de \mathbb{C}^3

$$W = \{\vec{x} \in \mathbb{C}^3 : A\vec{x} = \vec{0}\} \quad \text{con} \quad A = \begin{pmatrix} 2 & 1 & 5+i \\ 3 & 2i & 4+2i \\ 4 & -1+4i & 3+3i \end{pmatrix}.$$

Un truco (no obligatorio) para evitar a toda costa que aparezcan fracciones al aplicar eliminación de Gauss es multiplicar la primera fila por 3 y la segunda por 2 antes de empezar, tras lo cual procedemos con

$$\begin{array}{ccc} \xrightarrow{\substack{f_2 \mapsto f_2 - f_1 \\ f_3 \mapsto f_3 - 2f_1/3}} & \begin{pmatrix} 6 & 3 & 15+3i \\ 0 & -3+4i & -7+i \\ 0 & -3+4i & -7+i \end{pmatrix} & \xrightarrow{\substack{f_1 \mapsto f_1/3 \\ f_3 \mapsto f_3 - f_2}} & \begin{pmatrix} 2 & 1 & 5+i \\ 0 & -3+4i & -7+i \\ 0 & 0 & 0 \end{pmatrix}. \end{array}$$

De aquí $\text{rg}(A) = 2$ y el Corolario 2.2.3 implica $\dim W = 3 - 2 = 1$. Una base estará formada por cualquier vector no nulo de W . Tomando⁸ $x_3 = 1$ se deduce de la forma escalonada $x_2 = -1 - i$ y $x_1 = -2$ que produce $\mathcal{B} = \{(-2, -1 - i, 1)^t\}$.

Vamos ahora a presentar un enunciado con aspecto bien diferente en el que se utilizan las mismas herramientas. Queremos comprobar que los siguientes subespacios $W_1, W_2 \subset \mathbb{R}^4$ son en realidad iguales:

$$W_1 = \left\{ \vec{x} \in \mathbb{R}^4 : \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 - x_3 + 2x_4 = 0 \end{array} \right\} \quad \text{y} \quad W_2 = \mathcal{L}(\{\vec{v}_1, \vec{v}_2\})$$

con $\vec{v}_1 = (3, 1, -2, -2)^t$ y $\vec{v}_2 = (0, 1, -1, 0)^t$.

Un paso de eliminación de Gauss muestra que la matriz A del sistema en W_1 tiene $\text{rg}(A) = 2$ por tanto $\dim W_1 = 4 - 2 = 2$. Además $A\vec{v}_1 = A\vec{v}_2 = \vec{0}$ implica $\vec{v}_1, \vec{v}_2 \in W_1$. Si mostramos que \vec{v}_1 y \vec{v}_2 son linealmente independientes, serán base tanto de W_1

⁸Aquí están las cosas preparadas para que $-7 + i$ sea "divisible" por $-3 + 4i$. Si no fuera así y quisiéramos evitar las fracciones y minimizar las cuentas con números complejos, una elección natural sería $x_3 = -3 + 4i$.

como de W_2 y se tendrá $W_1 = W_2$. Ahora bien, \vec{v}_1 y \vec{v}_2 no son uno múltiplo del otro, por tanto son linealmente independientes.

El siguiente ejemplo es un problema de un antiguo examen. Es muy sencillo y lo único que podría asustarte es que habla de un espacio de matrices. Se considera

$$W = \left\{ \begin{pmatrix} \lambda - \mu & 2\lambda \\ \mu & \lambda + \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

Lo que se pedía en el examen era mostrar que es un subespacio de $\mathcal{M}_2(\mathbb{R})$ y calcular su dimensión. Se tiene

$$\begin{pmatrix} \lambda - \mu & 2\lambda \\ \mu & \lambda + \mu \end{pmatrix} = \lambda A_1 + \mu A_2 \quad \text{con} \quad A_1 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Es decir $W = \mathcal{L}(\{A_1, A_2\})$ y el subespacio generado por dos vectores (en nuestro caso con aspecto de matriz) es obviamente subespacio (si te empeñas apela a la Proposición 2.1.1). Está claro que A_1 y A_2 son linealmente independientes ya que $\lambda A_1 + \mu A_2 = O$ implica $\lambda = \mu = 0$, por tanto $\{A_1, A_2\}$ es base y $\dim W = 2$.

Siguiendo con problemas de examen que no involucran subespacios de \mathbb{R}^n , en uno de ellos se pedía mostrar que $\mathcal{B} = \{1, x+1, (x+1)^2\}$ es una base de $\mathbb{R}_2[x]$, el espacio vectorial de polinomios reales en x de grado a lo más 2, y calcular las coordenadas de $P = 7 + 4x + x^2$ en dicha base. Claramente $\mathbb{R}_2[x] = \mathcal{L}(\{1, x, x^2\})$ y $1, x, x^2$ son linealmente independientes (un polinomio es nulo cuando los coeficientes de x^j lo son) por tanto $\dim \mathbb{R}_2[x] = 3$. Entonces usando la Proposición 2.2.4, basta probar que los elementos de \mathcal{B} son linealmente independientes. Operando y agrupando por grados se tiene

$$\lambda_1 + \lambda_2(x+1) + \lambda_3(x+1)^2 = \lambda_1 + \lambda_2 + \lambda_3 + (\lambda_2 + 2\lambda_3)x + \lambda_3x^2.$$

Este es el polinomio nulo solo si $\lambda_1 + \lambda_2 + \lambda_3 = \lambda_2 + 2\lambda_3 = \lambda_3 = 0$, lo que implica $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Para hallar las coordenadas de P lo igualamos a esta combinación lineal, resultando el sistema $\lambda_1 + \lambda_2 + \lambda_3 = 7$, $\lambda_2 + 2\lambda_3 = 4$, $\lambda_3 = 1$ cuya solución es $(\lambda_1, \lambda_2, \lambda_3) = (4, 2, 1)$. Estas son las coordenadas de P en la base \mathcal{B} .

Terminamos con las pruebas que faltaban.

Demostración de la Proposición 2.2.2. Veamos que es imposible que haya dos bases $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ y $\mathcal{B}' = \{\vec{c}_1, \dots, \vec{c}_m\}$ con $m > n$.

