Notas de Álgebra Superior I

Gerardo Miguel Tecpa Galván

A todos mis alumnos, los que fueron, los que son y los que serán,

por darle un sentido real y profundo a mi vida

...y por soportar todas mis locuras en el salón de clase.

Índice general

Capítulo 1

Conjuntos

1.1. Agrupaciones y conjuntos

El primer objeto matemático que analizaremos es el de conjunto. Actualmente no existe rama matemática que no use en algún grado los conjuntos y, de hecho, la historia detrás del uso de los conjuntos como base de la matemática moderna es muy profunda e interesante, pero no será explorada por ahora en este trabajo. La importancia de los conjuntos como herramienta radica en varias cosas, entre ellas destacan su versatilidad de trabajo, su relación con el lenguaje lógico y su capacidad natural de agrupar objetos.

La idea básica de los conjuntos es cómo agrupar de manera abstracta los objetos, es decir, siempre que se quiera trabajar con agrupaciones de objetos, tendremos que utilizar al menos un conjunto. Aunque no existe una definición *universalmente aceptada* de conjunto, nosotros usaremos una provisional que nos ayudará a trabajar este tipo de estructuras matemáticas.

Un **conjunto** es una agrupación de objetos vista como un todo en la que, dado un objeto cualquiera, siempre es verdad que el objeto pertenece a la agrupación o no pertenece a la agrupación.

Primero revisemos qué quiere decir de manera intuitiva la definición anterior. Un conjunto A, en primer lugar, es una agrupación de objetos. Esos objetos generalmente son llamados **elementos** y pueden ser cualquier tipo objeto que desees: personas, números, ecuaciones, otros conjuntos, etc. Además, tendremos la capacidad de considerar a la agrupación como un solo objeto, sin necesidad de utilizar los elementos que la conforman, por lo que los conjuntos también son considerados como objetos. Y tercero, el conjunto A determina una clasificación de todos los objetos existentes. Se pueden clasificar solamente en dos tipos: los que *pertenecen al conjunto* y los que *no pertenecen al conjunto*.

Esto último es tan importante que lo volveremos a decir: cuando hablemos de un conjunto A y un objeto arbitrario x, siempre tenemos dos posibilidades para x: (i) que x *pertenezca* al conjunto A o (ii) que x *no pertenezca* al conjunto A. Para evitarnos escribir tanto, usaremos simplemente la notación $x \in A$ cuando x

pertenece a A y usaremos $x \notin A$ cuando x no pertenezca a A. Gracias a lo anterior, tendremos nuestra primera tautología:

Nota 1.1.1. *Dado un conjunto A y un objeto arbitrario x, siempre es verdadero $x \in A$ o $x \notin A$.*

Es importante mencionar que, derivado de la nota 1.1.1, ahora tenemos una contradicción lógica. Dado un conjunto A y un objeto cualquiera x, siempre es falso que: $a \in A$ **y** $a \notin A$. La anterior es útil pues algunas de las demostraciones por contradicción se pueden reducir a una contradicción como la anterior.

1.2. Escribiendo y describiendo conjuntos

Si ya tenemos un poco aterrizada la idea de conjunto, ahora vamos a ver cómo utilizarlos. Para ello vamos a introducir su notación estándar. Por lo general los objetos de un conjunto se escriben entre llaves, por ejemplo $\{cama, mesa, pizza\}$. Lo anterior indica que el conjunto mostrado es la agrupación que consta de tres objetos: cama, mesa y pizza.

Cuando podemos dar un listado explícito de los miembros de un conjunto, diremos que el conjunto está **escrito por extensión** o simplemente que es un **conjunto por extensión.** Más ejemplos de ello son los conjuntos $\{1, 3, 4, 5\}$, $\{mato, meto, mito, moto, muto\}$, $\{a, b, c, d, e, f, g, h\}$.

La principal ventaja de los conjuntos por extensión es que enlistar todos sus elementos puede facilitar mucho el determinar si un objeto cualquiera forma parte o no del conjunto, por ejemplo, el conjunto $\{1\}$ es un conjunto que sólo tiene un elemento (a los conjuntos que sólo constan de un elemento les llamamos **conjuntos unitarios)**. Sin embargo, enlistar todos los elementos de un conjuntos tiene un problema, por ejemplo, ¿qué pasaría si quisiera enlistar todos los números naturales? Eso sería bastante problemático, pues yo podría empezar a enlistarlos, es decir, podría empezar escribiendo

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9...$$

El problema es que hay *demasiados* números naturales: podría empezar a enlistarlos yo, morirme y no acabar, podría dejar que mi hijo siguiera con la lista donde yo la dejé, pero seguramente moriría antes de terminarla, es más, seguramente primero se termina la vida en la Tierra antes siquiera de ir a la *mitad* de la lista. Esto sí es un problema serio ¿cómo es posible que no se pueda enlistar todos los números naturales en un conjunto por extensión? pero no te preocupes, que hay una segunda variante de los conjuntos por extensión. En algunas ocasiones, basta con escribir algunos elementos del conjunto para que se intuya la totalidad de estos. Por ejemplo, el conjunto de los número naturales lo podemos escribir como

$$\{0, 1, 2, 3, 4, \dots\}$$

Donde los tres puntos suspensivos indican que los otros elementos están determinados por la secuencia inicial. Otro ejemplo:

$$\{0, 2, 4, 6, 8, 10, \dots\}$$

El conjunto anterior indica que hay una cantidad infinita de objetos, pero con los cinco miembros exhibidos es posible determinar los otros. De hecho este último conjunto indica el conjunto de los número pares. Un último ejemplo:

$$\{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4 \dots\}$$

Este conjunto son todos aquellos números enteros, es decir, los números naturales junto con los negativos, como se suele decir en la bachillerato.

Aunque la representación anterior nos ayuda a escribir más conjuntos por extensión, no nos bastaría. Por ejemplo, ¿qué pasaría si te diera el siguiente conjunto?:

$$\{1, \phi, 4, 1/2, e, 27, -5, \dots\}$$

Éste ya no es tan intuitivo determinar de qué conjunto se trata, así que nuevamente los conjuntos por extensión presentan problemas. Pero ese no es el mayor de todos: ¿qué pasaría si hay agrupaciones tan, pero tan grandes, que no es posible enlistarlos? Ese sí que es un problema porque entonces los conjuntos por extensión ya no podrían ser utilizables en todos los casos. Y si consideras que no existen agrupaciones tan grandes que es imposible enlistarlas, entonces tu profesor de Cálculo I te dará un par de sorpresas en unos cuantos días.

Para evitar el problema antes dicho, usaremos la segunda forma de escribir un conjunto y, de hecho, es la que usaremos casi todo el tiempo en la Facultad: un conjunto está **escrito por comprensión** o simplemente decimos que es un **conjunto por comprensión**, cuando se describe con base en una o varias propiedades que solamente los elementos del conjunto satisfacen.

Vamos más despacio con este último término porque es muy importante que se comprenda. Como ya habíamos dicho antes, la propiedad importante de los conjuntos es poder determinar si un objeto pertenece o no al conjunto. Derivado de ello, cualquier conjunto tiene al menos una propiedad que solamente los objetos del conjunto satisfacen y cualquier objeto que no esté en el conjunto no lo satisface. Dicha propiedad es la que *determina* qué objetos sí pertenecen al conjunto. Los conjuntos por comprensión se escriben generalmente de la siguiente manera:

$$\{x \in X : P(x)\}$$

Notemos que nuevamente los objetos se escriben entre llaves, pero dentro de las llaves hay un enunciado lógico, en lugar de objetos, como con los conjuntos por extensión. Ahora vamos a analizar y desmenuzar la notación anterior. La estructura del conjunto va a estar dividida por dos puntos que se leen *tal que* o *tales que* [1]. Antes de

[1] En algunas ocasiones, en lugar de dos puntos se usa el símbolo | que también se lee *tal que*

dichos puntos se pone una variable, en este caso particular es x, que indica principalmente qué clase de objetos pertenecen al conjunto, por ejemplo, números, ecuaciones, personas, etc. Este primer enunciado usualmente, pero no exclusivamente, es de la forma $x \in X$, donde X es llamado **conjunto de referencia** y sirve para delimitar cuáles objetos entran dentro del conjunto que estamos definiendo. La parte derecha del conjunto por comprensión es la o las propiedades que satisfacen los objetos del conjunto y debe estar escrito en términos de la variable usada previamente. Vamos a un ejemplo:

$$\{x \in \mathbb{N} : x \text{ es par }\}$$

Antes de los dos puntos notemos que está el enunciado $x \in \mathbb{N}$, esto indica que los objetos del conjunto son números naturales, pero no *todos* los números naturales. Después de los dos puntos está la propiedad que deben cumplir todos los objetos del conjunto y solamente ellos: *son pares*. Esta propiedad *debe* estar en términos de la variable que usamos antes de los dos puntos. Así, el conjunto anterior es *el conjunto de los números naturales tales que son pares.*

Nota 1.2.1. *La variable puede tener varios nombres y, siempre que la propiedad esté en términos del nombre de la variable, el conjunto no cambia. Por ejemplo, los conjuntos* $\{x \in \mathbb{N} : x \text{ es par }\}$, $\{n \in \mathbb{N} : n \text{ es par }\}$, $\{\alpha \in \mathbb{N} : \alpha \text{ es par }\}$ *describen al mismo conjunto: el que contiene los naturales pares.*

Igualmente importante es tener en consideración la siguiente nota:

Nota 1.2.2. *Al escribir un conjunto por comprensión, la propiedad que sigue a los dos puntos debe estar en término de la variable previa a los dos puntos. Es decir, escrituras como* $\{x \in \mathbb{Z} : \alpha \text{ es impar}\}$ *no tienen sentido.*

Antes de continuar quiero hacer un paréntesis en la importancia de los conjuntos de referencia. Si no colocamos un conjunto de referencia cuando escribimos un conjunto por comprensión es probable que surjan ambigüedades al momento de interpretar quién es el conjunto que estamos describiendo. Tomemos el siguiente ejemplo:

$$A = \{x : x \text{ es múltiplo de } 2\}$$

Este conjunto, de hecho, es posible que se pueda escribir como un conjunto por extensión:

$$A = \{0, 2, 4, 6, 8, \dots\}$$

Pero algunos, muy listos, también podrían presentar el siguiente conjunto:

$$A = \{\dots, -4, -2, 0, 2, 4, \dots\}$$

Entonces estamos en un problema realmente serio, pues nuestro lenguaje matemático no debería prestarse a ambigüedades y aquí ya tenemos dos posibles interpretaciones de un mismo enunciado. Tenemos que salvar esto de alguna forma y para ello primero necesitamos saber dónde está el problema....y el problema es, justamente, la falta de un conjunto de referencia en nuestro conjunto por comprensión. Sin embargo, notemos que si especificamos quién es el conjunto de referencia, entonces podemos tener una mejor valoración de nuestro conjunto. Por ejemplo, si consideramos que $\mathbb{N}$ es el conjunto de referencia, entonces es preferible escribir el conjunto A como

$$A = \{x \in \mathbb{N} : x \text{ es múltiplo de } 2\}.$$

Y esto obliga a que A sea el conjunto:

$$A = \{0, 2, 4, 6, 8, \dots\}$$

Sin embargo, ahora consideremos que nuestro conjunto de referencia es el conjunto $\mathbb{Z}$, en tal caso, A podría quedarse definido por:

$$A = \{x \in \mathbb{Z} : x \text{ es múltiplo de } 2\}.$$

Y con esta especificación resulta que A es el conjunto:

$$A = \{\dots, -4, -2, 0, 2, 4, \dots\}$$

Lo importante de lo anterior es sencillo:

Nota 1.2.3. *Cada que sea posible, cuando escribamos un conjunto por comprensión, antes de los dos puntos es recomendable que se indique un conjunto de referencia.*

Ahora que conocemos la estructura básica de un conjunto por comprensión, es muy importante saber cuándo un objeto pertenece o no a dicho conjunto. Si $B = \{a \in A : P(a)\}$ es un conjunto y x es un objeto arbitrario, el objeto x debe cumplir dos cosas para pertenecer a B: (i) el objeto x debe estar en el conjunto de referencia, es decir, $x \in A$ y (ii) el objeto x debe satisfacer la propiedad del conjunto B, es decir, $P(x)$ es verdadera. Si x cumple las dos propiedades, entonces x debe estar en B. Y de manera análoga, si un objeto z está en B, entonces dicho objeto debe cumplir simultáneamente que (i) está en el conjunto de referencia, es decir, $z \in A$ y (ii) z debe satisfacer la propiedad que determina a B, es decir, $P(z)$ debe ser verdadera.

Lo anterior es tan, pero tan importante que lo resaltaremos en una nota:

Nota 1.2.4. *Dado un conjunto por comprensión, digamos $\{x \in A : P(x)\}$, y un objeto arbitrario, digamos z, los siguientes enunciados son equivalentes.*

a) $z \in \{x \in A : P(x)\}$.

b) $z \in A$ *y* $P(z)$ *es verdad.*

Debido a que la observación previa será usada con frecuencia en distintos modos, también queremos recalcar la importancia de su negación.

Nota 1.2.5. *Dado un conjunto por comprensión, digamos $\{x \in A : P(x)\}$, y un objeto arbitrario, digamos z, los siguientes enunciados son equivalentes.*

a) $z \notin \{x \in A : P(x)\}$.

b) $z \notin A$ *o* $P(z)$ *es falso.*

Las notas anteriores son realmente importantes, al punto de que muchas demostraciones se reducen a ellas dos, por lo que es conveniente tenerlas siempre presentes. Además de ello, es importante señalar que la variable no necesariamente es un elemento del conjunto escrito por comprensión. Por lo que cosas como $x \in \{x \in A : P(x)\}$ podrían no ser convenientes de escribir, por lo que es preferible que el nombre de la variable no coincida con el nombre de los objetos en el conjunto.

Nota 1.2.6. *Dado un conjunto por comprensión, digamos $\{x \in A : P(x)\}$, en medida de lo posible el nombre de la variable no debe coincidir con el nombre de ninguno de los objetos en el conjunto.*

La nota anterior está enfocada no sólo a un buen uso de la redacción en las demostraciones, sino que algunos de los primeros errores interpretativos de las demostraciones surgen a partir de la mala escritura antes mencionada.

Para terminar esta sección, introduciremos dos Axiomas de la teoría de conjuntos. Son enunciados lógicos que tomaremos por *verdaderos*. El primero de ellos viene a sustentar lo que ya hemos mencionado sobre la escritura por comprensión en conjuntos.

Axioma de Especificación

si A es un conjunto y P es un predicado, entonces la familia $\{x \in A : P(x)\}$ es un conjunto

Una consecuencia inmediata del Axioma de especificación, entre otras cosas importantes, es que todos los conjuntos puedan ser escritos por comprensión, pues si A es un conjunto, éste puede ser descrito trivialmente por $\{x \in A : x \in A\}$. Otra consecuencia de este axioma nos permitirá mostrar que existen agrupaciones tan grandes que no es posible que sean conjuntos.

Teorema 1.2.7. *No existe un conjunto que tenga como elementos a todos los conjuntos.*

Demostración. La demostración es un enunciado simple y la realizaremos por contradicción. Supondremos que $\mathscr{C}$ es una familia de objetos que tiene como elementos a todos los conjuntos. Es decir, si C es un conjunto, entonces $C \in \mathscr{C}$. De acuerdo al Axioma de especificación, se tiene que la familia:

$$B = \{a \in \mathscr{C} : a \notin a\}$$

es un conjunto. Así, sabemos por la nota 1.1.1 que $B \in B$ o $B \notin B$. Si $B \in B$, entonces B debe satisfacer que $B \in \mathscr{C}$ y que $B \notin B$, lo cual concluiría que $B \in B$ y $B \notin B$, lo cual no es posible. Por otro lado, si $B \notin B$, entonces $B \notin \mathscr{C}$ o $B \in B$. Pero como B es un conjunto, por hipótesis $B \in \mathscr{C}$, lo que implica que $B \in B$, es decir, concluimos que $B \notin B$ y $B \in B$.

En ambos casos, tenemos una contradicción, lo cual indica que no existe una familia que contenga a todos los conjuntos. $\square$

Otro conjunto que obtendremos a partir del axioma de especificación es el conjunto vacío. Pero para ello necesitamos el siguiente axioma:

Axioma de Existencia

Existe un conjunto.

Notemos que hasta este punto nada nos asegura que los conjuntos realmente existieran, de hecho es tan problemático mostrar que los conjuntos existen que tenemos que darle la categoría de axioma: no podemos deducir que existan, tenemos que suponerlo. Y aunque el axioma de existencia nos asegura la existencia de sólo un conjunto, podemos usarlo para obtener otros conjuntos. Por ejemplo, supongamos que A es el conjunto que existe gracias al axioma de existencia, entonces, aplicando el Axioma de especificación tenemos que:

$$\{x : x \neq x\}$$

es un conjunto. Y no sólo un conjunto, es un conjunto que se usa muy frecuentemente y es llamado el **conjunto vacío**. Más aun, tiene su propia notación y se denota universalmente por $\emptyset$. Este conjunto es el conjunto *al que no le pertenece nadie*.

Nota 1.2.8. *Dado cualquier objeto x, siempre es verdadero que $x \notin \emptyset$ y siempre es falso que $x \in \emptyset$*

1.3. Subconjuntos e igualdad de conjuntos

Pues bien, si ya tenemos claro la escritura de conjuntos por comprensión, pasemos a algo más serio y formal. Existen dos cualidades fundamentales de los conjuntos y éstas son la contención de conjuntos y la igualdad de conjuntos.

Vamos primero con la contención. Dados dos conjuntos A y B, decimos que **A está contenido en B** o que **A es subconjunto de B**, si todos los elementos de A son elementos de B. Lo anterior se simboliza por $A \subseteq B$.

Lo primero que hay que destacar es que una frase de la forma $A \subseteq B$ es un enunciado lógico, es decir, se le puede atribuir un valor de verdad. Mucho de lo que veremos en este capítulo es verificar la veracidad de este tipo de enunciados. Como recordatorio, pareciera más sencillo el poder determinar el valor de verdad del enunciado $A \subseteq B$ si encontramos una condicional adecuada a dicho enunciado. Notemos que el enunciado *todo elemento de A es elemento de B* es equivalente a la condicional *Si $x \in A$, entonces $x \in B$*. Lo anterior es tan importante que lo volveremos a escribir:

Nota 1.3.1. *Dados dos conjuntos, si queremos demostrar que $A \subseteq B$, basta verificar que la condicional Si $x \in A$, entonces $x \in B$ es verdadera.*

Lo siguiente que hay que destacar es su negación, es decir, decimos que **A no es subconjunto de B** o que **A no está contenido en B** si existe un elemento de A que no está en B. A lo anterior se le denotará simplemente por $A \nsubseteq B$. Al igual que con lo anterior, esto lo pondremos en una nota:

Nota 1.3.2. *Dados dos conjuntos, si queremos demostrar que $A \nsubseteq B$, basta verificar que el enunciado existe $a \in A$ tal que $a \notin B$ es verdadero.*

Para empezar a agarrar calor en esto de las demostraciones, probaremos los siguientes resultados:

Lema 1.3.3. *Si A es un conjunto, entonces $A \subseteq A$.*

Demostración. Ahora sí, vamos a iniciar con lo bueno. Primero que nada vamos a identificar el tipo de enunciado lógico que hay que demostrar. Es claro que es una condicional, tiene una sola hipótesis: A *es un conjunto* y un solo consecuente: $A \subseteq A$.

Vamos a demostrar esta condicional de forma directa. Es decir, tenemos que demostrar que $A \subseteq A$....y ¿cómo hacemos eso? Pues si nos fijamos en la nota 1.3.1 ésta dice que para demostrar que $A \subseteq A$, basta ver que la condicional: *Si* $x \in A$, *entonces* $x \in A$ es verdadera....y pff....otra condicional más, pero esta está de regalo, pues es una tautología, por lo que siempre es verdadera y con eso acabamos nuestra demostración. □

Lema 1.3.4. *Si A es un conjunto, entonces $\emptyset \subseteq A$*

Demostración. Primero identificamos el enunciado a demostrar. Como es una condicional, tenemos varias opciones para demostrarla. Esta en particular la demostraremos por contradicción. es decir, supondremos que es falso que $\emptyset \subseteq A$. En tal caso, es verdadero que $\emptyset \nsubseteq A$ y esto significaría que es verdadero que existe $x \in \emptyset$ tal que $x \notin A$. Y ya encontramos una contradicción, pues sabemos que siempre es falso que $x \in \emptyset$. Con lo anterior, es verdadero que $\emptyset \subseteq A$. □

Lema 1.3.5. *Sean A, B y C conjuntos. Si $A \subseteq B$ y $B \subseteq C$, entonces $A \subseteq C$.*

Demostración. Este lema tiene un poco de más jugo que los anteriores dos. Nuevamente estamos frente a una condicional, pero ahora tenemos dos hipótesis: $A \subseteq B$ y $B \subseteq C$. Lo que debemos demostrar es que $A \subseteq C$.

Como queremos demostrar la condicional $A \subseteq C$, iniciamos tomando un elemento de A y la demostración concluirá cuando probemos que dicho objeto pertenezca a C.

Sea $x \in A$. Por hipótesis, $A \subseteq B$, por lo que $x \in B$. Sin embargo, por hipótesis tenemos que $B \subseteq C$, por lo que $x \in C$, concluyendo nuestra demostración. □

La segunda cosa importante que tendremos es la igualdad entre conjuntos, ésta está dada por el siguiente axioma

Axioma de Extensión.

Dos conjuntos A y B son iguales si todos los elementos de A son elementos de B y todos los elementos de B son elementos de A.

Notemos que el concepto de igualdad entre conjuntos es, en realidad, dos contenciones: $A \subseteq B$ y $B \subseteq A$. Esta propiedad es muy importante, tanto que hasta tiene nombre:

Nota 1.3.6 (Doble Contención)**.** *Dados dos conjuntos A y B, se tiene que $A = B$ si y sólo si $A \subseteq B$ y $B \subseteq A$.*

La anterior es una propiedad fundamental de las Matemáticas e invitamos a cualquier lector que no desprestigie la nota anterior. Más aun, es posible que se percaten en semestres más avanzados que muchas demostraciones son, en el fondo, una doble contención.

Además, es importante notar que un conjunto siempre es subconjunto de él mismo, como lo dice el Lema 1.3.3. En algunos casos es frecuente que, dado un conjunto A, consideremos subconjuntos de A que son distintos del mismo conjunto A. A este tipo de subconjuntos les llamamos **subconjuntos propios**. En tal caso, si queremos verificar que B es un subconjunto propio de A, denotado por $A \subset B$, entonces debemos verificar dos cosas: B es subconjunto de A y B es distinto de A. Si queremos, podemos reescribir usando la propiedad de doble contención como sigue:

Nota 1.3.7. *Dados dos conjuntos A y B, diremos que B es subconjunto propio de A si se cumple que $B \subseteq A$ y existe $a \in A$ tal que $a \notin B$.*

Para terminar esta sección introduciremos un conjunto muy similar a los conjuntos de referencia. El **conjunto universal** es un conjunto que limita todo el marco de trabajo, todos los objetos con los que se trabajen deben ser elementos del conjunto universal. y todos los conjuntos que trabajemos están relacionados de alguna forma con el conjunto universal, principalmente, todos los conjuntos que trabajemos son subconjunto del conjunto universal.

La condición importante del conjunto universal es que, cuando lo establecemos, cualquier objeto que consideremos debe pertenecer a ÉL. Por ejemplo, si establecemos que nuestro conjunto universal es el conjunto $\mathbb{N}$ entonces automáticamente establecemos que nosotros no vamos a trabajar con objetos de la forma π, $1/2$, -5, etc. Pues aunque sí son números, ninguno de ellos es un número natural. De esto, lo que me interesa que sepan es que:

Nota 1.3.8. *Si X es nuestro conjunto universal y a es un objeto cualquiera de nuestro marco de trabajo, entonces $a \in X$.*

1.4. Operaciones de conjuntos

Esta sección será posiblemente una de las más amplias del capítulo y en la que recaerán muchas demostraciones. Es importante destacar al alumno que por el momento no es necesario que conozca de memoria todas las propiedades aquí descritas, lo realmente importante es que conozca los métodos de demostraciones y que pueda llevarlos exitosamente a cabo. Es preferible aprender una habilidad *comprendiendo* cómo se hace, que simplemente hacer por hacerla. Tenemos varias operaciones con conjuntos, así que iniciaremos con las más simples.

1.4.1. Intersección de conjuntos

La siguiente operación que analizaremos es la intersección de conjuntos. Dados dos conjuntos A y B, la **intersección de $\boldsymbol{A}$ y $\boldsymbol{B}$**, denotada por $A \cap B$, es el conjunto

$$A \cap B = \{x \in A : x \in B\}.$$

De manera similar a la unión, a los conjuntos A y B se les suele llamar **intersecandos**. Notemos que, derivada directamente de la definición de intersección, tenemos lo siguiente:

Nota 1.4.1. *Dados dos conjuntos A y B y un objeto x, se tiene que $x \in A \cap B$ si y sólo si $x \in A$ y $x \in B$.*

De igual forma, es necesario saber qué significa la negación de que un objeto pertenezca a una intersección.

Nota 1.4.2. *Dados dos conjuntos A y B y un objeto x se tiene que $x \notin A \cap B$ si y sólo si $x \notin A$ o $x \notin B$.*

La intersección de dos conjuntos tiene dos posibles interpretaciones en nuestro lenguaje. Un objeto pertenece a una intersección si está en *todos* los intersecandos. O visto de otra manera, la intersección de conjuntos son los objetos que todos los intersecandos tienen en común. De esto último es importante una definición nueva. Dados dos conjuntos A y B, decimos que **$\boldsymbol{A}$ y $\boldsymbol{B}$ son ajenos** si $A \cap B = \emptyset$, es decir, dos conjuntos son ajenos si no tienen elementos en común.

Bastará el uso de las definiciones anteriores para obtener los siguientes resultados.

Lema 1.4.3. *Si A, B y C son conjuntos, entonces los siguientes incisos se satisfacen:*

a) $A \cap B \subseteq A$ *y* $A \cap B \subseteq B$. *Es decir, la intersección siempre es subconjunto de los intersecandos.*

b) $A \cap B = B \cap A$

c) $A \cap \emptyset = \emptyset$.

d) $A \cap A = A$.

e) $A \cap (B \cap C) = (A \cap B) \cap C$.

f) $A \cap B = A$ *si y sólo si* $A \subseteq B$.

g) *Si* $A \subseteq B$, *entonces* $A \cap C \subseteq B \cap C$.

h) *Si* $C \subseteq A$ *y* $C \subseteq B$, *entonces* $C \subseteq A \cap B$.

Demostración. (a) Mostraremos que $A \cap B \subseteq A$ y una prueba análoga muestra que $A \cap B \subseteq B$.

Sea $x \in A \cap B$. Por definición de intersección, $x \in A$ y $x \in B$, en particular, $x \in A$, concluyendo que $A \cap B \subseteq A$.

(g) La demostración de este enunciado pueden verla en https://youtu.be/URiIgrGbE6c

(h) Si $C \subseteq A$ y $C \subseteq B$, entonces $C \subseteq A \cap B$.

Sea $x \in C$. Por hipótesis, $C \subseteq A$, y como $x \in C$, entonces $x \in A$. Nuevamente, por hipótesis $C \subseteq B$ y como $x \in C$, entonces $x \in B$. De lo anterior, como $x \in A$ y $x \in B$, entonces por definición de intersección, $x \in A \cap B$.

□

1.4.2. Unión de conjuntos

Dados dos conjuntos A y B, la **unión de A con B**, denotado por $A \cup B$, es el conjunto:

$$A \cup B = \{x : x \in A \text{ o } x \in B\}.$$

Dada la unión $A \cup B$, a los conjuntos A y B se les llamará **uniendos**. Antes de continuar es importante mencionar que, debido a la escritura por comprensión de la unión, no es posible determinar hasta ahora si dicha agrupación es un conjunto mediante el Axioma de especificación. Tal es así que tenemos el siguiente axioma.

Axioma de la unión

si A y B son conjuntos, entonces $A \cup B$ es un conjunto

Por otro lado, derivada de la definición de unión, podemos enunciar la siguiente nota:

Nota 1.4.4. *Dados dos conjuntos A y B y un objeto x, se tiene que $x \in A \cup B$ si y sólo si $x \in A$ o $x \in B$.*

Igualmente importante que la definición, está la negación de ésta, es decir qué significa que un objeto no sea elemento de una unión.

Nota 1.4.5. *Dados dos conjuntos A y B y un objeto x, se tiene que $x \notin A \cup B$ si y sólo si $x \notin A$ y $x \notin B$.*

Lo anterior será suficiente para poder enunciar nuestro siguiente resultado.

Lema 1.4.6. *Si A, B y C son conjuntos, entonces:*

a) $A \subseteq A \cup B$ *y* $B \subseteq A \cup B$. *Es decir, los uniendos siempre son subconjunto de la unión.*

b) $A \cup B = B \cup A$.

c) $A = A \cup A$.

d) $A = A \cup \emptyset$.

e) $(A \cup B) \cup C = A \cup (B \cup C)$.

f) $A = A \cup B$ *si y sólo si* $B \subseteq A$.

g) *Si* $A \subseteq B$, *entonces* $A \cup C \subseteq B \cup C$.

h) *Si* $A \subseteq C$ *y* $B \subseteq C$, *entonces* $A \cup B \subseteq C$.

Demostración. En estas notas no realizaremos todas las demostraciones de este resultado. Sólo realizaremos algunas de ellas que consideremos representativas.

(a) Mostraremos que $B \subseteq A \cup B$ y una prueba análoga mostrará que $A \subseteq A \cup B$.

Sea $x \in B$, como $x \in B$, entonces es verdad que $x \in A$ o $x \in B$, con lo cual, por definición de unión, $x \in A \cup B$.

(g) Si $A \subseteq B$, entonces $A \cup C \subseteq B \cup C$.

Como demostraremos una contención, iniciaremos tomando un objeto $x \in A \cup C$. Por definición de unión, $x \in A$ o $x \in C$. Ahora procederemos a revisar ambos casos.

Caso 1. $x \in A$.

En este caso, como $A \subseteq B$ y $x \in A$, entonces $x \in B$ y con ello, por definición de unión, $x \in B \cup C$.

Caso 1. $x \in C$.

En este caso, como $x \in C$, por definición de unión, $x \in B \cup C$.

Por los casos anteriores, $x \in B \cup C$ y con ello, $A \cup C \subseteq B \cup C$.

(h) La demostración de este enunciado pueden verlo en https://youtu.be/YO7OnGdV-Aw

$\square$

Estas primeras dos operaciones entre conjuntos tienen diversas propiedades por si solas, pero también es posible considerar propiedades de ambas operaciones simultáneamente:

Teorema 1.4.7. *Si A, B y C son conjuntos, entonces los siguientes incisos se satisfacen:*

a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

1.4.3. Diferencia de conjuntos

La siguiente operación que consideraremos es la diferencia entre conjuntos. Dados dos conjuntos A y B, la **diferencia de A y B**, denotada por $A \setminus B$, es el conjunto:

$$A \setminus B = \{x \in A : x \notin B\}.$$

La diferencia entre conjuntos $A \setminus B$ son los elementos que están en A pero no están en B. En algunas ocasiones, la diferencia de conjuntos también de denota como $A - B$, por lo que es común llamar a la diferencia de A y B como **A menos B**. Notemos que, derivada directamente de la definición de diferencia, tenemos lo siguiente:

Nota 1.4.8. *Dados dos conjuntos A y B y un objeto x, se tiene que $x \in A \setminus B$ si y sólo si $x \in A$ y $x \notin B$.*

De igual forma, es necesario saber qué significa la negación de que un objeto pertenezca a una diferencia.

Nota 1.4.9. *Dados dos conjuntos A y B y un objeto x se tiene que $x \notin A \setminus B$ si y sólo si $x \notin A$ o $x \in B$.*

A diferencia de las intersecciones y las uniones, la diferencia no necesariamente es conmutativa, es decir, $A \setminus B$ no necesariamente es igual que $B \setminus A$. Considérese, por ejemplo, $A = \{1, 2, 3, 4\}$ y $B = \{3, 4\}$. Con los conjuntos anteriores es sencillo verificar que $A \setminus B \neq B \setminus A$.

La diferencia de conjuntos tampoco resulta ser una operación asociativa, es decir, si A, B y C son conjuntos, no necesariamente es cierto que $A \setminus (B \setminus C) = (A \setminus B) \setminus C$. Nuevamente, considerando $A = \{1, 2, 3, 4\}$, $B = \{3, 4\}$ y $C = \{2\}$, es sencillo verificar que $A \setminus (B \setminus C) \neq (A \setminus B) \setminus C$.

Sin embargo, sí tenemos algunas propiedades verdaderas sobre esta operación:

Lema 1.4.10. *Si A, B y C son conjuntos, entonces los siguientes incisos se satisfacen:*

a) $A \setminus B \subseteq A$.

b) $(A \setminus B) \cap (B \setminus A) = \emptyset$.

c) $A \setminus \emptyset = A$ *y* $\emptyset \setminus A = \emptyset$.

d) $A \setminus A = \emptyset$.

e) $(A \setminus B) \cap B = \emptyset$.

f) $A \setminus B = A$ *si y sólo si* $A \cap B = \emptyset$.

g) Si $A \subseteq B$, *entonces* $A \setminus C \subseteq B \setminus C$.

h) Si $A \subseteq B$, *entonces* $C \setminus B \subseteq C \setminus A$.

i) $A \setminus B = A \setminus (A \cap B)$.

j) $A \setminus B = \emptyset$ *si y sólo si* $A \subseteq B$

Demostración. (b) Para mostrar que $(A \setminus B) \cap (B \setminus A) = \emptyset$, procederemos por contradicción, es decir, supondremos que $(A \setminus B) \cap (B \setminus A) \neq \emptyset$. Ya que $(A \setminus B) \cap (B \setminus A)$ es no vacío, podemos considerar $x \in (A \setminus B) \cap (B \setminus A)$.

Por definición de intersección, $x \in A \setminus B$ y $x \in B \setminus A$. Como $x \in A \setminus B$, entonces por definición de diferencia, $x \in A$ y $x \notin B$, en particular, $x \in A$. Por otro lado, como $x \in B \setminus A$, entonces por definición de diferencia, $x \in B$ y $x \notin A$, en particular, $x \notin A$. Así podemos concluir que $x \in A$ y $x \notin A$, lo cual es una contradicción.

Con lo anterior, $(A \setminus B) \cap (B \setminus A) = \emptyset$.

(f) La parte suficiente de este enunciado pueden consultarla en `https://youtu.be/3V4rYajOvRw`. La parte necesaria pueden verla en `https://youtu.be/pP-Jg9pzgzs`.

(h) Sea $x \in C \setminus B$. Por definición de diferencia de conjuntos, $x \in C$ y $x \notin C$. Ya que por hipótesis, $A \subseteq B$, entonces $x \notin A$ y nuevamente, como $x \in C$ y $x \notin A$, por definición de diferencia de conjuntos, $x \in C \setminus A$.

$\square$

Existen otras dos propiedades que relacionan a las tres operaciones que ya hemos visto, dichas propiedades son llamadas las leyes de D'Morgan y se enuncian en el siguiente teorema:

Teorema 1.4.11 (Leyes de D'Morgan). *Si A, B y C son conjuntos, entonces:*

a) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.

b) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

Demostración. (a) Para mostrar que $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ procederemos por doble contención.

Primero mostraremos que $A \setminus (B \cup C) \subseteq (A \setminus B) \cap (A \setminus C)$. Sea $x \in A \setminus (B \cap C)$, por definición de diferencia de conjuntos, se tiene que $x \in A$ y $x \notin B \cup C$. Por la negación de la definición de unión $x \notin B$ y $x \notin C$.

Ya que $x \in A$ y $x \notin B$, entonces por definición de diferencia, $x \in A \setminus B$. Por otro lado, ya que $x \in A$ y $x \notin C$, entonces por definición de diferencia, $x \in A \setminus C$. Como $x \in A \setminus B$ y $x \in A \setminus C$, entonces por definición de intersección, $x \in (A \setminus B) \cap (A \setminus C)$.

Ahora mostraremos que $(A \setminus B) \cap (A \setminus C) \subseteq A \setminus (B \cup C)$. Sea $x \in (A \setminus B) \cap (A \setminus C)$. Por definición de intersección, $x \in A \setminus B$ y $x \in A \setminus C$.

Como $x \in A \setminus B$, entonces por definición de diferencia, $x \in A$ y $x \notin B$. Del mismo modo, como $x \in A \setminus C$, entonces por definición de diferencia, $x \in A$ y $x \notin C$. De manera análoga, como $x \in A \setminus C$, entonces por definición de diferencia de conjuntos $x \in A$ y $x \notin C$. Como $x \notin B$ y $x \notin C$, entonces por la negación de la definición de unión $x \notin B \cup C$. Así, como $x \in A$ y $x \notin B \cup C$, entonces por definición de diferencia de conjuntos $x \in A \setminus (B \cup C)$. $\square$

1.4.4. Complemento de conjuntos

El complemento de un conjunto está definido en términos de la diferencia entre conjuntos. Si A es un conjunto y X es el conjunto universal, el **complemento de A**, denotado por A^c, es el conjunto

$$A^c = \{x \in X : x \notin A\}.$$

Notemos que $A^c = X \setminus A$. Es decir, el complemento de un conjunto son todos aquellos objetos que no están en dicho conjunto. En esta operación siempre es necesario tener claro quién es el conjunto universal, pues con base en este se determina el complemento. Por ejemplo, si consideramos el conjunto $B = \{1, 2, 3\}$ y tenemos como conjunto universal a $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, entonces:

$$A^c = \{4, 5, 6, 7, 8, 9, 10\}$$

Pero si nuestro conjunto universal es $X = \{1, 2, 3, 4, 5\}$, entonces:

$$A^c = \{5\}$$

Con lo anterior, se puede usar la notación A^c siempre que se tenga claro quién es el conjunto universal, en otro caso, es recomendable usar la notación $X \setminus A$ cuando se quiera enfatizar quién es nuestro conjunto universal.

Ahora bien, dejado claro lo anterior y de la misma manera que en las operaciones anteriores, es bueno destacar que:

Nota 1.4.12. *Si A es un conjunto y x es un objeto, entonces $x \in A^c$ si y sólo si $x \notin A$*

Del mismo modo, en el caso de la negación:

Nota 1.4.13. *Si A es un conjunto y x es un objeto, entonces $x \notin A^c$ si y sólo si $x \in A$.*

Además podemos considerar las siguientes propiedades de esta operación:

Lema 1.4.14. *Si A y B son conjuntos y X es el conjunto universal, entonces los siguientes enunciados se satisfacen:*

a) $(A^c)^c = A$.

b) $(A \cup B)^c = A^c \cap B^c$.

c) $(A \cap B)^c = A^c \cup B^c$.

d) $X^c = \emptyset$.

e) *Si* $A^c = \emptyset$, *entonces* $A = X$.

f) $A \setminus B = A \cap B^c$.

g) *Si* $A \subseteq B$, *entonces* $B^c \subseteq A^c$

1.5. Familias de conjuntos

Una **familia de conjuntos** es un conjunto tal que todos sus elementos son conjuntos. Hay que tener en cuenta es que si $\mathscr{A}$ es una familia de conjuntos y x es un objeto, entonces el hecho de que $x \in \mathscr{A}$, implica que x *también* es un conjunto, es decir, cosas como $a \in x$ o $a \notin x$ tienen sentido, lo cual hasta ahora no necesariamente era cierto. Ejemplos de familias de conjuntos son las siguientes:

$$\{\{1,2\},\emptyset,\{3\}\} \qquad \{\{a\},\{a,b\}\} \qquad \{\emptyset\} \qquad \{\{1\},\{2\},\{3\},\dots\}$$

La primera familia de conjuntos que trabajaremos puede verse como una operación particular que no se revisó en la sección anterior porque consideré que era mejor introducirla aquí. Dado un conjunto A, la **potencia de $\boldsymbol{A}$**, denotada por $\mathcal{P}(A)$, es la familia que contiene a todos los subconjuntos de A, es decir:

$$\mathcal{P}(A) = \{x : x \subseteq A\}.$$

Nótese que, derivado directamente de la definición de potencia, tenemos la siguiente nota:

Nota 1.5.1. *Si A es un conjunto y x es un objeto, entonces $x \in \mathcal{P}(A)$ si y sólo si $x \subseteq A$.*

Y del mismo modo, es importante recalcar su negación:

Nota 1.5.2. *Si A es un conjunto y x es un objeto, entonces $x \notin \mathcal{P}(A)$ si y sólo si $x \nsubseteq A$*

En el siguiente teorema mostramos algunas de las propiedades del conjunto potencia.

Lema 1.5.3. *Si A y B son conjuntos, entonces:*

a) $\emptyset \in \mathcal{P}(A)$.

b) $A \in \mathcal{P}(A)$.

c) $\mathcal{P}(A) \neq \emptyset$.

d) $A \subseteq B$ *si y sólo si* $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.

e) $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$.

f) $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.

Demostración. d) Primero mostraremos la parte suficiente del enunciado, es decir, mostraremos que si $A \subseteq B$, entonces $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.

Consideremos un elemento $x \in \mathcal{P}(A)$, notemos que de acuerdo a la definición de potencia, se tiene que $x \subseteq A$. Para mostrar que $x \in \mathcal{P}(B)$, basta verificar que $x \subseteq B$. Lo siguiente lo mostraremos en una afirmación.

Afirmación. $x \subseteq B$.

Sea $t \in x$. Ya que sabemos que $x \subseteq A$, entonces podemos concluir que $t \in A$. Además, por hipótesis, podemos afirmar que $t \in B$, lo cual concluye la demostración de la afirmación.

Gracias a la afirmación anterior, tenemos que $x \subseteq B$ y por definición de potencia, se sigue que $x \in \mathcal{P}(B)$. Así concluye la demostración de la parte suficiente del enunciado.

Ahora procederemos a demostrar la parte necesaria, es decir, si $\mathcal{P}(A) \subseteq \mathcal{P}(B)$, entonces $A \subseteq B$.

Consideremos un objeto $x \in A$. Debido a lo anterior, tenemos que $\{x\} \subseteq A$, y por definición de potencia, $\{x\} \in \mathcal{P}(A)$. Debido a nuestras hipótesis podemos concluir que $\{x\} \subseteq \mathcal{P}(B)$, y por definición de potencia, entonces $\{x\} \subseteq B$. De lo anterior, $x \in B$, concluyendo la demostración.