Por la definición de base, $\mathcal{L}(\mathcal{B}) = V$, en particular existen coeficientes a_{ij} tales que

$$\vec{c}_j = a_{1j}\vec{b}_1 + a_{2j}\vec{b}_2 + \dots + a_{nj}\vec{b}_n, \quad 1 \leq j \leq m.$$

Sea $A = (a_{ij})_{i,j=1}^{n,m}$, necesariamente $\text{rg}(A) \leq n$ (a lo más un pivote por fila). De este modo, si $n < m$ el sistema $A\vec{\lambda} = \vec{0}$ tiene infinitas soluciones y existen $\lambda_1, \dots, \lambda_m$ no todos nulos con $\sum_j a_{ij}\lambda_j = 0$. Esto implica $\sum_j \lambda_j \vec{c}_j = \vec{0}$ y contradice la independencia lineal de los \vec{c}_j . \square

Demostración de la Proposición 2.2.4. Si $\vec{v}_1, \dots, \vec{v}_n$ son linealmente independientes hay que ver que para cualquier $\vec{u} \in V$ se tiene $\vec{u} \in \mathcal{L}(\{\vec{v}_1, \dots, \vec{v}_n\})$. Procediendo como en la demostración anterior con $m = n + 1$ y $\vec{c}_1 = \vec{v}_1, \dots, \vec{c}_n = \vec{v}_n, \vec{c}_{n+1} = \vec{u}$, se obtiene $\sum_j \lambda_j \vec{c}_j = \vec{0}$ con λ_j no todos nulos. Si $\lambda_{n+1} \neq 0$ se deduce $\vec{u} = -\lambda_{n+1}^{-1} \sum_{j=1}^n \lambda_j \vec{c}_j \in \mathcal{L}(\{\vec{v}_1, \dots, \vec{v}_n\})$. El caso $\lambda_{n+1} = 0$ es imposible porque implicaría que $\vec{v}_1, \dots, \vec{v}_n$ son linealmente dependientes. \square

Demostración de la Proposición 2.2.5. El Teorema 1.2.2 asegura que si A' es la matriz formada por las $\text{rg}(A)$ columnas pivote de A , el sistema $A'\vec{\lambda} = \vec{0}$ tiene solución única trivial y si añadimos alguna otra columna hay soluciones no triviales, por tanto $\text{rg}(A)$ es la dimensión del espacio generado por las columnas.

Por otro lado, las transformaciones elementales preservan el espacio generado por las filas, no cambian el conjunto de todas sus combinaciones lineales, por tanto su dimensión es la misma que la del espacio generado por las filas de la forma escalonada reducida. Esta claro que tal dimensión es el número de filas no nulas, la cual coincide con el número de escalones $\text{rg}(A)$. \square

Exprimiendo el silicio [opcional]. Veamos cómo resolver computacionalmente los dos primeros ejemplos que siguen a la Proposición 2.2.4. En la salida se muestra una fila de guiones separando ambos ejemplos.

Con `matlab/octave` una posibilidad es el siguiente código:

```

1 % Matriz (compleja) A
2 A = [2,1,5+1; 3, 2*I, 4+2*I; 4,4*I-1,3+3*I ];
3 % La dimensión es n-rg(A). n = size(A,1)
4 size(A,1)-rank(A)
5 % Una base es la única columna de
6 null(A)
7
8 disp('-----')
9
10 % Matriz que define W_1
11 A = [1, 1, 1, 1; 1, -1, -1, 2];
12 % Vectores que definen W_2
13 v1 = [3;1;-2;-2]
14 v2 = [0;1;-1;0]
15 % Este producto es la matriz nula por tanto
16 % W_2 está incluido en W_1
17 A*[v1, v2]
18 % Una base de W_1 son las columnas de
19 null(A)
20 % Si añadimos v1 y v2 el rango es dos
21 rank( [null(A) v1 v2])
22 % Por tanto v1 y v2 son combinaciones lineales
23 % de la base de W_1 y se concluye que
24 % W_1 está incluido en W_2.

```

Todo está explicado en los comentarios y los comandos ya han aparecido antes con la excepción de `size(A,1)` que da el número de filas de la matriz A , el de columnas es `size(A,2)`, y de `disp` que muestra (*display*) algo en la salida.

Quizá las líneas finales merezcan una explicación: Las columnas \vec{b}_1, \vec{b}_2 de la matriz $M = \text{null}(A)$ dan una base de W_1 por tanto $M\vec{\lambda} = \vec{0}$ tiene solución única y $\text{rg}(M) = 2$. Si el rango de la línea 21 se mantiene en 2, entonces $M\vec{\lambda} = \vec{v}_1$ y $M\vec{\lambda} = \vec{v}_2$ también tendrán solución y estos vectores se expresan como combinación lineal de \vec{b}_1 y \vec{b}_2 .

En `sagemath` hay comandos especiales para calcular dimensiones y bases.

```

1 # Matriz (compleja) A
2 A = matrix(3,3, [2,1,5+I, 3, 2*I, 4+2*I, 4,4*I-1,3+3*I ])
3 # Construye el espacio de los que anulan A
4 W = A.right_kernel()
5 # Calcula la dimensión
6 print(W.dimension())
7 # Calcula una base
8 print(W.basis())
9
10 print('-----')
11
12 # Matriz que define W_1
13 A = matrix(QQ,2,4,[1, 1, 1, 1, 1, -1, -1, 2])
14 # Construye W_1
15 W1 = A.right_kernel()
16 # Construye W_2
17 Q4 = VectorSpace(QQ,4)
18 v1 = vector([3,1,-2,-2])
19 v2 = vector([0,1,-1,0])
20 W2 = Q4.span([v1, v2])
21 # ¿Son iguales?
22 print(W1==W2)

```

Incluso, como se muestra en la parte final, es posible preguntar directamente si los dos subespacios son iguales (línea 22). Para la construcción de los subespacios se ha usado `right_kernel` (línea 15), que da el conjunto de soluciones de un sistema homogéneo y `span` que da $\mathcal{L}(C)$ con C un conjunto de vectores (línea 20). Se ha utilizado \mathbb{Q}^4 en lugar de \mathbb{R}^4 para que no se introduzcan decimales, eso es lo que indica `QQ`. También podría reemplazarse en las líneas 13 y 17 por `SR` que indica *symbolic ring*, algo así como no tengo ganas de especificar dónde trabajo y quiero que las cuentas sean exactas.

2.3. Aplicaciones lineales

Los espacios vectoriales son la estructura subyacente del álgebra lineal, el marco sobre el que se construye, sin embargo el concepto importante que da sentido a este cimientamiento es el de aplicación lineal.

Dados espacios vectoriales V y W sobre un mismo cuerpo K , una función $f : V \rightarrow W$ se dice que es una *aplicación lineal* si preserva las combinaciones lineales, es decir, si para $\vec{u}, \vec{v} \in V$, $\lambda, \mu \in K$ cualesquiera $f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v})$.

Igual que la Tierra a nuestra escala nos parece plana o al observar una curva con una lupa de muchos aumentos la confundimos con una recta, los incrementos de cualquier función decente de una o varias variables a pequeña escala se aproxima por una aplicación lineal. De esta forma el álgebra lineal se convierte en una teoría bastante universal de las pequeñas variaciones. Por ejemplo, $f(x, y) = \sin(x - y) + e^{\cos(xy)}$ es una función muy rara pero si nos movemos de $\vec{0} = (0, 0)$ a los puntos cercanos $\vec{u} = (0.03, 0.02)$ y $\vec{v} = (0.05, 0.01)$ se tiene que la función incremento $g(x, y) = f(x, y) - f(0, 0)$ verifica $g(\vec{u}) + g(\vec{v}) = 0.049988\dots$, $g(\vec{u} + \vec{v}) = 0.049971\dots$ y $g(2\vec{v}) = 0.079909\dots$, $2g(\vec{v}) = 0.079977\dots$, los números son muy similares.

Supongamos $V = \mathbb{R}^n$, $W = \mathbb{R}^m$. Si $\{\vec{e}_1, \dots, \vec{e}_n\}$ es la base canónica de V , cada $\vec{x} \in V$ es de la forma $\vec{x} = (x_1, \dots, x_n)^t = \sum_j x_j \vec{e}_j$ y se cumplirá $f(\vec{x}) = \sum_j x_j f(\vec{e}_j)$. Dada f , los vectores $f(\vec{e}_j) \in W$ son constantes $f(\vec{e}_j) = (a_{1j}, a_{2j}, \dots, a_{mj})^t$. Por tanto la i -ésima coordenada de $f(\vec{x})$ es $\sum_j a_{ij} x_j$. Con ello hemos probado el siguiente

resultado que es la razón de ser de las matrices en álgebra lineal:

Proposición 2.3.1. *Cualquier aplicación lineal $f : K^n \rightarrow K^m$ es de la forma $f(\vec{x}) = A\vec{x}$ con $A \in \mathcal{M}_{m \times n}(K)$. Las columnas de A son los vectores $f(\vec{e}_j)$ con $\{\vec{e}_1, \dots, \vec{e}_n\}$ la base canónica de K^n .*

Una vez más, de cara a los ejemplos de este curso $K = \mathbb{R}, \mathbb{C}$ pero nada impide considerar otros cuerpos, siendo $\mathcal{M}_{m \times n}(K)$ las matrices con elementos en dicho cuerpo.

Ahora estamos en condiciones de entender plenamente por qué las matrices se multiplican de manera tan rara. La única manera de usar una aplicación lineal f sobre el resultado de otra aplicación g , lo que se llama *composición* de funciones $f \circ g$, es que g acabe donde comienza f . Digamos $K^{n_1} \xrightarrow{g} K^{n_2} \xrightarrow{f} K^{n_3}$. Si $A \in \mathcal{M}_{n_3 \times n_2}(K)$ y $B \in \mathcal{M}_{n_2 \times n_1}(K)$ son las matrices de f y g , respectivamente,

$$(2.2) \quad (f \circ g)(\vec{x}) = f(g(\vec{x})) = A(B\vec{x}).$$

La función $f \circ g$ es también una aplicación lineal, por tanto $(f \circ g)(\vec{x}) = C\vec{x}$. ¿Qué relación hay entre A , B y C ? La i -ésima coordenada de $(f \circ g)(\vec{x})$ es $\sum_k a_{ik} m_k$ con m_k la k -ésima coordenada de $B\vec{x}$, que es $\sum_j b_{kj} x_j$. Por consiguiente $c_{ij} = \sum_k a_{ik} b_{kj}$. Esta es la manera de operar matrices para que indique la composición. Dejándonos llevar por la notación (2.2) quitando paréntesis, decimos que esto es una multiplicación, aunque no comparta muchas propiedades con la habitual.

Para que la prueba de la Proposición 2.3.1 funcione, todo lo que necesitamos es que haya coordenadas y las tendremos siempre que fijemos una base. Es decir, por el mismo precio tenemos:

Proposición 2.3.2. *Sea $f : V \rightarrow W$ una aplicación lineal entre espacios vectoriales de dimensión finita. Fijadas bases \mathcal{B}_V y \mathcal{B}_W de V y W , las coordenadas de \vec{x} en la base \mathcal{B}_V y las de $f(\vec{x})$ en \mathcal{B}_W están relacionadas mediante la multiplicación por cierta matriz. Sus columnas son las coordenadas en \mathcal{B}_W de las imágenes de los elementos de \mathcal{B}_V .*

Esta matriz se dice que es la *matriz de la aplicación lineal*, aunque en la letra pequeña de nuestra mente deberíamos añadir “una vez fijadas bases”.

Veamos un ejemplo en el espacio $\mathbb{R}_2[x]$ de polinomios reales en x de grado a lo más 2 en el que fijamos la base $\{1, x, x^2\}$. Consideremos la función $f : \mathbb{R}_2[x] \rightarrow \mathbb{R}_2[x]$ dada por $f(P) = (x-1)P' + P$, con P' la derivada. Es una aplicación lineal porque para $P, Q \in \mathbb{R}_2[x]$, $\lambda, \mu \in \mathbb{R}$ se tiene que $f(\lambda P + \mu Q)$ es

$$(x-1)(\lambda P' + \mu Q') + \lambda P + \mu Q = \lambda((x-1)P' + P) + \mu((x-1)Q' + Q) = \lambda f(P) + \mu f(Q).$$

Calculemos la matriz A de f con la base elegida. Según la Proposición 2.3.2 tenemos que poner en columnas las coordenadas de $f(1)$, $f(x)$ y $f(x^2)$:

$$\begin{array}{lll} f(1) & = 1 & = 1 + 0x + 0x^2 \\ f(x) & = x - 1 + x & = -1 + 2x + 0x^2 \\ f(x^2) & = (x-1)2x + x^2 & = 0 - 2x + 3x^2 \end{array} \implies A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Asociados a una aplicación lineal $f : V \rightarrow W$ se define el *núcleo* de f y la *imagen* de f como los subespacios de V y W dados respectivamente por

$$\text{Ker}(f) = \{\vec{x} \in V : f(\vec{x}) = \vec{0}\}, \quad \text{Im}(f) = \{\vec{y} \in W : \vec{y} = f(\vec{x}) \text{ con } \vec{x} \in V\}.$$

Este “Ker” es la abreviatura de *kernel*, núcleo en inglés. A veces en textos en español se escribe Nuc(f) en vez de Ker(f).

Estos subespacios tienen su traducción en unas propiedades de funciones que seguramente ya conozcas. Si no es el caso, están definidas en el enunciado entre paréntesis. El resultado en sí es muy sencillo, en parte trivial, es más una excusa para recordar o aprender las definiciones.