$\square$

La familia anterior es una familia muy específica: la familia de subconjuntos de un conjunto dado. Sin embargo, existen muchas familias de conjuntos diferentes a la potencia y serán las que abordaremos a continuación.

1.6. Familias indexadas de conjuntos

A las familias de conjuntos las denotaremos por letras mayúsculas cursivas, esto para para distinguirlas de los conjuntos comunes. Además, también será usual que los elementos de la familia sean conjuntos escritos por comprensión y sea necesario introducir un nuevo elemento en la familia: el uso de índices. Por ejemplo, consideremos los siguientes conjuntos:

$$A = \{x \in \mathbb{N} : x \leq 1\} \qquad B = \{x \in \mathbb{N} : x \leq 2\} \qquad C = \{x \in \mathbb{N} : x \leq 3\} \qquad D = \{x \in \mathbb{N} : x \leq 4\}$$

Nótese que la escritura por comprensión de los cuatro conjuntos se parece mucho. Además, la familia que tiene a los cuatro conjuntos es muy sencilla de describir: $\{A, B, C, D\}$. Pero, ¿qué pasaría si quiero agregar los siguientes 2 conjuntos del mismo tipo a la familia? Es decir, quiero agregar a la familia los conjuntos:

$$E = \{x \in \mathbb{N} : x \leq 5\} \qquad F = \{x \in \mathbb{N} : x \leq 6\}$$

Y con ello, tenemos la familia $\{A, B, C, D, E, F\}$. Y aun no tenemos problemas. Pero ¿qué pasa si ahora quiero agregar los siguientes 27 conjuntos a la familia? Ahora sí hay un problema, pues no tenemos suficiente alfabeto para nombrar a los nuevos elementos que quiero en la familia. Para solucionar dicho problema, usaremos el uso de familias indexadas.

Primero notemos que la escritura por comprensión de los conjuntos A, B, C, D, E y F son muy parecidas entre ellos, de hecho sólo varían en el numero natural 1,2,3,4,5, y 6, respectivamente. Eso lo usaremos para renombrar los conjuntos anteriores usando índices y poniendo su escritura por comprensión en términos del índice. Por ejemplo:

$$A_k = \{x \in \mathbb{N} : x \leq k\}$$

Nótese que el conjunto arriba descrito es un conjunto escrito por comprensión, pero la propiedad que lo describe está en términos de dos variables: el objeto genérico x y el índice k. En ese sentido, cada que el índice cambia, la escritura por comprensión del conjunto también se modifica, por ejemplo:

$$A_1 = \{x \in \mathbb{N} : x \leq 1\} \qquad A_2 = \{x \in \mathbb{N} : x \leq 2\} \qquad A_3 = \{x \in \mathbb{N} : x \leq 3\}$$

Este uso de los índices nos permite tener un nombre para todos los conjuntos que nosotros queramos, es decir, aunque los primeros nombres que usamos: A, B, C, D, E y F no son malos, estamos limitados porque no tenemos *suficientes nombres* para los conjuntos que estamos describiendo, considerando que sólo tenemos 27 letras en nuestro alfabeto. Pero con el uso de subíndices, tenemos nombres para todos los conjuntos del ejemplo:

$$A_{34} = \{x \in \mathbb{N} : x \leq 34\} \qquad A_{1000} = \{x \in \mathbb{N} : x \leq 1000\} \qquad A_{10^{100}} = \{x \in \mathbb{N} : x \leq 10^{100}\}$$

Con esto, podemos usar los índices para tener tantos conjuntos como nosotros queramos. Además de lo anterior, nos servirá para poder introducir la escritura de familias de conjuntos mediante índices. Por ejemplo, teniendo nuevamente en cuenta los conjuntos A_i, podemos describir las siguientes familias:

$\mathscr{A} = \{A_1, A_2, A_3, A_4, A_5, A_6\}$ $\mathscr{B} = \{A_2, A_4, A_6, A_8, \dots\}$ $\mathscr{C} = \{A_1, A_2, A_3, A_4, \dots\}$

Las familias antes descritas están por extensión, pero al igual que en la sección de escritura de conjuntos, dicha escritura puede presentar varios problemas. En el caso de la escritura de conjuntos generales, la escritura por comprensión es la que ayuda a resolver dichos problemas. En el caso de las familias de conjuntos, usaremos una variación de la escritura por comprensión:

$$\{A_i : i \in I\}$$

Antes de continuar, será necesario aprender a interpretar este tipo de notaciones. La notación del conjunto está dividido por el signo de dos puntos, éstos se leen igual que en la notación de los conjuntos por comprensión: *tal que*. Sin embargo, el objeto antes de los dos puntos en dicha escritura no necesariamente requiere un conjunto de referencia, al contrario que en la notación por comprensión usual. Esto se debe a que el objeto genérico A_i es un conjunto y puede estar descrito en términos del índice. La parte importante de esta notación es la segunda condición: $i \in I$. La variable i es un objeto genérico de un conjunto I, a dicho conjunto se le suele llamar **conjunto de índices.** Dicho conjunto nos indica quiénes son los índices que vamos a considerar para la familia y, por consiguiente, qué conjuntos sí forman parte de la familia. Nuevamente retomando el ejemplo anterior, los índices que consideramos en general son los naturales, así, la familia $\mathscr{C} = \{A_1, A_2, A_3, A_4, \dots\}$ puede ser reescrita como $\mathscr{C} = \{A_i : i \in \mathbb{N}\}$. En este caso, los objetos que están en la familia $\mathscr{C}$ son todos los A_i con i variando en los naturales, es decir, $A_1, A_2, A_3, \dots$

También puede ser que no necesitemos a todos los índices posibles, sólo algunos de ellos. Por ejemplo, para la familia $\mathscr{B} = \{A_2, A_4, A_6, A_8, \dots\}$ sólo necesitamos los índices pares, en tal caso, podemos definir $J = \{n \in \mathbb{N} : n \text{ es par}\}$ y con ello, se puede reescribir a $\mathscr{B}$ como $\mathscr{B} = \{A_i : i \in J\}$.

Cabe señalar que la escritura estándar de las familias de conjuntos tiene una variante que en algunos casos es más sencilla de usar. Si tenemos una familia de conjuntos $\{A_i : i \in I\}$, ésta también se puede escribir como $\{A_i\}_{i \in I}$. Cualquiera de las dos notaciones es válida y podemos usar ambas de manera indistinta.

La siguiente nota es una forma sencilla de verificar una igualdad de conjuntos mediante su indexado.

Nota 1.6.1. *Si A_i y A_j son conjuntos y además $i = j$, entonces $A_i = A_j$.*

Por último, es importante señalar que el conjunto de índices no necesariamente son conjuntos de números. Por ejemplo, consideremos la potencia del conjunto de los naturales, $\mathcal{P}(\mathbb{N})$. Ahora, consideremos el conjunto $I = \{A \in \mathcal{P}(\mathbb{N}) : A \text{ es un conjunto de números pares}\}$ y, para cada $A \in I$, podemos definir los siguientes conjuntos $R_A = \{S \in \mathcal{P}(\mathbb{N}) \colon S \subseteq A \text{ y } S \text{ tiene 2 elementos}\}$.

Para tener una idea de quién es la familia $\{R_A \colon A \in I\}$ podemos describir algunos de sus elementos.

Tomemos un índice en el conjunto I, digamos $\{0,2,4\}$. En tal caso, el conjunto $R_{\{0,2,4\}}$ es el conjunto

$$R_{\{0,2,4\}} = \{S \in \mathcal{P}(\mathbb{N}) : S \subseteq \{0,2,4\} \text{ y } S \text{ tiene 2 elementos}\}$$

Así, la escritura por comprensión del conjunto nos permite en muchos casos comprender al conjunto en sí, pero en este ejemplo, además podemos dar su escritura por extensión.

$$R_{\{0,2,4\}} = \{\{0,2\},\ \{0,4\},\ \{2,4\}\}$$

1.6.1. Unión de familias

Uno de los intereses principales de las familias de conjuntos será el poder tener una definición más amplia de algunas operaciones entre conjuntos, en nuestro caso, de la unión y la intersección. Nosotros hemos definido a la unión e intersección entre dos conjuntos. ¿Pero qué pasaría si necesitamos unir o intersecar más conjuntos? Por ejemplo, ¿qué tal si necesitamos la unión de tres conjuntos A_1, A_2 y A_3? Bueno, en este caso no es tan difícil, pues podríamos intersecar primero dos de ellos y luego intersecar un tercero: $A_1 \cup (A_2 \cup A_3)$. ¿Y si queremos unir cinco conjuntos? Pues es un poco de lo mismo, ¿no? primero unimos dos de los cinco, luego le unimos un tercero, luego un cuarto y luego el último:

$$A_5 \cup (\,A_4 \cup (\,A_3 \cup (A_2 \cup A_1)\,)\,).$$

Así pareciera que la cosa no es tan difícil, podemos unir uno por uno hasta obtener la unión de todos ellos. Pero aquí viene un pequeño problema: ¿qué pasaría si tengo que unir una infinidad tan grande de conjuntos que no es posible unirlos uno por uno? Por ejemplo, si definimos $B_r = \{x \in \mathbb{R} : |x| = r\}$, entonces no es posible unir uno por uno los elementos de la familia $\{B_r : r \in \mathbb{R}\}$, pues el conjunto de índices es demasiado grande y, en particular, no es numerable.

En los casos anteriores sí estamos en un problema. Y además en un problema que necesitamos resolver porque será muy frecuente el tener que unir o intersecar cantidades enormes de conjuntos. Para esto nos van a servir también las familias de conjuntos.

Sea $\mathcal{A}$ una familia de conjuntos. La **unión de $\mathcal{A}$**, denotada por $\cup\mathcal{A}$, es el conjunto:

$$\cup\mathcal{A} = \{x : x \in A \text{ para algún } A \in \mathcal{A}\}.$$

Intentemos rescatar cosas importantes de esta definición. Primero que nada, los elementos en $\mathcal{A}$ son los **uniendos**, por lo que notemos que un elemento x pertenece a la unión si está en alguno de los uniendos...que de hecho es exactamente la misma que usamos cuando definimos la unión de dos conjuntos.

Aunque la anterior es la definición formal, en general nosotros trabajaremos con familias de conjuntos indexadas, por lo que podemos usar una variación de la notación. Si $\mathcal{A} = \{A_i : i \in I\}$ es una familia de conjuntos, entonces podemos denotar a la unión de la familia $\mathcal{A}$ de otra forma:

$$\bigcup_{i\in I} A_i$$

Y como bien parece, está notación es un poco más escalofriante, pero muchas veces más versátil cuando se usa correctamente. Vamos a ver primero cómo interpretarla. Noten que el primer símbolo es un indicativo de unión. El objeto bajo este símbolo es de la forma $i \in I$ e indica *quién es el conjunto de índices* y, por consiguiente, quiénes son los conjuntos que funcionarán como uniendos. Los términos A_i indican el nombre de los uniendos. Esta notación nos permite obtener de manera más precisa el significado de la pertenencia a una unión:

Nota 1.6.2. *Sean $\{A_i\}_{i\in I}$ una familia de conjuntos y x un objeto cualquiera.*

$$x \in \bigcup_{i\in I} A_i \text{ si y sólo si existe } i \in I \text{ tal que } x \in A_i.$$

Lo que dice el enunciado anterior es que un objeto está en una unión si y sólo si está en alguno de los uniendos. En este caso, utilizamos el cuantificador existencial sobre el conjunto de índices para garantizar la existencia de un uniendo al que pertenece x. Una vez que tenemos la afirmación, tenemos la negación del enunciado.

Nota 1.6.3. *Sean $\{A_i\}_{i\in I}$ una familia de conjuntos y x un objeto cualquiera.*

$$x \notin \bigcup_{i\in I} A_i \text{ si y sólo si para todo } i \in I \text{ sucede que } x \notin A_i.$$

Este tipo de uniones son un caso mucho más general de las que estudiamos previamente y, de hecho, las uniones que vimos en secciones previas son un caso particular cuando el conjunto de índices tiene sólo dos índices. Primero veremos un ejemplo concreto sobre cómo podemos trabajar este tipo de operaciones, principalmente la pertenencia o no pertenencia de un objeto en la unión de la familia de conjuntos.

Ejemplo 1.6.4. *Sea $\mathbb{R}$ el conjunto de los números reales. Para cada $k \in \mathbb{R}$ definimos el conjunto $Q_k = \{p \in \mathbb{R} \colon |p| = k\}$ e $I = \{r \in \mathbb{R} \colon 0 \le r \le 1\}$. Muestra lo siguiente:*

a) $\frac{1}{2} \in \bigcup_{i\in I} Q_i$.

b) $-\frac{3}{4} \in \bigcup_{i\in I} Q_i$.

c) $-2 \notin \bigcup_{i\in I} Q_i$.

Demostración. Antes de iniciar la demostración, es de suma importancia que comprendas quiénes son todos los uniendos de nuestra unión. En este caso, los uniendos están descritos en términos de los índices y, por ello, tenemos una cantidad considerable de uniendos[2], pero que pueden ser comprendidos mediante su escritura por comprensión. De acuerdo al conjunto de índices, sólo consideraremos índices que sean al menos 0 y a lo mucho 1. Por ejemplo los índices $0, \frac{1}{2}, \frac{\pi}{4}$ etc. sí son índices válidos en I, por lo que tenemos que los siguientes conjuntos sí son uniendos de dicha unión:

Q_0 $\qquad$ $Q_{\frac{1}{2}}$ $\qquad$ $Q_{\frac{\pi}{4}}$

Pero, más que eso, podemos tener una escritura por comprensión de dichos conjuntos, como se muestra a continuación.

$Q_0 = \{p \in \mathbb{R} \colon |p| = 0\}$ $\qquad$ $Q_{\frac{1}{2}} = \{p \in \mathbb{R} \colon |p| = \frac{1}{2}\}$ $\qquad$ $Q_{\frac{\pi}{4}} = \{p \in \mathbb{R} \colon |p| = \frac{\pi}{4}\}$

Y aunque es importante que ya puedas dominar dichos conjuntos escritos por comprensión, aun se puede ser más explícito en los conjuntos anteriores y escribirlos por extensión.

$Q_0 = \{0\}$ $\qquad$ $Q_{\frac{1}{2}} = \{-\frac{1}{2}, \frac{1}{2}\}$ $\qquad$ $Q_{\frac{\pi}{4}} = \{-\frac{\pi}{4}, \frac{\pi}{4}\}$

Así pues, debes tener en consideración que los tres conjuntos anteriores son los conjuntos asociados a tres índices: 0,$\frac{1}{2}$ y $\frac{\pi}{4}$, pero tenemos todavía una infinidad de índices, pues por cada número r entre 0 y 1, tenemos un conjunto Q_r, que son bastantes. Lo importante aquí es que puedas comprender cuáles conjuntos son los que forman parte de la unión o te será difícil avanzar en lo que sigue. Pues bien, si ya estás listo, pasemos a lo siguiente.

a) $\frac{1}{2} \in \bigcup_{i \in I} Q_i$.

Para esta demostración, tenemos que ver que existe un índice j en I tal que $\frac{1}{2} \in Q_j$. En este caso, como los conjuntos Q_i están descritos por comprensión y el objeto $\frac{1}{2}$ es un objeto concreto, tenemos que determinar el valor exacto para j.

Sin embargo, del análisis previamente realizado, es sencillo ver que $\frac{1}{2} \in Q_{\frac{1}{2}}$, y como $\frac{1}{2} \in I$, entonces podemos concluir que existe $\frac{1}{2} \in I$ tal que $\frac{1}{2} \in Q_{\frac{1}{2}}$. Luego, por definición de unión de familia se concluye que $\frac{1}{2} \in \bigcup_{i \in I} Q_i$.

b) $-\frac{3}{4} \in \bigcup_{i \in I} Q_i$.

[2]De hecho, es una cantidad tan grade que no se pueden enlistar

Para esta demostración nuevamente probaremos que existe un índice j en I tal que $-\frac{3}{4} \in Q_j$. Al igual que en el inciso anterior, como los conjuntos Q_i están descritos por comprensión y el objeto $\frac{3}{4}$ es un objeto concreto, tenemos que determinar el valor exacto para j.

Notemos que, del análisis previo a la demostración, se tiene que $-\frac{3}{4} \in Q_{\frac{3}{2}}$. Más aun, como $\frac{3}{2} \in I$, se concluye que existe $\frac{3}{2} \in I$ tal que $-\frac{3}{2} \in Q_{\frac{3}{2}}$. Luego, por definición de unión de familia se sigue que $-\frac{3}{2} \in \bigcup_{i\in I} Q_i$.

c) $-2 \notin \bigcup_{i\in I} Q_i$.

Para este inciso probaremos que para todo $k \in I$, $-2 \notin Q_k$.

Sea $k \in I$. Por definición de I se tiene que $k \in \mathbb{R}$ y además, $0 \leq k$ y $k \leq 1$. Así, dado que $|-2| > 1$, podemos concluir que $|-2| \neq k$, siguiéndose de ello que -2 no satisfaga las propiedades que definen al conjunto Q_k y así, $-2 \notin Q_k$.

□

Lo anterior es un ejemplo concreto sobre pertenencia o no pertenencia de un objeto a una unión de familias donde los uniendos y el conjunto de índices son conjuntos muy concretos. Sin embargo, también tenemos algunas propiedades generales de esta operación, como las siguientes.

Lema 1.6.5. *Si $\{A_i\}_{i\in I}$ es una familia de conjuntos y B es un conjunto, entonces:*

a) Para toda $j \in I$ se tiene que $A_j \subseteq \cup_{i\in I} A_i$, es decir, cualquier uniendo es subconjunto de la unión.

b) $B \cap (\cup_{i\in I} A_i) = \cup_{i\in I}(B \cap A_i)$

c) Si para toda $m \in I$, $A_m \subseteq B$, entonces $\cup_{i\in I} A_i \subseteq B$.

d) $\cup_{i\in I} A_i = \emptyset$ si y sólo si para todo $j \in I$, $A_j = \emptyset$.

Demostración. a) La demostración de este inciso pueden verla en https://youtu.be/nly8cAGAqBo.

b) En este inciso procederemos por doble contención. Primero mostraremos que $B \cap (\cup_{i\in I} A_i) \subseteq \cup_{i\in I}(B \cap A_i)$.

Sea $x \in B \cap (\cup_{i\in I} A_i)$. Por definición de intersección se tiene que $x \in B$ y que $x \in \cup_{i\in I} A_i$. Luego, por definición de unión de familia se tiene que existe $k \in I$ tal que $x \in A_k$. En tal caso, por definición de intersección se puede concluir que $x \in B \cap A_k$. Con lo anterior se muestra que existe $k \in I$ tal que $x \in B \cap A_k$, siguiéndose de la definición de unión de familia que $x \in \cup_{i\in I}(B \cap A_i)$, mostrando así la contención deseada.

Ahora procederemos a mostrar la otra contención, es decir $\cup_{i\in I}(B \cap A_i) \subseteq B \cap (\cup_{i\in I} A_i)$.

Sea $z \in \cup_{i\in I}(B \cap A_i)$, por definición de unión de familias, existe $m \in I$ tal que $z \in B \cap A_m$. Luego, por definición de intersección se tiene que $x \in B$ y $x \in A_m$. Así, existe $m \in I$ tal que $z \in A_m$, por lo que por definición de unión de familia se sigue que $z \in \cup_{i\in I}(B \cap A_i)$. Dado que $z \in B$ y $z \in \cup_{i\in I}(B \cap A_i)$, por definición de intersección concluimos que $z \in B \cap \cup_{i\in I}(B \cap A_i)$, mostrando la contención deseada.

c) Para esta demostración, consideremos $x \in \bigcap_{i\in I} A_i$. De acuerdo a la definición de unión se tiene que existe $k \in I$ tal que $x \in A_k$. De acuerdo a nuestras hipótesis, como $k \in I$, podemos garantizar que $A_k \subseteq B$ y con ello, $x \in B$.

$\square$

1.6.2. Intersección de familias

Dada una familia de conjuntos $\mathcal{A}$, la **intersección de $\mathcal{A}$**, denotada por $\cap\mathcal{A}$, es el conjunto:

$$\cap\mathcal{A} = \{x : \text{para toda } A \in \mathcal{A} \text{ sucede que } x \in A\}.$$

A los términos A_i se les llama **intersecandos**. Nótese que, de acuerdo a la definición, un objeto está en una intersección si y sólo si está en todos los intersecandos.

Al igual que en la unión de familias, vamos a introducir una notación más versátil en el caso de la intersección. Si $\mathcal{A} = \{A_i : i \in I\}$, podemos usar la siguiente notación para indicar la intersección de conjuntos:

$$\bigcap_{i\in I} A_i.$$

Usando esta forma de notación tenemos que:

Nota 1.6.6. *Sean $\{A_i\}_{i\in I}$ una familia de conjuntos y x un objeto cualquiera.*

$$x \in \bigcap_{i\in I} A_i \text{ si y sólo si para todo } i \in I \text{ sucede que } x \in A_i.$$

Y a partir de ella tenemos una forma más simple de analizar la negación de lo anterior:

Nota 1.6.7. *Sean $\{A_i\}_{i\in I}$ una familia de conjuntos y x un objeto cualquiera.*

$$x \notin \bigcap_{i\in I} A_i \text{ si y sólo si existe } i \in I \text{ tal que } x \notin A_i.$$

Al igual que en el caso de la unión, es importante poder determinar la pertenencia o no pertenencia de un objeto a la intersección de una familia. Para ello, vamos a considerar el siguiente ejemplo.

Ejemplo 1.6.8. *Sea* $\mathbb{R}$ *el conjunto de los números reales. Para cada* $k \in \mathbb{R}$ *definimos el conjunto* $S_k = \{x \in \mathbb{R} \colon 0 \leq x \leq k\}$ *y* $J = \{r \in \mathbb{N} \colon 1 \leq r\}$*. Muestra lo siguiente:*

a) $1 \in \bigcap_{i \in J} S_i$.

b) $\frac{1}{2} \in \bigcap_{i \in J} S_i$.

c) $-2 \notin \bigcap_{i \in J} S_i$.

Demostración. Antes de iniciar la demostración de estos incisos, es importante entender quiénes son los conjuntos involucrados en esta intersección. En este caso, los conjuntos S_i son conjuntos indexados, por lo que su escritura por comprensión está relacionada con el índice de cada conjunto. Sin embargo, en este ejemplo no será posible describir el conjunto por extensión. Sin embargo, sí es posible comprender dichos conjuntos, por ejemplo, para los valores $2, \pi, \sqrt{5}$, todos ellos en J, tenemos que

$$S_2 = \{x \in \mathbb{R} \colon 0 \leq x \leq 2\} \qquad S_\pi = \{x \in \mathbb{R} \colon 0 \leq x \leq \pi\} \qquad S_{\sqrt{2}} = \{x \in \mathbb{R} \colon 0 \leq x \leq \sqrt{5}\}$$

Y, aunque el conjunto de índices es un conjunto infinito, teniendo claridad de quiénes son los intersecandos, será más sencillo determinar quién es la intersección de dicha familia. Por el momento, bastará saber cuándo es que un objeto está o no en la intersección y hay una manera formal de determinarlo, misma que revisaremos en los siguientes incisos.

a) $1 \in \bigcap_{i \in J} S_i$.

Para este inciso, mostraremos que para todo $m \in J$, $1 \in S_m$.

Sea $m \in J$. Por la definición de J se tiene que $m \in \mathbb{N}$ y además, $1 \leq m$. Por otro lado, de acuerdo a la definición de la familia, se tiene que $S_m = \{x \in \mathbb{R} \colon 0 \leq x \leq m\}$. Así, es sencillo notar que $1 \in \mathbb{R}$ y además, $0 \leq 1 \leq m$. Con lo anterior, $1 \in S_m$. Por la definición de intersección de familia podemos concluir que $1 \in \bigcap_{i \in J} S_i$.

b) $\frac{1}{2} \in \bigcap_{i \in J} S_i$.

Nuevamente, para este inciso mostraremos que para todo $n \in J$, $\frac{1}{2} \in S_n$.

Sea $n \in J$. Por la definición de J se tiene que $n \in \mathbb{N}$ y además, $1 \leq n$. De esto último en particular podemos garantizar que $\frac{1}{2} \leq n$. Por otro lado, de acuerdo a la definición de la familia, se tiene que $S_n = \{x \in \mathbb{R} \colon 0 \leq x \leq n\}$. Así, es sencillo notar que $\frac{1}{2} \in \mathbb{R}$ y además, $0 \leq \frac{1}{2} \leq n$. Con lo anterior, $\frac{1}{2} \in S_n$. Por la definición de intersección de familia podemos concluir que $\frac{1}{2} \in \bigcap_{i \in J} S_i$.

c) $-2 \notin \bigcap_{i \in J} S_i$.

Para este inciso, basta ver que existe $t \in J$ tal que $-2 \notin S_t$.

Así, proponemos el índice 3. Notemos que efectivamente $3 \in J$ pues 3 satisface que $3 \in \mathbb{N}$ y $1 \leq 3$. Además, por la definición de la familia se tiene que $S_3 = \{x \in \mathbb{R} \colon 0 \leq x \leq 3\}$, siguiéndose de ello el que $-2 \notin S_3$. Con lo anterior, existe $3 \in J$ tal que $-2 \notin S_3$, concluyéndose de la definición de intersección de familias que $-2 \notin \bigcap_{i \in J} S_i$.

□

Es importante poder acostumbrarse con la notación de uniones e intersecciones porque se trabajarán mucho en cualquier carrera que use matemáticas en la Facultad de Ciencias. Las siguientes son propiedades básicas de la intersección de familias.

Lema 1.6.9. *Si $\{A_i\}_{i \in I}$ es una familia de conjuntos y B es un conjunto, entonces:*

a) Si para todo $n \in I$, $B \subseteq A_n$, entonces $B \subseteq \cap_{i \in I} A_i$.

b) Para toda $j \in I$ se tiene que $\cap_{i \in I} A_i \subseteq A_j$, es decir, la intersección siempre es subconjunto de cualquier intersecando.

c) $B \cup (\cap_{i \in I} A_i) = \cap_{i \in I}(B \cup A_i)$.

Demostración. (a) La demostración de este video pueden verla en https://youtu.be/-YELY4NYu4k

(b) Sea $k \in I$ arbitrario. Mostraremos que $\bigcap_{i \in I} A_i \subseteq A_k$ mediante la siguiente afirmación:

Afirmación. $\bigcap_{i \in I} A_i \subseteq A_k$.

Sea $x \in \bigcap_{i \in I} A_i$, con ello, para todo $t \in I$, $x \in A_t$. En particular, como $k \in I$, se tiene que $x \in A_k$, concluyendo la demostración de la afirmación.

Por la afirmación anterior, $\bigcap_{i \in I} A_i \subseteq A_k$.

(c) Esta demostración la realizaremos por doble contención, y para ello, primero mostraremos la contención $B \cup (\bigcap_{i \in I} A_i) \subseteq \bigcap_{i \in I}(B \cup A_i)$.

Sea $x \in B \cup (\bigcap_{i \in I} A_i)$. Por definición de unión tenemos que $x \in B$ o $x \in \bigcap_{i \in I} A_i$. Ahora procederemos por casos.

Caso 1. $x \in B$.

Dado que tenemos que demostrar que para todo $j \in I$, $x \in B \cup A_j$, realizaremos dicha demostración mediante la siguiente afirmación.

Afirmación 1. Para todo $j \in I$, $x \in B \cup A_j$.

Sea $k \in I$ arbitrario. Debido a el caso en el que estamos, sabemos que $x \in B$, siguiéndose de la definición de unión que $x \in B \cup A_k$, lo que muestra la afirmación de la demostración.

Gracias a la afirmación 1, podemos garantizar por la definición de intersección de familias el que $x \in \bigcap_{i \in I}(B \cup A_i)$, concluyendo este caso.

Caso 2. $x \in \bigcap_{i \in I} A_i$.

En este caso tenemos, por la definición de intersección de familias, el que para toda $l \in I$, $x \in A_l$. Ahora, como debemos demostrar que $j \in I$, $x \in B \cup A_j$, realizaremos la siguiente afirmación.

Afirmación 2. Para todo $j \in I$, $x \in B \cup A_j$.

Sea $k \in I$ arbitrario. Como sabemos que para todo $l \in I$, $x \in B \cup A_l$, entonces $x \in A_k$ y por definición de unión, $x \in B \cup A_k$, concluyendo la demostración de la afirmación.

Gracias a la afirmación 2, y por la definición de intersección de familias, se tiene que $x \in \bigcap_{i \in I}(B \cup A_i)$, concluyendo este caso.

Gracias al caso 1 y 2 anteriores, podemos decir que $x \in \bigcap_{i \in I}(B \cup A_i)$, mostrando así la contención deseada.

Ahora procederemos a realizar la segunda contención, es decir, $\bigcap_{i \in I}(B \cup A_i) \subseteq B \cup (\bigcap_{i \in I} A_i)$. Para ello, consideremos $x \in \bigcap_{i \in I}(B \cup A_i)$.

Debido a la definición de intersección de familias, podemos decir que para todo $j \in I$, $x \in B \cup A_j$. Debido a que queremos demostrar una disyunción, a saber, $x \in B$ o para todo $l \in I$, $x \in A_l$, podemos suponer que $x \notin B$ y procederemos a demostrar que para todo $l \in I$, $x \in A_l$. Dicha demostración la realizaremos mediante una afirmación.

Afirmación. Para todo $l \in I$, $x \in A_l$.

Sea $k \in I$ arbitrario. De acuerdo a lo que sabemos sobre x, podemos decir que $x \in B \cup A_k$, y por definición de unión se tiene que $x \in B$ o $x \in A_k$. Sin embargo, por la suposición antes realizada, sabemos que $x \notin B$, por lo que $x \in A_k$, concluyendo nuestra demostración.

Gracias a la afirmación anterior, podemos afirmar que $x \in \bigcap_{i \in I} A_i$. Por definición de unión, se sigue que $x \in B \cup \bigcap_{i \in I} A_i$, concluyendo la segunda contención. $\square$

Además de las propiedades anteriores, es importante recalcar un resultado más que, de hecho, relaciona la unión e intersección:

Lema 1.6.10. *Si* $\{A_i\}_{i\in I}$ *es una familia de conjuntos y* B *es un conjunto, entonces:*

a) $B \setminus (\cap_{i\in I} A_i) = \cup_{i\in I}(B \setminus A_i)$.

b) $B \setminus (\cup_{i\in I} A_i) = \cap_{i\in I}(B \setminus A_i)$.

Demostración. (a) Esta demostración la realizaremos por doble contención, por lo que primero mostraremos que $B \setminus (\bigcap_{i\in I} A_i) \subseteq \bigcup_{i\in I}(B \setminus A_i)$.

Sea $x \in B \setminus (\bigcap_{i\in I} A_i)$. Gracias a la definición de diferencia se tiene que $x \in B$ y $x \notin \bigcap_{i\in I} A_i$. De esto último, y de acuerdo a la negación de la definición de intersección, existe $k \in I$ tal que $x \notin A_k$. Notemos que por definición de diferencia, en particular se tiene que $x \in B \setminus A_k$. Así, existe $k \in I$ tal que $x \in B \setminus A_k$. Por definición de unión de familias, podemos concluir que $x \in \bigcup_{i\in I}(B \setminus A_i)$, mostrando la primera contención.

Ahora mostraremos que $\bigcup_{i\in I}(B \setminus A_i) \subseteq B \setminus (\bigcap_{i\in I} A_i)$. Sea $x \in \bigcup_{i\in I}(B \setminus A_i)$. Por definición de unión se sigue que existe $k \in I$ tal que $x \in B \setminus A_k$, y por la definición de diferencia, $x \in B$ y $x \notin A_k$.

Como existe $x \in I$ tal que $x \notin A_k$, se tiene por la negación de la definición de intersección que $x \notin \bigcap_{i\in I} A_i$. De esto último y del hecho de que $x \in B$, podemos decir que $x \in B \setminus (\bigcap_{i\in I} A_i)$, concluyendo nuestra segunda contención. □

Corolario 1.6.11. *Si* $\{A_i\}_{i\in I}$ *es una familia de conjuntos, entonces:*

a) $(\cap_{i\in I} A_i)^c = \cup_{i\in I}(A_i^c)$.

b) $(\cup_{i\in I} A_i)^c = \cap_{i\in I}(A_i^c)$.

Demostración. Ambos incisos se siguen inmediatamente del teorema 1.6.10 al sustituir a B por el conjunto universo. □

Para terminar con este tema quiero mencionar algunas variaciones de la notación de intersecciones y uniones. Ya sabemos que si tenemos una familia de conjuntos, por decir: $\{A_i\}_{i\in I}$, la unión de esta familia está denotada por:

$$\bigcup_{i\in I} A_i$$

Así como la intersección está indicada por:

$$\bigcap_{i\in I} A_i$$

Sin embargo, es frecuente el encontrar que las familias son finitas, es decir, que $I = \{1, 2, \ldots, n\}$ para cierto número n. En este caso, es más sencillo usar la notación:

$$\bigcup_{i=1}^{n} A_i$$

Y su análoga para intersección:

$$\bigcap_{i=1}^{n} A_i$$

Este tipo de notación es más específica y suele usarse frecuentemente con familias de conjuntos finitas. No dejemos de lado que con esta notación el subíndice debajo del símbolo de unión o intersección indica en qué número inicia nuestro conteo y el superíndice indica hasta cuál número detendremos el conteo.

Del mismo modo, cuando $I = \mathbb{N}$ es frecuente que se deje indicada la notación:

$$\bigcup_{i\in\mathbb{N}} A_i$$

Aunque tampoco es raro que se use la siguiente:

$$\bigcup_{i=1}^{\infty} A_i$$

De la misma manera, cuando tenemos la intersección de conjuntos:

$$\bigcap_{i\in\mathbb{N}} A_i$$

Se usa de manera indistinta respecto a esta notación:

$$\bigcap_{i=1}^{\infty} A_i$$

Lo importante de todas ellas es no tenerles miedo y considerar que todas indican una unión o una intersección. Si nos remitimos a las definiciones de ambas no habrá problemas al momento de interpretarlas.

1.7. Particiones

El último tema que veremos de esta sección de conjuntos está relacionado con un hecho cotidiano. Supongamos que tenemos un grupo de personas, digamos G, y queremos llevarlas a visitar diferentes lugares. Debido a la cantidad de personas que tenemos es preferible dividirlas en grupos más pequeños para que sea más accesible el viaje, pero hay que ver qué condiciones requeriremos para estos subgrupos [3].

En primer lugar, queremos que efectivamente haya personas en los subgrupos, es decir, (i) que todos los subgrupos estén conformados por al menos una persona. Además, (ii) cada persona de G debe estar en

[3] En estas notas, cuando se habla de *grupos* o *subgrupos*, no nos referimos a las estructuras matemáticas de la Teoría de grupos, sino que hacemos alusión al uso coloquial de dichas palabras.

exactamente uno de los subgrupos, pues no es conviene que una misma persona esté simultáneamente en dos subgrupos diferentes. Por otro lado, vamos a requerir que (iii) no haya personas fuera de nuestro grupo original G. Es decir, no admitiremos a más personas que aquellas que están en G. Por último, debido a que todas las personas quieren participar en este viaje, es necesario que (iv) cualquier participante en G esté en al menos uno de los subgrupos que vamos a formar. Todo lo anterior será lo que nosotros, matemáticamente hablando, aceptaremos como una buena forma de separar al grupo G en subgrupos.

El problema anterior es muy frecuente en nuestra vida cotidiana, tiene una interpretación matemática que aparecerá frecuentemente a lo largo de la carrera y es la siguiente. Dado un conjunto no vacío A, una **partición de A** es una familia de de conjuntos, digamos $\mathscr{Q}$, que satisface lo siguiente:

a) Para todo $B \in \mathscr{Q}$, $B \neq \emptyset$.

b) Para todo $\{B, C\} \subseteq \mathscr{Q}$ tal que $B \cap C \neq \emptyset$, se satisface que $B = C$.

c) $\cup\mathscr{Q} = A$.

A los elementos de la familia $\mathscr{Q}$ les llamaremos **partes** o **clases** de la partición. Noten que el inciso (a) de la definición de partición coincide con el pinto (i), el inciso (b) coincide con el punto (ii) y el inciso (c) se corresponde simultáneamente con los incisos (iii) y (iv), como lo mostraremos en el siguiente lema.

Lema 1.7.1. *Sean A un conjunto no vacío y $\mathcal{Q}$ una familia de conjuntos. Los siguientes enunciados son equivalentes:*

a) $\cup\mathcal{Q} = A$.

b) La familia $\mathcal{Q}$ satisface que:

b.1) Para todo $B \in \mathcal{Q}$, $B \subseteq A$.

b.2) Para todo $x \in A$ existe $B \in \mathcal{Q}$ tal que $x \in B$.

Antes de continuar, es importante saber que el inciso (b) puede venir en una versión distinta, pero lógicamente equivalente, a la aquí presentada:

Nota 1.7.2. *Si A es un conjunto no vacío y $\mathcal{Q}$ es una partición de A, en muchas ocasiones se usa el recíproco del enunciado (b) de la definición de partición. Es decir, no es de extrañar que en algunos libros y demostraciones se use el siguiente enunciado en lugar del aquí propuesto:*

b) Para toda $\{B, C\} \subseteq \mathscr{Q}$ tal que $B \neq C$, se satisface que $B \cap C = \emptyset$.

Por otro lado, cabe mencionar que es posible poner el concepto de partición en términos de familias de conjuntos indexadas, las cuales serán las que utilizaremos principalmente en el curso.

Sean A un conjunto y $\{B_i : i \in I\}$ una familia de conjuntos, decimos que $\{B_i : i \in I\}$ es una **partición de** A si satisface lo siguiente:

a) Para todo $i \in I$, $B_i \neq \emptyset$.

b) Para todo $\{i, j\} \subseteq I$ tal que $B_i \cap B_j \neq \emptyset$, se satisface que $B_i = B_j$.

c) $\cup_{i \in I} B_i = A$.

Y al igual que con la nota pasada, también es posible que el inciso (b) tenga una presentación equivalente:

Nota 1.7.3. *Si A es un conjunto no vacío y $\{B_i : i \in I\}$ es una partición de A, en muchas ocasiones se usa el recíproco del enunciado (b) de la definición de partición. Es decir, no es de extrañar que en algunos libros y demostraciones se use el siguiente enunciado en lugar del aquí propuesto:*
b) Para toda $\{i, j\} \subseteq I$ tal que $B_i \neq B_j$, se satisface que $B_i \cap B_j = \emptyset$.

Ahora mostraremos un par de enunciados referentes a particiones. El primero será con familias indexadas de conjuntos explícitas y el segundo con familias de conjuntos más abstractas.

Ejemplo 1.7.4. *Sea $\mathbb{N}$ el conjunto de los números naturales. Para toda $n \in \mathbb{N}$ definimos el conjunto $R_n = \{(x, z) \in \mathbb{N} \times \mathbb{N} : x = n\}$. Muestra que la familia $\{R_n : n \in \mathbb{N}\}$ es una partición de $\mathbb{N} \times \mathbb{N}$.*

Demostración. Para esta demostración, debemos verificar que la familia $\{R_n : n \in \mathbb{N}\}$ satisface los tres incisos de la definición de familia.

a) Para toda $j \in \mathbb{N}$, $R_j \neq \emptyset$.

   Sea $k \in \mathbb{N}$ un índice arbitrario. Notemos que la pareja ordenada $(k, 0)$ satisface las condiciones del la escritura por comprensión del conjunto R_k, por lo que $(k, 0) \in R_k$ y, en particular, $R_k \neq \emptyset$.

b) Para toda $\{i, j\} \subseteq \mathbb{N}$ tal que $R_i \cap R_j \neq \emptyset$, se satisface que $R_i = R_j$.

   Para este inciso, si bien es posible realizar una doble contención, preferiremos ver que $i = j$ y con ello, gracias a la nota 1.6.1, podremos concluir que $R_i = R_j$. Para mostrar que $i = j$ nos auxiliaremos de la hipótesis $R_i \cap R_j \neq \emptyset$.

   Sea $(x, z) \in R_i \cap R_j$. Por definición de intersección, $(x, z) \in R_i$ y $(x, z) \in R_j$. Debido a la definición por comprensión de dichos conjuntos, se tiene que $(x, z) \in \mathbb{N} \times \mathbb{N}$, $x = i$ y $z = j$. De lo anterior, $i = j$ y por la nota 1.6.1, podemos concluir que $R_i = R_j$.

c) $\cup_{i\in\mathbb{N}} R_i = \mathbb{N}\times\mathbb{N}$.

Esta demostración la haremos por doble contención. Primero mostraremos que $\cup_{i\in\mathbb{N}} R_i \subseteq \mathbb{N}\times\mathbb{N}$.

Sea $(x,z) \in \cup_{i\in\mathbb{N}} R_i$. Por definición de unión de familias, existe $k \in \mathbb{N}$ tal que $(x,z) \in R_k$. Debido a la definición por comprensión de dichos conjuntos, se tiene que $(x,z) \in \mathbb{N}\times\mathbb{N}$ y $x = k$. En particular, $(x,z) \in \mathbb{N}\times\mathbb{N}$, mostrando la primera contención.

Ahora probaremos que $\mathbb{N}\times\mathbb{N} \subseteq \cup_{i\in\mathbb{N}} R_i$. Sea $(x,z) \in \mathbb{N}\times\mathbb{N}$. Por definición de producto cartesiano, $x \in \mathbb{N}$ y $z \in \mathbb{N}$. Además, notemos que (x,z) satisfacen las condiciones de la escritura por comprensión de R_x, por lo que $(x,z) \in R_x$. De lo anterior, existe $x \in \mathbb{N}$ tal que $(x,z) \in R_x$. Por definición de familia de conjuntos, se sigue que $(x,z) \in \cup_{i\in\mathbb{N}} R_i$, concluyendo la segunda contención.

De acuerdo a los puntos anteriores, por definición de partición se tiene que $\{R_n : n \in \mathbb{N}\}$ es una partición de $\mathbb{N}\times\mathbb{N}$. □

La demostración anterior es un ejemplo de cómo se puede proceder cuando, dada una familia de conjuntos indexada cuyos miembros estén descritos por comprensión, debemos demostrar que la familia es una partición de un conjunto. Sin embargo, en algunas ocasiones también podremos realizar demostraciones de particiones sin necesidad de que los miembros de la familia tengan una descripción por extensión. Para ello tenemos el siguiente ejemplo.

Ejemplo 1.7.5. *Sean A y B conjuntos tales que $B \subseteq A$ y $\{S_i : i \in I\}$ una partición de A. Muestra que si para todo $j \in I$, $B \cap S_j \neq \emptyset$, entonces $\{B \cap S_i : i \in I\}$ es una partición de B.*

Demostración. Para verificar que $\{B \cap S_i : i \in I\}$ es una partición de B, mostraremos que $\{B \cap S_i : i \in I\}$ satisface la definición de partición.

a) Para todo $i \in I$, $B \cap S_i \neq \emptyset$.

Sea $k \in I$. De acuerdo a las hipótesis del problema, se tiene que $B \cap S_k \neq \emptyset$.

b) Para todo $\{m,n\} \subseteq I$ tal que $(B \cap S_n) \cap (B \cap S_m) \neq \emptyset$, se satisface que $B \cap S_n = B \cap S_m$.

Sea $\{l,k\} \subseteq I$ tal que $(B \cap S_l) \cap (B \cap S_k) \neq \emptyset$. Para mostrar que $B \cap S_l = B \cap S_k$, utilizaremos que $\{S_i : i \in I\}$ es una partición de A y en particular cumple que para todo $\{i,j\} \subseteq I$ tal que $S_i \cap S_j \neq \emptyset$ se satisface que $S_i = S_j$.

Dado que por suposición $(B\cap S_l)\cap(B\cap S_k) \neq \emptyset$, podemos considerar un objeto $x \in (B\cap S_l)\cap(B\cap S_k)$. Por definición de intersección se tiene que $x \in B$, $x \in S_l$ y $x \in S_k$. En particular se tiene que $x \in S_l \cap S_k$, concluyendo que $S_l \cap S_k \neq \emptyset$. Con lo anterior, como $\{l,k\} \subseteq I$ satisface que $S_l \cap S_k \neq \emptyset$, entonces por lo antes mencionado en (i) podemos asegurar que $S_l = S_k$ y con ello, $(B \cap S_l) = (B \cap S_k)$.

c) $\bigcup_{i\in I}(B\cap S_i)=B$.

La demostración la realizaremos por doble contención. Primero mostraremos que $\bigcup_{i\in I}(B\cap S_i)\subseteq B$.

Sea $x\in\bigcup_{i\in I}(B\cap S_i)$. Por definición de unión se tiene que existe $k\in I$ tal que $x\in B\cap S_k$. De acuerdo a la definición de intersección, $x\in B$ y $x\in S_k$, en particular, $x\in B$, concluyendo la primera contención.

Ahora mostraremos que $B\subseteq\bigcup_{i\in I}(B\cap S_i)$. Sea $x\in B$. De acuerdo a nuestras hipótesis, $B\subseteq A$ y con ello, se tiene que $x\in A$. Debido a que $\{S_i : i\in I\}$ es una partición de A, en particular satisface que $\bigcup_{i\in I}S_i=A$, por lo que podemos garantizar que $x\in\bigcup_{i\in I}S_i$.

Gracias a la definición de unión, podemos afirmar que existe $k\in I$ tal que $x\in S_k$. Como además sabemos que $x\in B$, entonces $x\in B\cap S_k$. Así, existe $k\in K$ tal que $x\in B\cap S_k$, concluyendo por la definición de unión de familias el que $x\in\bigcup_{i\in I}(B\cap S_i)$, terminando la segunda contención.

Gracias a los puntos anteriores, podemos garantizar por la definición de partición el que la familia $\{B\cap S_i : i\in I\}$ es una partición de B. $\square$

Por último, es importante recalcar que la idea de partir un conjunto nos será muy importante para establecer conteos durante el capítulo referente a dicho tema. Para ello, dado un conjunto no vacío A y $\{B_i : i\in I\}$ una familia de conjuntos, diremos que $\{B_i : i\in I\}$ es una **pseudo-partición de A** si se satisface que:

a) Para todo $\{i,j\}\subseteq I$ tal que $B_i\cap B_j\neq\emptyset$, se satisface que $B_i=B_j$.

b) $\cup_{i\in I}B_i=A$.

1.8. Apéndice del capítulo 1

Sean A y B conjuntos y x un objeto cualquiera. Los incisos (a) y (b) de cada recuadro son equivalentes:

Pertenencia a un conjunto por comprensión.

a) $x \in \{z \in A : P(z)\}$

b) $x \in A$ y $P(x)$ es verdadero

a) $x \notin \{z \in A : P(z)\}$

b) $x \notin A$ o $P(x)$ es falso

Contención.

a) $A \subseteq B$

b) Si $x \in A$, entonces $x \in B$

a) $A \nsubseteq B$

b) Existe $x \in A$ tal que $x \notin B$

Contención propia.

a) $A \subset B$

b) $A \subseteq B$ y existe $x \in B$ tal que $x \notin A$

a) $A \not\subset B$

b) $A \nsubseteq B$ o para todo $x \in B$, $x \in A$

Unión.

a) $x \in A \cup B$

b) $x \in A$ o $x \in B$

a) $x \notin A \cup B$

b) $x \notin A$ y $x \notin B$

Intersección.

a) $x \in A \cap B$

b) $x \in A$ y $x \in B$

a) $x \notin A \cap B$

b) $x \notin A$ o $x \notin B$

Diferencia.

a) $x \in A \setminus B$

b) $x \in A$ y $x \notin B$

a) $x \notin A \setminus B$

b) $x \notin A$ o $x \in B$

Igualdad de parejas ordenadas. Si a, b, x y z son objetos, los incisos (a) y (b) de los siguientes recuadros son equivalentes

a) $(a, b) = (x, z)$

b) $a = x$ y $b = z$

a) $(a, b) \neq (x, z)$

b) $a \neq x$ o $b \neq z$

Complemento.

a) $x \in A^c$

b) $x \notin A$

a) $x \notin A^c$

b) $x \in A$

Potencia

a) $x \in P(A)$

b) $x \subseteq A$

a) $x \notin P(A)$

b) $x \nsubseteq A$

Unión (de familias)

a) $x \in \cup_{i \in I} A_i$

b) Existe $k \in I$ tal que $x \in A_k$

a) $x \notin \cup_{i \in I} A_i$

b) Para todo $k \in I$, $x \notin A_k$

Intersección (de familias).

a) $x \in \cap_{i \in I} A_i$

b) Para todo $k \in I$, $x \in A_k$

a) $x \notin \cap_{i \in I} A_i$

b) Existe $k \in I$ tal que $x \notin A_k$

Partición.

Sean A un conjunto y $\{A_i : i \in I\}$ una familia de subconjuntos de A.

a) $\{A_i : i \in I\}$ es una partición de A

b) Los tres enunciados siguientes se satisfacen:

b.1) Para todo $k \in I$, $A_k \neq \emptyset$

b.2) Para todo $\{n, m\} \subseteq I$ con $A_n \cap A_m \neq \emptyset$, se satisface que $A_n = A_m$

b.3) $\cup_{i \in I} A_i = A$

Capítulo 2

Relaciones binarias

2.1. Producto cartesiano

Hasta este punto hemos trabajado con agrupaciones de objetos, su escritura, propiedades de contención y algunas operaciones. Sin embargo, hasta ahora dichas estructuras no nos han permitido establecer un orden, pues, por ejemplo, los conjuntos $\{a,b,c\}$, $\{a,c,b\}$, $\{b,a,c\}$, $\{b,c,a\}$, $\{c,a,b\}$ y $\{c,b,a\}$ son todos iguales entre sí, es decir, no importa el orden en el que se escriban sus elementos. Sin embargo, para muchas situaciones es importante poder contar con estructuras matemáticas para las cuales sí importe el orden de los objetos y esa es la intención de la operación que veremos a continuación. Pero para ello primero necesitaremos un par de definiciones.

Dados dos objetos a y b, definimos la **pareja no ordenada de a y b** simplemente como el conjunto $\{a,b\}$. Como bien lo dice su nombre, es una pareja de objetos en la que es indistinto el orden en el que estén el objeto a y b, ya que $\{a,b\} = \{b,a\}$. Sin embargo, para definir la última operación de esta sección, es necesario introducir una estructura matemática en la que sí sea importante el orden en el que aparezcan los objetos.

Dados dos objetos a y b, la **pareja ordenada de a y b**, denotado por (a,b), es el conjunto:

$$(a,b) = \{\{a\},\{a,b\}\}.$$

Lo primero que veremos es el siguiente resultado.

Teorema 2.1.1. *Sean a, b, c y d objetos cualesquiera. Los siguientes incisos se satisfacen:*

a) Si $a \neq b$ entonces $(a,b) \neq (b,a)$.

b) $(a,b) = (c,d)$ si y sólo si $a = c$ y $b = d$.

Demostración. Primero mostraremos (a) procediendo por contraposición, es decir, demostraremos que: si

$(a,b)=(b,a)$, entonces $a=b$.

Como $(a,b)=(b,a)$, de acuerdo a la definición de pareja ordenada, se tiene que $\{\{a\},\{a,b\}\}=\{\{b\},\{b,a\}\}$. En particular, $\{a\}$ debe estar en el conjunto $\{\{b\},\{b,a\}\}$, por lo que sólo hay dos posibles casos:

Caso 1. $\{a\}=\{b\}$.

En este caso, se sigue inmediatamente que $a=b$.

Caso 2. $\{a\}=\{b,a\}$.

Como el conjunto $\{b,a\}$ sólo tiene un elemento, pues es igual al conjunto $\{a\}$, entonces $b=a$.

De los casos anteriores $a=b$, concluyendo la demostración del inciso (a).

Ahora mostraremos la parte necesaria del inciso (b), es decir, mostraremos que: si $a=c$ y $b=d$, entonces $(a,b)=(c,d)$.

Debido a que $a=c$ y $b=d$, podemos concluir que $\{a\}=\{c\}$ y que $\{a,b\}=\{c,d\}$, por lo que el conjunto $\{\{a\},\{a,b\}\}$ y el conjunto $\{\{c\},\{c,d\}\}$ deben ser iguales, es decir, $(a,b)=(c,d)$.

Por último, mostraremos la parte suficiente del inciso (b), es decir, si $(a,b)=(c,d)$, entonces $a=c$ y $b=d$.

Procediendo por contradicción, supongamos que $a\neq c$ o $b\neq d$. Con lo anterior, consideremos los siguientes casos:

Caso 1. $a\neq c$.

Por hipótesis, tenemos que $(a,b)=(c,d)$, es decir, de acuerdo a la definición de pareja ordenada, $\{\{a\},\{a,b\}\}=\{\{c\},\{c,d\}\}$. En tal caso, el conjunto $\{a\}$ debe estar en $\{\{c\},\{c,d\}\}$. Debido al caso en el que nos encontramos, no es posible que $\{a\}=\{c\}$, por lo que $\{a\}=\{c,d\}$ y ello en particular implicaría que $a=c$, lo cual no es posible, de acuerdo al caso en el que nos encontramos.

Caso 2. $b\neq d$.

Para este caso, podremos suponer que $a=c$, pues en caso contrario, nos remitimos al primer caso para generar una contradicción. Por hipótesis, tenemos que $(a,b)=(c,d)$, es decir, de acuerdo a la definición de pareja ordenada, $\{\{a\},\{a,b\}\}=\{\{c\},\{c,d\}\}$. Con la igualdad anterior, $\{a,b\}$ debe ser un elemento de $\{\{c\},\{c,d\}\}$. Consideremos los siguientes subcasos para $\{a,b\}$:

Subcaso 2.1 $\{a,b\}=\{c\}$.

En este subcaso, podemos afirmar que $a=c$ y que $c=b$, por lo que $a=b$ y con ello, el conjunto $\{\{a\},\{a,b\}\}$ es igual a $\{\{b\}\}$. Como ya habíamos mencionado anteriormente, $\{\{a\},\{a,b\}\}=\{\{c\},\{c,d\}\}$, y de lo recientemente dicho, tenemos que $\{\{b\}\}=\{\{c\},\{c,d\}\}$. En particular, $\{b\}=\{c,d\}$ y de ello, $b=d$, lo cual no es posible de acuerdo al caso en que nos encontramos.

Subcaso 2.1 $\{a, b\} = \{c, d\}$.

En este subcaso tengamos en consideración que por suposición, $b \neq d$, por lo que derivado del hecho $\{a, b\} = \{c, d\}$, debe ser que $b = c$. Además, como habíamos supuesto que $a = c$, entonces obtenemos nuevamente que $a = b$. Como ya habíamos mencionado anteriormente, $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$, y de lo recientemente dicho, tenemos que $\{\{b\}\} = \{\{c\}, \{c, d\}\}$. En particular, $\{b\} = \{c, d\}$ y de ello, $b = d$, lo cual no es posible de acuerdo al caso en que nos encontramos.

Debido a que en todos los casos anteriores obtenemos una contradicción, podemos concluir que $a = c$ y que $b = d$. $\square$

La definición de pareja ordenada antes mostrada fue propuesta por Kuratowski y efectivamente satisface lo deseado, es decir, es una pareja en la que no es lo mismo el orden a y después b, que el orden b y después a.

Ligado al concepto de pareja ordenada, podemos establecer el concepto de **triada ordenada** como sigue: $(a, b, c) = (a, (b, c))$ No es difícil verificar que $(a_1, a_2, a_3) = (b_1, b_2, b_3)$ si y sólo si $a_1 = b_1$, $a_2 = b_2$ y $a_3 = b_3$.

En general, si n es un número natural distinto de cero, una **n-ada ordenada**, denotada por $(a_1, \ldots, a_n)$ se define como: $(a_1, \ldots, a_n) = (a_1, (a_2, \ldots, a_n))$. De lo anterior sólo resaltaremos el hecho de que dos n-adas ordenadas son iguales, es decir, $(a_1, \ldots, a_n) = (b_1, \ldots, b_n)$, si y sólo si para toda $j \in \{1, \ldots, n\}$ se tiene que $a_j = b_j$.

Dados dos conjuntos A y B, el **producto cartesiano de A y B**, denotado por $A \times B$, es el conjunto:

$$A \times B = \{(a, b) : a \in A \text{ y } b \in B\}$$

El producto cartesiano tiene un par de novedades respecto de las operaciones ya vistas. Notemos que los objetos de un producto cartesiano de dos conjuntos son parejas ordenadas. Derivado de ello es preferible escribir cosas como $(x, y) \in A \times B$ en lugar de $x \in A \times B$. Esto lo dejaremos asentado en la siguiente nota:

Nota 2.1.2. *Los objetos de un producto cartesiano $A \times B$ son parejas ordenadas, por lo que es preferible usar enunciados de la forma $(x, y) \in A \times B$ en lugar de enunciados como $x \in A \times B$*

Lo siguiente que hay que recalcar es cuándo una pareja ordenada pertenece o no a un producto cartesiano.

Nota 2.1.3. *Sean A y B conjuntos y a y b dos objetos. $(a, b) \in A \times B$ si y sólo si $a \in A$ y $b \in B$.*

Con esto podemos deducir cuándo una pareja ordenada no pertenece a un producto cartesiano.

Nota 2.1.4. *Sean A y B conjuntos y a y b dos objetos.* $(a,b) \notin A \times B$ *si y sólo si* $a \notin A$ *o* $b \notin B$.

El siguiente es un resultado son propiedades básicas que cumple el producto cartesiano de conjuntos.

Lema 2.1.5. *Sean A, B, C y D conjuntos. Los siguientes enunciados se satisfacen:*

a) *Si A, B, C y D son no vacíos, entonces* $A \times B = C \times D$ *si y sólo si* $A = C$ *y* $B = D$.

b) $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

c) $(A \cup B) \times C = (A \times C) \cup (B \times C)$.

d) $A \times (B \cap C) = (A \times B) \cap (A \times C)$

e) $(A \cap B) \times C = (A \times C) \cap (B \times C)$

f) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$.

g) $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$

Demostración. b) Para mostrar que $A \times (B \cup C) = (A \times B) \cup (A \times C)$, procederemos por doble contención.

Primero probaremos que $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$. Sea $(x,z) \in A \times (B \cup C)$. Por definición de producto cartesiano, $x \in A$ y $z \in B \cup C$. De esto último, y por definición de unión, se sigue que $z \in B$ o $z \in C$. Procedemos por casos.

Caso 1. $z \in B$.

En este caso, como sabemos que $x \in A$ y $z \in B$, entonces por definición de pareja ordenada, $(x,z) \in A \times B$, y por definición de unión, $(x,z) \in (A \times B) \cup (A \times C)$.

Caso 2. $z \in C$.

En este caso, como sabemos que $x \in A$ y $z \in C$, entonces por definición de pareja ordenada, $(x,z) \in A \times C$, y por definición de unión, $(x,z) \in (A \times B) \cup (A \times C)$.

De los casos anteriores, podemos concluir que $(x,z) \in (A \times B) \cup (A \times C)$, concluyendo la demostración de la primera contención.

Ahora mostraremos la segunda contención, es decir, $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$. Sea $(x,z) \in (A \times B) \cup (A \times C)$. Por definición de unión, $(x,z) \in A \times B$ o $(x,z) \in A \times C$. Ahora procederemos por casos.

Caso 1. $(x,z) \in A \times B$.

En este caso, por definición de producto, se sigue que $x \in A$ y $z \in B$. Gracias a la definición de unión, podemos afirmar que $z \in B \cup C$. Ya que $x \in A$ y $z \in B \cup C$, entonces por definición de producto cartesiano, $(x,z) \in A \times (B \cup C)$.

Caso 2. $(x,z) \in A \times C$.

En este caso, por definición de producto, se sigue que $x \in A$ y $z \in C$. Gracias a la definición de unión, podemos afirmar que $z \in B \cup C$. Ya que $x \in A$ y $z \in B \cup C$, entonces por definición de producto cartesiano, $(x, z) \in A \times (B \cup C)$.

De los casos anteriores concluimos que $(x, z) \in A \times (B \cup C)$, mostrando así la segunda contención. $\square$

2.2. Relaciones

Uno de los grandes beneficios derivados del producto cartesiano son las relaciones. Éstas guardan gran importancia en todo lo que se refiere a Matemáticas elementales debido a su uso intuitivo en conjuntos: la interacción entre objetos. En este curso veremos 3 tipos de relaciones que son de interés: las relaciones de orden, las relaciones de equivalencia y las funciones.

Sean A y B conjuntos. Una **relación de A en B** es cualquier subconjunto del producto cartesiano $A \times B$. Por ejemplo, si $A = \{a, b, c, d\}$ y $B = \{1, 2, 3, 4\}$, las siguientes son relaciones de A en B.

$$R_1 = \emptyset \qquad R_2 = \{(a, 1), (b, 2), (c, 3), (d, 4)\} \qquad R_3 = \{(b, 1), (b, 2), (d, 1), (d, 2)\}$$

Lo primero que hay que recordar es que, dado que una relación es un subconjunto del producto cartesiano, entonces sus elementos son parejas ordenadas.

Nota 2.2.1. *Sean A y B conjuntos y R una relación de A en B. Dado que $R \subseteq A \times B$, es preferible usar enunciados de la forma $(x, y) \in R$ en lugar de enunciados de la forma $x \in R$, siempre teniendo en cuenta que R es una relación.*

Otra situación importante que hay que rescatar es que, si R es una relación de A en B, al conjunto B se le suele llamar **codominio de R**, aunque cabe señalar que el concepto de codominio es más frecuente al usar funciones que relaciones en general.

Si R es una relación de A en B y a es un elemento de A y b es un elemento de B, diremos que **a está relacionado con b bajo R** si $(a, b) \in R$. Esto es importante porque las relaciones se pueden interpretar como una *interacción* de los objetos dentro de un conjunto o interacciones entre conjuntos distintos y justamente es importante destacar quiénes están interactuando entre ellos.

Una relación arbitraria R de A en B se puede representar mediante un dibujo en el plano de la siguiente manera: Los elementos de $A \cup B$ se representarán en el plano mediante puntos, y si la pareja (a, b) forma parte de R, entonces uniremos con un segmento dirigido al punto asociado a a hacia el punto asociado a b. Al dibujo obtenido le llamaremos **gráfica dirigida de la relación** o simplemente **digráfica de la relación**. Por ejemplo, retomando la relación R_3 mostrada previamente, un dibujo de dicha relación es la siguiente:

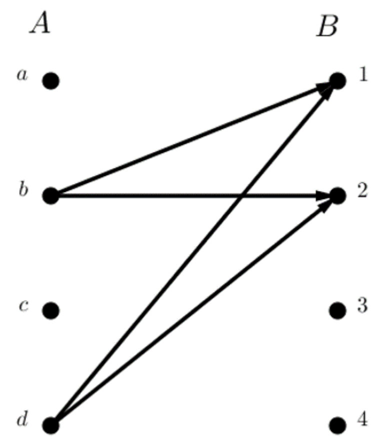

Figura 2.1: Representación de la relación $R_3 = \{(b,1),(b,2),(d,1),(d,2)\}$

También existe una forma alterna para la notación de las relaciones. Ya sabemos que dada una relación de A en B, digamos R, y dos objetos $a \in A$ y $b \in B$, se tiene que la frase a *está relacionado con* b significa simplemente que $(a,b) \in R$, pero ¿qué pasaría si a y b son también parejas ordenadas? por ejemplo, si quisiéramos decir: (x_1,x_2) está relacionada con (y_1,y_2) bajo R, entonces esto se interpretaría como

$$((x_1,x_2),(y_1,y_2)) \in R$$

Lo cual ya empieza a ser un poco engorroso en cuanto a paréntesis. En estos casos es mejor utilizar una notación más sencilla: Dada una relación de A en B, digamos R, y $a \in A$ y $b \in B$, se puede denotar el hecho que a esté relacionado con b bajo R simplemente por:

$$aRb$$

Notemos que bajo esta situación tenemos que:

Nota 2.2.2. *Si R es una relación de A en B y $a \in A$ y $b \in B$, entonces aRb si y sólo si $(a,b) \in R$.*

Así, cuando queremos decir que dos parejas ordenadas están relacionadas bajo una relación R, es más sencillo decir que $(x_1,x_2)R(y_1,y_2)$ en lugar de $((x_1,x_2),(y_1,y_2)) \in R$.

Recordemos que, de acuerdo a la definición de relación, toda relación es un conjunto y por ello es posible escribirlo por comprensión. Esto último puede facilitar la escritura en gran medida porque en lo que refiere a relaciones de orden y funciones muchas veces es muy importante decir *cómo* es que se relacionan los objetos sin necesidad de dar un listado de las parejas ordenadas.

Por otro lado, dado que los objetos en una relación son parejas ordenadas, la escritura por comprensión de dicha relación depende de dos objetos, es decir, si R es una relación de A en B, entonces R se puede escribir como $R = \{(a,b) \in A \times B : P(a,b)\}$. En tal caso, un objeto está en R si y sólo si es un elemento de $A \times B$ y satisface la propiedad $P(a,b)$, aunque ésta última propiedad será la que usaremos más frecuentemente, es decir dada $(x,y) \in A \times B$, $(x,y) \in R$ si y sólo si $P(x,y)$, pero recordemos que es lo mismo escribir $(x,y) \in R$ que escribir xRy, entonces simplemente tendríamos que: xRy si y sólo si $P(x,y)$. Con lo anterior, es frecuente que en cuestiones de relaciones tengamos enunciados del siguiente estilo:

Sean A y B conjuntos, consideremos la relación R de A en B definida por: xRy si y sólo si $P(x,y)$.

El anterior es un enunciado típico de relaciones y por ello es importante que podamos interpretarlo. El enunciado presupone que vamos a trabajar una relación llamada R, no hay que perder de vista que R es un símbolo genérico que puede variar y ser muy extraño: $\trianglelefteq$, $\leq$ $\preceq$, $\ll$, φ, Ω, etc., por lo que no se espanten en cuanto vean este tipo de símbolos, recuerden que éstos son sólo una forma de llamar a los conjuntos. La parte a la que hay que prestar atención es la bicondicional, y en particular, a la propiedad $P(x,y)$, pues ésta es un enunciado lógico que se debe satisfacerse para que los objetos pertenezcan a la relación. En este sentido, si dos objetos satisfacen la propiedad P, entonces estarán relacionados bajo R y viceversa: si dos objetos están relacionados bajo R, entonces deben satisfacer la propiedad P.

Aunque parezca una notación rebuscada, en realidad es una notación que se presta a mucha versatilidad, por lo que la debemos tener muy en cuenta cuando trabajemos las siguientes secciones.

Dados dos conjuntos A y B y R una relación de A en B, la **relación inversa de $\boldsymbol{R}$**, denotada por R^{-1}, es la relación de B en A dada por:

$$R^{-1} = \{(b,a) \in B \times A : (a,b) \in R\}$$

Esta relación casi no la trabajaremos, salvo en la sección de funciones, pero es importante dejar como nota qué significa que algo pertenezca a dicha relación.

Nota 2.2.3. *Sean A y B conjuntos, R una relación de A en B y (x,z) una pareja ordenada arbitraria. $(x,z) \in R^{-1}$ si y sólo si $(z,x) \in R$.*

Del mismo modo, es importante indicar qué significa que una pareja ordenada no pertenezca a dicha relación.

Nota 2.2.4. *Sean A y B conjuntos, R una relación de A en B y (x,z) una pareja ordenada arbitraria. $(x,z) \notin R^{-1}$ si y sólo si $(z,x) \notin R$.*

2.2.1. Conjuntos asociados a las relaciones

Existen algunos conjuntos asociados a las relaciones y que serán de importancia en lo consecuente. Dada una relación de A en B, digamos R, definimos el **dominio de $\boldsymbol{R}$**, denotado por $Dom(R)$, como el conjunto:

$$Dom(R) = \{a \in A : \text{ existe } b \in B \text{ tal que } (a, b) \in R\}$$

Hay dos cosas importantes que recalcar sobre el dominio de una relación, la primera de ellas es que el dominio es el conjunto de todos los elementos de A que tienen una pareja bajo R, por lo que el dominio de la relación es un subconjunto de A. En tal caso, es más frecuente que los elementos del dominio sean objetos y no necesariamente parejas ordenadas. Lo siguiente que hay que recalcar es que el dominio no necesariamente es igual a todo el conjunto A, por ejemplo, retomando la relación R_3 descrita anteriormente, $Dom(R_3) = \{b, d\}$. Además de lo anterior, tenemos la siguiente notas que nos ayudará a determinar cuándo un objeto está en el dominio de una función.

Nota 2.2.5. *Sean A y B conjuntos, R una relación de A en B y x un objeto arbitrario. $x \in Dom(R)$ si y sólo si existe $b \in B$ tal que $(x, b) \in R$.*

Y de igual forma, podemos determinar cuándo un objeto no pertenece al dominio de una relación.

Nota 2.2.6. *Sean A y B conjuntos, R una relación de A en B y x un objeto arbitrario. $x \notin Dom(R)$ si y sólo si para todo $b \in B$, $(x, b) \notin R$.*

El segundo conjunto importante que consideraremos es la imagen de la relación. Para ello, consideremos una relación R de A en B. Si bien el dominio es el conjunto de todos los elementos de A que tienen una pareja bajo la relación R, la imagen de la relación será su análoga respecto del conjunto B, es decir, la imagen de R será el conjunto de todos aquellos objetos en B que son pareja de alguien. De manera formal, dada una relación de A en B, digamos R, la **imagen de $\boldsymbol{R}$**, denotada por $Im(R)$, es el conjunto:

$$Im(R) = \{b \in B : \text{ existe } a \in A \text{ tal que } (a, b) \in R\}$$

Es importante recalcar que al igual que en el caso del dominio, no necesariamente se satisface que $Im(R)$ sea igual a todo el conjunto B, por ejemplo, considerando la relación R_3 mostrada previamente se tiene que $Im(R_3) = \{1, 2\}$. Ahora tendremos un par de notas que nos ayudarán a determinar cuándo un objeto está en la imagen de una relación.

Nota 2.2.7. *Sean A y B conjuntos, R una relación de A en B y z un objeto arbitrario. $z \in Im(R)$ si y sólo si existe $a \in A$ tal que $(a, z) \in R$.*

Y de igual forma, podemos determinar cuándo un objeto no pertenece al dominio de una relación.

Nota 2.2.8. *Sean A y B conjuntos, R una relación de A en B y z un objeto arbitrario. $z \notin Im(R)$ si y sólo si para todo $a \in A$, $(a, z) \notin R$.*

Antes de revisar algunas propiedades generales de dominio e imagen de relaciones, veremos un ejemplo concreto de estos conceptos.

Nota 2.2.9. *Sea A el conjunto de números naturales pares y R la relación de A en $\mathbb{Z}$ dada por $(n, m) \in R$ si y sólo si m es múltiplo de n. Demuestra lo siguiente.*

a) $2 \in Dom(R)$.

b) $-12 \in Im(R)$.

c) $3 \notin Dom(R)$.

d) $-11 \notin Im(R)$.

Demostración. Ahora mostraremos los cuatro incisos anteriores.

a) $2 \in Dom(R)$.

Para este inciso basta demostrar existe $x \in \mathbb{Z}$ tal que $(0, x) \in R$. Para ello, notemos que $4 \in \mathbb{Z}$ y 4 es múltiplo de 2, luego, por la definición de la relación R se tiene que $(2, 4) \in R$. Así, como existe $4 \in \mathbb{Z}$ tal que $(2, 4) \in R$, se concluye de la definición de dominio que $2 \in Dom(R)$.

b) $-12 \in Im(R)$.

Para este inciso basta demostrar que existe $x \in A$ tal que $(x, -12) \in R$. Para ello, notemos que $6 \in A$ y -12 es múltiplo de 6, luego, por la definición de la relación R se tiene que $(6, -12) \in R$. Así, como existe $6 \in A$ tal que $(6, -12) \in R$, se concluye de la definición de imagen que $-12 \in Im(R)$.

c) $3 \notin Dom(R)$.

Para esta demostración veremos que para todo $b \in \mathbb{Z}$, $(3, b) \notin R$. Sea $b \in \mathbb{Z}$. Notemos que como $3 \notin A$ y R es una relación de A en $\mathbb{Z}$, entonces $(3, b) \notin R$.

d) $11 \notin Im(R)$.

Para este inciso mostraremos que para todo $a \in A$, $(a, 11) \notin R$. Sea $a \in A$. Por la definición de A tenemos que a es par, en cuyo caso, 11 no es múltiplo de a y con ello, $(a, 11) \notin R$.

□

Ahora procederemos a demostrar algunas propiedades generales respecto al concepto de dominio e imagen.

Lema 2.2.10. *Sean A y B conjuntos y R y S dos relaciones de A en B. Los siguientes enunciados se satisfacen:*

a) Si $R \subseteq S$, entonces $Dom(R) \subseteq Dom(S)$.

b) $Dom(R \cup S) = Dom(R) \cup Dom(S)$.

c) $Dom(R \cap S) \subseteq Dom(R) \cap Dom(S)$.

d) Si $R \subseteq S$, entonces $Im(R) \subseteq Im(S)$.

e) $Im(R \cup S) = Im(R) \cup Im(S)$.

f) $Im(R \cap S) \subseteq Im(R) \cap Im(S)$.

Demostración. (a) Sea $x \in Dom(R)$. Por definición de dominio, existe $t \in B$ tal que $(x,t) \in R$. Por hipótesis sabemos que $R \subseteq S$, por lo que $(x,t) \in S$. Así, existe $t \in B$ tal que $(x,t) \in S$, concluyendo de la definición de dominio que $x \in Dom(S)$.

(c) Este inciso pueden consultarlo en el video `https://youtu.be/a5_BH2yP9u0`

(d) Este inciso pueden consultarlo en el video `https://youtu.be/BM2q70QxFwo`

(f) Sea $w \in Im(R \cap S)$. Por definición de imagen existe $a \in A$ tal que $(a,w) \in R \cap S$. Gracias a la definición de intersección, se tiene que $(a,w) \in R$ y $(a,w) \in S$. Así, dado que existe $a \in A$ tal que $(a,w) \in R$, se sigue por definición de imagen que $w \in Im(R)$. De manera análoga, como existe $a \in A$ tal que $(a,w) \in S$, entonces por definición de imagen, $w \in Im(S)$. Por la definición de intersección, se puede concluir que $w \in Im(R) \cap Im(S)$. $\square$

2.3. Relaciones de equivalencia

Las relaciones de equivalencia son relaciones que tratan de recuperar el comportamiento de la igualdad entre objetos. La relación en sí da criterios para establecer cuándo dos objetos de un conjunto son lo *suficientemente parecido entre ellos*. Dado un conjunto A y $\sim$ una relación sobre A, diremos que $\sim$ **es una relación de equivalencia sobre A** si $\sim$ satisface las siguientes condiciones:

a) Para todo $x \in A$ se satisface que $x \sim x$

b) Para todo $\{x,z\} \subseteq A$ tales que $x \sim z$, se satisface que $z \sim x$.

c) Para todo $\{x,z,w\}$ tales que $x \sim z$ y $z \sim w$, se satisface que $x \sim w$.

Para que una relación sea una relación de equivalencia, es necesario que satisfaga las tres condiciones anteriores y viceversa.

Ahora procederemos a demostrar un par de ejemplos sobre relaciones de equivalencia. El primero de ellos es un ejemplo con un conjunto y objetos concretos y el segundo es un ejemplo más abstracto.

Ejemplo 2.3.1. *Sea $\mathbb{N}$ el conjunto de los números naturales. Definimos la relación $\equiv$ en $\mathbb{N} \times \mathbb{N}$ dada por $(x, z) \equiv (a, b)$ si y sólo si $x = a$. Muestra que la relación $\equiv$ es una relación de equivalencia sobre $\mathbb{N} \times \mathbb{N}$.*

Demostración. Para demostrar que $\equiv$ es una relación de equivalencia sobre $\mathbb{N}\times\mathbb{N}$, mostraremos que $\equiv$ satisface las tres condiciones de la definición de relación de equivalencia.

a) Para todo $(a, b) \in \mathbb{N} \times \mathbb{N}$, se satisface que $(a, b) \equiv (a, b)$.

   Sea $(a, b) \in \mathbb{N} \times \mathbb{N}$. Como $a = a$, se sigue de la definición de la relación $\equiv$ el que $(a, b) \equiv (a, b)$.

b) Para todo $\{(x, z), (a, b)\} \subseteq \mathbb{N} \times \mathbb{N}$ tales que $(x, z) \equiv (a, b)$, se satisface que $(a, b) \equiv (x, z)$.

   Sea $\{(x, z), (a, b)\} \subseteq \mathbb{N} \times \mathbb{N}$ tales que $(x, z) \equiv (a, b)$. Por la definición de la relación $\equiv$, se sigue que $x = a$, o lo que es lo mismo, $a = x$, por lo que, nuevamente por la definición de la relación $\equiv$ podemos concluir que $(a, b) \equiv (x, z)$.

c) Para todo $\{(a, b), (x, z), (r, s)\} \subseteq \mathbb{N} \times \mathbb{N}$ tales que $(a, b) \equiv (x, z)$ y $(x, z) \equiv (r, s)$, se satisface que $(a, b) \equiv (r, s)$.

   Sea $\{(a, b), (x, z), (r, s)\} \subseteq \mathbb{N} \times \mathbb{N}$ tales que $(a, b) \equiv (x, z)$ y $(x, z) \equiv (r, s)$. Como $(a, b) \equiv (x, z)$, entonces por definición de la relación $\equiv$ se tiene que $a = x$. Por otro lado, como $(x, z) \equiv (r, s)$, se sigue de la definición de $\equiv$ el que $x = r$. Ya que $a = x$ y $x = r$, concluimos que $a = r$, y nuevamente por la definición de la relación $\equiv$ se concluye que $(a, b) \equiv (r, s)$.

$\square$

Ejemplo 2.3.2. *Sean A un conjunto no vacío y $\cong$ una relación de equivalencia en A. Definimos la relación $\sim$ en A dada por: $x \sim z$ si y sólo si $z \cong x$. Muestra que la relación $\sim$ es una relación de equivalencia sobre A.*

Demostración. Para esta demostración, primero analizaremos lo que queremos demostrar y lo que sí sabemos de este enunciado en este video `https://youtu.be/-_5OoBEdWWs`. Ahora tendremos que validar que la relación $\sim$ satisface las tres condiciones de relación de equivalencia.

a) Para todo $x \in A$, $x \sim x$

   La demostración de este inciso pueden verla en `https://youtu.be/QRcPSYHV4RM`

b) Para todo $\{x, z\} \subseteq A$ tales que $x \sim z$, se satisface que $z \sim x$.

   La demostración de este inciso pueden verla en `https://youtu.be/Y1I3hbuVVK0`

c) Para todo $\{x, z, w\} \subseteq A$ tales que $x \sim z$ y $z \sim w$ se satisface que $x \sim w$.

La demostración de este inciso pueden verla en https://youtu.be/6ezgImIP3pU

□

Dado un conjunto A, $\sim$ una relación de equivalencia sobre A y $a \in A$, la **clase de equivalencia de a bajo la relación $\sim$**, denotada por $[a]_\sim$, es el conjunto: $[a]_\sim = \{x \in A : x \sim a\}$. Al objeto a se le conoce como **representante de la clase** y es importante mencionar qué significa que un objeto pertenezca a una clase de equivalencia.

Nota 2.3.3. *Sean A un conjunto, $\sim$ una relación de equivalencia en A y $z \in A$. si w es un objeto en A, entonces: $w \in [a]_\sim$ si y sólo si $w \sim a$.*

Antes de continuar, primero mostraremos con un ejemplo concreto el comportamiento de las clases de equivalencia.

Ejemplo 2.3.4. *Sea $\mathbb{N}$ el conjunto de los números naturales. Definimos la relación de equivalencia $\equiv$ en $\mathbb{N} \times \mathbb{N}$ dada por $(x, z) \equiv (a, b)$ si y sólo si $x = a$. Muestra que $[(1, 2)]_\sim = \{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\}$*

Demostración. Para realizar esta demostración procederemos por doble contención y primero mostraremos que $[(1, 2)]_\sim \subseteq \{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\}$.

Sea $(a, b) \in [(1, 2)]_\sim$. Por definición de clase de equivalencia se tiene que $(a, b) \sim (1, 2)$ y, de la definición de la relación $\sim$ se tiene que $a = 1$. Por lo anterior, como (a, b) satisface las condiciones de pertenencia al conjunto $\{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\}$, se concluye que $(a, b) \in \{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\}$.

Ahora mostraremos la otra contención, es decir, $\{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\} \subseteq [(1, 2)]_\sim$. Sea $(u, v) \in \{(x, z) \in \mathbb{N} \times \mathbb{N} \colon x = 1\}$, por definición de pertenencia a un conjunto por comprensión se tiene que $(u, v) \in \mathbb{N} \times \mathbb{N}$ y $u = 1$. Luego, por definición de la relación $\sim$ se tiene que $(u, v) \sim (1, 2)$, en cuyo caso, por definición de clase de equivalencia se tiene que $(u, v) \in [(1, 2)]_\sim$. □

Ahora veremos propiedades generales de las clases de equivalencia,

Lema 2.3.5. *Sean A un conjunto y $\sim$ una relación de equivalencia sobre A.*

(a) Si $x \in A$, entonces $x \in [x]_\sim$

(b) $[x]_\sim = [z]_\sim$ si y sólo si $x \sim z$.

Demostración. (a) Sea $x \in A$. Como $\sim$ es una relación de equivalencia sobre A, entonces $x \sim x$, siguiéndose de ello el que $x \in [x]_\sim$.

(b) Primero mostraremos la parte suficiente, es decir, si $[x]_\sim = [z]_\sim$, entonces $x \sim z$.

Para ello, notemos que $x \in [x]_\sim$, pero como $[x]_\sim = [z]_\sim$, entonces $x \in [z]_\sim$, concluyendo que $x \sim z$.

Ahora mostraremos la parte necesaria, es decir, si $x \sim z$, entonces $[x]_\sim = [z]_\sim$. Procederemos por doble contención y primero mostraremos que $[x]_\sim \subseteq [z]_\sim$.

Sea $t \in [x]_\sim$. Por definición de clase de equivalencia, se tiene que $t \sim x$. Ya que $x \sim z$, entonces por la transitividad de $\sim$ podemos asegurar que $t \sim z$, y nuevamente por definición de clase de equivalencia, $t \in [z]_\sim$, concluyendo la contención deseada. La segunda contención es análoga. □

Aunque en este curso no se usará la siguiente notación, es importante mencionar que si $\sim$ es una relación de equivalencia sobre un conjunto A, entonces a la familia de las clases de equivalencia de $\sim$ se denota por $A/\sim$, es decir, $A/\sim = \{[a]_\sim : a \in A\}$. Con la notación anterior se puede establecer un teorema que es fundamental para muchas situaciones matemáticas y que muestra que hay una relación matemática entre las relaciones de equivalencia y las particiones.

Teorema 2.3.6. *Sea A un conjunto y $\sim$ una relación sobre A. Si $\sim$ es una relación de equivalencia sobre A, entonces el conjunto $A/\sim$ es una partición de A.*

Recíprocamente, si A es un conjunto y $\mathcal{Q}$ es una partición de A, entonces existe una relación de equivalencia sobre A, digamos $\sim$, tal que $\mathcal{Q} = A/\sim$.

Demostración. Primero mostraremos que si $\sim$ es una relación de equivalencia sobre A, entonces el conjunto $A/\sim$ es una partición de A, para ello, verificaremos que $A/\sim$ satisface la definición de relación de equivalencia.

a) Para toda $x \in A$, $[x]_\sim \neq \emptyset$.

   Sea $x \in A$, por el lema 2.3.5 (a), se tiene que $x \in [x]_\sim$, por lo que $[x]_\sim \neq \emptyset$.

b) Para toda $\{x, z\} \subseteq A$ tal que $[x]_\sim \cap [z]_\sim \neq \emptyset$, se satisface que $[x]_\sim = [z]_\sim$.

   Sean $\{x, z\} \subseteq A$ tales que $[x]_\sim \cap [z]_\sim \neq \emptyset$. Consideremos $b \in [x]_\sim \cap [z]_\sim$, por definición de intersección, $b \in [x]_\sim$ y $b \in [z]_\sim$, siguiéndose de la definición de clase de equivalencia el que $x \sim b$ y $b \sim z$. Ya que $\sim$ es una relación transitiva, entonces $x \sim z$, y por el lema 2.3.5 (b), podemos concluir que $[x]_\sim = [z]_\sim$.

c) $\bigcup_{x \in A} [x]_\sim = A$.

   Para este inciso, dado que para todo $z \in A$, $[z]_\sim \subseteq A$, podemos garantizar por el teorema 1.6.5 (c) que $\bigcup_{x \in A} [x]_\sim \subseteq A$, por lo que resta demostrar la segunda contención.

Sea $w \in A$ arbitrario. Como $w \in [w]_\sim$, entonces podemos decir que existe $w \in A$ tal que $w \in [w]_\sim$, siguiéndose de la definición de unión el que $w \in \bigcup_{x\in A}[x]_\sim$, concluyendo la segunda contención.