Proposición 2.3.3. *Sea $f : V \rightarrow W$ una aplicación lineal. Se cumple:*

- a) $\text{Ker}(f) = \{\vec{0}\}$ si y solo si f es inyectiva (valores distintos de \vec{x} dan $f(\vec{x})$ distintos).
- b) $\text{Im}(f) = W$ si y solo si f es sobreyectiva (todo $\vec{w} \in W$ es $f(\vec{v})$ con $\vec{v} \in V$).
- c) $\text{Im}(f) = W$ y $\text{Ker}(f) = \{\vec{0}\}$ si y solo si f es biyectiva (inyectiva y sobreyectiva).

Demostración. Por supuesto, b) y c) son obvias. Para a) basta notar que $f(\vec{x}_1) = f(\vec{x}_2)$ con $\vec{x}_1, \vec{x}_2 \in V$ distintos implica $f(\vec{v}) = \vec{0}$ con $\vec{v} = \vec{x}_1 - \vec{x}_2 \neq \vec{0}$. Recíprocamente, si $\vec{v} \neq \vec{0}$ está en $\text{Ker}(f)$ podríamos tomar $\vec{x}_1 = \vec{v}$, $\vec{x}_2 = \vec{0}$. \square

El cálculo de la dimensión de $\text{Ker}(f)$ e $\text{Im}(f)$ se deduce del rango de la matriz de f independientemente de las bases elegidas. Esencialmente lo único que hay que hacer para demostrar el siguiente resultado es utilizar coordenadas para pasar todo a K^n donde tenemos el Corolario 2.2.3 y la Proposición 2.2.5.

Proposición 2.3.4. *Sea A la matriz de una aplicación lineal $f : V \rightarrow W$ para ciertas bases (supuestas finitas) de V y W . Se cumple*

$$\dim \text{Ker}(f) = \dim V - \text{rg}(A) \quad \text{y} \quad \dim \text{Im}(f) = \text{rg}(A).$$

Además los vectores cuyas coordenadas en la base elegida para W corresponden a columnas pivote de A son una base de $\text{Im}(f)$.

Demostración. Sean $\mathcal{B}_V = \{\vec{v}_1, \dots, \vec{v}_n\}$ y $\mathcal{B}_W = \{\vec{w}_1, \dots, \vec{w}_m\}$ las bases escogidas de V y W . Dado $\vec{u} \in V$ con coordenadas $\vec{c} \in K^n$, esto es $\vec{u} = \sum_j c_j \vec{v}_j$, se cumple $f(\vec{u}) = \vec{0}$ si y solo si $\vec{c} \in V' = \{\vec{x} \in K^n : A\vec{x} = \vec{0}\}$. Por el Lema 2.2.1 y el Corolario 2.2.3 se deduce $\dim \text{Ker}(f) = \dim V' = n - \text{rg}(A)$.

De la misma forma, $\vec{w} \in W$ con coordenadas $\vec{d} \in K^m$ cumple $\vec{w} = f(\vec{u})$ si y solo si $\vec{d} = A\vec{c}$, por la Proposición 2.3.1. Equivalentemente si y solo si \vec{d} está en el espacio generado por las columnas de A y basta emplear la Proposición 2.2.5. \square

Notando $A \in \mathcal{M}_{m \times n}$ con $n = \dim V$ y $m = \dim W$ y que es $\text{rg}(A)$ es a lo más la menor de estas cantidades, es fácil ver que f no puede ser inyectiva si $\dim V > \dim W$ ni sobreyectiva si $\dim V < \dim W$. Un poco más curioso, y todavía directo a partir del resultado anterior, es:

Corolario 2.3.5. *Suponiendo $\dim V = \dim W = n < \infty$, una aplicación lineal $f : V \rightarrow W$ es biyectiva si y solo si su matriz A verifica $\text{rg}(A) = n$.*

Por la Proposición 1.3.2, $\text{rg}(A) = n$ es el criterio para que A sea invertible. Una aplicación lineal biyectiva $f : V \rightarrow W$ tiene inversa $f^{-1} : W \rightarrow V$ que también es aplicación lineal. Una vez fijadas bases, su matriz es la matriz inversa de la matriz de f . De nuevo esto radica en la relación entre composición y producto de matrices como en (2.2).

Practicemos sobre un ejemplo proveniente de un examen pasado. Queremos hallar bases de $\text{Ker}(f)$ e $\text{Im}(f)$ para

$$f : \mathbb{R}^4 \rightarrow \mathbb{R}^4 \quad \text{dada por} \quad f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 4 & 2 & -1 & 3 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 3 & -2 \\ 9 & 2 & -1 & 8 \end{pmatrix}$$

y decidir si es inyectiva o sobreyectiva.

Intercambiando la primera y la tercera fila la eliminación de Gauss no induce fracciones en la primera columna:

$$A \xrightarrow{f_1 \leftrightarrow f_3} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 2 & 1 & 0 & 1 \\ 4 & 2 & -1 & 3 \\ 9 & 2 & -1 & 8 \end{pmatrix} \xrightarrow{\substack{f_2 \mapsto f_2 - 2f_1 \\ f_3 \mapsto f_3 - 4f_1 \\ f_4 \mapsto f_4 - 9f_1}} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 2 & -13 & 11 \\ 0 & 2 & -28 & 26 \end{pmatrix}$$

y los otros pasos son también sencillos

$$\xrightarrow{\substack{f_3 \mapsto f_3 - 2f_2 \\ f_4 \mapsto f_4 - 2f_2}} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -16 & 16 \end{pmatrix} \xrightarrow{f_4 \mapsto f_4 - 16f_3} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Por consiguiente $\text{rg}(A) = 3$ y la Proposición 2.3.4 da las dimensiones $\dim \text{Ker}(f) = 1$ y $\dim \text{Im}(f) = 3$. De ello se deduce que f no es inyectiva ni sobreyectiva, esto último requeriría $\dim \text{Im}(f) = \dim \mathbb{R}^4 = 4$. Eligiendo $x_4 = 1$ en el sistema $A\vec{x} = \vec{0}$ se obtiene $x_2 = x_3 = 1$, $x_1 = -1$, por tanto $\{(-1, 1, 1, 1)\}$ es una base de $\text{Ker}(f)$. Las columnas pivote de A son las tres primeras y de nuevo por la Proposición 2.3.4 conforman una base de $\text{Im}(f)$.

Vayamos ahora a nuestro viejo amigo el espacio de polinomios $\mathbb{R}_2[x]$ con la base $\mathcal{B} = \{1, x, x^2\}$ y hallemos bases de $\text{Ker}(f)$ e $\text{Im}(f)$ para la aplicación lineal $f(P) = xP'' + 6x \int_0^1 P$ con P'' la derivada segunda.

Comenzamos formando la matriz de f con las coordenadas de $f(1)$, $f(x)$ y $f(x^2)$ en columna, como en un ejemplo anterior:

$$\begin{aligned} f(1) &= 6x \int_0^1 1 \, dx &= 0 + 6x + 0x^2 \\ f(x) &= 6x \int_0^1 x \, dx &= 0 + 3x + 0x^2 \\ f(x^2) &= 2x^2 + 6x \int_0^1 x^2 \, dx &= 0 + 2x + 2x^2 \end{aligned} \quad \implies \quad A = \begin{pmatrix} 0 & 0 & 0 \\ 6 & 3 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Moviendo la primera fila al final ya tenemos la forma escalonada, por tanto $\text{rg}(A) = 2$ y se deduce $\dim \text{Ker}(f) = 1$, $\dim \text{Im}(f) = 2$. Se cumple $A\vec{c} = \vec{0}$ con $\vec{c} = (1, -2, 0)^t$ y estas son las coordenadas de $1 - 2x + 0x^2$, entonces $\{1 - 2x\}$ es una base de $\text{Ker}(f)$. Las columnas pivote son la primera y la tercera, entonces una base de $\text{Im}(f)$ está formada por los vectores (polinomios) cuyas coordenadas son $(0, 6, 0)^t$ y $(0, 2, 2)^t$, es decir, $\{6x, 2x + 2x^2\}$. Evidentemente la podemos “simplificar” a $\{x, x + x^2\}$. También sería válida $\{x, x^2\}$, pues genera lo mismo. Por la Proposición 2.3.3, f no es ni inyectiva ni sobreyectiva.

Exprimiendo el silicio [opcional]. A pesar de que esta sección hay un concepto nuevo importante, el de aplicación lineal, los cálculos son similares a los que ya hemos hecho y por tanto no hay demasiado que decir aquí.

Con `matlab/octave` solo recordaremos que si capturamos la salida de `rref` con dos variables, la segunda será asignada con la lista de los números de las columnas pivote. Así en el ejemplo de la aplicación $f : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ produciríamos una base de $\text{Im}(f)$ con las líneas:

```
1 % Matriz A
2 A = [4,2,-1,3; 2,1,0,1; 1,0,3,-2; 9,2,-1,8];
3 [E, p] = rref(A);
4 % Una base de Im(f) es:
5 A(:,p)
```

y lo mismo funcionaría con cualquier otra matriz. La sintaxis de `matlab/octave` permite escribir `A(:,p)` en la última línea para indicar que las filas son arbitrarias y las columnas las que corresponden a los índices de `p`.

Con `sagemath` ya habíamos usado `right_kernel` y ahora encontramos sentido al nombre (por cierto `left_kernel` es esencialmente el núcleo de la traspuesta). Para la imagen se emplea `column_space` que es coherente con la Proposición 2.3.1. De esta forma, el ejemplo de la aplicación $f : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ se resolvería con:

```
1 # Matriz A
2 A = matrix(QQ,4,4,[4,2,-1,3, 2,1,0,1, 1,0,3,-2, 9,2,-1,8])
3 # Núcleo
4 nuc = A.right_kernel()
5 # Imagen
6 ima = A.column_space()
7 # Dimensiones
8 print('dim_Ker =', nuc.dimension(), 'dim_Im =', ima.dimension())
9 # Base del núcleo
10 print(nuc.basis())
11 # Base de la imagen
12 print(ima.basis())
```

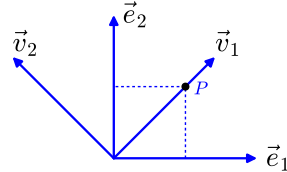
La base que produce de $\text{Im}(f)$ es más bonita que la nuestra.

2.4. Cambio de base

Corriendo el riesgo de propiciar el fenómeno psicológico de la profecía autocumplida, diré que el tema de esta sección es uno de los que más confunde a los estudiantes de álgebra lineal. La buena noticia es que, al menos a mi juicio, la confusión desapa-

rece cuando en vez de empeñarnos en memorizar una fórmula nos limitamos a tener en mente una idea sencilla.

La motivación del cambio de base es muy natural. Ya habíamos incidido sobre ello diciendo que una base es un sistema de medidas que especificamos, siendo las medidas los valores de las coordenadas en la base escogida, y a veces es conveniente pasar de un sistema de medidas a otro. Si me sitúo de pie en el origen mirando a la derecha, encontraré lógico expresar los puntos o vectores con $a(1, 0)^t + b(0, 1)^t$, a unidades a la derecha y b hacia arriba, las coordenadas cartesianas habituales. Pero si giro la cabeza 45° en sentido positivo (antihorario) mi nueva derecha es lo que mi yo sin la cabeza girada llamaría $(1/\sqrt{2}, 1/\sqrt{2})^t$ y mi nuevo arriba lo que llamaría $(-1/\sqrt{2}, 1/\sqrt{2})^t$.

$$\begin{aligned} \vec{e}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \vec{e}_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \vec{v}_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & \vec{v}_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$


Así tenemos la base antigua, la canónica $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ y la base nueva $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2\}$ con los vectores indicados. Si antes veía un pájaro en $(1/2, 1/2)^t$, ahora estará en $(1/\sqrt{2}, 0)^t$ como es fácil de comprobar en un dibujo o algebraicamente con la identidad $\frac{1}{2}\vec{e}_1 + \frac{1}{2}\vec{e}_2 = \frac{1}{\sqrt{2}}\vec{v}_1 + 0\vec{v}_2$. En realidad todo esto, de manera más o menos implícita lo sabíamos ya, la novedad es que vamos a examinar la fórmula general de un cambio de base y sobre todo vamos a ver cómo afecta a la matriz de una aplicación lineal $f : V \rightarrow V$. Por ejemplo, si el pájaro cae hacia el eje X reduciendo su altura a la mitad en cada unidad de tiempo, la primera base nos llevaría a considerar la aplicación lineal $(x_1, x_2)^t \mapsto (x_1, x_2/2)^t$ que representa cada una de las reducciones de altura pero ¿cómo las veríamos si hemos girado la cabeza?