Por lo anterior, $A/\sim$ es una partición de A.

Ahora procederemos a mostrar que si $\mathcal{Q} = \{B_i : i \in I\}$ es una partición de A, entonces existe una relación de equivalencia sobre A, digamos $\approx$, tal que $\mathcal{Q} = A/\approx$.

Para ello, definiremos la relación $\approx$ en A dada por $x \approx z$ si y sólo si existe $k \in I$ tal que $x \in B_k$ y $w \in B_k$.

Ahora mostraremos que la relación $\approx$ es una relación de equivalencia sobre A

(i) Para todo $x \in A$, $x \approx x$.

Sea $x \in A$. Como $\{B_i : i \in I\}$ es una partición de A, entonces $x \in \cup_{i\in I}B_i$. En particular, existe $t \in I$ tal que $x \in B_i$. Por la definición de $\approx$ se concluye que $x \approx x$.

(ii) Para todo $x, z \in A$ tal que $x \approx z$, se satisface que $z \approx x$.

Sean $x, z \in A$ tal que $x \approx z$. Por definición de la relación $\approx$ se tiene que existe $m \in I$ tal que $x, z \in B_m$, en cuyo caso, $z, x \in B_m$ y de acuerdo a la definición de $\approx$ se concluye que $z \approx x$.

(iii) Para todo $x, z, w \in A$ tal que $x \approx z$ y $z \approx w$ se satisface que $x \approx w$.

Sean $x, z, w \in A$ tal que $x \approx z$ y $z \approx w$. Dado que $x \approx z$, se sigue por definición de la relación se tiene que existe $r \in I$ tal que $x, z \in B_r$. Del mismo modo, como $z \approx w$ se tiene que existe $s \in I$ tal que $z, w \in B_s$. Notemos que $z \in B_r \cap B_s$ y por ser $\{B_i : i \in I\}$ una partición, se concluye que $B_r = B_s$. Así, existe $r \in I$ tal que $x, z \in B_r$, siguiéndose de la definición de la relación que $x \approx z$.

Ahora que hemos visto que la relación $\approx$ es de equivalencia, sólo resta mostrar que la igualdad $\{B_i : i \in I\} = \{[x]_\approx : x \in A\}$ es verdadera.

Para simplificar la demostración, tenemos la siguiente afirmación:

Afirmación. Si $x \in B_k$, entonces $B_k = [x]_\approx$.

Primero mostraremos que $B_k \subseteq [x]_\approx$. Sea $w \in B_k$. Como $w, x \in B_k$, por definición de la relación $\approx$ se concluiría que $w \approx x$, luego, por definición de clase de equivalencia se concluye que $w \in [x]_\approx$.

Ahora mostraremos que $[x]_\approx \subseteq B_k$. Sea $t \in [x]_\approx$. Por definición de clase de equivalencia se tiene que $t \approx x$ y por definición de la relación se seguiría que existe $s \in I$ tal que $t, x \in B_s$. Notemos que $x \in B_k \cap B_s$, siguiéndose del hecho de que $\{B_i : i \in I\}$ es una partición el que $B_k = B_s$, concluyendo que $t \in B_k$.

De las contenciones previas se demuestra que $B_k = [x]_\approx$, concluyendo la demostración del lema.

Gracias a la afirmación anterior, podremos ver de manera más sencilla que $\{B_i : i \in I\} = \{[x]_\approx : x \in A\}$. Procederemos por doble contención y primero veremos que $\{B_i : i \in I\} \subseteq \{[x]_\approx : x \in A\}$.

Sea $B_j \in \{B_i : i \in I\}$. Dado que la familia $\{B_i : i \in I\}$ es una partición de A, en particular se tiene que $B_j \neq \emptyset$, por lo que consideramos $b \in B_j$. Por la afirmación anterior, $B_j = [b]_\approx$, concluyendo que $B_j \in \{[x]_\approx : x \in A\}$, mostrando así la primera contención.

Ahora procederemos a mostrar la segunda contención, es decir, $\{[x]_\approx : x \in A\} \subseteq \{B_i : i \in I\}$.

Sea $[b]_\approx \in \{[x]_\approx : x \in A\}$. Dado que b es un elemento de A y $\{B_i : i \in I\}$ es una partición de A, podemos garantizar que $b \in \bigcup_{i\in I} B_i$, por lo que por definición de unión, existe $k \in I$ tal que $b \in B_k$. Así, por la afirmación anterior, $B_k = [b]_\approx$, concluyendo que $[b]_\approx \in \{B_i : i \in I\}$, mostrando la segunda contención. $\square$

2.4. Relaciones de orden

Las relaciones que trabajaremos en esta sección pretenden introducir y formalizar en términos matemáticos una *jerarquía* dentro de los objetos de un conjunto, es decir, queremos determinar, dado un conjunto A y dos objetos de A, digamos a y b, cuándo es *más importante* el objeto a que el objeto b. A este tipo de relaciones les llamaremos de manera genérica relaciones de orden. Para llegar a ellas, es necesario tener en cuenta los siguientes conceptos

Dada un conjunto A y una relación $\trianglelefteq$ sobre A, decimos que $\trianglelefteq$ **es un orden parcial sobre A** si $\trianglelefteq$ satisface las siguientes propiedades:

a) Para todo $x \in A$, $x \trianglelefteq x$.

b) Para todo $\{x, z\} \subseteq A$ tales que $x \trianglelefteq z$ y $z \trianglelefteq x$, se satisface que $x = z$.

c) Para todo $\{x, z, w\} \subseteq A$ tales que $x \trianglelefteq z$ y $z \trianglelefteq w$, se satisface que $x \trianglelefteq w$.

Hay que tener en cuenta que algunos libros consideran una definición distinta de orden parcial: algunos autores consideran que un orden parcial es una relación antirreflexiva[1] y transitiva. Dicho orden es mejor conocido como un **orden parcial estricto** y nosotros no lo trabajaremos en estas notas.

Es importante recalcar que el símbolo más usado para los ordenes parciales suele ser el símbolo $\leq$, aunque no se extrañen si se usan otros símbolos como $\lll$, $\trianglelefteq$, $\preceq$, $\sqsubseteq$, etc.

El siguiente ejemplo es una relación de orden parcial.

Ejemplo 2.4.1. *Sea $\mathbb{Z}$ el conjunto de los números enteros. Definimos la relación $\preceq$ en $\mathbb{Z} \times \mathbb{Z}$ dada por: $(x, z) \preceq (a, b)$ si y sólo si $x \leq a$ y $z \leq b$. Muestra que la relación $\preceq$ es un orden parcial sobre $\mathbb{Z} \times \mathbb{Z}$.*

[1] Una relación R es antirreflexiva si para todo $x \in A$, $(x, x) \notin A$

Demostración. Para esta demostración, primero analizaremos el enunciado previo a realizar su demostración en https://youtu.be/cKnby0uogAY

a) Para todo $(x,z) \in \mathbb{Z} \times \mathbb{Z}$, $(x,z) \preceq (x,z)$.

La demostración de este inciso pueden consultarla en https://youtu.be/caWTjvpe_5c

b) Para todo $\{(x,z),(a,b)\} \subseteq \mathbb{Z} \times \mathbb{Z}$ tales que $(x,z) \preceq (a,b)$ y $(a,b) \preceq (x,z)$, se satisface que $(x,z) = (a,b)$.

La demostración de este inciso pueden consultarla en https://youtu.be/o4CTa0iYSCk

c) Para todo $\{(x,z),(a,b),(r,s)\} \subseteq \mathbb{Z} \times \mathbb{Z}$ tales que $(x,z) \preceq (a,b)$ y $(a,b) \preceq (r,s)$, se satisface que $(x,z) \preceq (r,s)$.

La demostración de este inciso pueden consultarla en https://youtu.be/8801-bcb9eQ

□

Ahora procederemos a demostrar el siguiente resultado, el cual es de gran importancia porque establece que cualquier familia de conjuntos se puede ordenar mediante la contención.

Teorema 2.4.2. *Sea $\mathscr{A}$ una familia de conjuntos no vacía. La relación $\trianglelefteq$ en $\mathscr{A}$ definida por $C \trianglelefteq D$ si y sólo si $C \subseteq D$ es un orden parcial en $\mathscr{A}$.*

Demostración. Para verificar que $\trianglelefteq$ es un orden parcial en $\mathscr{A}$, mostraremos que dicha relación satisface la definición de orden parcial.

a) Para todo $B \in \mathscr{A}$, $B \trianglelefteq B$.

Sea $B \in \mathscr{A}$, sabemos que por ser B un conjunto, entonces $B \subseteq B$, por lo que por definición de la relación $\trianglelefteq$, se tiene que $B \trianglelefteq B$.

b) Para todo $\{A,B\} \subseteq \mathscr{A}$ tales que $A \trianglelefteq B$ y $B \trianglelefteq A$, se satisface que $A = B$.

Sea $\{A,B\} \subseteq \mathscr{A}$ tales que $A \trianglelefteq B$ y $B \trianglelefteq A$. Como $A \trianglelefteq B$, entonces por definición de $\trianglelefteq$ se sigue que $A \subseteq B$. Análogamente, como $B \trianglelefteq A$, por la definición de $\trianglelefteq$ se sigue que $B \subseteq A$. Gracias a que $A \subseteq B$ y $B \subseteq A$, podemos concluir que $A = B$.

c) Para todo $\{A,B,C\} \subseteq \mathscr{A}$ tales que $A \trianglelefteq B$ y $B \trianglelefteq C$, se satisface que $A \trianglelefteq C$.

Sea $\{A,B,C\} \subseteq \mathscr{A}$ tales que $A \trianglelefteq B$ y $B \trianglelefteq C$. Como $A \trianglelefteq B$, entonces por definición de $\trianglelefteq$ se sigue que $A \subseteq B$. Análogamente, como $B \trianglelefteq C$, por la definición de $\trianglelefteq$ se sigue que $B \subseteq C$. Gracias a que $A \subseteq B$ y $B \subseteq C$, podemos concluir que $A \subseteq C$, lo cual, nuevamente por definición de $\trianglelefteq$, concluye que $A \trianglelefteq C$.

□

La idea del orden parcial es poder establecer una jerarquía entre los objetos de un conjunto, pero ello no implica que, dados dos objetos, siempre nos sea posible determinar cuál es más importante que el otro, por ejemplo, si tomamos los conjuntos $A = \{1, 2\}$ y $B = \{a, 1, 2, c\}$, entonces bajo el orden descrito en el lema 2.4.2, se tiene que $A \preceq B$, es decir, B es mayor que A. Sin embargo, si ahora tomamos nuevamente al conjunto A y al conjunto $C = \{3\}$, no es posible determinar bajo dicho orden quién es mayor que quién, pues ambos objetos no son comparables.

Para introducir un nuevo tipo de orden en la que todos los objetos son comparables, tenemos la siguiente definición. Dado un conjunto A y $\ll$ un orden parcial sobre A, decimos que $\ll$ es un **orden total** si satisface que: para todo $\{a, b\} \subseteq A$, $a \ll b$ o $b \ll a$.

La propiedad anterior dice que cualquier par de objetos es *comparable* en los ordenes totales, es decir, siempre es posible determinar, dado un par de objetos a y b cuándo uno es prioritario sobre el otro bajo la relación $\ll$.

2.4.1. Jerarquías de objetos

Sean A un conjunto, $B \subseteq A$, $\preceq$ un orden parcial sobre A y $x \in A$. Decimos que **x es minimal en B** si satisface que $x \in B$ y para todo $b \in B$ tal que $b \preceq x$, $x = b$. Decimos que **x es cota inferior de B** si se satisface que para todo $b \in B$, $x \preceq b$. Si B tiene cota inferior, diremos que B está **acotado inferiormente**. Si x es una cota inferior de B y además $x \in B$, diremos que **x es mínimo de B**

Sean A un conjunto, $B \subseteq A$, $\preceq$ un orden parcial sobre A y $x \in A$. Decimos que **x es maximal en B** si satisface que $x \in B$ y para todo $b \in B$ tal que $x \preceq b$, $x = b$. Decimos que **x es cota superior de B** si se satisface que para todo $b \in B$, $b \preceq x$. Si B tiene cota superior, diremos que B está **acotado superiormente**. Si x es una cota superior en B y además, $x \in B$, diremos que **x es máximo de B**.

Los tres conceptos anteriores son distintos entre ellos en general, y para aclararlos un poco, usaremos el siguiente ejemplo. Consideremos los siguientes conjuntos:

$$
\begin{array}{lllll}
A_1 = \{c, d\} & A_3 = \{c\} & A_5 = \{a\} & A_7 = \{e, f, h\} & A_9 = \{e, f\} \\
A_2 = \{a, b, c\} & A_4 = \{a, b\} & A_6 = \{e, f, h, i\} & A_8 = \{e, g\} & A_{10} = \{e\}
\end{array}
$$

Consideramos $\mathscr{A} = \{A_i : i \in \{1, \dots 10\}\}$ el orden $\ll$ en $\mathscr{A}$ dado por $A_i \ll A_j$ si y sólo si $A_i \subseteq A_j$. La relación queda representada como sigue:

Considerando el conjunto $R_1 = \{A_1, A_2, A_3, A_4\}$ tenemos que éste no tiene cotas superiores, pues no hay nadie en $\mathscr{A}$ que sea mayor a todos los elementos en R_1. Derivado de lo anterior, tampoco tiene máximo.

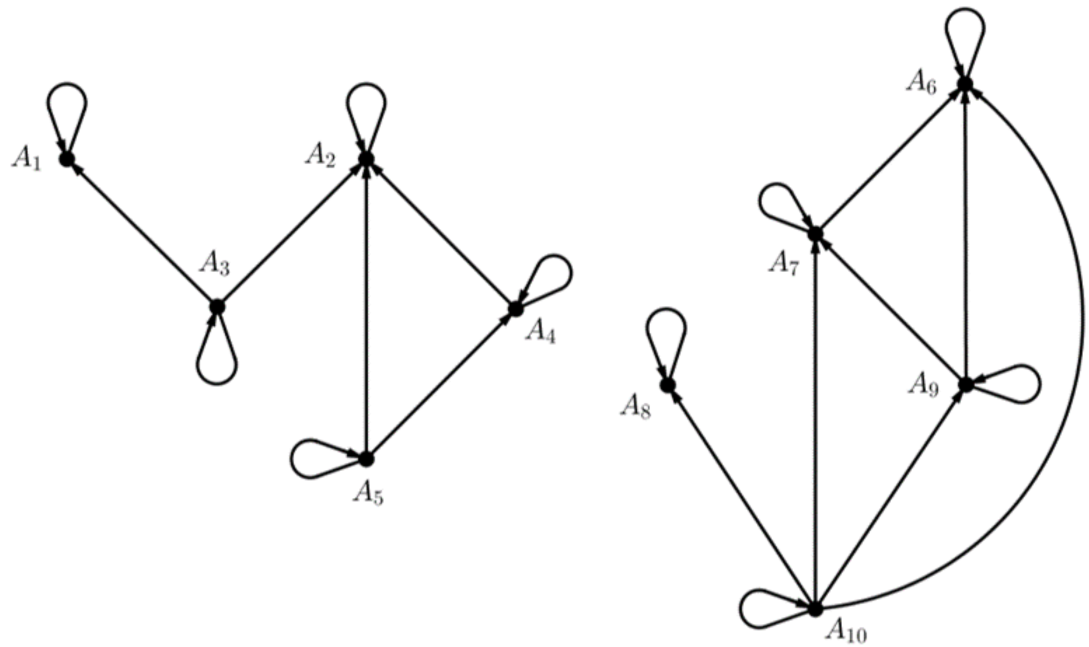

Figura 2.2: Representación del orden $\ll$ en la familia $\mathscr{A}$

Sin embargo, R_1 sí tiene elementos maximales: A_1, pues no hay nadie en R_1 distinto de A_1 que sea mayor que A_1. Lo mismo para A_2.

El conjunto sí tiene elementos minimales: A_3, pues no hay nadie en R_1 distinto de A_3 que sea menor que A_3. Lo mismo para A_4.

El conjunto R_1 no tiene cota inferior, pues no hay nadie en $\mathscr{A}$ que sea menor que todos los elementos de $\mathscr{A}$. Derivado de lo anterior, tampoco tiene mínimo.

Si ahora prestamos atención al conjunto $R_2 = \{A_7, A_8, A_9\}$, éste tampoco tiene cotas superiores, ni máximos, pero sí elementos maximales: A_8 y A_7. El conjunto sí tiene cotas inferiores: A_{10}, pues A_{10} es menor que todos los elementos de R_2, sin embargo, el conjunto R_2 no tiene mínimo, pues no hay nadie en R_2 que sea menor que todos los elementos de R_2.

Por último, si consideramos $R_3 = \{A_6, A_7, A_9\}$, éste tiene como elemento máximo y maximal a A_6, como cotas inferiores a A_9 y A_{10}, y como elemento mínimo y minimal, a A_9.

Ahora mostraremos un ejemplo un poco más abstracto sobre cómo se pueden manejar estos objetos.

Ejemplo 2.4.3. *Sea $\mathbb{N}$ el conjunto de los números enteros. Definimos la relación $\preceq$ en $\mathbb{N}\times\mathbb{N}$ dada por: $(x,z)\preceq(a,b)$ si y sólo si $x\leq a$ y $z\leq b$.*

a) Muestra que $(2,2)$ es un elemento minimal del conjunto $\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$.

b) Muestra que $(0,0)$ es un mínimo de $\mathbb{N}\times\mathbb{N}$.

c) Muestra que $(5,8)$ es una cota superior del conjunto $\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon x=5 \text{ y } z\leq 5\}$.

Demostración. Ahora procederemos a demostrar todos los incisos de este enunciado.

a) Muestra que $(2,2)$ es un elemento minimal del conjunto $\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$.

Es sencillo ver que $(2,2)$ pertenece al conjunto de referencia de $\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$ y satisface las propiedades que definen a dicho conjunto, por lo que por definición de pertenencia a un conjunto escrito por comprensión se tiene que $(2,2)\in\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$.

Ahora mostraremos que para todo $(a,b)\in\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$ tal que $(a,b)\preceq(2,2)$, se satisface que $(a,b)=(2,2)$.

Como $(a,b)\in\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon 2\leq x \text{ y } 2\leq z\}$, entonces tenemos en particular que $2\leq a$ y $2\leq z$. Por otro lado, como $(a,b)\preceq(2,2)$ se tiene de la definición de $\preceq$ que $a\leq 2$ y $b\leq 2$, concluyendo que $a=2$ y que $b=2$, por lo cual $(a,b)=(2,2)$.

b) Muestra que $(0,0)$ es un mínimo de $\mathbb{N}\times\mathbb{N}$.

Es claro que $(0,0)\in\mathbb{N}\times\mathbb{N}$, por lo que mostraremos que para todo $(a,b)\in\mathbb{N}\times\mathbb{N}$, $(0,0)\preceq(a,b)$.

Como $(a,b)\in\mathbb{N}\times\mathbb{N}$ entonces $a\in\mathbb{N}$ y $b\in\mathbb{N}$, siguiéndose de ello que $0\leq a$ y que $0\leq b$. Así, por la definición de $\preceq$ se puede concluir que $(0,0)\preceq(a,b)$.

c) Muestra que $(5,8)$ es una cota superior del conjunto $\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon x=5 \text{ y } z\leq 5\}$.

Para este inciso basta ver que para todo $(a,b)\in\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon x=5 \text{ y } z\leq 5\}$ se satisface que $(a,b)\preceq(5,8)$.

Como $(a,b)\in\{(x,z)\in\mathbb{N}\times\mathbb{N}\colon x=5 \text{ y } z\leq 5\}$, entonces tenemos que $a=5$ y que $b\leq 5$, en tal caso tenemos que $a\leq 5$ y $b\leq 8$, siguiéndose de la definición de $\preceq$ que $(a,b)\leq(5,8)$.

$\square$

Ahora procederemos a mostrar algunas propiedades más generales de los elementos mínimos, minimales, máximos y maximales.

Lema 2.4.4. *Sean A un conjunto, $\preceq$ un orden parcial sobre A, $B \subseteq A$ y $a \in A$.*

a) Si a es máximo en B, entonces a es maximal en B.

b) Si a es mínimo en B, entonces a es minimal en B.

c) Si a es máximo en B, entonces es único.

d) Si a es mínimo en B, entonces es único.

e) Si $\preceq$ es un orden total en A, entonces: a es máximo en B si y sólo si a es maximal en B.

f) Si $\preceq$ es un orden total en A, entonces: a es mínimo en B si y sólo si a es minimal en B.

Demostración. (a) Sea $t \in B$ tal que $a \preceq t$. Como a es máximo en B, entonces $t \preceq t$. Ya que $a \preceq t$ y $t \preceq a$, podemos concluir de la antisimetría de $\preceq$ que $a = t$.

(d) Supongamos que a' es un elemento mínimo de B y mostraremos que $a' = a$. Dado que a es mínimo en B y $a' \in B$, entonces $a \preceq a'$. Y de manera análoga, como a' es mínimo en B y $a \in B$, entonces $a' \preceq a$. Así, como $a \preceq a'$ y $a' \preceq a$, podemos concluir que $a = a'$.

(f) La demostración de la condición suficiente de este enunciado puede verse en `https://youtu.be/VY6BTLb2mFM`

La condición necesaria de este enunciado puede verse en `https://youtu.be/Go84QVldgYo` □

Cabe señalar que dado un conjunto A en el que está establecido un orden parcial, entonces no todos los subconjuntos de A tienen mínimo o máximo, pero cuando podemos asegurar que todos sus subconjuntos tienen un mínimo, dichos ordenes recibirán un nombre especial. Sean A un conjunto y $\preceq$ un orden parcial sobre A, decimos que $\preceq$ **es un buen orden** si todo subconjunto no vacío de A tiene mínimo.

Lema 2.4.5. *Sean A un conjunto y $\preceq$ un orden parcial sobre A. Si $\preceq$ es un buen orden, entonces $\preceq$ es un orden total.*

Demostración. Dado que $\preceq$ es un orden parcial, para demostrar que $\preceq$ es un orden total, solo resta mostrar que para todo $\{a, b\} \subseteq A$, $a \preceq b$ o $b \preceq a$.

Sea $\{a, b\} \subseteq A$. Como $\preceq$ es un buen orden, entonces el conjunto $\{a, b\}$ tiene máximo, es decir, a es máximo de $\{a, b\}$ o b es máximo de $\{a, b\}$, por lo que consideremos los siguientes casos:

Caso 1. a es máximo de $\{a, b\}$.

En este caso, por definición de máximo, $b \preceq a$ y con ello, $a \preceq b$ o $b \preceq a$.

Caso 2. b es máximo de $\{a, b\}$.

En este caso, por definición de máximo, $a \preceq b$ y con ello, $a \preceq b$ o $b \preceq a$.

En cualquiera de los dos casos concluimos que $a \preceq b$ o $b \preceq a$, por lo que $\preceq$ es un orden total sobre A. □

2.5. Funciones

Sean A y B conjuntos y R una relación de A en B. Decimos que R es una **función de A en B**, denotado por $R : A \to B$, si satisface que $Dom(R) = A$ y todo elemento del dominio está relacionado con un único elemento bajo R. Al conjunto B le llamaremos el **codominio** de la función.

Nota 2.5.1. *Sean A y B conjuntos y f una relación de A en B. f es una función de A en B si y sólo si la relación f satisface lo siguiente:*

a) $Dom(f) = A$.

b) Si $(a, b) \in f$ y $(c, d) \in f$ y $a = c$, entonces $b = d$.

Si $f : A \to B$ es una función y $a \in A$, la **imagen de a bajo** f, denotado por $f(a)$, es el único elemento $b \in B$ tal que $(a, b) \in f$. En general, siempre que tengamos una función f y un objeto en su dominio, digamos a, usaremos la notación $f(a) = x$ para indicar que los objetos a y x están relacionados bajo la relación f, es decir, preferiremos la notación $f(a) = x$ en lugar de la notación $(a, x) \in f$. Esto es similar cuando trabajamos con relaciones de equivalencia y de orden, pues preferíamos usar la notación alterna $a \ll b$ en lugar de $(a, b) \in \ll$. Esto es importante y lo resaltaremos en una nota.

Nota 2.5.2. *Si $f : A \to B$ es una función y $a \in A$, entonces: $f(a) = b$ si y sólo si $(a, b) \in f$.*

Otra nota que es importante resaltar, es que la imagen de un objeto siempre está en el codominio de la función.

Nota 2.5.3. *Si $f : A \to B$ es una función y $a \in A$, entonces: $f(a) \in B$.*

Del mismo modo que en las relaciones, las funciones son conjuntos, por lo que tendremos algunas propiedades importantes que podemos recuperar de los conjuntos y que heredarán las funciones, entre ellas destacaremos el que las funciones tengan (i) una escritura por comprensión y (ii) condiciones de igualdad entre funciones.

Respecto a la notación por comprensión, tenemos la escritura por comprensión conjuntista. Es decir, si f

es una función de A en B, entonces f puede escribirse como:

$$f = \{(a, b) \in A \times B : P(a, b)\}$$

Sin embargo, al igual que en las relaciones de orden y de equivalencia, usaremos una redacción distinta a la escritura por comprensión conjuntista para poder definir funciones. Por ejemplo, será más frecuente encontrar algo como:

Sean A y B conjuntos, definimos la función $f : A \to B$ dada por: $f(x) = z$ si y sólo si $P(x, z)$ es verdadera.

En estas circunstancias, al enunciado $P(a, b)$ se le suele llamar **regla de correspondencia**[2] (pueden ver https://youtu.be/n4Z01r152OA) .

También es importante mencionar que el tipo de enunciado que describimos anteriormente lo podemos usar cuando sabemos que f es una función, pues la notación $f(a)$ es exclusiva de las funciones.

Con lo anterior podemos establecer un primer ejemplo de una función.

Ejemplo 2.5.4. *Sea A un conjunto arbitrario. Definimos la relación 1_A de A en A dada por $(x, z) \in 1_A$ si y sólo si $x = z$. Muestra que la relación 1_A es una función de A en A.*

Demostración. Para verificar que 1_A es una función, veremos que satisface la definición de función, es decir, mostraremos los siguientes incisos.

(i) $Dom(1_A) = A$.

En este caso, realizaremos una doble contención. Sin embargo, de la definición de dominio, se sigue que $Dom(1_A) \subseteq A$, por lo que resta demostrar que $A \subseteq Dom(1_A)$.

Sea $w \in A$. De acuerdo a la definición de 1_A, se tiene que $(w, w) \in 1_A$, es decir, existe $w \in A$ tal que $(w, w) \in 1_A$, por lo que, por definición de dominio, $w \in Dom(1_A)$, concluyendo la demostración deseada.

(ii) Si $(x, z) \in 1_A$ y $(s, t) \in 1_A$ y $x = s$, entonces $z = t$.

Sean $(x, z) \in 1_A$ y $(s, t) \in 1_A$ tales que $x = s$. Ya que $(x, z) \in 1_A$, se sigue de la definición de dicha relación que $x = z$. De manera análoga, como $(s, t) \in 1_A$, se sigue que $s = t$. Con lo anterior, como $x = s$, podemos concluir que $z = t$.

[2] En muchas ocasiones se le suele llamar regla de correspondencia al elemento $f(a)$, no dejemos de lado que una *regla de correspondencia* es un enunciado lógico y no necesariamente un objeto.

De los incisos previos, la relación 1_A es una función de A en A. □

La relación descrita en el ejemplo 2.5.4 es de suma importancia y tiene nombre. Si A es un conjunto, la **función identidad sobre A**, denotada por 1_A, es la función $1_A : A \to A$ dada por $1_A(x) = x$. Esta función, aunque sencilla, tiene una relación interesante con algunos conceptos matemáticos que veremos posteriormente.

Ahora tendremos un segundo ejemplo de cómo demostrar que una relación es una función.

Ejemplo 2.5.5. *Sea $\mathbb{Z}$ el conjunto de los números enteros. Definimos la relación φ de $\mathbb{Z} \times \mathbb{Z}$ en $\mathbb{Z}$ dada por: $((x, z), a) \in \varphi$ si y sólo si $a = x$. Muestra que la relación φ es una función de $\mathbb{Z} \times \mathbb{Z}$ en $\mathbb{Z}$.*

Demostración. Para ver que φ es una función de $\mathbb{Z} \times \mathbb{Z}$ en $\mathbb{Z}$, primero analizaremos el enunciado anterior en el siguiente video https://youtu.be/EScmJ9vc9HE. Ahora, basta ver que φ satisface la definición de función.

(i) $Dom(\varphi) = \mathbb{Z} \times \mathbb{Z}$.

La demostración de este inciso pueden verla en https://youtu.be/MT5mMoxLCrw

(ii) Si $((x, z), a) \in \varphi$ y $((r, s), b) = \varphi$ y $(x, z) = (r, s)$, entonces $a = b$.

La demostración de este inciso pueden verla en https://youtu.be/6W2gLBVX7SY

□

La siguiente condición que tendremos en consideración para las funciones será la igualdad de funciones. Si bien las funciones son conjuntos y se puede demostrar que dos funciones son iguales mediante una doble contención, en ocasiones nos será más sencillo demostrar que dos funciones son iguales mediante el siguiente teorema:

Teorema 2.5.6 (Igualdad de Funciones.). *Si $f : A \to B$ y $g : C \to D$ son dos funciones, entonces: $f = g$ si y sólo si se satisfacen las siguientes condiciones:*

a) $Dom(f) = Dom(g)$.

b) Para todo $x \in A$, $f(x) = g(x)$.

Demostración. Primero mostraremos la parte suficiente del enunciado, es decir, si $f = g$, entonces (a) $Dom(f) = Dom(g)$ y (b) para todo $x \in A$, $f(x) = g(x)$.

Para demostrar (a) basta ver que como $f = g$, entonces $Dom(f) = Dom(g)$.

Ahora mostraremos (b) considerando $x \in A$ arbitrario. Nótese que debido a que f es una función de A en B, entonces $Dom(f) = A$, por lo que $x \in Dom(A)$ y por definición de dominio, existe $w \in B$ tal que

$f(x) = w$, es decir, $(x, w) \in f$. Ya que $f = g$, podemos decir que $(x, w) \in g$, por lo que $g(x) = w$, concluyendo que $f(x) = g(x)$.

Ahora procederemos a mostrar la parte necesaria del enunciado, es decir, mostraremos que si $f : A \to B$ y $g : C \to D$ son funciones tales que (a) $Dom(f) = Dom(g)$ y (b) para todo $x \in A$, $f(x) = g(x)$, entonces $f = g$.

Para esto realizaremos una doble contención. Primero mostraremos que $f \subseteq g$.

Sea $(s, t) \in f$. Con lo anterior, podemos garantizar que $f(s) = t$. Nótese además que como $s \in Dom(f)$ y f es una función de A en B, entonces $s \in A$. Gracias a la condición (b) que sigue que $g(s) = t$, es decir, $(s, t) \in g$, concluyendo la contención deseada.

Ahora mostraremos que $g \subseteq f$. Para ello consideremos $(a, b) \in g$. Nótese que $g(a) = b$. Por definición de dominio se tiene que $a \in Dom(g)$. Gracias a la hipótesis (a) tenemos en particular que $a \in Dom(f)$ y como f es una función de A en B, entonces $a \in A$. Ahora, aplicando la condición (b) se seguiría que $f(a) = g(a)$ y como anteriormente vimos que $g(a) = b$, podemos concluir que $f(a) = b$, así, $(a, b) \in f$, mostrando la contención restante. $\square$

2.5.1. Composición de funciones

Para terminar esta primera sección referente a funciones, primero definiremos la composición de relaciones. Sean f una relación de A en B y g una relación de B en C. La **composición de f seguida de g**, denotada por $g \circ f$, es la relación de A en C tal que $(x, z) \in g \circ f$ si y sólo si existe $w \in B$ tal que $(x, w) \in f$ y $(w, z) \in g$.

El siguiente teorema muestra que la composición de dos funciones necesariamente es una función.

Teorema 2.5.7. *Si $f : A \to B$ y $g : B \to C$ son funciones, entonces la composición $g \circ f$ es una función de A en C*

Demostración. Para mostrar que $g \circ f$ es una función de A en C, debemos mostrar los siguientes incisos:

(i) $Dom(g \circ f) = A$.

Para este inciso, se sigue directamente de la definición de dominio que $Dom(g \circ f) \subseteq A$, por lo que sólo resta demostrar la segunda contención, es decir, $A \subseteq Dom(g \circ f)$.

Sea $t \in A$. Ya que $Dom(f) = A$, entonces $t \in Dom(A)$ y se tiene que existe $s \in B$ tal que $(t, s) \in f$. Por otro lado, como $s \in B$ y $Dom(g) = B$, se sigue que $s \in Dom(g)$, concluyendo que existe $r \in C$ tal que $(s, r) \in g$. De lo anterior, existe $s \in B$ tal que $(t, s) \in f$ y $(s, r) \in g$, por lo que por definición de composición, $(t, r) \in g \circ f$. Así, dado que existe $r \in C$ tal que $(t, r) \in g \circ f$, podemos garantizar de la definición de dominio de una relación que $t \in Dom(g \circ f)$.

(ii) Si $(a,b) \in g \circ f$, $(x,z) \in g \circ f$ y $a = x$, entonces $b = z$.

Sean $(a,b) \in g \circ f$ y $(x,z) \in g \circ f$ tales que $a = x$. Como $(a,b) \in g \circ f$, se tiene de la definición de composición que existe $t \in B$ tal que $(a,t) \in f$ y $(t,b) \in g$. Por otro lado, como $(x,z) \in g \circ f$, se sigue de la definición de composición que existe $t' \in B$ tal que $(x,t') \in f$ y $(t',z) \in g$.

Debido a que $(a,t) \in f$ y $(x,t') \in f$ y $a = x$, entonces se sigue del hecho de que f es una función el que $t = t'$. De manera análoga, como $(t,b) \in g$, $(t',z) \in g$, $t = t'$ y g es una función, podemos concluir que $b = z$.

De los puntos anteriores, se sigue que $g \circ f$ es una función de A en C. □

El teorema anterior es bastante accesible, pues ahora podemos hablar de la composición de funciones como una función. Más aun, tenemos lo siguiente:

Corolario 2.5.8. *Sean $f : A \to B$ y $g : B \to C$ dos funciones. Si $a \in A$, entonces $(g \circ f)(a) = g(\, f(a)\,)$.*

Demostración. Sea $a \in A$ arbitrario. Dado que $g \circ f$ es una función de A en C, consideremos $t \in C$ tal que $g \circ f(a) = t$. Se sigue que $(a,t) \in g \circ f$. Por definición de composición, existe $z \in B$ tal que $(a,z) \in f$ y $(z,t) \in g$, y por definición de función, $f(a) = z$ y $g(z) = t$. Así,

$$\begin{aligned} g \circ f(a) =& t \\ =& g(z) \\ =& g(f(a)) \end{aligned}$$

De lo anterior concluimos que $g \circ f(a) = g(f(a))$. □

Además de lo anterior, tenemos un par de resultados que es importante recalcar:

Lema 2.5.9. *Si $f : A \to B$, $g : B \to C$ y $h : C \to D$ son funciones, entonces:*

a) $f \circ 1_A = f$.

b) $1_B \circ f = f$.

c) $h \circ (g \circ f) = (h \circ g) \circ f$.

Demostración. Primero demostraremos el inciso (a) mediante el teorema de igualdad de funciones. Para ello, necesitamos demostrar los siguientes incisos:

(i) $Dom(f \circ 1_A) = Dom(f)$.

En este caso, por el teorema 2.5.7 se tiene que $f \circ 1_A$ es una función de A en B, por lo que $Dom(f \circ 1_A) = A$, y como f es una función de A en B, entonces $Dom(f) = A$, concluyendo que $Dom(f \circ 1_A) = Dom(f)$.

(ii) Para todo $x \in A$, $f \circ 1_A(x) = f(x)$.

Sea $x \in A$, por definición de la función identidad, $1_A(x) = x$, y con ello, $f(1_A(x)) = f(x)$, siguiéndose del corolario 2.5.8 el que $f \circ 1_A(x) = f(x)$.

Con lo anterior, y gracias el teorema de igualdad de funciones, se concluye que $f \circ 1_A = f$.

Ahora procederemos a mostrar que $1_B \circ f = f$, nuevamente mediante el teorema de igualdad de funciones.

(i) $Dom(1_B \circ f) = Dom(f)$.

En este caso, por el teorema 2.5.7 se tiene que $1_B \circ f$ es una función de A en B, por lo que $Dom(1_B \circ f) = A$, y como f es una función de A en B, entonces $Dom(f) = A$, concluyendo que $Dom(1_B \circ f) = Dom(f)$.

(ii) Para todo $x \in A$, $1_B \circ f(x) = f(x)$.

Sea $x \in A$, por definición de la función, podemos considerar $b \in B$ tal que $f(x) = b$. En tal caso, notemos que $1_B(f(x)) = 1_B(b) = b$, así $1_B(f(x)) = f(x)$, concluyendo del corolario 2.5.8 que $1_B \circ f(x) = f(x)$.

Con lo anterior, y gracias el teorema de igualdad de funciones, se concluye que $1_B \circ f = f$.

Por último, sólo resta demostrar el inciso (c) de este enunciado. Nuevamente, la demostración la realizaremos mediante el teorema de igualdad de funciones. La demostración del enunciado puede verse en `https://youtu.be/11RVZEoC9Ic` □

2.5.2. Conjuntos asociados a funciones

En esta sección veremos tres conjuntos importantes relacionados con funciones: la imagen de la función, la imagen directa de un conjunto y la imagen inversa de un conjunto.

Si $f : A \to B$ es una función, la **imagen de f**, denotada por $Im(f)$, es el conjunto:

$$Im(f) = \{b \in B : \text{ existe } a \in A \text{ tal que } f(a) = b\}.$$

Esta definición de imagen de una función es exactamente la misma que la definición de imagen de una relación, con la salvedad que en este caso preferiremos la notación $f(a) = b$ en lugar de la notación $(a, b) \in f$.

Nota 2.5.10. *Sean $f : A \to B$ una función y b un objeto arbitrario. Los siguientes enunciados son equivalentes:*

a) $b \in Im(f)$.

b) $b \in B$ *y existe* $a \in A$ *tal que* $f(a) = b$.

Ahora consideremos una función $f : A \to B$ y $C \subseteq A$. La **imagen directa de C bajo f**, denotada por $f[C]$, es el conjunto:

$$f[C] = \{b \in B : \text{ existe } a \in C \text{ tal que } f(a) = b\}$$

Notemos que la imagen directa es un subconjunto del codominio de una función. Es importante recalcar esto pues la imagen directa de un conjunto es otro conjunto, mientras que la imagen de un objeto es un elemento del codominio. La siguiente nota será de utilidad para realizar demostraciones.

Nota 2.5.11. *Sean $f : A \to B$ una función, $C \subseteq A$ y b un objeto. Los siguientes enunciados son equivalentes:*

a) $b \in f[C]$.

b) $b \in B$ *y existe* $c \in C$ *tal que* $f(c) = b$.

Es importante señalar que estas funciones tienen algunas propiedades que es importante destacar.

Lema 2.5.12. *Sean $f : A \to B$ una función, $A_1 \subseteq A$, $A_2 \subseteq A$, $B_1 \subseteq B$ y $B_2 \subseteq B$.*

a) Si $A_1 \subseteq A_2$, *entonces* $f[A_1] \subseteq f[A_2]$.

b) $f[A_1 \cup A_2] = f[A_1] \cup f[A_2]$.

c) $f[A_1 \cap A_2] \subseteq f[A_1] \cap f[A_2]$.

d) $f[A_1] \setminus f[A_2] \subseteq f[A_1 \setminus A_2]$.

Demostración. Ahora procederemos a demostrar algunos de los incisos previos.

(c) $f[A_1 \cap A_2] \subseteq f[A_1] \cap f[A_2]$.

Sea $w \in f[A_1 \cap A_2]$. Por definición de imagen directa, existe $x \in A_1 \cap A_2$ tal que $f(x) = w$. Luego, por definición de intersección, podemos afirmar que $x \in A_1$ y $x \in A_2$.

Dado que existe $x \in A_1$ tal que $f(x) = w$, por definición de imagen directa, podemos concluir que $w \in f[A_1]$. De manera análoga, como existe $x \in A_2$ tal que $f(x) = w$, se concluye de la definición de imagen directa el que $w \in f[A_2]$. Por la definición de intersección, se concluye que $w \in f[A_1] \cap f[A_2]$.

(d) $f[A_1] \setminus f[A_2] \subseteq f[A_1 \setminus A_2]$.

Esta demostración pueden verla en el video https://youtu.be/gu6s5qPExyI

□

Ahora consideremos una función $f : A \to B$ y $D \subseteq B$, la **imagen inversa de D bajo f**, denotada como $f^{-1}[D]$, como el conjunto:

$$f^{-1}[D] = \{a \in A : \text{ existe } c \in D \text{ tal que } f(a) = c\}$$

Del mismo modo, se ha de tener cuidado en la notación: cada que usemos corchetes cuadrados de la forma $f^{-1}[D]$ se está tratando como un conjunto en particular, de hecho, de un subconjunto del dominio. No hay que confundir esta notación con el de función inversa. En caso de que $D = \{x\}$ para algún $x \in B$, preferiremos la notación $f^{-1}[x]$ en lugar de $f^{-1}[\{x\}]$. Es muy importante recalcar que esta notación no indica que f sea una función invertible, sino que $f^{-1}[x]$ son todos aquellos elementos del dominio de f que están asociados a x, los cuales pueden ser bastantes.

Nota 2.5.13. *Sean $f : A \to B$ una función, $C \subseteq B$ y a un objeto. Los siguientes enunciados son equivalentes:*

a) $a \in f^{-1}[C]$.

b) $a \in A$ y existe $c \in C$ tal que $f(a) = c$.

El siguiente resultado puede usarse para entrar en calor y manejar de mejor manera estos conjuntos:

Lema 2.5.14. *Sean $f : A \to B$ una función, $A_1 \subseteq A$, $A_2 \subseteq A$, $B_1 \subseteq B$ y $B_2 \subseteq B$.*

a) Si $B_1 \subseteq B_2$, entonces $f^{-1}[B_1] \subseteq f^{-1}[B_2]$.

b) $f^{-1}[B_1 \cup B_2] = f^{-1}[B_1] \cup f^{-1}[B_2]$.

c) $f^{-1}[B_1 \cap B_2] = f^{-1}[B_1] \cap f^{-1}[B_2]$.

d) $f^{-1}[B_1 \setminus B_2] = f^{-1}[B_1] \setminus f^{-1}[B_2]$.

e) $A_1 \subseteq f^{-1}[f[A_1]]$.

f) $f[f^{-1}[B_1]] \subseteq B_1$.