Respecto a la fórmula general del cambio de base, si queremos hallar las coordenadas en la base \mathcal{B}' de un vector genérico $\vec{x} = (x_1, x_2)^t = x_1\vec{e}_1 + x_2\vec{e}_2$, debemos resolver $(x_1, x_2)^t = x'_1\vec{v}_1 + x'_2\vec{v}_2$ en x'_1 y x'_2 . Esta ecuación se escribe como

$$(2.3) \quad \vec{x} = C\vec{x}' \quad \text{con} \quad C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{donde} \quad \begin{cases} \vec{x} = \text{coordenadas en } \mathcal{B}, \\ \vec{x}' = \text{coordenadas en } \mathcal{B}'. \end{cases}$$

En general, sean cuales sean \mathcal{B} y \mathcal{B}' , en dimensión finita siempre tendremos el esquema (2.4)

$$\vec{x} = C\vec{x}' \quad \text{con} \quad \begin{cases} \vec{x} = \text{coord. en } \mathcal{B} \\ \vec{x}' = \text{coord. en } \mathcal{B}' \end{cases} \quad \text{colum. de } C = \begin{cases} \text{coord. de los vectores} \\ \text{de } \mathcal{B}' \text{ en la base } \mathcal{B}. \end{cases}$$

La idea que no debes olvidar es que el cambio de base se lleva a cabo multiplicando por una matriz, la cual tiene que ser cuadrada por la Proposición 2.2.2 y de hecho invertible ya que cambiando de \mathcal{B} a \mathcal{B}' y después de \mathcal{B}' a \mathcal{B} recuperaremos las coordenadas iniciales.

Las aplicaciones lineales $f : V \rightarrow W$ con $W = V$ tienen una mayor relevancia y reciben el nombre de *endomorfismos*, son las que actúan en los vectores de un espacio sin salirse de él. Por ejemplo, para $V = \mathbb{R}^2$, un giro alrededor del origen es un endomorfismo y también lo es la aplicación que envía $(x_1, x_2)^t$ a su sombra $(x_1, 0)^t$ en el eje X . Si $\dim V = n < \infty$, una vez fijada una base \mathcal{B} , las coordenadas conforman un vector de K^n y de acuerdo con la Proposición 2.3.2, f queda representada mediante la multiplicación por una matriz: en términos de las coordenadas la función es $\vec{y} = \vec{y}(\vec{x})$ dada por $\vec{y} = A\vec{x}$. ¿Cómo hallar la matriz A' que da la relación $\vec{y}' = A'\vec{x}'$ correspondiente a la base \mathcal{B}' ? Simplemente sustituyendo la relación (2.4) entre las coordenadas en \mathcal{B} y \mathcal{B}' :

$$\vec{y} = A\vec{x} \quad \begin{array}{l} \vec{x} = C\vec{x}' \\ \vec{y} = C\vec{y}' \end{array} \quad \vec{y}' = C^{-1}AC\vec{x}'.$$

Es decir, $A' = C^{-1}AC$. El cambio de base de la matriz de un endomorfismo es por tanto un sencillo cambio de variable, ¿dónde está la dificultad? Si miras tu libro favorito de álgebra lineal quizá leas allí $A' = CAC^{-1}$ porque parece más natural cambiar en (2.4) de lo antiguo a lo nuevo (¿evolución?) o habrá una fórmula que involucra memorizar notaciones más o menos arbitrarias. Mi consejo es que en vez de tratar de recordar la fórmula de un libro o de estas notas, pienses en términos de cambios de variable.

Como primer ejemplo, resolvamos el problema antes mencionado de averiguar en qué se transformaría $(x_1, x_2)^t \mapsto (x_1, x_2/2)^t$ cuando usamos la base $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = \frac{1}{\sqrt{2}}(1, 1)^t$, $\vec{v}_2 = \frac{1}{\sqrt{2}}(-1, 1)^t$. Tenemos que la aplicación lineal en la base canónica \mathcal{B} actúa como

$$\vec{x} \in \mathbb{R}^2 \mapsto \vec{y} = A\vec{x} \in \mathbb{R}^2 \quad \text{con} \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Esta A es la matriz correspondiente a $y_1 = x_1$, $y_2 = x_2/2$. Por (2.3) tenemos

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \vec{y}' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \vec{x}' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1/2 & 1/2 \end{pmatrix} \vec{x}'.$$

Así pues

$$\vec{y}' = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ 1/2 & 1/2 \end{pmatrix} \vec{x}' = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1/2 & 1/2 \end{pmatrix} \vec{x}' = \begin{pmatrix} 3/4 & -1/4 \\ -1/4 & 3/4 \end{pmatrix} \vec{x}'.$$

Esta última matriz es la que buscábamos del endomorfismo cuando consideramos la base \mathcal{B}' . Es decir, al girar la cabeza 45° el endomorfismo lo observamos como $(x'_1, x'_2)^t \mapsto \frac{1}{4}(3x'_1 - x'_2, -x'_1 + 3x'_2)^t$.

En un problema como el del cambio de base que involucra el producto de tres matrices vamos ahora con un ejemplo en el que el propio endomorfismo está definido mediante un producto de tres matrices. Todo un reto para comprobar que uno tiene las ideas claras.

Consideramos en $V = \{X \in \mathcal{M}_2(\mathbb{R}) : X = X^t\}$, el endomorfismo definido por

$$f(X) = TXT \quad \text{con} \quad T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

y las bases de V

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad \text{y} \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

Queremos calcular la matriz de f en la base \mathcal{B} y después cambiarla a la base \mathcal{B}' . Según la Proposición 2.3.2 para la primera tarea hay que poner en columna las coordenadas de las imágenes de los elementos de \mathcal{B} , estas son:

$$(2.5) \quad \begin{cases} f(B_1) = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} = 4B_1 + 2B_2 + B_3 \longrightarrow (4, 2, 1)^t, \\ f(B_2) = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix} = 4B_1 + 3B_2 + 2B_3 \longrightarrow (4, 3, 2)^t, \\ f(B_3) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = B_1 + B_2 + B_3 \longrightarrow (1, 1, 1)^t, \end{cases}$$

donde B_j son los elementos de \mathcal{B} . Además los de \mathcal{B}' son $B_1 + B_2$, $2B_1 + B_3$, $-B_1$. Por tanto la matriz A de f en la base \mathcal{B} y la matriz C que pasa coordenadas en \mathcal{B}' a coordenadas en \mathcal{B} son:

$$A = \begin{pmatrix} 4 & 4 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Seguramente ayuda a comprender que C lleva a cabo el cambio $\vec{x} = C\vec{x}'$ pensar en que la primera matriz de \mathcal{B}' es $(1, 0, 0)^t$ en su propia base y multiplicado por C debe dar $(1, 1, 0)^t$ que son las coordenadas en \mathcal{B} y entonces este vector conformará la primera columna. Lo mismo se aplica a las otras dos columnas.

La inversa de C es muy sencilla de calcular por reducción de Gauss-Jordan. Partiendo de $(C|I_3)$ poniendo la primera fila al final

$$\xrightarrow{f_3 \rightarrow f_3 - f_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & -1 & 1 & -1 & 0 \end{array} \right) \xrightarrow{f_3 \rightarrow f_3 - 2f_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & -1 & -2 \end{array} \right)$$

y basta cambiar de signo la última fila para obtener la inversa en la mitad derecha. La relación $\vec{y} = A\vec{x}$ equivale a $C\vec{y}' = AC\vec{x}'$ y así la matriz en la base \mathcal{B}' es

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

En problemas que conlleven cierto número de cálculos siempre es conveniente disponer de una comprobación, aunque sea parcial, del resultado. Si nuestros cálculos son correctos, según la Proposición 2.3.2 aplicada a la primera columna se debe tener

$$T \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} T = 5 \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + 3 \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} + 3 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

y unas sencillas cuentas muestran que es así.

Un comentario final es que resulta asequible (y esto está relacionado con que C sea fácil de invertir) llevar a cabo cálculos como los de (2.5) pero con los elementos de la base \mathcal{B}' lo que nos conduciría a la matriz final sin necesidad de pasar por la matriz en la base \mathcal{B} pero justamente el propósito del ejemplo era practicar con el cambio de base.

Exprimiendo el silicio [opcional]. El siguiente código para matlab/octave lleva a cabo los cálculos del cambio de base en el último ejemplo.

```

1 % Matriz T
2 T = [2,1; 1,1];
3 % Matrices de la base B
4 B1 = [1,0; 0,0];
5 B2 = [0,1; 1,0];
6 B3 = [0,0; 0,1];
7 % Imágenes concatenadas en una matriz
8 fB = [T*B1*T, T*B2*T, T*B3*T];
9 % Matriz de zeros (vacía)
10 A = zeros(3);
11 for j = 1:3
12     % Cada submatriz empieza en la columna 2j-1
13     % y las coordenadas en la base B son los
14     % valores de a_11, a_12 y a_22
15     coord = [fB(1,2*j-1); fB(1,2*j); fB(2,2*j)];
16     % Rellena la columna j de A
17     A(:,j) = coord;
18 end
19 % Matriz de cambio de base
20 C = [1,2,-1; 1,0,0; 0,1,0];
21 % Matriz del endomorfismo cambiada de base
22 C^-1*A*C

```

Lo único reseñable es el bucle para extraer las coordenadas como en (2.5). Estas se almacenan en el vector `coord` y se asignan a la columna j con `A(:,j) = coord`.

Aunque no tenga que ver mucho con el álgebra lineal, para el que tenga curiosidad aquí está el código en `sagemath` que utilicé para hacer la figura de esta sección:

```

1 # Punto inicial fantasma (invisible, hace espacio)
2 P = point([(0,-0.1)], size=0)
3
4 # Bases
5 B = [(1,0), (0,1)]
6 Bp = [(1/sqrt(2),1/sqrt(2)), (-1/sqrt(2),1/sqrt(2))]
7 for k in range(2):
8     P += arrow( (0,0), B[k], arrowsize=8, width=5)
9     P += text('␣$\\vec{e}_'+str(k+1)+'$', B[k], fontsize=60,
10             ↪ horizontal_alignment='left', color='black')
11     P += arrow( (0,0), Bp[k], arrowsize=8, width=5)
12     P += text('␣$\\vec{v}_'+str(k+1)+'$', Bp[k],
13             ↪ vertical_alignment='bottom', fontsize=60,
14             ↪ color='black')
15
16 # Punto

```

```

14 P += point([(1/2,1/2)], size=300, color='black', zorder=100)
15 P += line([(1/2,0),(1/2,1/2),(0,1/2)], thickness=3, linestyle='--')
16 P += text('$P$',(0.6,0.48), fontsize=40)
17
18 P.set_aspect_ratio(1)
19 P.axes(False)
20 show(P)

```

Cualquiera puede obtener algo similar en un instante con *software* de diseño gráfico. La única ventaja de tenerlo en un programa es la posibilidad de hacer mil y una variaciones ajustando números.

2.5. Suma e intersección de subespacios

Por alguna razón, el tema de esta sección no aparece en la guía docente de la asignatura aunque es común en los cursos de álgebra lineal. Incluyo aquí una discusión con ejemplos en K^n . En relación con esto, supondremos en esta sección sin decirlo cada vez que V y W son subespacios vectoriales de un cierto K^n . Como siempre, en este curso $K = \mathbb{R}, \mathbb{C}$ aunque todo funcionaría igual en otros cuerpos.

Las definiciones de la *suma* $V + W$ y la *intersección* $V \cap W$ de V y W son las obvias. Concretamente, $V + W$ está compuesto por los vectores $\vec{v} + \vec{w}$ con $\vec{v} \in V$ y $\vec{w} \in W$ mientras que $V \cap W$ lo está por los vectores que pertenecen simultáneamente a V y W . Es evidente que $V \cap W$ es un subespacio apelando a la Proposición 2.1.1 y algo menos evidente pero similar que $V + W$ también lo es, todo lo que hay que pensar es que $V + W$ coincide con el conjunto de todas las combinaciones lineales posibles de vectores de V y W . Esta última frase nos da un procedimiento para “calcular” $V + W$: si obtenemos bases \mathcal{B}_V y \mathcal{B}_W de V y W entonces $V + W = \mathcal{L}(\mathcal{B}_V \cup \mathcal{B}_W)$.

Hay dos maneras principales de presentar un subespacio de K^n , que más o menos se corresponden con los conceptos de imagen y núcleo, una es la que acabamos de usar de dar un conjunto que lo genera, preferiblemente una base, la otra es dar unas ecuaciones $A\vec{x} = \vec{0}$ con $A \in \mathcal{M}_{m \times n}(K)$ que determinan los vectores del subespacio. Si tenemos la primera presentación, para obtener $V + W$ solo necesitamos unir los conjuntos que generan V y W , mientras que hallar $V \cap W$ no es tan fácil. Por otro lado, dadas ecuaciones lineales que determinan V y W basta imponerlas simultáneamente para obtener las que determinan $V \cap W$, mientras que no está claro cómo hallar $V + W$.

Después de que leas el número suficiente de veces el párrafo anterior deberías sacar la conclusión de que, con lo que sabemos hasta ahora, el único cabo suelto es hallar un procedimiento para obtener las ecuaciones de un subespacio a partir de una base o de un conjunto finito de vectores que lo genera. Este procedimiento consiste en aplicar eliminación de Gauss a las combinaciones lineales igualadas a un vector genérico y estudiar qué condiciones surgen de que haya solución para los coeficientes. En vez de intentar desentrañar esta frase, es mejor ir directamente a un ejemplo.