Demostración. Ahora procederemos a demostrar algunos de los incisos previos.

(a) Si $B_1 \subseteq B_2$, entonces $f^{-1}[B_1] \subseteq f^{-1}[B_2]$.

Esta demostración pueden verla en el video https://youtu.be/KPh37JVb-6M

(e) $A_1 \subseteq f^{-1}[f[A_1]]$.

Sea $w \in A_1$. Supongamos que $f(w) = b$ para algún $b \in B$. Notemos que de la definición de imagen directa, $b \in f[A_1]$. Así como existe $b \in f[A_1]$ tal que $f(w) = b$, se sigue de la definición de imagen inversa que $w \in f^{-}[f[A_1]]$.

□

2.6. Tipos de funciones

Existen muchos tipos de funciones que serán de interés a lo largo de la carrera. Cada área matemática tiene funciones particulares que estudiarán en sus cursos, pero existen tres que son muy frecuentes en todas las ramas matemáticas y que son básicas que las comprendan: las funciones inyectivas, sobreyectivas y las biyectivas. Este tipo de funciones buscan, en cierta medida, replicar en el codominio el comportamiento de las funciones sobre su dominio.

Por ejemplo, uno de los puntos de la definición de función es que todos los elementos de su dominio estén relacionados con alguien del codominio, pero no necesariamente todos los elementos del codominio están relacionados. Cuando esto último sí pasa, diremos que la función es sobreyectiva.

Otro ejemplo, el segundo punto de la definición de función indica que las parejas de los elementos en el dominio deben ser únicas, pero los elementos del codominio pueden haberse relacionado desde varios objetos del dominio. Cuando esto último no pasa, diremos que la función es inyectiva. Estas funciones, junto con las biyectivas, serán los objetos de estudio de esta sección.

2.6.1. Funciones inyectivas

Las primeras funciones que analizaremos serán las funciones inyectivas. Estas funciones satisfacen que aquellos elementos del codominio que estén relacionados con alguien del dominio, se relacionen también de manera única. Formalmente hablando, dada una función $f : A \to B$, decimos que **f es inyectiva** si se satisface que para todo $\{x, z\} \subseteq A$ tales que $f(x) = f(z)$, se satisface que $x = z$.

En la figura 2.3 se muestra el ejemplo de una función inyectiva, a saber, f. Nótese que en esta función todos los elementos del codominio que sí están relacionados, se relacionan con un único objeto del dominio, que es justamente la definición de inyectividad. Por otro lado, en la misma figura se muestra una función no inyectiva, a saber, g. Nótese que en dicha función hay un objeto del codominio que está relacionado con dos objetos distintos del dominio.

Aunque con un modelo no tan abstracto como el anterior es relativamente sencillo determinar si una función es inyectiva o no, en general no se dispondrá de una representación como la anterior, o incluso no será necesaria

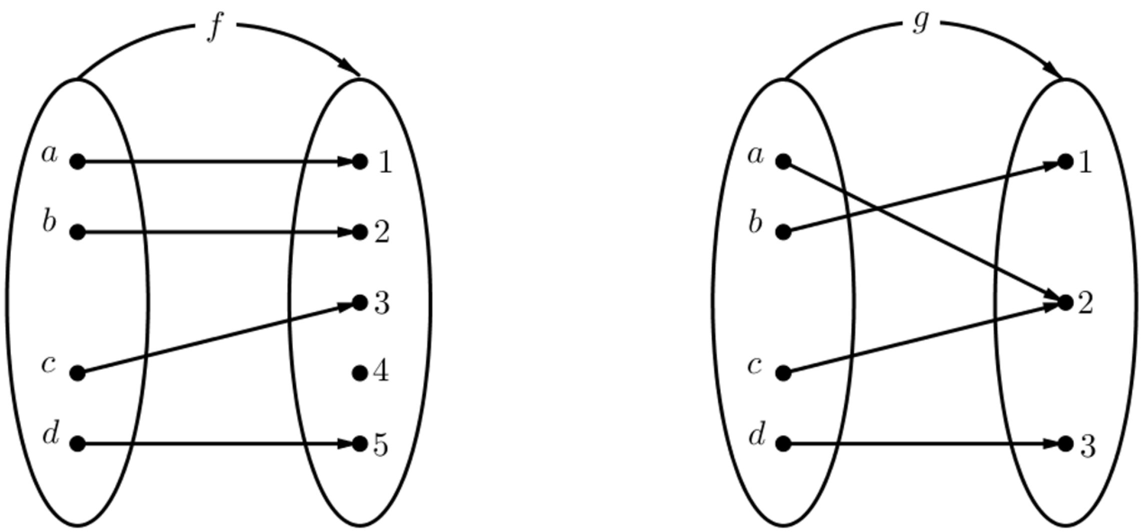

Figura 2.3: Ejemplo de una función inyectiva y una función no inyectiva

para poder usar la definición de inyectividad. Así, para poder verificar que algo es inyectivo podemos usar su definición, por lo que la definición de inyectividad será lo primero que debemos aprender para esta sección (véase https://youtu.be/0qVPa9JVAzE) y para ello, tenemos las siguientes notas.

Nota 2.6.1. *Sean A y B conjuntos y $f : A \to B$ una función. Los siguientes incisos son equivalentes:*

a) *f es inyectiva.*

b) *Si $\{x, z\} \subseteq A$ y $f(x) = f(z)$, entonces $x = z$*

En algunos casos también podemos usar una contraposición del enunciado (b) de la nota anterior, es decir, en lugar de dicho inciso pueden encontrar lo siguiente: si $\{x, z\} \subseteq A$ y $x \neq z$, entonces $f(x) \neq f(z)$. Aunque dicho enunciado es útil, nosotros preferiremos el mostrado en la nota previa debido a que usualmente es más sencillo usar una igualdad de objetos por encima de una no igualdad. También es importante saber cuándo una función no es inyectiva mediante la definición.

Nota 2.6.2. *Sean A y B conjuntos y $f : A \to B$ una función. Los siguientes incisos son equivalentes:*

a) *f no es inyectiva.*

b) *Existe $\{x, z\} \subseteq A$ tal que $f(x) = f(z)$ y $x \neq z$.*

Ahora que hemos establecido el concepto de función inyectiva, vamos a demostrar que una función particular es inyectiva.

Ejemplo 2.6.3. *Sean $\mathbb{Z}$ el conjunto de números enteros y $\mathbb{R}$ el conjunto de los números reales. Definimos la función η de $\mathbb{Z}$ en $\mathbb{R} \times \mathbb{R}$ dada por: $\eta(k) = (\frac{k}{3}, (\frac{k}{3})^2)$. Muestra que la función η es inyectiva.*

Demostración. Para verificar que η es una función inyectiva, veremos que dicha función satisface la definición de inyectividad, es decir: si $\{m, n\} \subseteq \mathbb{Z}$ y $\eta(n) = \eta(m)$, entonces $n = m$.

Para mostrar la condicional anterior, consideremos $\{m, n\} \subseteq \mathbb{Z}$ tales que $\eta(n) = \eta(m)$. De acuerdo a la definición de η se tiene que $\eta(n) = (\frac{n}{3}, (\frac{n}{3})^2)$ y que $\eta(m) = (\frac{m}{3}, (\frac{m}{3})^2)$. Ya que por suposición $\eta(n) = \eta(m)$, se sigue que $(\frac{n}{3}, (\frac{n}{3})^2) = (\frac{m}{3}, (\frac{m}{3})^2)$ y por igualdad de parejas ordenadas, en particular se sabe que $\frac{n}{3} = \frac{m}{3}$, concluyendo que $n = m$. $\square$

El anterior es un ejemplo clásico de cómo demostrar mediante la definición que una función es inyectiva. Ahora veremos un par de propiedades más sobre funciones inyectivas.

Lema 2.6.4. *Sean A, B y C tres conjuntos y $f : A \to B$ y $g : B \to C$ dos funciones. Los siguientes enunciados se satisfacen:*

a) Si f y g son inyectivas, entonces $g \circ f$ es inyectiva

b) Si $g \circ f$ es inyectiva, entonces f es inyectiva.

Demostración. (a) La demostración de este inciso puede verse en https://youtu.be/-dfKfHFv-u8

(b) Si $g \circ f$ es inyectiva, entonces f es inyectiva.

Para verificar que f es inyectiva, consideremos $\{r, s\} \subseteq A$ tales que $f(r) = f(s)$. Ya que $f(r)$ y $f(s)$ son objetos en el dominio de g, podemos considerar $g(f(r))$ y $g(f(s))$. Debido a que $f(r) = f(s)$, podemos garantizar que $g(f(r)) = g(f(s))$, es decir, $g \circ f(r) = g \circ f(s)$, y como $g \circ f$ es inyectiva, se sigue que $r = s$. $\square$

Muy ligado al concepto de función inyectiva está el de inverso izquierdo. Dada una función $f : A \to B$, un **inverso izquierdo de f** es cualquier función $g : B \to A$ tal que $g \circ f = 1_A$.

Aunque no nos será frecuente el concepto de inverso izquierdo en estas notas hay dos cosas importantes que rescatar sobre dicho concepto. En primer lugar, la existencia de un inverso izquierdo no asegura que éste sea único, es decir, pueden existir muchos inversos izquierdos para una función dada. En segundo lugar, es importante no confundir el inverso izquierdo con la función inversa, misma que aun no hemos definido, pero que será importante considerar más adelante. Como resultado general, tenemos el siguiente teorema:

Teorema 2.6.5. *Sea $f : A \to B$ una función. La función f es inyectiva si y sólo si f tiene al menos un inverso izquierdo.*

Demostración. Primero mostraremos la parte suficiente del teorema, es decir, si f es inyectiva, entonces f tiene al menos un inverso izquierdo.

Sea f una función inyectiva y consideremos $a_0 \in A$ un punto fijo. Definimos la relación g de B en A dada por: $(s,t) \in g$ si y sólo si (i) $s \in Im(f)$ y $f(t) = s$ o (ii) $t = a_0$ si $s \notin Im(f)$.

Ahora mostraremos que g es una función de B en A. Para ello tenemos los siguientes incisos:

(i) $Dom(g) = B$.

Esta demostración la realizaremos por doble contención. Debido a la definición de dominio, se sigue que $Dom(g) \subseteq B$, por lo que sólo resta demostrar la segunda contención.

Sea $w \in B$. Consideremos los siguientes dos casos para w.

Caso 1. $w \in Im(f)$.

En este caso, por definición de imagen, existe $r \in A$ tal que $f(r) = w$. Por el punto (i) de la definición de g, se concluye que $(w,t) \in g$, mostrando así que $w \in Dom(g)$.

Caso 2. $w \notin Im(f)$.

En este caso, por el punto (ii) de la definición de g, se sigue que $(w, a_0) \in g$, mostrando así que $w \in Dom(g)$.

De los casos previos, se sigue que $B \subseteq Dom(g)$, concluyendo la igualdad deseada.

(ii) Si $(x,z) \in g$, $(s,t) \in g$ y $x = s$, entonces $z = t$.

Sean $(x,z) \in g$ y $(s,t) \in g$ tales que $x = s$. Tenemos dos posibles casos a considerar sobre x:

Caso 1. $x \notin Im(f)$.

En este caso, por el punto (ii) de la definición de g, podemos asegurar que $z = a_0$. Por otro lado, ya que $s = x$, se sigue que $s \notin Im(f)$, y nuevamente por el punto (ii) de la definición de g, se sigue que $t = a_0$. De lo anterior, $z = t$.

Caso 2. $x \in Im(f)$.

En este caso, por el punto (i) de la definición de g se sigue que $f(z) = x$. Además, ya que $x = s$, se tiene que $s \in Im(f)$ y por el punto (i) de la definición de g, podemos garantizar que $f(t) = s$. Así, nuevamente del hecho de que $x = s$, se sigue que $f(z) = f(t)$, y como f es inyectiva, entonces $z = t$, como era deseado.

De los casos anteriores, se sigue que $z = t$.

Por el punto (i) y (ii) aquí demostrados, g es una función de B en A, por lo que podremos empezar a usar la notación de funciones en g como sigue:

$$g(x) = \begin{cases} t & \text{Si } x \in Im(f) \text{ y } f(t) = x. \\ a_0 & \text{En otro caso.} \end{cases}$$

Ahora sólo resta mostrar que g es un inverso izquierdo de f, es decir, $g \circ f = 1_A$. Ya que tanto $g \circ f$ como 1_A son funciones con dominio en A, entonces sólo resta verificar la condición (b) del teorema de igualdad de funciones (teorema 2.5.6).

Sea $x \in A$ arbitrario. Notemos que $g \circ f(x) = g(f(x))$ por definición de composición. Sin embargo, como $f(x) \in Im(f)$, se sigue de la definición de g el que $g(f(x)) = x$, es decir, $g(f(x)) = 1_A(x)$, concluyendo que $g \circ f(x) = 1_A(x)$. Así, por el teorema de igualdad de funciones, $g \circ f = 1_A$, mostrando que g es un inverso izquierdo de f.

Ahora mostraremos la parte necesaria del teorema, es decir, mostraremos que si f tiene un inverso izquierdo, entonces f es inyectiva.

Sea $g : B \to A$ un inverso izquierdo de f, es decir, $g \circ f = 1_A$. Ya que la función 1_A es inyectiva, entonces $g \circ f$ es inyectiva, por lo que de acuerdo al lema 2.6.4 (b) podemos garantizar que f es inyectiva.

□

2.6.2. Funciones sobreyectivas

El segundo tipo de función que estudiaremos será el de función sobreyectiva[3]. Dada una función $f : A \to B$, decimos que **f es sobreyectiva** si para todo $b \in B$ existe $a \in A$ tal que $f(a) = b$.

En la figura 2.4 se muestra una función sobreyectiva, a saber, f. Nótese que en dicha función todos los elementos del codominio están relacionados con alguien del dominio. En la misma figura se muestra además a una función no sobreyectiva, a saber, g. En dicha función, al menos un elemento del codominio no es´ta relacionado con nadie del dominio.

De manera análoga que en varias ocasiones pasadas, tener un modelo visual de las funciones puede ayudar a determinar si es una función sobreyectiva o no. Sin embargo, se tendrán los casos en los que no será posible usar dicho modelo y es necesario acceder a algo más formal. Para ello, tendremos que usar la definición por lo que esto será lo primero que tendrás que aprender para poder utilizar dicho tipo de funciones.

[3] A éstas también se les puede llamar suprayectiva.

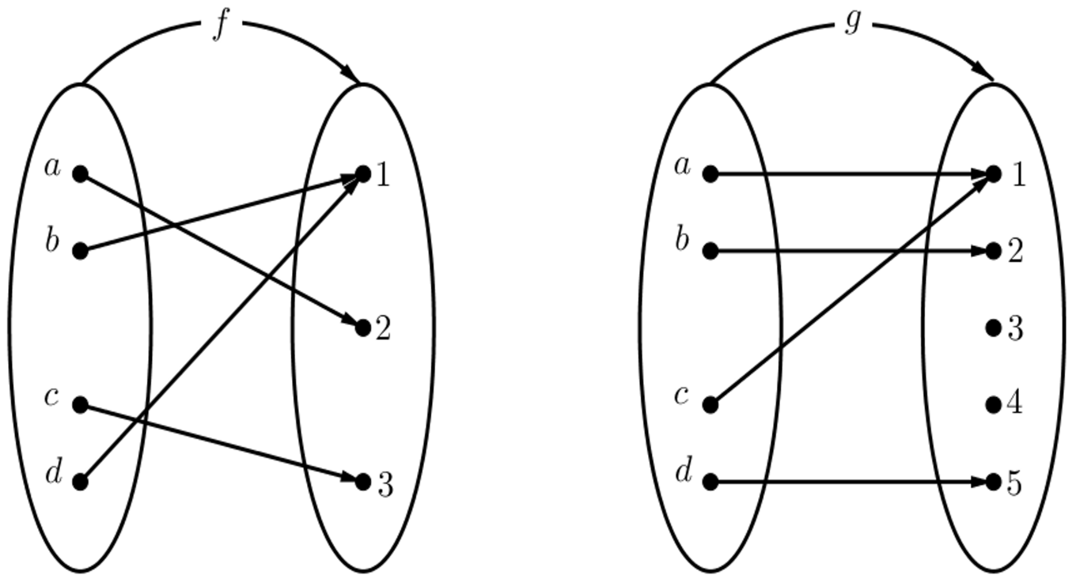

Figura 2.4: Ejemplo de una función sobreyectiva y una función no sobreyectiva

Nota 2.6.6. *Sea $f : A \to B$ una función. Los siguientes incisos son equivalentes:*

a) *La función f es sobreyectiva*

b) *Para todo $b \in B$, existe $a \in A$ tal que $f(a) = b$.*

Del mismo modo en que hemos venido trabajando, es importante recalcar también la negación de la definición anterior.

Nota 2.6.7. *Sea $f : A \to B$ una función. Los siguientes incisos son equivalentes:*

a) *La función f no es sobreyectiva*

b) *Existe $b \in B$ tal que para todo $a \in A$, $f(a) \neq b$.*

Ahora mostraremos un ejemplo sobre cómo demostrar que una función es sobreyectiva utilizando solamente la definición de sobreyectividad.

Ejemplo 2.6.8. *Sean $\mathbb{R}$ el conjunto de números reales. Definimos la función α de $\mathbb{R} \times \mathbb{R}$ en $\mathbb{R}$ dada por: $\alpha(r, s) = r + s - rs$. Muestra que la función α es sobreyectiva.*

Demostración. Para demostrar que α es sobreyectiva, basta verificar que α satisface la definición de sobreyectividad, es decir: para todo $x \in \mathbb{R}$ existe $(m, n) \in \mathbb{R} \times \mathbb{R}$ tal que $\alpha(n, m) = x$.

Sea $x \in \mathbb{R}$. Notemos que como $0 \in \mathbb{R}$, entonces $(x, 0) \in \mathbb{R}$ y además, por definición de α:

$$\begin{aligned}\alpha(x,0) =& x + 0 - x \cdot 0 \\ =& x\end{aligned}$$

Así, existe $(x, 0) \in \mathbb{R} \times \mathbb{R}$ tal que $\alpha(x, 0) = x$, concluyendo que α es sobreyectiva. □

El siguiente lema son resultados generales de las funciones sobreyectivas y composición de las mismas.

Lema 2.6.9. *Sean $f : A \to B$ y $g : B \to C$ dos funciones. Los siguientes incisos se satisfacen:*

a) Si f y g son sobreyectivas, entonces $g \circ f$ es sobreyectiva.

b) Si $g \circ f$ es sobreyectiva, entonces g es sobreyectiva.

Demostración. (a) Para mostrar que $g \circ f : A \to C$ es una función sobreyectiva, verificaremos que satisface la definición de sobreyectividad.

Sea $z \in C$. Debido a que g es sobreyectiva, existe $w \in B$ tal que $g(w) = z$. De manera análoga, como $w \in B$ y f es sobreyectiva, entonces existe $t \in A$ tal que $f(t) = w$. Ya que $f(t) = w$ y $g(w) = z$, se concluye que $g(f(t)) = z$, y por definición de composición tenemos que $g \circ f(t) = z$. Con lo anterior, existe $t \in A$ tal que $g \circ f(t) = z$, mostrando así que $g \circ f$ es sobreyectiva.

(b) La demostración de este inciso pueden verla en https://youtu.be/_60pmlEAu0A

□

Del mismo modo que en las funciones inyectivas, existe una definición de inversos que está íntimamente ligada a las funciones sobreyectivas. Sea $f : A \to B$ una función. Un **inverso derecho de f** es cualquier función $g : B \to A$ tal que $f \circ g = 1_B$.

Los inversos derechos de una función no necesariamente son únicos, sin mencionar que es importante no confundir el hecho de que una función tenga inverso derecho con el que dicha función sea invertible, pues si bien ambos conceptos están relacionados, no son equivalentes. El siguiente es un resultado clásico referente a inversos derechos.

Teorema 2.6.10. *Sea $f : A \to B$ una función. La función f es sobreyectiva si y sólo si f tiene al menos un inverso derecho.*

Demostración. Primero mostraremos la parte suficiente del enunciado, es decir, mostraremos que si f es una función sobreyectiva, entonces f tiene al menos un inverso derecho.

Para ello, notemos que por ser f una función sobreyectiva, entonces para todo $t \in B$ existe $x \in A$ tal que $f(x) = t$. Aunque el objeto x no necesariamente es único, de todos los objetos que satisfacen la propiedad anterior, consideremos sólo uno de ellos y lo nombraremos x_t. Así, para cada $t \in B$ seleccionamos un único $x_t \in A$ tal que $f(x_t) = t$. Ahora definimos la relación g de B en A dada por $(r, s) \in g$ si y sólo si $s = x_r$ y mostraremos que g es una función de B en A. Para ello, probaremos los siguientes incisos.

(i) $Dom(g) = B$.

Realizaremos esta demostración por doble contención. Sin embargo, de la definición de dominio, tenemos que $Dom(g) \subseteq B$, por lo que sólo resta demostrar que $B \subseteq Dom(g)$.

Sea $t \in B$, por lo mencionado previo a la definición de g, se sigue que $x_t \in A$ y por la definición de g, $(t, x_t) \in g$, mostrando que $t \in Dom(g)$.

(ii) Si $(r, s) \in g$, $(m, n) \in g$ y $r = m$, entonces $s = n$.

Sean $(r, s) \in g$ y $(m, n) \in g$ tales que $r = m$. Debido a que $(r, s) \in g$, se tiene de la definición de g que $s = x_r$. De manera análoga, como $(m, n) \in g$, por la definición de g se puede concluir que $n = x_m$. Sin embargo, debido a que $r = m$, entonces $x_r = x_m$ y así, $s = n$.

Gracias a los incisos anteriores, podemos garantizar que g es una función de B en A, por lo que podemos empezar a utilizar la notación de función sobre g. Es decir, $g : B \to A$ es una función tal que $g(r) = x_r$.

Ahora sólo resta mostrar que $f \circ g = 1_B$. Para ello utilizaremos el teorema de igualdad de funciones. Por un lado, como $f \circ g$ y 1_B tienen dominio en B, sólo resta mostrar que para todo $w \in B$, $f \circ g(w) = 1_B(w)$.

Sea $w \in B$. Por definición de g, $g(w) = x_w$. Ya que $g(w)$ es un objeto de A y f es una función de A en B, entonces podemos asegurar que $f(g(w)) = f(x_w)$. Sin embargo, por la definición de x_w, se tiene que $f(x_w) = w$, es decir, $f(g(w)) = w$, por lo que $f \circ g(w) = w$ y por definición de función identidad, concluimos que $f \circ g(w) = 1_B(w)$.

Por lo antes dicho, y de acuerdo al teorema de igualdad de funciones, se concluye que $f \circ g = 1_B$, es decir, g es un inverso derecho de f.

Ahora sólo resta demostrar la parte necesaria del enunciado, es decir, si f tiene un inverso derecho, entonces f es sobreyectiva.

Para ellos, supongamos que g es un inverso derecho de f, es decir, $g : B \to A$ es una función tal que $f \circ g = 1_B$. Ya que 1_B es una función sobreyectiva, entonces $f \circ g$ es una función sobreyectiva y por el lema 2.6.9 (b) podemos concluir que f es sobreyectiva.

□

2.6.3. Funciones biyectivas

El último tipo de función que analizaremos en esta sección serán las funciones biyectivas. Dada una función $f : A \to B$, decimos que **f es biyectiva** si es inyectiva y sobreyectiva.

Pareciera que la definición no ofrece nada nuevo pues las funciones biyectivas son, simultáneamente, inyectivas y sobreyectivas, mismas que ya estudiamos previamente, sin embargo la combinación de estas dos propiedades hace que este tipo de funciones sean muy importantes en distintos campos de las Matemáticas. El primer resultado que tendremos será el siguiente:

Lema 2.6.11. *Sea $f : A \to B$ y $g : B \to C$ funciones.*

a) Si f y g son biyectivas, entonces $g \circ f$ es biyectiva.

b) Si $g \circ f$ es biyectiva, entonces g es sobreyectiva y f es inyectiva.

Demostración. (a) Para este inciso, como f y g son inyectivas, entonces por el lema 2.6.4 (a) tenemos que $g \circ f$ es inyectiva. Del mismo modo, como f y g son sobreyectivas, entonces por el lema 2.6.9 (a), $g \circ f$ es sobreyectiva. Así, como $g \circ f$ es inyectiva y sobreyectiva, dicha función es biyectiva.

(b) Ya que $g \circ f$ es biyectiva, en particular es inyectiva, y por el lema 2.6.4 (b) tenemos que f es inyectiva. De manera análoga, como $g \circ f$ es sobreyectiva, por el lema 2.6.9 (b), g es inyectiva. $\square$

Además de lo anterior, por fin podremos usar un término que es conocido por muchos de sus cursos de preparatoria: la función inversa. Dada una función f, decimos que **f es invertible** si f tiene al menos un inverso izquierdo y un inverso derecho. Claramente, de los teoremas 2.6.5 y 2.6.10 se sigue el siguiente corolario:

Corolario 2.6.12. *Sea $f : A \to B$ una función. La función f es biyectiva si y sólo si f es invertible.*

Demostración. Supongamos que f es una función biyectiva. Dado que f es sobreyectiva, entonces f tiene al menos un inverso derecho. Por otro lado, como f es inyectiva, entonces tiene al menos un inverso izquierdo. De lo anterior, y de acuerdo a la definición de función invertible, f es invertible.

Ahora supondremos que f es una función invertible. Esto quiere decir que f tiene al menos un inverso izquierdo, por lo que f es inyectiva. Del mimo modo, como f tiene al menos un inverso derecho, entonces es sobreyectiva. De lo anterior, f es biyectiva. $\square$

Aunque el corolario 2.6.12 plantea una equivalencia de las funciones biyectiva, éstas vienen con un par de sorpresas más.

Lema 2.6.13. *Sea $f : A \to B$ una función. La función f es biyectiva si sólo si f^{-1} es una función de B en A.*

Demostración. Primero mostraremos la parte suficiente del enunciado, es decir, si f es biyectiva, entonces f^{-1} es una función de B en A.

Para ver que f^{-1} es una función de B en A, bastará revisar los siguientes incisos.

(i) $Dom(f^{-1}) = B$

Dado que f es una relación de A en B, entonces f^{-1} es una relación de B en A, por lo que $Dom(f^{-1}) \subseteq B$. Para demostrar la otra contención, consideremos un objeto $x \in B$. Dado que f es sobreyectiva, existe $a \in A$ tal que $f(a) = b$ o equivalentemente, existe $a \in A$ tal que $(a, b) \in f$. En tal caso, por definición de relación inversa, $(b, a) \in f^{-1}$, concluyendo que $b \in Dom(f^{-1})$. De lo anterior $Dom(f^{-1}) = B$

(ii) Si $(a, b) \in f^{-1}$ y $(c, d) \in f^{-1}$ y $a = c$, entonces $b = d$.

Dado que $(a, b) \in f^{-1}$ y $(c, d) \in f^{-1}$, entonces $(b, a) \in f$ y $(d, c) \in f$, es decir, $f(b) = a$ y $f(d) = c$. Sin embargo, dado que $a = c$, entonces tenemos que $f(b) = f(d)$ y por ser f inyectiva, entonces concluimos que $b = d$

De los puntos anteriores, f^{-1} es una función de B en A.

Ahora procederemos a demostrar la condición necesaria de la proposición,es decir, mostraremos que si f^{-1} es una función de B en A, entonces f es biyectiva.

Para mostrar que f efectivamente es biyectiva, verificaremos los siguientes incisos.

(i) f es suprayectiva.

Sea $b \in B$ un objeto cualquiera. Dado que f^{-1} es una función de B en A, entonces $Dom(f^{-1}) = B$ y con ello, $b \in Dom(f^{-1})$. Por definición de dominio de una relación, tenemos que existe $a \in A$ tal que $(b, a) \in f^{-1}$. Sin embargo, esto querría decir que existe $a \in A$ tal que $(a, b) \in f$ o equivalentemente, existe $a \in A$ tal que $f(a) = b$, concluyendo que f es sobreyectiva.

(ii) f es inyectiva.

Sean $x \in A$ y $z \in A$ tales que $f(x) = f(z)$. Solamente para simplificar notación, supongamos que $f(x) = f(z) = b$ con $b \in B$. En ta caso, tenemos que $(x, b) \in f$ y $(z, b) \in f$. Por definición de relación inversa se sigue que $(b, x) \in f^{-1}$ y $(b, z) \in f^{-1}$ y dado que f^{-1} es una función, entonces $x = z$.

De lo anterior concluimos que f es biyectiva. $\square$

La proposición que acabamos de demostrar es muy importante en muchos aspectos. En primer lugar ya habíamos visto que toda relación tiene una relación inversa, pero dada una función f de A en B, la única manera en que su relación inversa f^{-1} sea una función de B en A es cuando f es biyectiva. En tal caso, si f es biyectiva, podemos decir que f^{-1} es su **función inversa** en lugar de decir que f^{-1} es su *relación* inversa.

Por otro lado, podemos establecer la siguiente nota de acuerdo a lo que ya sabemos de relación inversa.

Nota 2.6.14. *Sean $f : A \to B$ una función biyectiva, $x \in A$ y $z \in B$.*

$$f(x) = z \text{ si y sólo si } f^{-1}(z) = x$$

Antes de continuar, quiero recalcar que la nota anterior es verdadera cuando la función es biyectiva. Gracias a lo anterior, tenemos el siguiente corolario.

Corolario 2.6.15. *Si $f : A \to B$ es una función biyectiva y $f^{-1} : B \to A$ es su función inversa, entonces los siguientes enunciados se satisfacen:*

a) *Si $x \in A$, entonces $f^{-1} \circ f(x) = x$*

b) *Si $z \in B$, entonces $f \circ f^{-1}(z) = z$*

Demostración. (a) Sea $x \in A$. Para simplificar la demostración, supondremos que $f(x) = b$. Nótese que en tal caso, $f^{-1}(b) = x$. Así, obtenemos que:

$$\begin{aligned} f^{-1} \circ f(x) &= f^{-1}(f(x)) \\ &= f^{-1}(b) \\ &= x \end{aligned}$$

De lo anterior, $f^{-1} \circ f(x) = x$, como era deseado.

(b) Este inciso es análogo al inciso previo. □

El corolario anterior tiene una forma alterna de ser descrito, y el cual está enunciado a continuación.

Corolario 2.6.16. *Si $f : A \to B$ es una función biyectiva y $f^{-1} : B \to A$ es su función inversa, entonces los siguientes enunciados se satisfacen:*

a) $f \circ f^{-1} = 1_B$

b) $f^{-1} \circ f = 1_A$

Retomemos un poco lo estudiado sobre funciones inyectivas y sobreyectivas. Ya sabemos que toda función sobreyectiva tiene al menos un inverso derecho, aunque no necesariamente es único. Y también sabemos que toda función inyectiva tiene al menos un inverso izquierdo, pero no necesariamente es único. Así, toda función biyectiva tiene al menos un inverso izquierdo y al menos un inverso derecho, pero con la novedad de que en el caso de las funciones biyectivas, sus inversos derechos e izquierdos sí son únicos y, de hecho, son iguales a f^{-1}. Esto lo mostraremos en el último enunciado de este capítulo.

Teorema 2.6.17. *Sea $f : A \to B$ una función biyectiva.*

a) Si $g : B \to A$ es un inverso izquierdo de f, entonces $g = f^{-1}$.

b) Si $h : B \to A$ es un inverso derecho de f, entonces $h = f^{-1}$.

Demostración. Primero demostraremos el punto (a), es decir, veremos que $g = f^{-1}$. Dicha igualdad la realizaremos usando nuevamente el teorema de igualdad de funciones.

Claramente $Dom(g) = Dom(f^{-1}) = B$, por lo que sólo resta demostrar que para todo $x \in B$ se satisface que $g(x) = f^{-1}(x)$.

Sea $x \in B$. Sabemos que $f \circ f^{-1}(x) = x$, por lo que $g(f \circ f^{-1}(x)) = g(x)$, es decir, $g(\,f(\,f^{-1}(x)\,)\,) = g(x)$ y en tal caso, como g es un inverso izquierdo de f, concluimos que $f^{-1}(x) = g(x)$. De lo anterior de demuestra que $f^{-1} = g$.

Ahora demostraremos el punto (b) utilizando el teorema de igualdad de funciones.

Nuevamente es sencillo verificar que $Dom(h) = Dom(f^{-1}) = B$, por lo que sólo resta demostrar que para todo $x \in B$, $h(x) = f^{-1}(x)$.

Sea $x \in B$. Dado que h es un inverso derecho de f, entonces $f \circ h(x) = x$. En tal caso, como $x \in Dom(f^{-1})$ entonces $f^{-1}(f \circ h(x)) = f^{-1}(x)$, es decir, $f^{-1}(\,f(\,h(x)\,)\,) = f^{-1}(x)$ y de acuerdo al corolario 2.6.15 se tiene que $h(x) = f^{-1}(x)$, concluyendo que $h = f^{-1}$. □

El teorema anterior tiene un par de cosas importantes a considerar, por un lado dice que el único inverso izquierdo de f es f^{-1}, lo cual no es cierto cuando f es simplemente inyectiva. De manera análoga, todos los inversos derechos de f son f^{-1}, lo cual tampoco pasa si f sólo es sobreyectiva. Más aun, una función que sea un inverso izquierdo y un inverso derecho de f tiene que ser necesariamente la función inversa de f.

Corolario 2.6.18. *Si $f : A \to B$ y $g : B \to A$ son funciones tales que $f \circ g = Id_B$ y $g \circ f = Id_A$, entonces f es biyectiva y $g = f^{-1}$.*

Demostración. Como $f \circ g = Id_B$, entonces g es un inverso derecho de f, por lo que f es sobreyectiva. De manera análoga, como $g \circ f = Id_A$, entonces g es un inverso izquierdo de f, concluyendo que f es inyectiva. De lo anterior, f es biyectiva.

Por último, como f es biyectiva y g es un inverso izquierdo de f, se sigue del teorema 2.6.17 (a) el que $g = f^{-1}$. □

Casi para finalizar este capítulo quiero aclara algo de notación. Previamente hemos usado la notación $f^{-1}[B]$ para hacer referencia a la imagen inversa de un objeto. Por lo general, cuando B sólo consta de un elemento, por ejemplo, $B = \{x\}$, no se suele usar la notación $f^{-1}[\{x\}]$, sino que es más frecuente omitir las llaves y escribir simplemente $f^{-1}[x]$. En esta situación hay que tener cuidado, pues $f^{-1}[x]$ es un *subconjunto* del dominio de f, mientras que $f^{-1}(x)$ hace referencia a un *elemento* del codominio de f^{-1} cuando f es invertible.

El último lema, enunciado a continuación, es un resultado técnico pero que será de gran utilidad en el capítulo siguiente.

Lema 2.6.19. *Sean A y B conjuntos y $f : A \to B$ una función biyectiva. Si $x \in A$ y $z \in B$, entonces existe una función biyectiva $g : A \to B$ tal que $g(x) = z$.*

Demostración. Dado que f es una función, consideremos $z' \in B$ tal que $f(x) = z'$ y por ser f sobreyectiva, consideremos $x' \in A$ tal que $f(x') = z$. En tal caso, la función $g : A \to B$ dada por:

$$g(a) = \begin{cases} z & \text{Si } a = x \\ z' & \text{Si } a = x' \\ f(a) & \text{En otro caso} \end{cases}$$

Es una función biyectiva de A en B y $g(x) = z$. □

2.7. Apéndice del Capítulo 2

Producto cartesiano.

a) $(x,z) \in A \times B$ b) $x \in A$ y $z \in B$	a) $(x,z) \notin A \times B$ b) $x \notin A$ o $z \notin B$

Dominio. Sean A y B conjuntos, R una relación de A en B y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $x \in Dom(R)$ b) $x \in A$ y existe $z \in B$ tal que $(x,z) \in R$	a) $x \notin Dom(R)$ b) $x \notin A$ o para todo $z \in B$, $(x,z) \notin R$

Imagen. Sean A y B conjuntos, R una relación de A en B y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $x \in Im(R)$ b) $x \in B$ y existe $z \in A$ tal que $(z,x) \in R$	a) $x \notin Im(R)$ b) $x \notin B$ o para todo $z \in A$, $(z,x) \notin R$

Relación inversa. Sean A y B conjuntos, R una relación de A en B y x y z objetos cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $(x,z) \in R^{-1}$ b) $(z,x) \in R$	a) $(x,z) \notin R^{-1}$ b) $(z,x) \notin R$

Orden parcial Sean A un conjunto y $\ll$ una relación sobre A. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) $\ll$ es un orden parcial en A

b) $\ll$ satisface las siguientes tres propiedades:

  b.1) Para todo $a \in A$, $a \ll a$.

  b.2) Para todo $\{a, b\} \subseteq A$ tales que $a \ll b$ y $b \ll a$, se satisface que $a = b$

  b.3) Para todo $\{a, b, c\} \subseteq A$ tales que $a \ll b$ y $b \ll c$ se satisface que $a \ll c$

Orden total Sean A un conjunto y $\ll$ una relación sobre A. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) $\ll$ es un orden total en A

b) $\ll$ satisface las siguientes cuatro propiedades:

  b.1) Para todo $a \in A$, $a \ll a$.

  b.2) Para todo $\{a, b\} \subseteq A$ tales que $a \ll b$ y $b \ll a$, se satisface que $a = b$.

  b.3) Para todo $\{a, b, c\} \subseteq A$ tales que $a \ll b$ y $b \ll c$ se satisface que $a \ll c$.

  b.4) Para todo $\{a, b\} \subseteq A$, $a \ll b$ o $b \ll a$.

Máximo Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es máximo en B

b) $x \in B$ y para todo $b \in B$, $b \ll x$.

a) x no es máximo en B.

b) $x \notin B$ o existe $b \in B$ tal que no es cierto que $b \ll x$.

Mínimo. Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es mínimo de B

b) $x \in B$ y para todo $b \in B$, $x \ll b$

a) x no es mínimo de B

b) $x \notin B$ o existe $b \in B$ tal que no es cierto que $x \ll b$

Cota superior Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es cota superior de B.

b) Para todo $b \in B$, $b \ll x$

a) x no es cota superior de B

b) Existe $b \in B$ tal que no es cierto que $b \ll x$

Cota inferior. Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es cota inferior de B.

b) Para todo $b \in B$, $x \ll b$.

a) x no es cota inferior de B.

b) Existe $b \in B$ tal que no es cierto que $x \ll b$.

Maximal Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es maximal de B

b) $x \in B$ y para todo $b \in B$ tal que $x \ll b$, se cumple que $x = b$

a) x no es maximal de B

b) $x \notin B$ o existe $b \in B$ tal que $x \ll b$ y $x \neq b$

Minimal Sean A un conjunto, $B \subseteq A$, $\ll$ un orden parcial sobre A y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) x es minimal de B

b) $x \in B$ y para todo $b \in B$ tal que $b \ll x$, se cumple que $x = b$

a) x no es minimal de B

b) $x \notin B$ o existe $b \in B$ tal que $b \ll x$ y $x \neq b$.

Relación de equivalencia Sean A un conjunto y $\sim$ una relación sobre A. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) $\sim$ es una relación de equivalencia en A

b) $\sim$ satisface las siguientes tres propiedades:

  b.1) Para todo $a \in A$, $a \sim a$.

  b.2) Para todo $\{a, b\} \subseteq A$ tal que $a \sim b$, se satisface que $b \sim a$.

  b.3) Para todo $\{a, b, c\} \subseteq A$ tal que $a \sim b$ y $b \sim c$, se satisface que $a \sim c$

Clase de equivalencia Sean A un conjunto, $\sim$ una relación de equivalencia sobre A y x y z dos objetos en A. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) $x \in [z]_\sim$

b) $x \sim z$

a) $x \notin [z]_\sim$

b) $x \not\sim z$

Función. Sean A y B conjuntos y η una relación de A en B. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) η es una función de A en B

b) η satisface las siguientes dos propiedades:

  b.1) $Dom(\eta) = A$

  b.2) Si $(a, b) \in \eta$, $(c, d) \in \eta$ y $a = c$, entonces $b = d$

Imagen (aplicado a funciones). Sean A y B conjuntos, $\eta : A \to B$ una función y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $x \in Im(\eta)$ b) $x \in B$ y existe $z \in A$ tal que $\eta(z) = x$	a) $x \notin Im(\eta)$ b) $x \notin B$ o para todo $z \in A$, $\eta(z) \neq x$.

Imagen directa de conjuntos. Sean A y B conjuntos, $\eta : A \to B$ una función, $C \subseteq A$ y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $x \in \eta[C]$ b) $x \in B$ y existe $z \in C$ tal que $\eta(z) = x$.	a) $x \notin \eta[C]$ b) $x \notin B$ o para todo $z \in C$, $\eta(z) \neq x$.

Imagen inversa de conjuntos. Sean A y B conjuntos, $\eta : A \to B$ una función, $D \subseteq B$ y x un objeto cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $x \in \eta^{-1}[D]$ b) $x \in A$ y existe $z \in D$ tal que $\eta(x) = z$.	a) $x \notin \eta^{-1}[C]$ b) $x \notin A$ o para todo $z \in C$, $\eta(x) \neq z$

Composición de funciones. Sean A, B y C conjuntos, $\eta : A \to B$ y $\alpha : B \to C$ dos funciones y x y z objetos cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $\alpha \circ \eta(x) = z$ b) $\alpha(\, \eta(x)\,) = z$	a) $\alpha \circ \eta(x) \neq z$ b) $\alpha(\, \eta(x)\,) \neq z$

Función inyectiva. Sean A y B conjuntos y $\eta : A \to B$ una función. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) η es inyectiva b) Si $\{x, z\} \subseteq A$ y $\eta(x) = \eta(z)$, entonces $x = z$	a) η no es inyectiva b) Existe $\{x, z\} \subseteq A$ tal que $\eta(x) = \eta(z)$ y $x \neq z$

Función sobreyectiva Sean A y B conjuntos y $\eta : A \to B$ una función. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) η es sobreyectiva b) Para todo $z \in B$, existe $x \in A$ tal que $\eta(x) = z$	a) η no es sobreyectiva b) Existe $z \in B$ tal que para todo $x \in A$, $\eta(x) \neq z$

Función biyectiva Sean A y B conjuntos y $\eta : A \to B$ una función. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) η es biyectiva b) η es inyectiva y η es sobreyectiva	a) η no es biyectiva b) η no es inyectiva o η no es sobreyectiva

Función inversa. Sean A y B conjuntos, $\eta : A \to B$ una función biyectiva y x y z objetos cualquiera. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $\eta^{-1}(x) = z$ b) $\eta(z) = x$	a) $\eta^{-1}(x) \neq z$ b) $\eta(z) \neq x$

Capítulo 3

Conteo

3.1. Números naturales

En esta sección introduciremos varios temas importantes, principalmente enfocados al conteo de objetos. Previo a ello, estableceremos los axiomas propuestos para la construcción de los números naturales. Dicho conjunto de axiomas fue propuesto por Peano en 1889 en *Aritmetices principia, nova methodo exposita*. Dado un conjunto 0, otro conjunto $\mathbb{N}$ cuyos elementos son llamados números naturales y una relación s sobre $\mathbb{N}$, los axiomas de Peano establecen que:

Axioma 1. $0 \in \mathbb{N}$.

Axioma 2. Si $n \in \mathbb{N}$, entonces hay un único natural m tal que $(n, m) \in s$. En particular, s es una función de $\mathbb{N}$ en $\mathbb{N}$.

Axioma 3. Para todo natural n, se satisface que $s(n) \neq 0$.

Axioma 4. Para cualquier par de naturales n y m, si $s(n) = s(m)$, entonces $n = m$.

Axioma 5. (Principio de Inducción) Si A es un subconjunto de los naturales y se cumple que (i) $0 \in A$ y que (ii) para todo $n \in \mathbb{N}$: si $n \in A$ implica que $s(n) \in A$, entonces $\mathbb{N} \subseteq A$.

Notemos que de los axiomas 2, 3 y 4, la relación s resulta ser una función inyectiva de $\mathbb{N}$ en $\mathbb{N} \setminus \{0\}$, a la cual llamaremos **función sucesor**. Aunado a lo anterior, von Neumann propuso una definición de los números naturales en términos de conjuntos y que además satisfacen los axiomas de Peano.

- $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$.
- Si n es un natural, definimos $s(n) = n \cup \{n\}$.

Nótese que en particular, $1 = s(0)$.

Para definir una suma y un producto en los números naturales utilizaremos funciones cuyo dominio es $\mathbb{N}\times\mathbb{N}$ y codominio es $\mathbb{N}$. Para la suma, dicha función será denotada por $+$ y para el producto dicha función será denotada ocasionalmente por $\cdot$. Además de lo anterior, será necesario aclarar que en el caso de operaciones será preferible la notación $a+b$ en lugar de la notación formal $+(a,b)$ para denotar la pareja de (a,b) bajo la función $+$. Del mismo modo, preferiremos la notación $a\cdot b$, o simplemente ab, en lugar de la notación $\cdot(a,b)$, para denotar la pareja de (a,b) bajo la función $\cdot$.

La definición de dichas operaciones será mediante un método conocido como **recursión** y tiene su fundamento en el teorema de recursión escrito a continuación.

Teorema 3.1.1 (Recursión). *Sean A un conjunto, x_0 un elemento de A y $f : A \to A$ una función. Existe una única función $g : \mathbb{N} \to A$ tal que:*

a) $g(0) = x_0$.

b) Para todo $n \in \mathbb{N}$ se satisface que $g \circ s(n) = f \circ g(n)$.

Nótese que, debido que que $g(0) = x_0$, se tiene entonces que $g(s(0)) = f(g(0))$, $g(s(1)) = f(g(1))$, $g(s(2)) = f(g(3))$ y así sucesivamente, es decir, la función g aplicada al sucesor de un natural n es lo mismo que la función f aplicada a $g(n)$.

Retomando la idea de definir una suma en $\mathbb{N}$, al utilizar el teorema de recursión para $A = \mathbb{N}$, $x_0 = 0$ y $f = s$, se tiene que existe una única función $g : \mathbb{N} \to \mathbb{N}$, denotada por $0+$, y a la que llamaremos **sumar cero**, tal que:

a) $0+(0) = 0$

b) Para todo n en $\mathbb{N}$, $0+(s(n)) = s(0+(n))$.

Es importante mencionar que el nombre de la función existente garantizada por el teorema de recursión se llama $0+$ y lo resaltamos con color magenta. Además, preferiremos la notación $0+n$ en lugar de $0+(n)$ para denotar a la pareja de n bajo la función $0+$.

La idea intuitiva de la función anterior es poder sumar un cero a cualquier número natural, pero también será importante poder sumar cosas no nulas, así usaremos nuevamente el teorema de recursión utilizando $A = \mathbb{N}$, $x_0 = k$ y $f = s$, para definir la función **sumar** k, denotada por $k+$ como la única función tal que:

a) $k+(0) = k$

b) Para todo n en $\mathbb{N}$, $k+(s(n)) = s(k+(n))$.

Al igual que en el caso de sumar cero, si n y k son naturales, preferiremos la notación $n+k$ en lugar de la notación $n+(k)$. Primero mostraremos el siguiente lema para después poder dar una definición más intuitiva de las funciones sumar k.

Lema 3.1.2. *Si n es un natural, entonces $s(n) = n+1$, es decir, el sucesor de n es $n+1$.*

Demostración. Consideremos un natural arbitrario n. Utilizando la definición de la función $n+$, se tiene que $n+(0) = n$ y, por otro lado, $n+(s(0)) = s(n+(0))$, con lo anterior, $n+(1) = s(n)$ y así, $s(n) = n+1$ □

Gracias al lema anterior podemos trabajar más intuitivamente la definición de sumar k. Dicha definición es una definición recursiva, es decir, inicia definiendo qué significa *sumar k a un primer natural*, en este caso, el cero, y lo define mediante el punto (a), es decir, $k+0 = k$. Después empezará a definir el sumar k al sucesor de un natural n mediante el significado de sumar k al valor n, lo cual lo establece en el punto (b), es decir, $k + s(n) = s(k+n)$. Sin embargo, por el lema 3.1.2, lo anterior puede quedar escrito de la siguiente manera: $k + (n+1) = (k+n)+1$. Así, la definición de sumar k puede reescribirse de la siguiente manera:

a) $k+(0) = k$

b) Para todo n en $\mathbb{N}$, $k+(n+1) = (k+(n))+1$.

Con lo anterior, definiremos la función suma $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por $+(n,m) = n+(m)$.

Lema 3.1.3. *Si n y m son naturales, entonces $m + (n+1) = (m+n)+1$.*

Demostración. Notemos que si n y m son naturales, entonces:

$$
\begin{aligned}
m + (n+1) &= m + (s(n)) && \text{Definición de sucesor}\\
&= s(m + (n)) && \text{Inciso (b) de la definición de } m+\\
&= (m+n)+1 && \text{Definición de sucesor}
\end{aligned}
$$

Esto muestra que si n y m son naturales, entonces $m + (n+1) = (m+n)+1$. □

Ahora, usando el mismo método, podremos definir la función producto en $\mathbb{N}$. Utilizaremos el teorema de recursión para $A = \mathbb{N}$, $x_0 = 0$ y $f = 0+$, se tiene que existe una única función de $\mathbb{N}$ en $\mathbb{N}$, denotada por $0\cdot$, y a la que llamaremos **producto por cero**, tal que:

a) $0\cdot(0) = 0$.

b) Para todo n en $\mathbb{N}$, $0\cdot(s(n)) = 0+(0\cdot(n))$.

Es importante mencionar que el nombre de la función existente garantizada por el teorema de recursión se llama $0\cdot$ y lo resaltamos con color violeta. Además, preferiremos la notación $0 \cdot n$ o simplemente $0n$, en lugar de $0\cdot(n)$. La idea intuitiva de la función anterior es poder realizar el producto por cero con cualquier número natural. Ahora procederemos a definir el producto por cualquier natural k. Usando nuevamente el teorema de recursión para $A = \mathbb{N}$, $x_0 = 0$ y $f = k+$, definimos la función **producto por** k, denotada por $k\cdot$ como la función tal que:

a) $k\cdot(0) = 0$.

b) Para todo n en $\mathbb{N}$, $k\cdot(s(n)) = k+(k\cdot(n))$.

Nuevamente, preferiremos la notación $n \cdot m$, o simplemente nm, en lugar de $n\cdot(m)$.

Al igual que la definición de sumar k, la definición de la función producto por k es una definición recursiva, es decir, inicia definiendo qué significa *producto por* k *a un primer natural*, en este caso, el cero, y lo define mediante el punto (a), es decir, $k \cdot 0 = 0$. Después empezará a definir el producto por k al sucesor de un natural n mediante el significado de producto por k al valor n, lo cual lo establece en el punto (b), es decir, $k \cdot s(n) = k + (k \cdot n)$. Sin embargo, por el lema 3.1.2, lo anterior puede quedar escrito de la siguiente manera: $k \cdot (n + 1) = k + (k \cdot n)$. Así, la definición de sumar k puede reescribirse de la siguiente manera:

a) $k\cdot(0) = 0$

b) Para todo n en $\mathbb{N}$, $k\cdot(n + 1) = k + (k\cdot n)$.

Con las definiciones anteriores, la función producto es la función $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ dada por $\cdot(n, m) = n\cdot(m)$.

El siguiente es un primer lema referente a este producto.

Lema 3.1.4. *Si n es un número natural, entonces $n \cdot 1 = n$.*

Demostración. Primero notemos que por la definición de producto por n se tiene que $n\cdot(s(0)) = n+(n\cdot\ (0))$. Sin embargo, ya que $s(0) = 1$ y que $n\cdot(0) = 0$, se sigue que $n\cdot(1) = n+(0)$ y con ello, $n \cdot 1 = n$. □

La función suma y producto en los naturales tiene muchas propiedades importantes, pero para poder demostrarlas, recurriremos a un nuevo tiempo de forma de demostración fundamentado en el Axioma 5 de Peano. Dicho método lo estudiaremos en la siguiente sección.

3.2. El Principio de Inducción Matemática

El Principio de Inducción Matemática genera un método de demostración matemático hasta ahora no usado en estas notas. Es una manera de verificar si objetos relacionados de alguna manera con los números naturales satisfacen una propiedad P. En este capítulo enunciaremos dos principios de inducción matemática y en esta sección desarrollaremos el primero de ellos. Recordemos que la versión axiomática del principio de inducción matemática es la siguiente:

Axioma 5. Si A es un subconjunto de los naturales y se cumple que (i) $0 \in A$ y que (ii) para todo $n \in \mathbb{N}$: si $n \in A$ implica que $s(n) \in A$, entonces $\mathbb{N} \subseteq A$.

El Principio de inducción matemática en su versión axiomática brinda condiciones suficientes para que un subconjunto de los naturales sea igual a todo el conjunto de los naturales. Sin embargo, existe una forma alterna de usarlo, mismo que será la que nos interese para este curso.

Teorema 3.2.1 (Principio de Inducción). *Sea P un predicado que toma valores sobre los números naturales. Para verificar que P es verdadero para todo natural mayor o igual que algún natural k_0 basta realizar lo siguiente:*

a) Base Inductiva. Demostrar que $P(k_0)$ es verdadero.

b) Hipótesis de Inducción. Suponemos que $P(n)$ es verdadera para algún $n \geq k_0$.

c) Paso Inductivo. Se demuestra que $P(n+1)$ es verdadera.

Entonces P es verdadera para todo natural mayor o igual que k_0.

El enunciado anterior lo usaremos en muchísimas ocasiones para hacer demostraciones matemáticas. En la siguiente subsección estableceremos el preámbulo previo a realizar una demostración inductiva.

3.2.1. Inducción de enunciados

Aunque en estas secciones utilizaremos inducción para demostrar diversas propiedades de la suma y el producto de naturales, los métodos inductivos sirven para realizar demostraciones de muy diversos tipos. Para ello necesitaremos tener presente las indicaciones del teorema 3.2.1 y un enunciado lógico en el que haya al menos un natural implícito o explícito y que sea variable. Bajo estas condiciones, es posible realizar una demostración inductiva para probar la veracidad del enunciado.

Para ejemplificar lo anterior presupondremos momentáneamente algunas propiedades de los naturales y utilizaremos el siguiente enunciado para realizar el proceso de inducción matemática en un enunciado más general.

Ejemplo 3.2.2. *Sean A un conjunto y $\trianglelefteq$ un orden total sobre A. Si B es un subconjunto de A, es no vacío y finito, entonces B tiene máximo bajo el orden $\trianglelefteq$.*

Al igual que en muchas demostraciones de nuestro curso, antes de iniciar la demostración desarrollaremos un preámbulo que será de utilidad al momento de iniciar nuestra demostración. Es importante que prestes atención a todos los pasos que realizaremos previo a iniciar la demostración.

Paso 1 **La preparación del enunciado a demostrar.**

Primero es importante saber escribir el enunciado de tal forma que sea evidente que hay un número natural variable en dicho enunciado. A dicho natural le llamaremos **variable inductiva**. En algunas ocasiones la variable inductiva estará de forma explícita en el enunciado, en otras ocasiones, como en el ejemplo 3.2.2, no será explícita. En tal caso lo primero que debemos hacer es reescribir el enunciado. Para nuestro ejercicio, el que un conjunto sea finito significa que tiene n elementos para algún $n \in \mathbb{N}$. Así, podemos reescribir el enunciado como sigue:

Si B es un subconjunto de A, es no vacío y tiene n elementos,
entonces B tiene máximo bajo el orden $\trianglelefteq$.

Así escrito nuestro enunciado, propondremos que la variable inductiva es el valor n.

Paso 2 **El inicio de tu redacción.**

Ahora que tenemos un natural explícitamente en el enunciado, vamos a indicar que iniciaremos una demostración por inducción sobre dicho enunciado, y además indicaremos quién será nuestra variable inductiva. Esto es importante porque (i) contextualiza a la persona que leerá tu inducción sobre qué es lo que vas a hacer para llevar a cabo la inducción y (ii) si existen varios naturales en tu enunciado, le indicas a la persona que está leyendo cuál de ellos será la variable inductiva. Así, tu demostración inductiva debe iniciar con una frase como:

Mostraremos por inducción sobre n que: si B es un subconjunto de A, es no vacío y tiene n elementos, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Paso 3 **La escritura de la base inductiva.**

El inicio de nuestra inducción es la base inductiva. En este caso volveremos a escribir todo el enunciado que queremos demostrar (el que obtuvimos del paso 1) y ahí en donde aparezca la variable inductiva, la sustituiremos por el primer natural posible de acuerdo a nuestro enunciado.

En este caso, como B *es un conjunto no vacío*, no es posible que B tenga 0 elementos, por lo que nuestra base inductiva no iniciará en cero, sino en 1. Es importante que prestes atención a este detalle porque no todas nuestras inducciones iniciarán desde cero. Así, el enunciado debe quedar escrito como:

Si B es un subconjunto de A, es no vacío y con 1 elemento, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Recuerda que la base de inducción sí se demuestra, tiene sus propias hipótesis y su propio objetivo a demostrar. Por ahora, no realizaremos la demostración, sólo estamos realizando el preámbulo de nuestra demostración inductiva.

Paso 4 **La escritura de la hipótesis de inducción.**

El siguiente paso de nuestra inducción es la hipótesis inductiva. En este caso, nuevamente volveremos a escribir todo el enunciado que queremos demostrar (el que obtuvimos del paso 1) y ahí en donde aparezca la variable inductiva, la sustituiremos por un natural al que le pondremos un nombre de variable, por ejemplo, k. Es importante decir que, aunque es un nombre de variable, una vez que lo establezcamos obtendrá categoría de objeto fijo dentro de nuestra demostración. También es importante indicar que dicho natural debe ser mayor o igual que el natural que elegimos en nuestra base inductiva.

Otro punto que es importante mencionar en este paso es que es conveniente cambiar el nombre de los objetos que dependen de la variable k. En este caso, el conjunto B depende de quién es k en el sentido de que un conjunto varía si varía el número de elementos que tiene, por ejemplo, un conjunto B que tiene k elementos no puede ser igual que un conjunto que tiene $k+1$ elementos. También recomiendo que cambie el nombre de aquellos objetos que son variables por sí mismos dentro del enunciado lógico. Este paso, aunque difícil, puede ayudar muchísimo al momento de usar la hipótesis inductiva durante el paso inductivo, por lo que te animo a hacerlo cuando te sea posible. En mi caso, pondré B' en lugar de B.

Por último, indicaremos en este paso que presupondremos que k es un natural que hace verdadero el enunciado que vamos a demostrar por inducción, así, nuestro enunciado en la hipótesis de inducción debe decir algo como:

Supondremos que k es un natural tal que si B' es un subconjunto de A, es no vacío y con $k \geq 1$ elementos, entonces B' tiene máximo bajo el orden $\trianglelefteq$.

Este paso no se demuestra, se presupone. Como tal, el enunciado obtenido a partir de este paso es una hipótesis y tiene su propia forma de ser usado de manera directa, ya sea que se trate de una condicional, un cuantificador universal, un cuantificador existencial, etc.

Paso 5 **La escritura del paso inductivo.**

El último paso es el paso inductivo. Nuevamente volveremos a escribir todo el enunciado que queremos demostrar (el que obtuvimos del paso 1) y ahí en donde aparezca la variable inductiva, la sustituiremos por el sucesor del natural que elegimos en el paso previo, en nuestro caso, $k+1$. Fuera de esto, no hay mucho más que cambiar en este punto, por lo que nuestro enunciado quedaría algo de la siguiente manera.

Si B es un subconjunto de A, es no vacío y con $k+1 \geq 1$ elementos, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Este último paso debe demostrarse. No pierdas de vista que dicho paso te dará sus propias hipótesis, sus propios objetos al iniciar la demostración y su propio objetivo a demostrar. Eventualmente, durante la demostración de este enunciado, deberás usar la hipótesis inductiva, por lo que debes estar muy atento al memento en que puedas usar dicha hipótesis.

Todos los pasos anteriores fueron el preámbulo para nuestra demostración. Ahora que ya lo hemos realizado, procederemos a demostrar el enunciado.

Ejemplo 3.2.3. *Sean A un conjunto y $\trianglelefteq$ un orden total sobre A. Si B es un subconjunto de A, es no vacío y finito, entonces B tiene máximo bajo el orden $\trianglelefteq$.*

Demostración. Mostraremos por inducción sobre n que: si B es un subconjunto de A, es no vacío y tiene n elementos, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Base de inducción. Si B es un subconjunto de A, es no vacío y tiene 1 elemento, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Para demostrar este paso, notemos que $B = \{x\}$ para algún $x \in A$, pues por hipótesis de este enunciado, B es un subconjunto de A y tiene un único elemento. Así, notemos que x es un máximo de B pues satisface que (i) $x \in B$ y (ii) para todo $z \in B$, $z \trianglelefteq x$, concluyendo la demostración de la base inductiva.

Hipótesis de inducción. Supondremos que k es un natural tal que si B' es un subconjunto de A, es no vacío y con $k \geq 1$ elementos, entonces B' tiene máximo bajo el orden $\trianglelefteq$.

Paso inductivo. Si B es un subconjunto de A, es no vacío y con $k+1 \geq 1$ elementos, entonces B tiene máximo bajo el orden $\trianglelefteq$.

Gracias a la base inductiva, podemos presuponer que $k+1 \geq 2$, es decir, $k \geq 1$.

Por otro lado, consideremos $x \in B$ arbitrario. Ahora sea $B' = B \setminus \{x\}$. Notemos que $B' \subseteq A$, B' es no vacío y B' tiene k elementos, por lo que por hipótesis de inducción, B' tiene un máximo bajo $\trianglelefteq$, digamos x'. Ya que $\trianglelefteq$ es un orden total sobre A, entonces $x' \trianglelefteq x$ o $x \trianglelefteq x'$, por lo que tenemos los siguientes casos.

Caso 1. $x \trianglelefteq x'$.

En este caso, x' es máximo de B.

Caso 2. $x' \trianglelefteq x$.

En este caso, x es máximo de B.

En cualquiera de los casos anteriores se concluye que B tiene máximo, probando así el paso inductivo.

Por el principio de inducción matemática, si B es un subconjunto de A, es no vacío y finito, entonces B tiene máximo bajo el orden $\trianglelefteq$. $\square$

El anterior es un ejemplo clásico de una demostración inductiva, por lo que te recomiendo que lo repases las veces que sean necesarias, así como los pasos que realizamos previo a la demostración.

Para finalizar esta subsección, quiero mencionar que los pasos descritos previamente son importantes de realizar, sin embargo, no son un método absoluto para escribir todas las demostraciones inductivas a las que te enfrentarás. En algunas ocasiones será necesario saber reescribir todo el enunciado para una mejor comprensión y uso de la inducción matemática, pero eso lo conseguirás poco a poco con la experiencia y con la práctica. Por mientras, considera estos pasos como un punto de referencia para futuras demostraciones inductivas.

Ahora realizaremos un par de inducciones más, ambas referentes a la función sucesor. Presta atención al uso de los pasos anteriores en estas demostraciones.

Lema 3.2.4. *Para todo* $n \in \mathbb{N}$, $s(n) \neq n$.

Demostración. Mostraremos por inducción matemática sobre n que: $s(n) \neq n$.

Base de inducción. $s(0) \neq 0$.

Por el axioma 3 de Peano, en particular $s(0) \neq 0$.

Hipótesis de inducción. Supongamos que k es un natural tal que $s(k) \neq k$.

Paso inductivo. $s(k+1) \neq k+1$.

Procediendo por contradicción, supongamos que $s(k+1) = k+1$. Sabemos que $s(k) = k+1$, por lo que $s(s(k)) = s(k)$. Ya que s es una función inyectiva, entonces $s(k) = k$, lo cual contradice la hipótesis de inducción. Con lo anterior, $s(k+1) \neq k+1$.

□

Teorema 3.2.5. *Para todo $n \in \mathbb{N}$ se tiene que $n = 0$ o existe $m \in \mathbb{N}$ tal que $s(m) = n$.*

Demostración. Mostraremos por inducción matemática sobre n que: $n = 0$ o existe $m \in \mathbb{N}$ tal que $s(m) = n$.

Base de inducción. $0 = 0$ o existe $m \in \mathbb{N}$ tal que $s(m) = 0$.

Como $0 = 0$, entonces la base de inducción se satisface.

Hipótesis de inducción. Supongamos que k es un natural tal que $k = 0$ o existe $m' \in \mathbb{N}$ tal que $s(m') = k$.

Paso inductivo. $k+1 = 0$ o existe $m \in \mathbb{N}$ tal que $s(m) = k+1$.

Para el paso inductivo, mostraremos que existe $m \in \mathbb{N}$ tal que $s(m) = k+1$. Como $s(k) = k+1$ y $k \in \mathbb{N}$, entonces por hipótesis de inducción se tiene que $k = 0$ o existe $m' \in \mathbb{N}$ tal que $s(m') = k$. Ahora revisaremos ambos casos.

Caso 1. $k = 0$.

En este caso, $k+1 = 1$, por lo que existe $0 \in \mathbb{N}$ tal que $s(0) = k+1$.

Caso 2. Existe $m' \in \mathbb{N}$ tal que $s(m') = k$.

Para este caso, $s(m') \in \mathbb{N}$ y satisface que $s(s(m')) = s(k)$, es decir, existe $s(m') \in \mathbb{N}$ tal que $s(s(m')) = k+1$.

De los casos anteriores, existe $m \in \mathbb{N}$ tal que $s(m) = k+1$.

□

El teorema anterior es importante debido a que establece que si n es un natural no nulo, entonces n es el sucesor de algún natural m. Debido al axioma 4 de Peano, m debe ser único por lo que recibirá el nombre de **antecesor de** n y se denotará por $n-1$. Es importante recalcar que solamente los naturales distintos de cero tienen antecesor.

3.3. Propiedades de la suma y el producto de naturales

Ahora veremos algunos ejemplos de inducción aplicados a propiedades de suma y producto de naturales. Aunque hay varios resultados inductivos en esta sección, quiero mencionar que los resultados importantes son los lemas 3.3.4, 3.3.5 y 3.3.11.

Lema 3.3.1. *Si* $n \in \mathbb{N}$, *entonces* $0 + n = n + 0$.

Demostración. Primero mostraremos por inducción sobre n que: $0 + n = n$.

Base de inducción. $0 + 0 = 0$

Hipótesis de inducción. Supongamos que k es un natural tal que $0 + k = k$.

Paso inductivo. $0 + (k + 1) = k + 1$.

Notemos que:

$$\begin{aligned}
0 + (k+1) &= 0 + (s(k)) && \text{Definición de sucesor} \\
&= s(0 + (k)) && \text{Inciso (b) de la definición de } 0+ \\
&= s(k) && \text{Hipótesis de inducción sobre } k \\
&= k + 1 && \text{Definición de sucesor de} k
\end{aligned}$$

Con lo anterior podemos concluir que $0 + (k + 1) = k + 1$.

Por el principio de inducción matemática, si $k \in \mathbb{N}$, entonces $0 + n = n$.

Por otro lado, se sigue del punto (a) de la definición de $n+$ que $n + 0 = n$, por lo que $n + 0 = n = 0 + n$, concluyendo lo deseado. □

Lema 3.3.2. *Si* $n \in \mathbb{N}$, *entonces* $1 + n = n + 1$.

Demostración. Mostraremos por inducción sobre n que $1 + n = n + 1$.

Base de inducción. $1 + 0 = 0 + 1$.

Esta demostración se sigue del lema 3.3.1.

Hipótesis de inducción. Supongamos que $n \in \mathbb{N}$ es tal que $1 + n = n + 1$.

Paso inductivo. $1 + (n + 1) = (n + 1) + 1$.

Para esta demostración, notemos lo siguiente:

$$\begin{aligned}
1+(n+1) &= 1+(s(n)) && \text{Definición de sucesor}\\
&= s(1+(n)) && \text{Inciso (b) de la definición de } 1+\\
&= s(n+1) && \text{Hipótesis de inducción}\\
&= (n+1)+1 && \text{Definición de sucesor}
\end{aligned}$$

Con lo anterior, $1+(n+1)=(n+1)+1$, concluyendo nuestro paso inductivo.

Por el Principio de inducción matemática podemos afirmar que si $n \in \mathbb{N}$, entonces $1+n=n+1$. $\square$

Lema 3.3.3. *Si n y m son naturales, entonces $(1+m)+n=1+(m+n)$.*

Demostración. Mostraremos por inducción sobre n que: $(1+m)+n=1+(m+n)$

Base de inducción. $(1+m)+0=1+(m+0)$.

En este caso, $(1+m)+0=1+m$ y $1+(m+0)=1+(m)$, concluyendo que $(1+m)+0=1+(m+0)$.

Hipótesis de inducción. Supongamos que k es un natural tal que $(1+m)+k=1+(m+k)$.

Paso inductivo. $(1+m)+(k+1)=1+(m+(k+1))$.

Para esta demostración, notemos lo siguiente:

$$\begin{aligned}
(1+m)+(k+1) &= (1+m)+(s(k)) && \text{Definición de sucesor}\\
&= s((1+m)+(k)) && \text{Inciso (b) de la definición de } (1+m)+\\
&= s(1+(m+k)) && \text{Hipótesis de inducción sobre } k\\
&= (1+(m+k))+1 && \text{Definición de sucesor}\\
&= 1+((m+k)+1) && \text{Lema 3.1.3}\\
&= 1+(m+(k+1)) && \text{Lema 3.1.3}
\end{aligned}$$

Con lo anterior, $(1+m)+(k+1)=1+(m+(k+1))$, concluyendo nuestro paso inductivo.

Por el Principio de inducción matemática podemos afirmar que si n y m son naturales, entonces $(1+m)+n=1+(m+n)$. $\square$

Por otro lado, tenemos las siguientes propiedades básicas de la suma de números enteros.

Lema 3.3.4 (Propiedades de la suma). *Para cualesquiera naturales m, n y k, se satisfacen los siguientes incisos:*

a) $0 + n = n + 0 = n$.

b) $m + n = n + m$.

c) $(m + n) + r = m + (n + r)$.

d) *Si* $m + n = m + k$, *entonces* $n = k$.

e) *Si* $n \neq 0$, *entonces* $n + m \neq 0$.

f) $n + m = 0$ *si y sólo si* $n = 0$ *y* $m = 0$.

Demostración. Las siguientes son demostraciones en su mayoría inductivas, por lo que en algunos enunciados reescribiremos lo que hay que demostrar para facilitar la demostración inductiva.

(a) Este enunciado se sigue del lema 3.3.1 y del punto (a) de la definición de $n+$.

(b) Mostraremos este enunciado por inducción sobre n, es decir, mostraremos que: si $m \in \mathbb{N}$, entonces $n + m = m + n$.

Base de inducción. Si $m \in \mathbb{N}$, entonces $0 + m = m + 0$.

Esta demostración se deduce del lema 3.3.1.

Hipótesis de inducción. Supongamos que n es un natural tal que si $m \in \mathbb{N}$, entonces $n+m = m+n$.

Paso inductivo. Si $m \in \mathbb{N}$, entonces $(n + 1) + m = m + (n + 1)$

Sean $m \in \mathbb{N}$. Notemos lo siguiente:

$$
\begin{aligned}
(n+1)+m &= (1+n)+m && \text{El natural } 1 \text{ conmuta}\\
&= 1+(n+m) && \text{Lema 3.3.3}\\
&= 1+(m+n) && \text{Hipótesis de inducción sobre } n\\
&= (m+n)+1 && \text{Lema 3.3.2}\\
&= m+(n+1) && \text{Lema 3.1.3}
\end{aligned}
$$

Por lo anterior, podemos afirmar que $(n + 1) + m = n + (m + 1)$.

Gracias al principio de inducción matemática, se concluye que si $m \in \mathbb{N}$, entonces $n + m = m + n$.

(c) Mostraremos este enunciado por inducción sobre m, es decir, mostraremos que: para todo $\{n, r\} \subseteq \mathbb{N}$, se satisface que $(m + n) + r = m + (n + r)$.

Base de inducción. Para todo $\{a, b\} \subseteq \mathbb{N}$, se satisface que $(0 + a) + b = 0 + (a + b)$.

En este caso, por definición de $0+$ se tiene que $0+a = a$, por lo que $(0+a)+b = a+b$. Por otro lado, nuevamente por definición de $0+$ se tiene que $0+(a+b) = a+b$. Así, $(0+a)+b = 0+(a+b)$.

Hipótesis de inducción. Supongamos que m es un natural tal que para todo $\{a', b'\} \subseteq \mathbb{N}$, se satisface que $(m + a') + b' = m + (a' + b')$.

Paso inductivo. Para todo $\{a, b\} \subseteq \mathbb{N}$, se satisface que $(m + 1) + (a + b) = ((m + 1) + a) + b$.

Para esta demostración notemos que:

$$\begin{aligned}
((m+1)+a)+b &= ((1+m)+a)+b && \text{El natural } 1 \text{ conmuta} \\
&= (1+(m+a))+b && \text{Lema } 3.3.3 \\
&= 1+((m+a)+b) && \text{Lema } 3.3.3 \\
&= 1+(m+(a+b)) && \text{Hipótesis de inducción sobre } m \\
&= (1+m)+(a+b) && \text{Lema } 3.3.3 \\
&= (m+1)+(a+b) && \text{El natural } 1 \text{ conmuta}
\end{aligned}$$

Con lo anterior, podemos concluir que $(m + 1) + (a + b) = ((m + 1) + a) + b$.

Por el Principio de inducción matemática, para todo $\{n, r\} \subseteq \mathbb{N}$, se satisface que $(m + n) + r = m + (n + r)$.

(d) Mostraremos por inducción sobre m este inciso, es decir, mostraremos que: para todo $\{k, n\} \subseteq \mathbb{N}$ tales que $m + n = m + k$, se satisface que $n = k$.

Base de inducción. Para todo $\{a, b\} \subseteq \mathbb{N}$ tales que $0 + a = 0 + b$, se satisface que $a = b$.

Sean $\{a, b\} \subseteq \mathbb{N}$ tales que $0 + a = 0 + b$, se sigue de la definición de $0+$ que $a = b$.

Hipótesis de inducción. Supongamos que m es un natural tal que para todo $\{k', n'\} \subseteq \mathbb{N}$ tales que $m + n' = m + k'$, se satisface que $n' = k'$.

Paso inductivo. Para todo $\{k, n\} \subseteq \mathbb{N}$ tales que $(m+1)+n = (m+1)+k$, se satisface que $n = k$.

Sean $\{k, n\} \subseteq \mathbb{N}$ tales que $(m + 1) + n = (m + 1) + k$. Gracias al inciso previo, se sigue que $m+(n+1) = m+(k+1)$ y por hipótesis de inducción, se sigue que $n+1 = k+1$, concluyendo del axioma 4 de Peano que $n = k$.

Gracias a lo anterior, por el principio de inducción matemática, se concluye que si $m + n = m + k$, entonces $n = k$.

(e) Dado que $n \neq 0$, entonces n es el sucesor de algún natural, es decir, $n = r + 1$ para algún $r \in \mathbb{N}$. Así, $n + m = (r + 1) + m = (r + m) + 1$, por lo que $n + m$ es el sucesor de $r + m$. Debido a que 0 no es sucesor de ningún natural, entonces $n + m \neq 0$.

(f) Primero mostraremos la parte necesaria, es decir, si $n = 0$ y $m = 0$, entonces $n + m = 0$.

Por la definición de suma, $n + m = 0 + (m) = m = 0$.

Ahora mostraremos la parte suficiente, es decir, si $n+m = 0$, entonces $n = 0$ y $m = 0$. Procederemos por contradicción al suponer que $n \neq 0$ o $m \neq 0$, y supongamos sin pérdida de generalidad que $n \neq 0$. Así, por el inciso previo, $n + m \neq 0$, lo cual no es posible, concluyendo que $n = 0$ y $m = 0$.

□

Antes de continuar con las propiedades del producto, tenemos el siguiente resultado.

Lema 3.3.5. *Si n y m son naturales, entonces existe $k \in \mathbb{N}$ tal que $n + k = m$ o $m + k = n$.*

Demostración. La demostración la realizaremos por inducción sobre n, es decir, mostraremos que: si $m \in \mathbb{N}$, entonces existe $k \in \mathbb{N}$ tal que $n + k = m$ o $m + k = n$.

Base de inducción. Si $m \in \mathbb{N}$, entonces existe $k \in \mathbb{N}$ tal que $0 + k = m$ o $m + k = 0$.

Para este inciso, notemos que $0 + m = m$, por lo que existe $m \in \mathbb{N}$ tal que $0 + m = m$.

Hipótesis de inducción. Supongamos que $n \in \mathbb{N}$ es tal que si $m \in \mathbb{N}$, entonces existe $k' \in \mathbb{N}$ tal que $n + k' = m$ o $m + k' = n$.

Paso inductivo. Si $m \in \mathbb{N}$, entonces existe $k \in \mathbb{N}$ tal que $(n + 1) + k = m$ o $m + k = (n + 1)$.

En este caso, directamente por hipótesis de inducción existe existe $k' \in \mathbb{N}$ tal que $n+k' = m$ o $m+k' = n$. Consideremos los siguientes casos.

Caso 1. $m + k' = n$.

Para este caso, notemos que $m+(k'+1) = n+1$, por lo que existe $k'+1 \in \mathbb{N}$ tal que $m+(k'+1) = n + 1$, como es deseado.

Caso 2. $n + k' = m$.

En este caso, si $k' = 0$, se seguiría que $n = m$ y así, existe $1 \in \mathbb{N}$ tal que $m + 1 = n + 1$.

Ahora veremos el caso en el que $k' \neq 0$. En estas circunstancias, k' es el sucesor de un natural, es decir, $k' = r + 1$ para algún $r \in \mathbb{N}$. Con ello:

$$\begin{aligned} m &= n + k' \\ &= n + (r + 1) \\ &= (n + 1) + r \end{aligned}$$

Así se concluye que existe $r \in \mathbb{N}$ tal que $(n + 1) + r = m$.

En cualquiera de los casos anteriores podemos garantizar que existe $k \in \mathbb{N}$ tal que $(n+1)+k = m$ o $m+k = (n+1)$.

Por el principio de inducción matemática, si n y m son naturales, entonces existe $k \in \mathbb{N}$ tal que $n+k = m$ o $m+k = n$. □

De manera análoga a los resultados anteriores, tenemos algunas propiedades del producto que es importante mencionar.

Lema 3.3.6. *Si n es un natural, entonces $0n = 0$.*

Demostración. La demostración la realizaremos por inducción sobre n, es decir, mostraremos que: si n es un natural, entonces $0n = 0$.

Base de inducción. Si 0 es un natural, entonces $0 \cdot 0 = 0 \cdot 0$.

Este inciso se sigue inmediatamente de la definición de $0\cdot$

Base de inducción. Supongamos que m es un natural tal que $0m = 0$.

Base de inducción. Si $m+1$ es un natural, entonces $0(m+1) = 0$.

Sea $m+1$ un natural. Notemos que:

$$
\begin{aligned}
0(m+1) &= 0 \cdot (m+1) && \text{Definición de producto}\\
&= 0 \cdot (s(m)) && \text{Definición de sucesor de } m\\
&= 0 + (0 \cdot (m)) && \text{Punto (b) de la definición de la función } 0\cdot\\
&= 0 + (0) && \text{Hipótesis de inducción sobre } m\\
&= 0 && \text{El natural } 0 \text{ es neutro aditivo}
\end{aligned}
$$

Con lo anterior, $0(m+1) = 0$

Gracias al principio de inducción matemática, podemos afirmar que si n es un natural, entonces $0n = 0$. □

Corolario 3.3.7. *Si n es un natural, entonces $n0 = 0n = 0$.*

Demostración. Sabemos que para todo natural n se cumple que $n0 = 0$ como lo establece la propiedad (a) de la definición de $n\cdot$. Por otro lado, el lema 3.3.6 establece que $0n = n$, mostrando así que $0n = n0$. □

Lema 3.3.8. *Si n es un natural, entonces $1n = n$.*

Demostración. La demostración la realizaremos por inducción sobre n, es decir, mostraremos que: si n es un natural, entonces $1n = n$.

Base de inducción. Si 0 es un natural, entonces $1 \cdot 0 = 0$.

Este inciso se sigue inmediatamente del corolario 3.3.7.

Hipótesis de inducción. Supongamos que m es un natural tal que $1m = m$.

Paso inductivo. Si $m+1$ es un natural, entonces $1(m+1) = m+1$.

Sea $m+1$ un natural. Notemos que:

$$\begin{aligned}
1(m+1) &= 1 \cdot (m+1) && \text{Definición de producto}\\
&= 1 \cdot (s(m)) && \text{Definición de sucesor de } m\\
&= 1 + (1 \cdot (m)) && \text{Punto (b) de la definición de la función } 1\cdot\\
&= 1 + (m) && \text{Hipótesis de inducción sobre } m\\
&= m + 1 && \text{La suma es conmutativa}
\end{aligned}$$

Con lo anterior, $1(m+1) = m+1$

Gracias al principio de inducción matemática, podemos afirmar que si n es un natural, entonces $1n = n$. □

Como consecuencia, tenemos el siguiente corolario.

Corolario 3.3.9. *Si n es un natural, entonces $n1 = 1n = n$.*

Demostración. Sea n un natural. Por el lema 3.1.4 tenemos que $n1 = n$, y por el lema 3.3.8 tenemos también que $1n = n$, concluyendo que $n1 = 1n$. □

Lema 3.3.10. *Si n y m son naturales, entonces $(n+1)m = nm + m$.*

Demostración. La demostración la realizaremos por inducción sobre m, es decir, mostraremos que: si m es un natural, entonces $(n+1)m = nm + m$.

Base de inducción. Si 0 es un natural, entonces $(n+1)0 = n0 + 0$.

Por definición de la función $(n+1)\cdot$ podemos garantizar que $(n+1)0 = 0$, y por otro lado, $n0 + 0 = 0$, concluyendo que $(n+1)0 = n0 + 0$.

Hipótesis de inducción. Supongamos que k es un natural tal que $(n+1)k = nk + k$.

Paso inductivo. Si $k+1$ es un natural, entonces $(n+1)(k+1) = n(k+1) + (k+1)$.

Para esta demostración, notemos lo siguiente:

$$\begin{aligned}
(n+1)(k+1) &= (n+1)\cdot(k+1) && \text{Definición de producto}\\
&= (n+1)\cdot(s(k)) && \text{Definición de sucesor de } k\\
&= (n+1) + ((n+1)\cdot(k)) && \text{Punto (b) de la definición de la función } (n+1)\cdot\\
&= (n+1) + (nk+k) && \text{Hipótesis de inducción sobre } k\\
&= (n+nm) + (m+1) && \text{La suma es conmutativa y asociativa}\\
&= n\cdot(s(m)) + (m+1) && \text{Punto (b) de la definición de la función } n\cdot\\
&= n(m+1) + (m+1) && \text{Definición de sucesor de } m
\end{aligned}$$

De las igualdades anteriores, se tiene que $(n+1)(k+1) = n(k+1) + (k+1)$.

Por el Principio de inducción matemática, se concluye que si n y m son naturales, entonces $(n+1)m = nm+m$. □

Ahora presentaremos algunas propiedades básicas del producto en los naturales. Algunas de estas demostraciones ya fueron realizadas previamente, pero las incluiremos en este enunciado a modo de recopilación.

Lema 3.3.11 (Propiedades del producto)**.** *Para cualesquiera naturales m, n y k, se satisfacen los siguientes incisos:*

a) $n0 = 0n = 0$.

b) $1n = n1 = n$.

c) $mn = nm$.

d) $m(n+k) = mn + mk$.

e) $(nm)k = n(mk)$

f) $mn = 0$ *si y sólo si* $m = 0$ *o* $n = 0$.

g) Si $mn = mk$ *y* $m \neq 0$, *entonces* $n = k$.

Demostración. La mayoría de estas demostraciones son demostraciones inductivas, por lo que en algunos incisos cambiaremos la redacción del enunciado para hacer más fácil su demostración inductiva.

(a) Este enunciado es el corolario 3.3.7.

(b) Este enunciado es el corolario 3.3.9.

(c) Esta demostración la realizaremos por inducción sobre n. Es decir, mostraremos que: si m es un natural, entonces $mn = nm$.

Base de inducción. Si m es un natural, entonces $m0 = 0m$.

En este caso, sabemos que $m0 = 0$ y $0m = 0$, concluyendo que $m0 = 0m$.

Hipótesis de inducción. Supongamos que k es un natural tal que si $m \in \mathbb{N}$, entonces $mk = km$.

Paso inductivo. Si m es un natural, entonces $m(k+1) = (k+1)m$.