Consideramos el subespacio $V \subset \mathbb{R}^4$ con base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (-1, 2, 2, -3)^t$ y $\vec{v}_2 = (5, -8, -6, 23)^t$ y queremos escribirlo como $V = \{\vec{x} \in \mathbb{R}^4 : A\vec{x} = \vec{0}\}$. Para ello aplicamos eliminación de Gauss a $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \vec{x}$ entendiéndolo como un sistema en

λ_1 y λ_2 .

$$\left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 2 & -8 & x_2 \\ 2 & -6 & x_3 \\ -3 & 23 & x_4 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 0 & 2 & 2x_1 + x_2 \\ 0 & 4 & 2x_1 + x_3 \\ 0 & 8 & -3x_1 + x_4 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 0 & 2 & 2x_1 + x_2 \\ 0 & 0 & -2x_1 - 2x_2 + x_3 \\ 0 & 0 & -11x_1 - 4x_2 + x_4 \end{array} \right).$$

Se sigue que la condición necesaria y suficiente para que existan λ_1 y λ_2 con $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \vec{x}$, o equivalentemente $\vec{x} \in V$, es que

$$\begin{cases} -2x_1 - 2x_2 + x_3 = 0, \\ -11x_1 - 4x_2 + x_4 = 0, \end{cases} \quad \text{esto es, } A = \begin{pmatrix} -2 & -2 & 1 & 0 \\ -11 & -4 & 0 & 1 \end{pmatrix}.$$

Para comprobar, es conveniente verificar que \vec{v}_1 y \vec{v}_2 cumplen las ecuaciones.

Consideremos ahora $W = \{\vec{x} \in \mathbb{R}^4 : x_1 + x_2 + x_3 + x_4 = 0\}$. Entonces

$$V \cap W = \{\vec{x} \in \mathbb{R}^4 : B\vec{x} = \vec{0}\} \quad \text{con } B = \begin{pmatrix} -2 & -2 & 1 & 0 \\ -11 & -4 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Por otro lado, al resolver la ecuación de W tomando $x_2 = \lambda$, $x_3 = \mu$, $x_4 = \nu$ parámetros arbitrarios, los vectores de W son de la forma $(-\lambda - \mu - \nu, \lambda, \mu, \nu)$, es decir, las combinaciones lineales de $\vec{v}_3 = (-1, 1, 0, 0)^t$, $\vec{v}_4 = (-1, 0, 1, 0)^t$ y $\vec{v}_5 = (-1, 0, 0, 1)^t$. De modo que se tiene $V + W = \mathcal{L}(\{\vec{v}_1, \dots, \vec{v}_5\})$.

Está claro que $V \cap W$ será mayor cuanto más tengan en común V y W y suena sensato que $V + W$ será menor porque habrá pocas sumas que no estén ya en V o en W . El resultado que cuantifica esta situación lleva el nombre del padre más reconocido del álgebra lineal.

Proposición 2.5.1 (fórmula de Grassman). *Si V y W son subespacios de un espacio vectorial de dimensión finita entonces⁹*

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W).$$

Para terminar, veamos un ejercicio de un examen pasado. Allí se definían los subespacios de \mathbb{R}^3

$$\begin{cases} V = \mathcal{L}(\{\vec{v}_1, \vec{v}_2\}), \\ W = \mathcal{L}(\{\vec{w}_1, \vec{w}_2\}) \end{cases} \quad \text{con } \vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}, \vec{w}_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \vec{w}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

⁹Esto tiene la misma pinta que la relación $\#(A \cup B) = \#A + \#B - \#(A \cap B)$ para conjuntos finitos donde $\#$ indica el número de elementos (no es un *hashtag*), la cual se prueba partiendo $A \cup B$ en $A \cap B$, en los que están en A pero no en B y en los que están en B pero no en A . En la demostración de la fórmula de Grassman, que no veremos aquí, se utiliza un argumento similar [8] pero con bases de los subespacios.

y se pedía hallar una base de $V \cap W$ y decidir si $V + W = \mathbb{R}^3$. Para lo primero buscamos las ecuaciones de V mediante eliminación de Gauss

$$\left(\begin{array}{cc|c} 1 & 3 & x_1 \\ 2 & 3 & x_2 \\ 1 & 1 & x_3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 3 & x_1 \\ 0 & -3 & -2x_1 + x_2 \\ 0 & -2 & -x_1 + x_3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 3 & x_1 \\ 0 & -3 & -2x_1 + x_2 \\ 0 & 0 & (x_1 - 2x_2 + 3x_3)/3 \end{array} \right).$$

Así $V = \{\vec{x} \in \mathbb{R}^3 : x_1 - 2x_2 + 3x_3 = 0\}$. De la misma forma, para W

$$\left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 3 & 1 & x_2 \\ 1 & 0 & x_3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & -2 & -3x_1 + x_2 \\ 0 & -1 & -x_1 + x_3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & -2 & -3x_1 + x_2 \\ 0 & 0 & (x_1 - x_2 + 2x_3)/2 \end{array} \right)$$

y se obtiene $V = \{\vec{x} \in \mathbb{R}^3 : x_1 - x_2 + 2x_3 = 0\}$. Resolviendo simultáneamente las ecuaciones de V y W se deduce que las soluciones son proporcionales a $(-1, 1, 1)^t$ por tanto este vector constituye una base. La manera larga de mostrar $V + W = \mathbb{R}^3$ es comprobar que la matriz formada por $\vec{v}_1, \vec{v}_2, \vec{w}_1, \vec{w}_2$ tiene rango 3 (recuérdese la Proposición 2.2.5) y la manera corta es usar la fórmula de Grassman para concluir $\dim(V + W) = 2 + 2 - 1 = 3$.

Expririendo el silicio [opcional]. En aras de la brevedad solo veremos un programa para `sagemath` que resuelve el último ejemplo dando bases de $V \cap W$ y de $V + W$. La única instrucción nueva está en la línea 12 y se explica por sí sola.

```

1 # Vectores
2 v1 = vector([1,2,1])
3 v2 = vector([3,3,1])
4 w1 = vector([1,3,1])
5 w2 = vector([1,1,0])
6 # Se utiliza Q^3 en lugar de R^3
7 # para evitar cálculos aproximados
8 Q3 = VectorSpace(QQ,3)
9 V = Q3.span([v1, v2])
10 W = Q3.span([w1, w2])
11 # Construye la intersección y la suma
12 interseccion = V.intersection(W)
13 suma = Q3.span([v1, v2, w1, w2])
14 # Bases
15 print(interseccion.basis())
16 print(suma.basis())
17 # Comprueba que la suma es todo el espacio
18 print(suma == Q3)

```

Como se indica en los comentarios, se utiliza \mathbb{Q}^3 en vez de \mathbb{R}^3 para forzar el cálculo simbólico.