Para esta demostración notemos lo siguiente:

$m(k+1)$	$= m \cdot (k+1)$	Definición de producto
	$= m \cdot (s(k))$	Definición de sucesor de n
	$= m + (m \cdot (k))$	Punto (b) de la definición de la función $m\cdot$
	$= m + (km)$	Hipótesis de inducción sobre k
	$= km + m$	La suma es conmutativa
	$= (k+1)m$	Lema 3.3.10

Por lo antes mostrado, se sigue que $m(k+1) = (k+1)m$.

Por el Principio de inducción matemática podemos concluir que si m es un natural, entonces $mn = nm$.

(d) Esta demostración la realizaremos por inducción sobre m, es decir, mostraremos que: si $\{n, k\} \subseteq \mathbb{N}$, entonces $m(n+k) = mn + mk$.

Base de inducción. Si $\{n, k\} \subseteq \mathbb{N}$, entonces $0(n+k) = 0n + 0k$.

Hipótesis de inducción. Supongamos que m es un natural tal que si $\{n, k\} \subseteq \mathbb{N}$, entonces $m(n+k) = mn + mk$.

Paso inductivo. Si $\{n, k\} \subseteq \mathbb{N}$, entonces $(m+1)(n+k) = (m+1)n + (m+1)k$.

Para esta demostración, notemos lo siguiente:

$(m+1)(n+k)$	$= m(n+k) + (n+k)$	Lema 3.3.10
	$= (mn + mk) + (n+k)$	Hipótesis de inducción sobre m
	$= (mn + n) + (mk + k)$	La suma es conmutativa y asociativa
	$= (m+1)n + (m+1)k$	Lema 3.3.10

De las igualdades anteriores se sigue que $(m+1)(n+k) = (m+1)n + (m+1)k$.

Por el Principio de inducción matemática podemos concluir que si $\{n, k\} \subseteq \mathbb{N}$, entonces $m(n+k) = mn + mk$.

(e) Esta demostración la realizaremos por inducción sobre k. Es decir, mostraremos que: si $\{m, n\} \subseteq \mathbb{N}$, entonces $(mn)k = m(nk)$.

Base de inducción. Si $\{m, n\} \subseteq \mathbb{N}$, entonces $(mn)0 = m(n0)$.

Para esta demostración, basta ver que $(mn)0 = 0$ y que $m(n0) = m0 = 0$, por lo que $(mn)0 = m(n0)$.

Hipótesis de inducción. Supongamos que a es un natural tal que para todo $\{m, n\} \subseteq \mathbb{N}$, $(mn)a = m(na)$.

Paso inductivo. Si $\{m, n\} \subseteq \mathbb{N}$, entonces $(mn)(a+1) = m\,(n(a+1))$.

Notemos lo siguiente:

$$\begin{aligned} (mn)(a+1) &= (mn)a + mn && \text{Inciso (d) de este lema} \\ &= m(na) + mn && \text{Hipótesis de inducción sobre } a \\ &= m((na) + n) && \text{Inciso (d) de este lema} \\ &= m(n(a+1)) && \text{Inciso (d) de este lema} \end{aligned}$$

Así, podemos concluir que $(mn)(a+1) = m\,(n(a+1))$.

Por el principio de inducción matemática, si $\{m, n\} \subseteq \mathbb{N}$, entonces $(mn)k = m(nk)$.

(f) La parte necesaria de este inciso se sigue inmediatamente del inciso (a) de este lema, por lo que solo resta mostrar la parte suficiente, es decir, si $mn = 0$, entonces $m = 0$ o $n = 0$. Para realizar dicha demostración supondremos que $n \neq 0$ y mostraremos que $m = 0$.

Como $n \neq 0$, podemos consideremos el antecesor de n, es decir, $n - 1$. Con ello, notemos que $mn = m + m(n-1)$ y por hipótesis podemos concluir que $m + m(n-1) = 0$. Gracias al inciso (f) del lema 3.3.4 podemos concluir que $m = 0$ y $m(n-1) = 0$, en particular, $m = 0$.

(g) Para esta demostración, usaremos primero el teorema 3.3.5 sobre los naturales n y k, por lo que existe $t \in \mathbb{N}$ tal que $n = k + t$ o $k = n + t$. Podemos suponer, sin pérdida de generalidad, que $n = k + t$.

Así, multiplicando m por ambos lados de la igualdad, tenemos que $nm = km + tm$. Luego, por hipótesis se sigue que $nm = nm + tm$ y así, por el inciso (f) del lema 3.3.4 podemos garantizar que $tm = 0$. Gracias al inciso (f) de este lema, entonces $t = 0$ o $m = 0$, sin embargo, por hipótesis se puede concluir que $t = 0$.

Con lo anterior, y debido a que $n = k + t$, se muestra que $n = k$.

□

Aun cuando ya tenemos definida una suma entre dos números naturales, es importante poder sumar varios números naturales. Para ello, consideremos los naturales $n_1, \ldots, n_k$, la **suma de los naturales $\boldsymbol{n_1, \ldots, n_k}$**, denotada por $\sum_{i=1}^{k} n_i$, se define como:

$$\sum_{i=1}^{k} n_i = \begin{cases} n_1 & \text{si } k = 1 \\ \left(\sum_{i=1}^{k-1} n_i\right) + n_k & \text{si } k \geq 2 \end{cases}$$

La suma anterior puede verse como la suma de varios naturales y tiene algunas propiedades que es importante mencionar.

Lema 3.3.12. *Sean $n_1, \ldots, n_k$ números naturales. Los siguientes enunciados se satisfacen:*

a) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i = t$, entonces $\sum_{i=1}^{k} n_i = kt$.

b) Si $j \in \{1, \ldots, k-1\}$, entonces $\sum_{i=1}^{k} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{k} n_i\right)$.

Demostración. (b) La demostración la realizaremos por inducción sobre k, es decir, mostraremos que: si $k \in \mathbb{N}$ y $j \in \{1, \ldots, k-1\}$, entonces $\sum_{i=1}^{k} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{k} n_i\right)$.

Base de inducción. Si $k = 2$ y $j \in \{1, \ldots, k-1\}$, entonces $\sum_{i=1}^{k} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{k} n_i\right)$.

En este caso, necesariamente $j = 1$. Por otro lado, por definición de suma. Además, $\sum_{i=1}^{k} n_i = n_1 + n_2$ y del mismo modo $\sum_{i=1}^{j} n_i = n_1$ y $\sum_{i=j+1}^{k} n_i = n_2$, concluyendo que $\sum_{i=1}^{k} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{k} n_i\right)$.

Base de inducción. Supongamos que m es un natural tal que si $j \in \{1, \ldots, m-1\}$, entonces $\sum_{i=1}^{m} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{m} n_i\right)$.

Base de inducción. Si $j \in \{1, \ldots, m\}$, entonces $\sum_{i=1}^{m+1} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{m+1} n_i\right)$.

Esta demostración la realizaremos por casos. Si $j = m$, entonces se tiene que:

$$\begin{aligned} \sum_{i=1}^{m+1} n_i &= \sum_{i=1}^{m} n_i + n_{m+1} && \text{Definición de suma} \\ &= \sum_{i=1}^{j} n_i + \sum_{i=j+1}^{m+1} n_i && \text{Suposición sobre } j \end{aligned}$$

Ahora podemos suponer que $j \neq m$ y por consiguiente, $j \in \{1, \ldots, m-1\}$. En tal caso, se tiene que:

$$\begin{aligned} \sum_{i=1}^{m+1} n_i &= \sum_{i=1}^{m} n_i + n_{m+1} && \text{Definición de suma} \\ &= \left(\sum_{i=1}^{j} n_i + \sum_{i=j+1}^{m+1} n_i\right) + n_{m+1} && \text{Hipótesis de inducción aplicada a } j \\ &= \sum_{i=1}^{j} n_i + \left(\sum_{i=j+1}^{m+1} n_i + n_{m+1}\right) && \text{La suma es asociativa} \\ &= \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{m+1} n_i\right) && \text{Definición de suma} \end{aligned}$$

□

Al igual que en el caso de la suma, es posible realizar el producto de varios números naturales. Para ello, consideremos los naturales $n_1, \ldots, n_k$, el **producto de los naturales $n_1, \ldots, n_k$**, denotada por $\prod_{i=1}^{k} n_i$, se define como:

$$\prod_{i=1}^{k} n_i = \begin{cases} n_1 & \text{si } k = 1 \\ \left(\prod_{i=1}^{k-1} n_i\right) \cdot n_k & \text{si } k \geq 2 \end{cases}$$

Lema 3.3.13. *Sean $n_1, \ldots, n_k$ y $b_1, \ldots, b_k$ números naturales. Los siguientes enunciados se satisfacen:*

a) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{k} n_i = t^k$.

b) Si $j \in \{1, \ldots, k\}$, entonces $\prod_{i=1}^{k} n_i = \left(\prod_{i=1}^{j} n_i\right)\left(\prod_{i=j+1}^{k} n_i\right)$.

Demostración. (a) Mostraremos por inducción sobre k que si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{k} n_i = t^k$.

Base de inducción. Si para todo $i \in \{1, \ldots, 1\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{1} n_i = t^1$

Para este caso, basta ver que $\prod_{i=1}^{1} n_i = n_1$ y que $n_1 = t = t^1$, concluyendo que $\prod_{i=1}^{1} n_i = t^1$.

Hipótesis de inducción. Supongamos que m es un natural tal que para todo $i \in \{1, \ldots, m\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{m} n_i = t^m$

Paso inductivo. Si para todo $i \in \{1, \ldots, m+1\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{m+1} n_i = t^{m+1}$.

□

A partir de este punto podemos empezar a utilizar las propiedades de las operaciones de suma y producto en los naturales. Es importante mencionar que, aunque hasta ahora no tenemos definida una resta y una división, podremos usarlas teniendo especial cuidado de ellas, considerando que estaremos trabajando en los naturales.

3.4. Orden en los naturales

Definimos la relación $\leq$ en $\mathbb{N}$ dada por: $n \leq m$ si y sólo si existe $k \in \mathbb{N}$ tal que $m = n + k$. Denotaremos por $n < m$ cuando $n \leq m$ y $n \neq m$. Es importante recalcar que la relación $<$ no es un orden parcial sobre $\mathbb{N}$[1], simplemente será una forma de denotar un hecho que usaremos frecuentemente.

Teorema 3.4.1. *La relación $\leq$ es un orden total en $\mathbb{N}$.*

Demostración. Primero mostraremos que la relación $\leq$ es un orden parcial sobre $\mathbb{N}$.

[1] Aunque sí es un orden parcial estricto.

(a) Para todo $n \in \mathbb{N}$, $n \leq n$.

Dado que existe $0 \in \mathbb{N}$ tal que $n + 0 = n$, se tiene por la definición de la relación que $n \leq n$.

(b) Para todo $\{n, m\} \subseteq \mathbb{N}$ tal que $n \leq m$ y $m \leq n$ se satisface que $n = m$.

Como $n \leq m$, entonces existe $k \in \mathbb{N}$ tal que $n + k = m$, y como $m \leq n$, entonces existe $t \in \mathbb{N}$ tal que $m + t = n$. Con lo anterior, sustituyendo una ecuación en la otra, obtenemos $n + k + t = n$, y por ello, $k + t = 0$, siguiéndose de esto el que $k = 0$ y $t = 0$, así, nuevamente sustituyendo en cualquiera de las primeras dos ecuaciones obtenemos que $n = m$.

(c) Para todo $\{n, m, k\} \subseteq \mathbb{N}$ tales que $n \leq k$ y $k \leq m$ se satisface que $n \leq m$.

Como $n \leq k$, entonces existe $s \in \mathbb{N}$ tal que $n + s = k$, y como $k \leq m$, entonces existe $t \in \mathbb{N}$ tal que $k + t = m$. Con lo anterior, sustituyendo una ecuación en la otra, obtenemos que $n + s + t = m$, y por ello, existe $s + t \in \mathbb{N}$ tal que $n + (s + t) = m$, siguiéndose de la definición de $\leq$ el que $n \leq m$.

Ahora que hemos demostrado que $\leq$ es un orden parcial, procederemos a demostrar que es un orden total.

(d) Si $\{n, m\} \subseteq \mathbb{N}$, entonces $n \leq m$ o $m \leq n$.

Sea $\{n, m\} \subseteq \mathbb{N}$. Por el teorema 3.3.5 se sigue que existe $k \in \mathbb{N}$ tal que $n = m + k$ o $m = n + k$. Si sucede que $n = m + k$, entonces por definición del orden, $m \leq n$. Si $m = n + k$, entonces por definición del orden, $n \leq m$. En cualquiera de los casos anteriores, se tiene que $n \leq m$ o $m \leq n$.

Por lo antes mostrado, podemos afirmar que $\leq$ es un orden total sobre $\mathbb{N}$. $\square$

El orden descrito previamente se le conoce como el **orden usual en los naturales** y tiene muchas propiedades importantes, pero aquí sólo mencionaremos algunas de ellas.

Lema 3.4.2. *Si m, n, k y l son naturales, entonces los siguientes enunciados se satisfacen.*

a) Si $m \leq n$, entonces $m + k \leq n + k$.

b) Si $m \leq n$ y $k \leq l$, entonces $m + k \leq n + l$.

c) Si $m \leq n$, entonces $mk \leq nk$

d) Si $m \leq n$ y $k \leq l$, entonces $mk \leq nl$.

e) $0 \leq n$.

f) $n \leq n + k$

g) Si $n \not\leq m$, entonces $m < n$.

h) Si $n < m$, entonces $n + 1 \leq m$.

i) Si $n \leq m$ y $m \leq n + 1$, entonces $m = n$ o $m = n + 1$.

Demostración. Para esta demostración sólo probaremos algunos incisos representativos.

(b) Como $m \leq n$, entonces existe $t \in \mathbb{N}$ tal que $m + t = n$. Del mismo modo, como $k \leq l$, entonces existe $s \in \mathbb{N}$ tal que $k + s = l$. Así, sumando ambas igualdades se tiene que $(m + t) + (k + s) = n + l$, con ello, existe $(t + s) \in \mathbb{N}$ tal que $(m + k) + (t + s) = n + l$, concluyendo que $m + k \leq n + l$.

(g) Dado que $\leq$ es un orden total, entonces $n \leq m$ o $m \leq n$, sin embargo, por hipótesis podemos concluir que $m \leq n$. Además, como $n \not\leq m$, entonces $n \neq m$, concluyendo que $m < n$.

(i) Como $n \leq m$, entonces existe $t \in \mathbb{N}$ tal que $n + t = m$. Del mismo modo, como $m \leq n + 1$, entonces existe $s \in \mathbb{N}$ tal que $m + s = n + 1$.

Como $n + t = m$ y $m + s = n + 1$, entonces $(n + t) + s = n + 1$. De lo anterior, $t + s = 1$ y así podemos suponer que (i) $t = 0$ y $s = 1$ o (ii) $t = 1$ y $s = 0$, por lo que tenemos los siguientes casos.

Caso 1. $t = 0$ y $s = 1$.

En este caso, como $n + t = m$, podemos concluir que $n = m$.

Caso 2. $t = 1$ y $s = 0$.

En este caso, como $n + t = m$, podemos concluir que $n + 1 = m$.

De los casos anteriores, $m = n$ o $m = n + 1$

□

Notemos que de la definición de orden en los naturales, tenemos que si $n \leq m$, entonces existe un natural k tal que $n + k = m$. Derivado del lema 3.3.4 (d) se tiene que dicho elemento k es único. Por lo anterior, si $n \leq m$, *denotaremos* por $m - n$ al único natural k que satisface $n + k = m$. Es importante recalcar que la notación $m - n$ no es una operación, sino simplemente una forma de denotar un objeto. A dicho número se le suele llamar la **diferencia de m menos n**.

Las siguientes son propiedades de la suma y el producto referentes al orden en los naturales.

Lema 3.4.3. *Sean $n_1, \ldots, n_k$ y $b_1, \ldots, b_k$ números naturales y $t \in \mathbb{N}$. Los siguientes enunciados se satisfacen:*

a) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq b_i$, entonces $\sum_{i=1}^{k} n_i \leq \sum_{i=1}^{k} b_i$.

b) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq t$, entonces $\sum_{i=1}^{k} n_i \leq kt$.

c) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq b_i$, entonces $\prod_{i=1}^{k} n_i \leq \prod_{i=1}^{k} b_i$.

d) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq t$, entonces $\prod_{i=1}^{k} n_i \leq t^k$.

Para finalizar la sección de orden, tenemos el siguiente teorema.

Teorema 3.4.4 (del Buen Orden en $\mathbb{N}$). *El orden usual en los naturales es un buen orden.*

Demostración. Sea $A \subseteq \mathbb{N}$ no vacío y definimos $B = \{n \in \mathbb{N} : \text{para toda } a \in A,\ n < a\}$. Notemos que $A \cap B = \emptyset$.

Por otro lado, si $0 \in A$, entonces 0 debe ser mínimo de A y habremos terminado, por lo que para el resto de la demostración, supondremos que $0 \notin A$. En tal caso, $0 < a$ para toda $a \in A$, por lo que por definición de B, tenemos que $0 \in B$.

Además de lo anterior, notemos que $B \neq \mathbb{N}$, pues por hipótesis, A es no vacío y considerando $a \in A$, entonces a es un natural tal que $a \notin B$, ya que $A \cap B = \emptyset$. Con lo anterior, por el principio de inducción en los naturales, existe $n \in B$ tal que $n + 1 \notin B$. Ahora procederemos a mostrar que $n + 1$ es mínimo de A.

Primero veremos que $n + 1 \in A$. Como $n + 1 \notin B$, entonces existe $a_0 \in A$ tal que $n + 1 \not< a_0$, es decir, $a_0 \leq n + 1$. Por otro lado, como $n \in B$ entonces por definición de B se tiene en particular que $n < a_0$ y así, $n + 1 \leq a_0$. Derivado de lo anterior se tiene que $a_0 = n + 1$, concluyendo que $n + 1 \in A$.

Ahora mostraremos que para todo $a \in A$, $n + 1 \leq a$. Para ello, sea $a \in A$ arbitrario. Como $n \in B$, se sigue que en particular $n < a$ y así $n + 1 \leq a$, mostrando lo deseado.

De lo anterior, $n + 1$ es un elemento mínimo de A. □

3.5. Principio de inducción modificado

Existe una versión alterna, pero no equivalente, al principio de inducción matemático propuesto en el axioma 5 de Peano, al cual llamaremos Principio de inducción modificado. Al igual que el Principio de inducción, establece condiciones para garantizar que un subconjunto de $\mathbb{N}$ es igual al conjunto de todos los números naturales.

Teorema 3.5.1 (Principio de inducción modificado.). *Si $A \subseteq \mathbb{N}$ satisface:*

a) $0 \in A$

b) Si $\{0, \ldots, n\} \subseteq A$, entonces $n + 1 \in A$

Entonces $A = \mathbb{N}$

Este principio inductivo se puede deducir de los Axiomas de Peano, en particular, del Principio de inducción.

Teorema 3.5.2. *El Principio de inducción matemática implica el Principio de inducción modificado.*

Demostración. Sea $A \subseteq \mathbb{N}$ es un conjunto tal que (i) $0 \in A$ y (ii) si $\{0, \ldots, n\} \subseteq A$, entonces $n+1 \in A$. Demostraremos que $A = \mathbb{N}$ mediante el Principio de inducción.

Ya que $0 \in A$, basta demostrar que si $n \in A$, entonces $n+1 \in A$. Para realizar dicha prueba, tomaremos $n \in A$ y veremos que $\{0, \ldots, n\} \subseteq A$, para que gracias a la propiedad (ii) podamos concluir que $n+1 \in A$. Consideremos $n \in A$ y sea $B = \{0, \ldots, n\} \setminus A$.

Afirmación 1. $B = \emptyset$.

Procederemos por contradicción al suponer que $B \neq \emptyset$. Por el Principio del buen orden en $\mathbb{N}$, el conjunto B tiene un mínimo, digamos k_0. Es importante notar que por definición de B, en particular $k_0 \notin A$ y así, $k_0 < n$. Notemos además que $k_0 \neq 0$ pues $0 \in A$ y $k_0 \notin A$. Teniendo en cuanta lo anterior, nótese que $\{0, \ldots, k_0 - 1\} \subseteq \{0, \ldots, n\}$.

Ahora veremos que $\{0, \ldots, k_0 - 1\} \subseteq A$. Sea $l \in \{0, \ldots, k_0 - 1\}$, en particular, se tiene que $l < k_0$ y como k_0 es mínimo en B, entonces $l \notin B$, lo que implica que $l \notin \{0, \ldots, n\}$ o $l \in A$. Ya que $l \in \{0, \ldots, k_0 - 1\}$ y $\{0, \ldots, k_0 - 1\} \subseteq \{0, \ldots, n\}$, la única posibilidad es que $l \in A$.

De lo anterior, por la propiedad (ii) del conjunto A, se seguiría que $k_0 \in A$, o cual no es posible por elección de k_0. Por lo tanto, $B = \emptyset$, mostrando la afirmación.

En tal caso, $\{0, \ldots, n\} \setminus A = \emptyset$, concluyendo que $\{0, \ldots, n\} \subseteq A$, por lo que por la propiedad (ii) del conjunto A, se concluye que $n+1 \in A$. Por el Principio de inducción matemática, tenemos que $A = \mathbb{N}$. $\square$

Al Principio de inducción matemática se le suele llamar coloquialmente como *inducción débil*, mientras que al Principio de inducción modificado se le suele llamar *inducción fuerte*. Aunque el Principio de inducción modificado tiene una versión conjuntista para ser descrito (teorema 3.5.1), también es importante decir que existe una manera de usarlo para realizar cierto timpo de demostraciones, mismo que mencionaremos ahora.

Teorema 3.5.3. *Sea P un predicado que toma valores sobre los números naturales. Para verificar que P es verdadero para todo natural mayor o igual que algún natural k_0 basta realizar lo siguiente:*

a) Base Inductiva. Demostrar que $P(k_0)$ es verdadero.

b) Hipótesis de Inducción. Supondremos que n es un natural tal que para todo $k \in \mathbb{N}$ tal que $k_0 \leq k \leq n$ se satisface que $P(k)$ es verdadera.

c) Paso Inductivo. Se demuestra que $P(n+1)$ es verdadera.

Entonces P es verdadera para todo natural mayor o igual que k_0.

Al igual que con la inducción débil, es necesario realizar un preámbulo a una demostración inductiva fuerte, mismo que analizaremos en la siguiente subsección.

3.5.1. Inducción fuerte de enunciados

Para poder llevar a cabo una inducción fuerte necesitaremos tener presente las indicaciones del teorema 3.5.3 y un enunciado lógico en el que haya al menos un natural implícito o explícito y que sea variable. Bajo estas condiciones, es posible realizar una demostración inductiva fuerte para probar la veracidad del enunciado.

Para ejemplificar lo anterior utilizaremos el siguiente enunciado para realizar el proceso de inducción fuerte.

Ejemplo 3.5.4. *Todo natural mayor que 5 es la suma de múltiplos de 2 y 7.*

Al igual que en muchas demostraciones de nuestro curso, antes de iniciar la demostración por inducción fuerte desarrollaremos un preámbulo que será de utilidad al momento de realizar nuestra demostración. Es importante que prestes atención a todos los pasos que haremos previo a iniciar la demostración y los compares con los pasos realizados en una inducción débil.

Paso 1 **La preparación del enunciado a demostrar.**

Primero es importante saber escribir el enunciado de tal forma que sea evidente que hay un número natural variable en dicho enunciado. En algunas ocasiones dicho natural estará de forma explícita en el enunciado y en otras ocasiones no será explícito. En este último caso lo primero que debemos hacer es reescribir el enunciado para mostrar explícitamente el natural que vamos a usar como variable inductiva.

En este caso el enunciado que demostraremos debe reescribirse para facilitar su demostración. Este paso requiere de práctica y experiencia al momento de realizar demostraciones, por lo que si no estás seguro de cómo hacerlo, puedes acercarte a alguna persona para que te auxilie. El enunciado que nosotros vamos a demostrar lo reescribiremos de la siguiente manera.

Si $n \in \mathbb{N}$ es un natural tal que $n \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = n$.

Paso 2 **El inicio de tu redacción.**

Ahora que tenemos un natural explícitamente en el enunciado vamos a indicar que iniciaremos una demostración por **inducción fuerte** sobre dicho enunciado y además indicaremos quién será nuestra variable inductiva. Esto es importante porque (i) contextualiza a la persona que leerá tu inducción sobre qué es lo que vas a hacer para llevar a cabo la inducción, (ii) le indica a la persona que está leyendo tu demostración que, de los dos tipos de inducción, tú realizarás una inducción fuerte y (iii) si existen varios naturales en tu enunciado, le indicas a la persona que está leyendo cuál de ellos será la variable inductiva. Así, tu demostración inductiva debe iniciar con una frase como:

Mostraremos por inducción fuerte sobre n que: si $n \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = n$.

Paso 3 **La escritura de la base inductiva.**

El inicio de nuestra inducción es la base inductiva. En este caso volveremos a escribir todo el enunciado que queremos demostrar (el que obtuvimos del paso 1) y ahí en donde aparezca la variable inductiva, la sustituiremos por el primer natural posible de acuerdo a nuestro enunciado. En este caso, como $n \geq 6$, nuestra base inductiva iniciará en 6. Es importante que prestes atención a este detalle porque, al igual que la inducción débil, no todas nuestras inducciones iniciarán desde cero. Así, el enunciado debe quedar escrito como:

Si $6 \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = 6$.

Recuerda que la base de inducción sí se demuestra, tiene sus propias hipótesis y su propio objetivo a demostrar. Por ahora no realizaremos la demostración, sólo estamos realizando el preámbulo de nuestra demostración inductiva.

Paso 4 **La escritura de la hipótesis de inducción fuerte.**

El siguiente paso de nuestra inducción es la hipótesis inductiva fuerte y es la única variación real respecto a la inducción débil. Su redacción requiere más trabajo que el de la inducción débil, por lo que ten mucho cuidado en este punto.

Primero decidiremos cuál será el nombre de nuestra variable inductiva, para este ejemplo, k será el nombre de nuestra variable inductiva en la hipótesis de inducción fuerte. Una vez realizado esto, el objeto k toma propiedades de objeto fijo en nuestra demostración. Declararemos este objeto en nuestra escritura con una frase como:

Supondremos que k es un natural tal que...

Lo siguiente es continuar la redacción de nuestra hipótesis eligiendo una nueva variable, a la que llamaremos m. A diferencia de la inducción débil, dicha variable m será la que hace verdadero al enunciado obtenido en el paso 1, es decir, el enunciado quedará escrito como sigue:

Si $m \in \mathbb{N}$*, entonces existe* $\{a, b\} \subseteq \mathbb{N}$ *tal que* $2a + 7b = m$*.*

Pero, además, agregaremos a dicho enunciado una hipótesis extra: que la variable m es mayor o igual que el natural de la base inductiva y menor o igual que la variable inductiva. Así, el enunciado que debemos demostrar debe ser algo como:

Si $m \in \mathbb{N}$ *y* $6 \leq m \leq k$*, entonces existe* $\{a, b\} \subseteq \mathbb{N}$ *tal que* $2a + 7b = m$*.*

Por último, juntando la declaración de nuestra variable inductiva y este último paso, nuestra hipótesis de inducción debe decir, hasta ahora, algo como:

Supondremos que k *es un natural tal que: si* $m \in \mathbb{N}$ *y* $6 \leq m \leq k$*, entonces existe* $\{a, b\} \subseteq \mathbb{N}$ *tal que* $2a + 7b = m$*.*

Pero aun no concluimos nuestra hipótesis de inducción, pues al igual que en la inducción débil, es conveniente cambiar el nombre de los objetos que varían en el enunciado que queremos demostrar, como aquellos que dependen de m o de k y los que son variables por sí mismos como objetos matemáticos dentro de un enunciado lógico. En este caso, los objetos a y b son variables, por lo que cambiaré momentáneamente su nombre, finalizando la escritura de nuestra hipótesis inductiva fuerte:

Supondremos que k *es un natural tal que: si* $m \in \mathbb{N}$ *y* $6 \leq m \leq k$*, entonces existe* $\{a', b'\} \subseteq \mathbb{N}$ *tal que* $2a' + 7b' = m$*.*

Este paso no se demuestra, se presupone. Como tal, el enunciado obtenido a partir de este paso es una hipótesis y tiene su propia forma de ser usado de manera directa, ya sea que se trate de una condicional (como en este ejemplo), un cuantificador universal, un cuantificador existencial, etc.

Paso 5 **La escritura del paso inductivo.**

El último punto es el paso inductivo y es similar al de la inducción débil. Nuevamente volveremos a escribir todo el enunciado que queremos demostrar (el que obtuvimos del paso 1) y ahí en donde aparezca la variable inductiva, la sustituiremos por el sucesor del natural que elegimos para la variable inductiva en el paso inductivo, en nuestro caso, k. Mucho ojo, el paso inductivo quedará escrito en términos del nombre de la variable inductiva que elegimos en la hipótesis de inducción, es decir, k. No queda en términos de la segunda variable que elegimos, es decir, no queda en términos de m. Fuera de esto, no hay mucho más que cambiar en este punto, por lo que nuestro enunciado quedaría algo de la siguiente manera.

Si $k+1 \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = k + 1$.

Este último paso debe demostrarse. No pierdas de vista que dicho paso te dará sus propias hipótesis, sus propios objetos al iniciar la demostración y su propio objetivo a demostrar. Eventualmente, durante la demostración de este enunciado, deberás usar la hipótesis inductiva fuerte, por lo que debes estar muy atento al momento en que puedas usar dicha hipótesis.

Todos los pasos anteriores fueron el preámbulo para nuestra demostración. Ahora que ya lo hemos realizado, procederemos a demostrar el enunciado. Nota cómo aplicamos los pasos antes descritos en este ejemplo.

Ejemplo 3.5.5. *Todo natural mayor que 5 es la suma de múltiplos de 2 y 7.*

Demostración. Mostraremos por inducción fuerte sobre n que: si $n \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = n$.

Base de inducción. Si $6 \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = 6$.

En este caso, como $\{3, 0\} \subseteq \mathbb{N}$ son tales que $2(3) + 7(0) = 6$, entonces existen $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = 6$.

Hipótesis de inducción. Supondremos que k es un natural tal que si $m \in \mathbb{N}$ y $6 \leq m \leq k$, entonces existe $\{a', b'\} \subseteq \mathbb{N}$ tal que $2a' + 7b' = m$.

Paso inductivo. Si $k + 1 \geq 6$, entonces existe $\{a, b\} \subseteq \mathbb{N}$ tal que $2a + 7b = k + 1$.

Notemos que en este caso, si $k + 1 = 6$, entonces existe $\{3, 0\} \subseteq \mathbb{N}$ tal que $2(3) + 7(0) = k + 1$. Del mismo modo, si $k + 1 = 7$, entonces existe $\{0, 1\} \subseteq \mathbb{N}$ tal que $2(0) + 7(1) = k + 1$.

Así, para lo que resta de la demostración, supondremos que $k+1 \geq 8$. En tal caso, notemos que $k-1 \in \mathbb{N}$ y $6 \leq k - 1 \leq k$, por lo que aplicando hipótesis de inducción fuerte sobre $k + 1$ obtenemos que existe $\{a', b'\} \subseteq \mathbb{N}$ tal que $2a' + 7b' = k - 1$.

De lo anterior, $2+(2a'+7b') = 2+(k-1)$, es decir, $2(a'+1)+7b' = k+1$. Con ello, existe $\{a'+1, b'\} \subseteq \mathbb{N}$ tal que $2(a' + 1) + 7b' = k + 1$.

Por el Principio de inducción matemática fuerte, podemos concluir que todo natural mayor que 5 es la suma de múltiplos de 2 y 7. $\square$

El anterior es un ejemplo clásico de una demostración inductiva fuerte, por lo que te recomiendo que lo repases las veces que sean necesarias, así como los pasos que realizamos previo a la demostración.

Para finalizar esta subsección, y de manera análoga a la subsección correspondiente con la inducción débil, quiero mencionar que los pasos descritos previamente para la elaboración del preámbulo son importantes de realizar, sin embargo, no son un método absoluto para escribir todas las demostraciones inductivas fuertes a las que te enfrentarás. En algunas ocasiones será necesario saber reescribir todo el enunciado para una mejor comprensión y uso de la inducción fuerte, pero eso lo conseguirás poco a poco con la experiencia y con la práctica. Por mientras, considera estos pasos como un punto de referencia para futuras demostraciones mediante inducción fuerte.

3.6. Conjuntos finitos

Sean A y B dos conjuntos, decimos que **A y B son equipotentes**, denotado por $A \approx B$, si existe una función biyectiva de A en B. El siguiente teorema es relativamente sencillo de verificar:

Teorema 3.6.1. *Los siguientes enunciados se satisfacen:*

a) Si A es un conjunto, entonces $A \approx A$.

b) Si A y B son conjuntos y $A \approx B$, entonces $B \approx A$.

c) Si A, B y C son conjuntos y $A \approx B$ y $B \approx C$, entonces $A \approx C$

Demostración. Ahora mostraremos los tres incisos de este teorema.

(a) Si A es un conjunto, entonces la función $1_A : A \to A$ es biyectiva, por lo que $A \approx A$.

(b) Como $A \approx B$, entonces existe una función biyectiva $f : A \to B$. En tal caso, la función $f^{-1} : B \to A$ es una función biyectiva, por lo que $B \approx A$.

(c) Como $A \approx B$, entonces existe una función biyectiva $f : A \to B$, y del mismo modo, dado que $B \approx C$, entonces existe una función biyectiva $g : B \to C$. Con lo anterior, $g \circ f : A \to C$ es una función biyectiva de A en C, por lo que $A \approx C$.

□

Dado un conjunto A, diremos que A es **infinito** si existe $B \subseteq A$ tal que $B \approx A$ y $A \neq B$. En otro caso, diremos que A es **finito**, es decir, para todo $B \subseteq A$ tal que $B \approx A$, se satisface que $B = A$.

Lema 3.6.2. *Los siguientes enunciados se satisfacen:*

a) El conjunto vacío es finito.

b) El conjunto $\mathbb{N}$ es infinito.

Demostración. Ahora mostraremos ambos incisos de este lema.

(a) Procederemos por contradicción y supondremos que $\emptyset$ es infinito. En tal caso, existe $B \subseteq \emptyset$ tal que $B \approx \emptyset$ y $B \neq \emptyset$, lo cual no es posible pues el único conjunto biyectable con $\emptyset$ es $\emptyset$.

(b) Para este inciso, basta definir $B = \{x \in \mathbb{N} : \text{ existe } z \in \mathbb{N} \text{ tal que } x = 2z\}$. Notemos que $B \subseteq \mathbb{N}$, $B \neq \mathbb{N}$ y la función $f : \mathbb{N} \to B$ dada por $f(n) = 2n$ es biyectiva, es decir, $\mathbb{N} \approx B$. Así, $\mathbb{N}$.

□

Para cada $n \in \mathbb{N}$ con $n \geq 1$, definimos $I_n = \{k \in \mathbb{N} : 1 \leq k \leq n\}$. Los conjuntos anteriores nos ayudarán a definir el cardinal de conjuntos, pero previo a ello, notemos que:

Lema 3.6.3. *Sean n y m naturales no nulos. Los siguientes enunciados se satisfacen.*

a) $I_n \approx I_m$ si y sólo si $n = m$.

b) $I_n \subseteq I_m$ si y sólo si $n \leq m$.

c) $I_n \subsetneq I_m$ si y sólo si $n < m$.

d) $I_n \subseteq I_m$ o $I_m \subseteq I_n$.

Demostración. Para este lema, sólo mostraremos el inciso (a). Los incisos restantes se deducen de las propiedades de los naturales.

(a) La parte suficiente es inmediata, por lo que mostraremos la parte suficiente. Para ello, mostraremos por inducción sobre n que si $I_n \approx I_m$, entonces $n = m$.

Base de Inducción. Si $I_1 \approx I_m$, entonces $1 = m$.

Como $Im(f) = I_m$ e $Im(f) = \{f(1)\}$, entonces $I_m = \{1\}$.

Hipótesis de Inducción. Si $I_k \approx I_r$, entonces $k = r$.

Paso Inductivo. Si $I_{k+1} \approx I_s$, entonces $k + 1 = s$.

Dado que $I_{k+1} \approx I_s$, entonces existe $f : I_{k+1} \to I_s$ biyectiva. Además podremos presuponer que $f(k + 1) = s$. Así, $f : I_k \to I_{s-1}$ es una función biyectiva, por lo que $I_k \approx I_{s-1}$ y por hipótesis de inducción, $k = s - 1$, concluyendo que $k + 1 = s$.

Por el Principio de inducción matemática podemos concluir que si $I_n \approx I_m$, entonces $n = m$ □

Con lo anterior, tendremos una manera alterna de determinar si un conjunto es finito o no.

Teorema 3.6.4. *Sea A un conjunto arbitrario. A es finito si y sólo si $A = \emptyset$ o existe $n \geq 1$ tal que $A \approx I_n$.*

Demostración. Se dará por hecho. □

Si A es un conjunto finito, la **cardinalidad de A**, denotada por $|A|$, se define como:

$$|A| = \begin{cases} 0 & \text{Si } A = \emptyset \\ n & \text{Si } A \approx I_n \text{ para algún natural } n \end{cases}$$

Un primer lema que podemos demostrar es el siguiente, en el cual nos muestran algunas propiedades de los subconjuntos de conjuntos finitos.

Lema 3.6.5. *Sean A un conjunto finito y B un conjunto arbitrario. Los siguientes enunciados se satisfacen.*

a) Si $B = A$, entonces B es finito y $|B| = |A|$.

b) Si $B \subseteq A$, entonces B es finito y $|B| \leq |A|$.

c) Si $B \subseteq A$ y $|A| = |B|$, entonces $A = B$.

Demostración. Ahora mostraremos ambos incisos de este lema.

(a) Es inmediato del hecho de que $A = B$.

(b) Esta es su tarea.

(c) Procederemos por contradicción al suponer que $A \neq B$. Para simplificar notación, supondremos que $|A| = n$. Notemos que como $|A| = n$ y $|B| = n$, entonces $A \approx I_n$ y $B \approx I_n$, por lo que $B \approx A$. En tal caso, $B \subseteq A$, $B \approx A$ y $A \neq B$, concluyendo que A es infinito, lo cual no es posible pues A es finito.

□

Como consecuencia del lema previo tenemos los siguientes resultados.

Corolario 3.6.6. *Sean A y B conjuntos. Los siguientes enunciados se satisfacen:*

a) Si $B \subseteq A$ y B es infinito, entonces A es infinito.

b) Si A es finito, entonces $A \cap B$ y $A \setminus B$ son conjuntos finitos.

Demostración. Ahora mostraremos ambos incisos de este corolario.

(a) Por el lema 3.6.5 (b) sabemos que si A es finito y $B \subseteq A$, entonces B es finito. En tal caso, este inciso es la contraposición de este hecho.

(b) Como $A \cap B \subseteq A$ y $A \setminus B \subseteq A$, entonces por el lema 3.6.5 (B), se concluye que $A \cap B$ y $A \setminus B$ son finitos.

□

El siguiente teorema es bastante útil para comparar los cardinales de dos conjuntos finitos mediante el uso de funciones.

Teorema 3.6.7. *Sean A y B conjuntos finitos. Los siguientes enunciados se satisfacen.*

a) Existe una función inyectiva de A en B si y sólo si $|A| \leq |B|$.

b) Existe una función sobreyectiva de A en B si y sólo si $|B| \leq |A|$.

c) Existe una función biyectiva de A en B si y sólo si $|B| = |A|$.

Demostración. Ahora mostraremos los incisos de este teorema.

(a) Para cuestiones de la demostración, como A y B son finitos, supondremos que $|A| = n$ y $|B| = k$, es decir, existen funciones biyectivas $g : I_n \to A$ y $h : B \to I_k$.

Primero mostraremos la parte suficiente del inciso (a), es decir, mostraremos que si existe una función inyectiva de A en B, digamos f, entonces $|B| \geq |A|$.

Denotamos por ρ a la función $h \circ f \circ g$ y con ello, es sencillo ver que $\rho : I_n \to I_k$ es una una función inyectiva, por lo que $\rho : I_n \to Im(\rho)$ es biyectiva, es decir, $Im(\rho) \approx I_n$, o equivalentemente, $|Im(\rho)| = n$. Por otro lado, como $Im(\rho) \subseteq I_k$, se sigue del lema 3.6.5 (b) que $|Im(\rho)| \leq |I_k|$, por lo que $n \leq k$, es decir, $|A| \leq |B|$.

Ahora mostraremos la condición necesaria de este inciso, es decir, si $|A| \leq |B|$, entonces existe una función inyectiva de A en B.

En este caso, como $|A| \leq |B|$, es decir, $n \leq k$, se tiene que $I_n \subseteq I_k$. Sea $f : I_n \to I_k$ dada por $f(x) = x$. Claramente, f es inyectiva. De lo anterior, $h^{-1} \circ f \circ g^{-1}$ es una función inyectiva de A en B.

(b) Primero mostraremos la parte parte suficiente de este inciso, es decir, si existe una función sobreyectiva de A en B, digamos f, entonces $|B| \leq |A|$.

Como f es sobreyectiva, podemos considerar un inverso derecho de f, digamos $g : B \to A$. En tal caso, g es inyectiva pues f es un inverso izquierdo de g, por lo que gracias al inciso (a) de este lema, $|B| \leq |A|$.

Ahora mostraremos la parte necesaria, es decir, si $|B| \leq |A|$, entonces existe una función sobreyectiva de A en B.

Como $|B| \leq |A|$, entonces por el inciso (a) existe una función inyectiva de B en A. Si g es un inverso izquierdo de dicha función, entonces g es una función de A en B y es sobreyectiva.

(c) Este inciso es inmediato de los incisos previos.

□

Como consecuencia inmediata del inciso (c) teorema anterior tenemos la siguiente nota que es importante recalcar.

Nota 3.6.8. *Sean A y B dos conjuntos finitos. $|A| = |B|$ si y sólo si $A \approx B$.*

La nota anterior, aunque sencilla, nos permite cambiar en los conjuntos finitos la noción de cardinal de conjunto con la de equipotencia. Así, en medida de lo posible, en lo consecuente preferiremos la notación $|A| = |B|$ en lugar de $A \approx B$. Otro resultado de conjuntos finitos y funciones es el siguiente.

Teorema 3.6.9. *Si A y B son dos conjuntos finitos tales que $|A| = |B|$ y $f : A \to B$ es una función. Los siguientes enunciados son equivalentes:*

a) f es inyectiva.

b) f es sobreyectiva.

c) f es biyectiva.

Demostración. Primero mostraremos que (a) implica (b), es decir, si f es inyectiva, entonces f es sobreyectiva.

Dado que f es inyectiva, entonces $f : A \to Im(f)$ es biyectiva, por lo que $|A| = |Im(f)|$. Por otro lado, como $|A| = |B|$, podemos garantizar que $|B| = |Im(f)|$ y debido a que $Im(f) \subseteq B$ y B es finito, entonces por el lema 3.6.5 (c) podemos concluir que $Im(f) = B$, mostrando que f es sobreyectiva.

Ahora mostraremos que (b) implica (c), es decir, si f es sobreyectiva, entonces f es biyectiva.

Como f es sobreyectiva, consideremos un inverso derecho de f, digamos F. Ahora mostraremos que F es biyectiva. Primero, por ser F un inverso derecho, entonces F debe ser inyectiva. Por otro lado, como $F : B \to A$ es inyectiva y $|A| = |B|$, entonces por la implicación (a) implica (b) de este enunciado, F debe ser sobreyectiva, concluyendo que F es biyectiva. Como f es un inverso izquierdo de F y F es biyectiva, entonces $f = F^{-1}$, por lo que f debe ser biyectiva.

El hecho de que (c) implique (a) es inmediato de la definición de biyectividad. □

Un último resultado en esta sección es el siguiente, el cual es una caracterización de los subconjuntos de naturales que son finitos en términos de máximos y mínimos.

Teorema 3.6.10. *Si $A \subseteq \mathbb{N}$ y A es no vacío, entonces A es finito si y sólo si A tiene máximo y mínimo.*

Demostración. Primero mostraremos la parte suficiente de este enunciado, es decir, mostraremos que si A es finito, entonces A tiene máximo y mínimo.

Por el Principio del buen orden, el conjunto A debe tener un elemento mínimo. Por otro lado, por el ejemplo 3.2.3, A debe tener máximo.

Ahora mostraremos la parte necesaria, es decir, si A tiene máximo y mínimo, entonces A es finito.

Supongamos que k es el máximo de A. En tal caso, $A \subseteq I_k$. Como I_k es finito, entonces por el lema 3.6.5, A debe ser finito. $\square$

3.7. Cardinales de operaciones con conjuntos

Ahora procederemos a demostrar algunas propiedades de cardinales finitos en algunas operaciones de conjuntos. Primero tenemos los siguientes lemas referentes a diferencia y unión de conjuntos respecto a un elemento.

Lema 3.7.1. *Si A es un conjunto finito no vacío y $x \in A$, entonces $|A \setminus \{x\}| = |A| - 1$.*

Demostración. Como A es finito, supondremos que $|A| = n$. Nótese que como A es no vacío, entonces $n \geq 1$. De acuerdo a lo anterior, existe una función $f : A \to I_n$ biyectiva, y asumiremos sin perdida de generalidad que $f(x) = n$. Con lo anterior, la función $g : A \setminus \{x\} \to I_n \setminus \{n\}$ dada por $g(t) = f(t)$ es una función biyectiva. Nótese que $I_n \setminus \{n\} = I_{n-1}$, por lo que $A \setminus \{x\} \approx I_{n-1}$, concluyendo que $|A \setminus \{x\}| = n - 1$, es decir, $|A \setminus \{x\}| = |A| - 1$. □

Y ahora un lema referente a unión.

Lema 3.7.2. *Si B es un conjunto finito y x es un objeto que no está en B, entonces $B \cup \{x\}$ es finito y $|B \cup \{x\}| = |B| + 1$.*

Demostración. Como B es finito entonces existe $n \in \mathbb{N}$ tal que $|B| = n$. Así, $B \approx I_n$, es decir, existe una función biyectiva $f : B \to I_n$. En tal caso, definimos la función $g : B \cup \{x\} \to I_{n+1}$ dada por:

$$g(z) = \begin{cases} f(z) & \text{Si } z \in B \\ n+1 & \text{Si } z = x \end{cases}$$

No es difícil verificar que g es una función biyectiva, por lo que $B \cup \{x\} \approx I_{n+1}$. De lo anterior, $B \cup \{x\}$ es finito y en particular $|B \cup \{x\}| = n + 1 = |B| + 1$. □

Los lemas previos serán de utilidad para poder realizar demostraciones inductivas sobre cardinales de conjuntos finitos.

Lema 3.7.3. *Si A es un conjunto finito y $B \subseteq A$, entonces $|A \setminus B| = |A| - |B|$.*

Demostración. Esta demostración la realizaremos por inducción sobre $|A|$. Es decir, mostraremos que si $|A| = n$ y $B \subseteq A$, entonces $|A \setminus B| = |A| - |B|$.

Base de inducción. Si $|A| = 0$ y $B \subseteq A$, entonces $|A \setminus B| = |A| - |B|$.

En este caso, como $|A| = 0$, entonces $A = B = \emptyset$, por lo que $|A \setminus B| = |A| - |B|$.

Hipótesis de inducción. Supongamos que k es un natural tal que si $|A'| = k$ y $B' \subseteq A'$, entonces $|A' \setminus B'| = |A'| - |B'|$.

Paso inductivo. Si $|A| = k+1$ y $B \subseteq A$, entonces $|A \setminus B| = |A| - |B|$.

Notemos que si $B = \emptyset$, entonces $|A \setminus B| = |A|$ y del mismo modo, $|A| - |B| = |A|$, concluyendo que $|A \setminus B| = |A| - |B|$. Para lo que resta de la demostración podemos suponer que $B \neq \emptyset$, en cuyo caso, sea $x \in B$.

Nótese que $|A \setminus \{x\}| = k$ y $B \setminus \{x\} \subseteq A \setminus \{x\}$, por lo que es posible aplicar hipótesis de inducción sobre $A \setminus \{x\}$ y $B \setminus \{x\}$, obteniendo que $|(A \setminus \{x\}) \setminus (B \setminus \{x\})| = |A \setminus \{x\}| - |B \setminus \{x\}|$.

Con lo anterior, tenemos que:

$$\begin{aligned}
|A \setminus B| &= |(A \setminus \{x\}) \setminus (B \setminus \{x\})| && \text{Igualdad conjuntista} \\
&= |A \setminus \{x\}| - |B \setminus \{x\}| && \text{Hipótesis de inducción sobre } A \setminus \{x\} \text{ y } B \setminus \{x\} \\
&= |A| - 1 - (|B| - 1) && \text{Lema 3.7.1} \\
&= |A| - |B|
\end{aligned}$$

Con lo anterior, $|A \setminus B| = |A| - |B|$.

Por el Principio de inducción matemática concluimos que si A es un conjunto finito y $B \subseteq A$, entonces $|A \setminus B| = |A| - |B|$. □

Aunque el siguiente resultado se deduce inmediatamente del corolario previo le otorgaremos rango de teorema por su importancia al momento de calcular cardinalidades de diferencias de conjuntos finitos.

Teorema 3.7.4. *Si A y B son conjuntos y A es finito, entonces $|A \setminus B| = |A| - |A \cap B|$.*

Demostración. Notemos que si A y B son conjuntos arbitrarios, entonces $A \setminus B = A \setminus (A \cap B)$. Si además, A es finito, entonces por el lema 3.7.3 se tiene que:

$$\begin{aligned}
|A \setminus B| &= |A \setminus (A \cap B)| && \text{Igualdad previa} \\
&= |A| - |A \cap B| && \text{Corolario 3.7.3}
\end{aligned}$$

Así, podemos concluir que $|A \setminus B| = |A| - |B \cap A|$. □

Ahora procederemos a demostrar un resultado referente a la unión de conjuntos.

Lema 3.7.5. *Si A y B son conjuntos finitos y ajenos, entonces $A \cup B$ es finito y $|A \cup B| = |A| + |B|$.*

Demostración. Esta demostración la realizaremos por inducción sobre $|A|$, es decir, mostraremos que: si $|A| = n$, B es finito y $A \cap B = \emptyset$, entonces $A \cup B$ es finito y $|A \cup B| = |A| + |B|$.

Base de Inducción. Si $|A| = 0$, B es finito y $A \cap B = \emptyset$, entonces $A \cup B$ es finito y $|A \cup B| = |A| + |B|$

Dado que $|A| = 0$, entonces $A = \emptyset$, por lo que $A \cup B = B$ y de nuestras hipótesis podemos concluir que $A \cup B$ es finito. Por otro lado, notemos :

$$\begin{aligned} |A \cup B| &= |\emptyset \cup B| \\ &= |B| \\ &= 0 + |B| \\ &= |A| + |B| \end{aligned}$$

Concluyendo de lo anterior que $|A \cup B| = |A| + |B|$.

Hipótesis de Inducción. Supongamos que k es un natural tal que si $|A'| = k$, B' es finito y $A' \cap B' = \emptyset$, entonces $A' \cup B'$ es finito y $|A' \cup B'| = |A'| + |B'|$.

Paso Inductivo. Si $|A| = k + 1$, B es finito y $A \cap B = \emptyset$, entonces $A \cup B$ es finito y $|A \cup B| = |A| + |B|$.

Notemos que como $|A| = k + 1$, entonces $A \neq \emptyset$ y consideremos $x \in A$. Es sencillo verificar que $|A \setminus \{x\}| = k$ y $(A \setminus \{x\}) \cap (B \cup \{x\}) = \emptyset$ y $B \cup \{x\}$ es finito, por lo que es posible usar hipótesis de inducción sobre $A \setminus \{x\}$ y $B \cup \{x\}$, obteniendo que $(A \setminus \{x\}) \cup (B \cup \{x\})$ es finito y $|(A \setminus \{x\}) \cup (B \cup \{x\})| = |A \setminus \{x\}| + |B \cup \{x\}|$.

Como $(A \setminus \{x\}) \cup (B \cup \{x\})$ es finito y $A \cup B = (A \setminus \{x\}) \cup (B \cup \{x\})$, entonces podemos concluir que $A \cup B$ es finito.

Ahora sólo resta mostrar que $|A \cup B| = |A| + |B|$. Para ello basta ver que:

$$\begin{aligned} |A \cup B| &= |(A \setminus \{x\}) \cup (B \cup \{x\})| && \text{Igualdad previa} \\ &= |A \setminus \{x\}| + |B \cup \{x\}| && \text{Hipótesis de inducción sobre } A \setminus \{x\} \text{ y } B \cup \{x\} \\ &= |A| - 1 + |B| + 1 && \text{Lemas 3.7.1 y 3.7.2} \\ &= |A| + |B| \end{aligned}$$

Así podemos concluir que $|A \cup B| = |A| + |B|$.

Por el Principio de inducción matemática, si A y B son conjuntos finitos y ajenos, entonces $A \cup B$ es finito y $|A \cup B| = |A| + |B|$. □

Como consecuencia de lo anterior tenemos el siguiente resultado. A pesar de que se deduce directamente, recibirá categoría de teorema debido a su importancia en el cálculo de cardinales de unión de conjuntos finitos.

Teorema 3.7.6. *Si A y B son conjuntos finitos, entonces $|A \cup B| = |A| + |B| - |A \cap B|$.*

Demostración. En este caso, notemos que $A \cup B = (A \setminus B) \cup B$. Por otro lado, es sencillo verificar que $A \setminus B$ y B son ajenos, por lo que por el lema 3.7.5 se tiene que $|A \cup B| = |A \setminus B| + |B|$ y por el teorema 3.7.4 se concluye que $|A \cup B| = |A| + |B| - |A \cap B|$. □

Ahora podemos establecer el siguiente teorema mediante un sencillo argumento inductivo.

Teorema 3.7.7. *Si $\{A_i : i \in \{1, \ldots, n\}\}$ es una familia de conjuntos finitos y ajenos dos a dos, entonces $\cup_{i=1}^n A_i$ es finito y además:*

$$|\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i|.$$

Demostración. La demostración la realizaremos por inducción sobre n, es decir, mostraremos que: si $\{A_i : i \in \{1, \ldots, n\}\}$ es una familia de conjuntos finitos y ajenos dos a dos, entonces $\cup_{i=1}^n A_i$ es finito y además:

$$|\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i|.$$

Base de inducción. si $\{A_i : i \in \{1\}\}$ es una familia de conjuntos finitos y ajenos dos a dos, entonces $\cup_{i=1}^1 A_i$ es finito y además $|\bigcup_{i=1}^1 A_i| = \sum_{i=1}^1 |A_i|$.

En este caso se tiene que $\bigcup_{i=1}^1 A_i = A_1$, por lo que dicha unión es finita y además $|\bigcup_{i=1}^1 A_i| = \sum_{i=1}^1 |A_i|$.

Hipótesis de inducción. Supongamos que k es un natural tal que si $\{A_i : i \in \{1, \ldots, k\}\}$ es una familia de conjuntos ajenos dos a dos, entonces $\cup_{i=1}^k A_i$ es finito y $|\bigcup_{i=1}^k A_i| = \sum_{i=1}^k |A_i|$.

Paso inductivo. Si $\{A_i : i \in \{1, \ldots, k+1\}\}$ es una familia de conjuntos ajenos dos a dos, entonces $\cup_{i=1}^{k+1} A_i$ es finito y $|\bigcup_{i=1}^{k+1} A_i| = \sum_{i=1}^{k+1} |A_i|$.

Sea $\{A_i : i \in \{1, \ldots, k+1\}\}$ una familia de conjuntos ajenos dos a dos. Dado que $\{A_i : i \in \{1, \ldots, k\}\}$ es una familia de conjuntos ajenos dos a dos, entonces por hipótesis de inducción se tiene que $\cup_{i=1}^k A_i$ es finito y además:

$$|\bigcup_{i=1}^{k} A_i| = \sum_{i=1}^{k} |A_i|$$

Por otro lado, es sencillo ver que $\bigcup_{i=1}^{k+1} A_i = \left(\bigcup_{i=1}^k A_i\right) \cup A_{k+1}$, por lo que $\bigcup_{i=1}^{k+1} A_i$ es finito y además:

$$\begin{aligned}
\left|\bigcup_{i=1}^{k+1} A_i\right| &= \left|\left(\bigcup_{i=1}^k A_i\right) \cup A_{k+1}\right| && \text{Igualdad previa} \\
&= \left|\bigcup_{i=1}^k A_i\right| + |A_{k+1}| - \left|\left(\bigcup_{i=1}^k A_i\right) \cap A_{k+1}\right| && \text{Teorema 3.7.6} \\
&= \left|\bigcup_{i=1}^k A_i\right| + |A_{k+1}| && \text{La familia es ajena dos a dos} \\
&= \left(\sum_{i=1}^k |A_i|\right) + |A_{k+1}| && \text{Hipótesis de inducción} \\
&= \sum_{i=1}^{k+1} |A_i| && \text{Definición de suma}
\end{aligned}$$

De lo anterior se concluye que $|\bigcup_{i=1}^{k+1} A_i| = \sum_{i=1}^{k+1} |A_i|$.

Gracias al Principio de inducción matemática podemos concluir que si $\{A_i : i \in \{1, \ldots, n\}\}$ es una familia de conjuntos finitos y ajenos dos a dos, entonces $\cup_{i=1}^{n} A_i$ es finito y $|\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i|$. □

Como consecuencia directa del resultado anterior tenemos un teorema que permite realizar el conteo de un conjunto finito con base en una partición. Nuevamente, aunque dicho resultado es consecuencia directa del enunciado previo, recibirá categoría de teorema debido a la importancia al momento de realizar el cálculo de cardinales.

Teorema 3.7.8 (De la Cajera). *Si A es un conjunto finito y no vacío y $\{A_i : i \in \{1, \ldots, n\}\}$ es una partición de A, entonces:*

$$|A| = \sum_{i=1}^{n} |A_i|.$$

Demostración. Como A es un conjunto finito y $\{A_i : i \in \{1, \ldots, n\}\}$ es una partición de A, en particular tenemos que $\{A_i : i \in \{1, \ldots, n\}\}$ es una familia de conjuntos finitos y ajenos dos a dos. Por el teorema 3.7.7 se tiene que $|\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i|$. Por último, como $A = \bigcup_{i=1}^{n} A_i$ podemos concluir que

$$|A| = \sum_{i=1}^{n} |A_i|.$$

□

Como consecuencia del teorema de la cajera, tenemos dos principios importantes en conteo.

Teorema 3.7.9 (Principio de las casillas). *Si A es un conjunto finito y no vacío y $\{A_i : i \in \{1, \ldots, n\}\}$ es una partición de A, entonces los siguientes enunciados se satisfacen:*

a) Existe $\alpha \in \{1, \ldots, n\}$ tal que $|A_\alpha| \leq \frac{|A|}{n}$.

b) Existe $\beta \in \{1, \ldots, n\}$ tal que $|A_\beta| \geq \frac{|A|}{n}$.

Demostración. Para esta demostración sólo probaremos el inciso (a) y el inciso (b) se realiza de manera análoga.

Procederemos por contradicción al suponer que para todo $\alpha \in \{1, \ldots, n\}$, $|A_\alpha| > \frac{|A|}{n}$. Así, tenemos lo siguiente.

$$\begin{aligned} |A| &= \sum_{i=1}^{n} |A_i| && \text{Teorema de la cajera (teorema 3.7.8)} \\ &> \sum_{i=1}^{n} \tfrac{|A|}{n} && \text{Suposición de la contradicción} \\ &= n(\tfrac{|A|}{n}) && \text{Propiedades de suma} \\ &= |A| \end{aligned}$$

De lo anterior se deduce que $|A| > |A|$, lo cual no es posible, concluyendo que existe $\alpha \in \{1, \ldots, n\}$ tal que $|A_\alpha| \leq \frac{|A|}{n}$. □

Corolario 3.7.10 (Principio del palomar). *Si A es un conjunto finito con n elementos y $\mathscr{A}$ es una partición de A con $n-1$ elementos, entonces existe $A_i \in \mathscr{A}$ tal que $2 \leq |A_i|$.*

Demostración. Este enunciado se sigue inmediatamente del inciso (b) del Principio de las casillas (teorema 3.7.9). □

Todos los resultados antes vistos brindan herramientas muy útiles para poder calcular el cardinal de conjuntos finitos, aun si son conjuntos no contemplados hasta ahora. Por ejemplo, tenemos el siguiente teorema sobre cardinal del Producto cartesiano.

Teorema 3.7.11. *Si A y B son finitos, entonces $|A \times B| = |A| \cdot |B|$*

Demostración. Notemos que si $A = \emptyset$ o $B = \emptyset$, entonces $A \times B = \emptyset$. En tal caso, la igualdad $|A \times B| = |A| \cdot |B|$ se satisface. Bastará demostrar que el teorema 3.7.11 es cierto cuando $A \neq \emptyset$ y $B \neq \emptyset$.

Para cada $a \in A$ definimos $B_a = \{(x, z) \in A \times B : x = a\}$. Notemos que si $a \in A$, entonces la función $\varphi : B \to B_a$ dada por $\varphi(b) = (a, b)$ es biyectiva, por lo que para toda $a \in A$ se tiene que:

$$|B| = |B_a| \tag{3.1}$$

Por otro lado, no es difícil verificar que $\{B_a : a \in A\}$ es una partición de $A \times B$. Con lo anterior podemos concluir que:

$$\begin{aligned} |A \times B| &= \sum_{a \in A} |B_a| && \text{Teorema de la cajera (teorema 3.7.8)} \\ &= \sum_{a \in A} |B| && \text{Ecuación 3.1} \\ &= |A| \cdot |B| && \text{Propiedades de la suma} \end{aligned}$$

Así se concluye que $|A \times B| = |A| \cdot |B|$. □

Previo a iniciar la siguiente sección tenemos un último resultado que será de utilidad para el cálculo combinatorio.

Lema 3.7.12. *Si A es un conjunto finito y no vacío y $A_0 = \{(a_i, a_j) \in A \times A : a_i \neq a_j\}$, entonces $|A_0| = |A|(|A| - 1)$.*

Demostración. Para cada $a \in A$, definimos $B_a = \{(x, z) \in A_0 : x = a\}$. Es claro que la familia $\{B_a : a \in A\}$ es una partición de A_0, por lo que por el teorema de la cajera podemos concluir que

$$|A_0| = \sum_{a \in A} |B_a|.$$

Por otro lado, es sencillo verificar que $|B_a| = |A \setminus \{a\}|$, por lo que $|A_0| = |A|(|A| - 1)$. □

3.8. Cálculo combinatorio

Dado un subconjunto A con n elementos y $m \leq n$, una **combinación de m elementos de A** es un subconjunto con m elementos de A. Cuando no es necesario especificar quién es el conjunto A, decimos que una **combinación de m elementos tomados de un conjunto con n elementos** es un subconjunto con m elementos de un conjunto con n elementos. Al número total de dichas combinaciones se le denota por $\binom{n}{m}$. En general, a dichas expresiones también se les conoce como **coeficiente binomial**.

Teorema 3.8.1. *Si n y m son naturales y $m \leq n$, entonces $\binom{n}{m} = \frac{n!}{m!(n-m)!}$.*

Demostración. Consideremos un conjunto A con n elementos y definimos $\mathcal{P}_k(A)$ como el conjunto de subconjuntos de A que tienen k elementos. Es claro por la definición de coeficiente binomial que $|\mathcal{P}_k(A)| = \binom{n}{k}$, por lo que basta demostrar que si A es un conjunto finito con n elementos y $k \leq |A|$, entonces $|\mathcal{P}_k(A)| = \frac{n!}{k!(n-k)!}$.

Lo anterior lo haremos por inducción sobre el cardinal de A, es decir, mostraremos que: si $|A| = n$ y $k \leq |A|$, entonces $|\mathcal{P}_k(A)| = \frac{n!}{k!(n-k)!}$.

Base de Inducción. Si $|A| = 0$ y $k \leq |A|$, entonces $|\mathcal{P}_k(A)| = \frac{n!}{k!(n-k)!}$.

Dado que $|A| = 0$, entonces $k = 0$ y así, $|\mathcal{P}_k(A)| = 1 = \frac{n!}{k!(n-k)!}$.

Hipótesis de Inducción. Si $|A'| = t$ y $l \leq |A'|$, entonces $|\mathcal{P}_l(A')| = \frac{t!}{l!(t-l)!}$.

Paso Inductivo. Si A es un conjunto finito, $|A| = t + 1$ y $k \leq |A|$, entonces $|\mathcal{P}_k(A)| = \frac{(t+1)!}{k!(t+1-k)!}$.

Primero hay que hacer notar que si $k = t + 1$, entonces $|\mathcal{P}_k(A)| = 1$ y $\frac{(t+1)!}{k!(t+1-k)!} = 1$, por lo que $|\mathcal{P}_k(A)| = \frac{(t+1)!}{k!(t+1-k)!}$.

Por otro lado, si $k = t$, entonces $|\mathcal{P}_k(A)| = t + 1$ y $\frac{(t+1)!}{k!(t+1-k)!} = t + 1$, por lo que nuevamente $|\mathcal{P}_k(A)| = \frac{(t+1)!}{k!(t+1-k)!}$.

Así, para lo que resta de la demostración supondremos que $k < t$. Consideremos ahora un elemento $x \in A$ arbitrario. Notemos que:

$$\mathcal{P}_k(A) = \{D \in \mathcal{P}_k(A) : x \notin D\} \cup \{D \in \mathcal{P}_k(A) : x \in D\}$$

Con lo anterior, procederemos a calcular los cardinales de cada uno de los uniendos mostrados en la igualdad previa. Ello lo mostraremos en las siguientes afirmaciones.

Afirmación 1. $|\{D \in \mathcal{P}_k(A) : x \notin D\}| = \frac{t!}{k!(t-k)!}$.

Notemos que $\{D \in \mathcal{P}_k(A) : x \notin D\} = \mathcal{P}_k(A \setminus \{x\})$ y con ello:

$$|\{D \in \mathcal{P}_k(A) : x \notin D\}| = |\mathcal{P}_k(A \setminus \{x\})|$$

Dado que $|A \setminus \{x\}| = t$ y $k \leq |A \setminus \{x\}|$, entonces por hipótesis de inducción podemos concluir que $|\mathcal{P}_k(A \setminus \{x\})| = \frac{t!}{k!(t-k)!}$. Y de ello, podemos garantizar que:

$$|\{D \in \mathcal{P}_k(A) : x \notin D\}| = \frac{t!}{k!(t-k)!}$$

Concluyendo la demostración de esta afirmación.

Afirmación 2. $|\{D \in \mathcal{P}_k(A) : x \in D\}| = \frac{t!}{(k-1)!(t-(k-1))!}$.

No es difícil verificar que la función $\varphi : \{D \in \mathcal{P}_k(A) : x \in D\} \to \mathcal{P}_{k-1}(A \setminus \{x\})$ dada por $\varphi(D) = D \setminus \{x\}$ es una función biyectiva, por lo que

$$|\{D \in \mathcal{P}_k(C) : x \in C\}| = |\mathcal{P}_{k-1}(C \setminus \{x\})|$$

Además, como $|A \setminus \{x\}| = t$ y $k - 1 \leq |A \setminus \{x\}|$, entonces por hipótesis de inducción tenemos que:

$$|\mathcal{P}_{k-1}(A \setminus \{x\})| = \frac{t!}{(k-1)!(t-(k-1))!}$$

Concluyendo que:

$$\{D \in \mathcal{P}_k(C) : x \in D\}| = \frac{t!}{(k-1)!(t-(k-1))!}$$

Concluyendo la demostración de esta afirmación.

Así, realizando la suma correspondiente para calcular el cardinal de $|P_k(A)|$, tenemos que:

$$\begin{aligned}
|P_k(A)| &= |\{D \in \mathcal{P}_k(A) : x \notin D\} \cup \{D \in \mathcal{P}_k(A) : x \in D\}| && \text{Ecuación previa}\\
&= |\{D \in \mathcal{P}_k(A) : x \notin D\}| + |\{D \in \mathcal{P}_k(A) : x \in D\}| && \text{Cardinal de unión ajena}\\
&= \tfrac{t!}{k!(t-k)!} + \tfrac{t!}{(k-1)!(t-(k-1))!} && \text{Afirmación 1 y afirmación 2}\\
&= \tfrac{(t+1)!}{k!(t+1-k)!} && \text{Cuentas}
\end{aligned}$$

De lo anterior, podemos decir que $|P_k(A)| = \frac{(t+1)!}{k!(t+1-k)!}$.

Por el principio de inducción matemática se tiene que si $|A| = n$ y $m \leq |A|$, entonces $|\mathcal{P}_k(A)| = \frac{n!}{m!(n-m)!}$ y por definición de combinaciones, podemos concluir que $\binom{n}{m} = \frac{n!}{m!(n-m)!}$. □

Los coeficientes binomiales son una herramienta útil para mostrar condiciones de conteo, por lo que ahora procederemos a demostrar algunas propiedades de los mismos.

Corolario 3.8.2. *Sean $\{n, m, k\} \subseteq \mathbb{N}$ tales que $m \leq n$ y $k \leq n$. Los siguientes se satisfacen:*

a) $\binom{n}{0} = 1$ *y* $\binom{n}{n} = 1$.

b) $\binom{n}{m} = \binom{n}{n-m}$.

c) $\binom{n}{m}\binom{n-m}{k} = \binom{n}{k}\binom{n-k}{m}$.

d) Si $k \geq 1$, *entonces* $n\binom{n-1}{k-1} = k\binom{n}{k}$.

e) Si $m \leq k$, *entonces* $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$.

f) Si $k < n$, *entonces* $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

Demostración. Para este corolario sólo realizaremos las demostraciones de algunos incisos.

(b) Si $k \geq 1$, entones $\binom{n}{m} = \binom{n}{n-m}$.

Para este inciso basta ver que:

$$\begin{aligned}
\binom{n}{n-m} &= \tfrac{n!}{(n-m)!(n-(n-m)!} && \text{Definición de coeficiente binomial}\\
&= \tfrac{n!}{m!(n-m)!} && \text{Cuentas}\\
&= \binom{n}{m} && \text{Definición de coeficiente binomial}
\end{aligned}$$

Con lo anterior podemos concluir que $\binom{n}{m} = \binom{n}{n-m}$.

(d) Si $k \geq 1$, entonces $n\binom{n-1}{k-1} = k\binom{n}{k}$.

Para este inciso basta ver que:

$$\begin{aligned}
n\binom{n-1}{k-1} &= n\left(\tfrac{(n-1)!}{(k-1)!((n-1)-(k-1))!}\right) && \text{Definición de coeficiente binomial}\\
&= \tfrac{n!}{(k-1)!(n-k)!} && \text{Cuentas}\\
&= \tfrac{k(n!)}{k(k-1)!(n-k)!} && \text{Multiplicación por 1}\\
&= k\left(\tfrac{n!}{k!(n-k)!}\right)\\
&= k\binom{n}{k}
\end{aligned}$$

Con lo anterior podemos concluir que $n\binom{n-1}{k-1} = k\binom{n}{k}$.

(f) Si $k < n$, entonces $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

En este inciso, notemos que:

$$\begin{aligned}
\tbinom{n}{k} + \tbinom{n}{k+1} &= \tfrac{n!}{k!(n-k)!} + \tfrac{n!}{(k+1)!(n-(k+1))!} && \text{Definición de coeficiente binomial} \\
&= \tfrac{(k+1)n!}{(k+1)k!(n-k)!} + \tfrac{(n-k)n!}{(k+1)!(n-k)(n-k-1)!} && \text{Multiplicación por 1 en ambos sumandos} \\
&= \tfrac{(k+1)n!}{(k+1)!(n-k)!} + \tfrac{(n-k)n!}{(k+1)!(n-k)!} && \text{Cuentas} \\
&= \tfrac{(k+1)n!+(n-k)n!}{(k+1)!(n-k)!} \\
&= \tfrac{n!(k+1+n-k)}{(k+1)!(n-k)!} \\
&= \tfrac{n!(n+1)}{(k+1)!((n+1)-(k+1))!} \\
&= \tfrac{(n+1)!}{(k+1)!((n+1)-(k+1))!} \\
&= \tbinom{n+1}{k+1} && \text{Definición de coeficiente binomial}
\end{aligned}$$

Con lo anterior se concluye que si $k < n$, entonces $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

□

Como consecuencia de lo anterior, podemos demostrar el siguiente teorema. Cabe mencionar que aunque dicho resultado muestra de manera explícita la forma de calcular los coeficientes de un binomio perfecto, este resultado tiene diversas formas de implementarse en diversos ámbitos.

Teorema 3.8.3 (Del Binomio). *Si* $\{a, b\} \subseteq \mathbb{R}$ *y* $n \in \mathbb{N}$, *entonces*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Demostración. La demostración de este teorema la realizaremos por inducción sobre n, es decir, mostraremos que: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$.

Base de inducción. $(a+b)^0 = \sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k$.

En este caso, es sencillo ver que $(a+b)^0 = 1$ y que $\sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k = 1$, por lo que $(a+b)^0 = \sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k$.

Hipótesis de inducción. Supongamos que n es un natural tal que $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$.

Paso inductivo. $(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k$.

Para el paso inductivo, basta ver lo siguiente:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b)\left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right) \\
&= a\left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right) + b\left(\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k\right) \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} \\
&= \binom{n}{0}a^{n+1}b^0 + \binom{n}{1}a^n b^1 + \ldots + \binom{n}{n-1}a^2b^{n-1} + \binom{n}{n}a^1b^n \\
&\quad +\binom{n}{0}a^n b^1 + \binom{n}{1}a^{n-1}b^2 + \ldots + \binom{n}{n-1}a^1b^n + \binom{n}{n}a^0b^{n+1} \\
&= \binom{n}{0}a^{n+1}b^0 + \left(\binom{n}{0}+\binom{n}{1}\right)a^nb^1 + \ldots + \left(\binom{n}{n}+\binom{n}{n-1}\right)a^1b^n + \binom{n}{n}a^0b^{n+1} \\
&= \binom{n+1}{0}a^{n+1}b^0 + \binom{n+1}{1}a^nb^1 + \ldots + \binom{n+1}{n}a^1b^n + \binom{n+1}{n+1}a^0b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k
\end{aligned}$$

Con lo anterior podemos concluir que $(a+b)^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n+1-k}b^k$.

Por el principio de inducción matemática obtenemos que si $\{a,b\} \subseteq \mathbb{R}$ y $n \in \mathbb{N}$, entonces $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k$. □

Un resultado interesante derivado del teorema del binomio es el siguiente:

Corolario 3.8.4. *Si A es un conjunto finito, entonces $|\mathcal{P}(A)| = 2^{|A|}$.*

Demostración. Supongamos que $|A| = n$. Es sencillo verificar que

$$\begin{aligned}
|\mathcal{P}(A)| &= \sum_{k=0}^{n}|\mathcal{P}_k(A)| && \text{Por definición de los conjuntos } \mathcal{P}_k(A) \\
&= \sum_{k=0}^{n}\binom{n}{k} && \text{Definición de coeficiente binomial}
\end{aligned}$$

Por otro lado, tenemos que:

$$\begin{aligned}
2^n &= (1+1)^n && \text{Cuentas} \\
&= \sum_{k=0}^{n}\binom{n}{k}1^{n-k}\cdot 1^k && \text{Teorema del binomio} \\
&= \sum_{k=0}^{n}\binom{n}{k}
\end{aligned}$$

Con lo anterior, concluimos que $|\mathcal{P}(A)| = 2^n$. □

Dado un conjunto A con n elementos y $m \in \mathbb{N}$, una **ordenación (con repetición) de m elementos de A** es una función $f : I_m \to A$. Cuando no es necesario especificar quién es el conjunto A, decimos que una **ordenación (con repetición) de m elementos tomados de un conjunto de n elementos** es una función de I_m en un conjunto con n elementos. Al número total de dichas ordenaciones con repetición se le denota por OR_n^m.

Teorema 3.8.5. *Si n y m son naturales, entonces $OR_n^m = n^m$.*

Demostración. Sea A un conjunto con n elementos y denotamos por $F(I_m; A)$ al conjunto de las funciones de I_m en A, es decir, $F(I_m; A) = \{f \subseteq I_m \times A : f \text{ es una función de } I_m \text{ en } A\}$.

Afirmación. $|F(I_m; A)| = |A^m|$.

Consideremos la relación φ de A^m en $F(I_m; A)$ dada por $((a_1, a_2, \ldots, a_m), f) \in \varphi$ si y sólo si para toda $k \in \{1, \ldots, m\}$, $f(k) = a_k$. Mostraremos que dicha relación es una función biyectiva.

a) φ es una función de A^m en $Fun(I_m; A)$.

Primero mostraremos por doble contención que $Dom(\varphi) = A^m$.

Claramente $Dom(\varphi) \subseteq A^m$. Pala mostrar la otra contención, consideremos $(b_1, \ldots, b_m) \in A^m$. La relación dada por $(n, b_k) \in f$ si y sólo si $k = n$ es una función de I_m en A, por lo que $f \in F(I_m; A)$ y además, $f(k) = b_k$ para toda $k \in \{1, \ldots, m\}$, por lo que $((b_1, \ldots, b_m), f) \in \varphi$ y así, $A^m \subseteq Dom(\varphi)$, concluyendo que $A^m = Dom(\varphi)$.

Ahora mostraremos que la asociación de parejas de φ es única.

Si $((a_1, \ldots, a_m), f) \in \varphi$ y $((b_1, \ldots, b_n), g) \in \varphi$ y $(a_1, \ldots, a_m) = (b_1, \ldots, b_m)$ entonces tenemos que para toda $k \in \{1, \ldots, m\}$ se tiene que $a_k = b_k$. Además, por definición de φ se sigue que $f(k) = a_k$ y $g(k) = b_k$, por lo cual, para toda $k \in I_m$ se cumple que $f(k) = g(k)$ y por el teorema 2.5.6 se sigue que $f = g$.

De lo anterior, φ es una función.

b) φ es inyectiva.

Sean $(a_1, a_2, \ldots, a_m)$ y $(b_1, \ldots, b_m)$ en A^m tales que $\varphi(a_1, a_2, \ldots, a_m) = g$ y que $\varphi(b_1, \ldots, b_m) = g$. Por definición de φ se sigue que para toda $k \in I_m$ se tiene que $g(k) = a_k$ y $g(k) = b_k$, por lo que para toda $k \in \{1, \ldots, m\}$ se obtiene que $a_k = b_k$ y así $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$.

c) φ es sobreyectiva.

Sea $f \in F(I_m; A)$. Dado que para toda $k \in I_m$ se tiene que $f(k) \in A$, entonces $(f(1), f(2), \ldots, f(m))$ es un elemento en A^m. Es inmediato de la definición de φ el que $\varphi(f(1), f(2), \ldots, f(m)) = f$.

Dado que φ es biyectiva, entonces $|A^m| = |F(I_m; A)|$, pero $|A^m| = |A|^m = n^m$, concluyendo que $|F(I_m; A)| = n^m$. De la definición de OR_n^m podemos afirmar que $OR_n^m = n^m$. De lo anterior se tiene que $OR_m^n = n^m$. □

Como consecuencia del resultado anterior obtenemos una forma de calcular el número de elementos del conjunto potencia.

Corolario 3.8.6. *Si n es un natural, entonces $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.*

Demostración. Sea A un conjunto con n elementos. Por definición sabemos que $\binom{n}{k}$ es el número de subconjuntos con k elementos de A. Así, $\sum_{i=0}^{n}\binom{n}{k} = |P(A)|$, pero por lo mostrado en el corolario 3.8.4 podemos concluir que $\sum_{k=0}^{n}\binom{n}{k} = 2^n$. □

Una **permutación** de un conjunto A es una función biyectiva de A en A. Al número de permutaciones de un conjunto con n elementos se le denota por P_n.

Lema 3.8.7. *Si n es un natural, entonces $P_n = n!$*

Demostración. Sean A y B conjuntos, denotamos por $Biy(A;B)$ al conjunto de funciones biyectivas de A en B. Demostraremos que si A y B con conjuntos con n elementos, entonces $|Biy(A;B)| = n!$ Dicha demostración la haremos por inducción sobre el número de elementos n.

Base de Inducción. Si A y B con conjuntos con 0 elementos, entonces $|Biy(A;B)| = n!$

Como A y B tienen 0 elementos, entonces tanto B como A son vacíos, por lo que la única función biyectiva entre ellos es la función vacía. En tal caso, $|Biy(A;B)| = 1$ pero como $0! = 1$, entonces concluimos lo deseado.

Hipótesis de Inducción. Si A' y B' son conjuntos con k elementos, entonces $|Biy(A';B')| = k!$

Paso Inductivo. Si A y B son conjuntos con $k+1$ elementos, entonces $|Biy(A;B)| = (k+1)!$

Supongamos que $A = \{a_1, \ldots, a_{k+1}\}$ y $B = \{b_1, \ldots, b_{k+1}\}$. Para cada $i \in \{1, \ldots, k+1\}$ definimos

$$F_i = \{f \in Biy(A;B) : f(a_1) = b_i\}.$$

Afirmación 1. Para toda $i \in \{1, \ldots, k+1\}$ se tiene que $|F_i| = |Biy(A \setminus \{a_1\}; B \setminus \{b_i\})|$.

Consideremos la relación η de F_i en $Biy(A \setminus \{a_1\}; B \setminus \{b_i\})$ dada por $(f,g) \in \eta$ si y sólo si $g = f \setminus \{(a_1, b_i)\}$.

Es simple ver que η es una función de F_i en $Biy(A \setminus \{a_1\}; B \setminus \{b_i\})$. Además, dicha función es biyectiva, por lo que el resultado de la afirmación se satisface.

A partir de lo anterior, aplicando Hipótesis de inducción se tiene que $|Biy(A \setminus \{a_1\}; A \setminus \{b_i\})| = k!$, por lo que para toda $i \in \{1, \ldots, k+1\}$, $|F_i| = k!$

Afirmación 2. $\{F_i : i \in \{1, \ldots, k+1\}\}$ es una partición de $Biy(A;B)$.

Es inmediata de la definición de partición.

A partir de lo anterior, por el teorema de la cajera, se concluye que:

$$|Biy(A;B)| = \sum_{i=1}^{k+1} |F_i| = \sum_{i=1}^{k+1} k! = (k+1)k! = (k+1)!$$

Por el principio de inducción matemática podemos concluir que si A y B con conjuntos con n elementos, entonces $|Biy(A;B)| = n!$, y por definición de P_n, se concluye que $P_n = n!$ □

Dado un conjunto finito A y $m \leq |A|$, una **ordenación de m elementos de A** es una función inyectiva $f : I_m \to A$. Cuando no es necesario especificar quién es el conjunto A, decimos que una **ordenación de m elementos tomados de un conjunto con n elementos** es una función inyectiva de I_m en un conjunto con n elementos. Al número total de dichas ordenaciones se le denota por O_n^m.

Teorema 3.8.8. *Si n y m son naturales y $m \leq n$, entonces* $O_n^m = \frac{n!}{(n-m)!}$

Demostración. Sean $A = \{a_1, \ldots, a_m\}$ y $B = \{b_1, \ldots, b_n\}$ dos conjuntos tales que $m \leq n$. Demostraremos que $|Iny(A;B)| = \frac{n!}{(n-m)!}$.

Para ello, supongamos que $\mathcal{P}_m(B) = \{C_1, C_2, \ldots, C_t\}$, donde $t = \frac{n!}{m!(n-m)!}$. Para cada $i \in \{1, \ldots, t\}$, definimos $F_i = \{f \in Iny(A;B) : Im(f) = C_i\}$. Notemos que $F_i = Biy(A;C_i)$, por lo que para toda $i \in \{1, \ldots, k\}$ se tiene que $|F_i| = m!$

Por otro lado, no es difícil de ver que $\{F_i : i \in \{1, \ldots, t\}\}$ es una partición de $Iny(A;B)$, por lo que por el teorema de la cajera:

$$|Iny(A;B)| = \sum_{i=1}^{t} |F_i| = \sum_{i=!}^{t} m! = tm! = \left(\frac{n!}{m!(n-m)!}\right) m! = \frac{n!}{(n-m)!}$$

□

3.9. Apéndice del Capítulo 3

Aunque en este capítulo se obtienen varias definiciones, serán los resultados los que cobran mayor importancia. Es importante mencionar que el orden de los resultados aquí presentados puede ser diferente al presentado a lo largo de este capítulo. Dentro de los resultados destacados y de mayor utilidad están los siguientes.

Principio de inducción débil. Sea P un predicado que toma valores sobre los números naturales. Para verificar que P es verdadero para todo natural mayor o igual que algún natural k_0 basta realizar lo siguiente:

a) Base Inductiva. Demostrar que $P(k_0)$ es verdadero.

b) Hipótesis de Inducción. Suponemos que $P(n)$ es verdadera para algún $n \geq k_0$.

c) Paso Inductivo. Se demuestra que $P(n+1)$ es verdadera.

Entonces P es verdadera para todo natural mayor o igual que k_0.

Principio de inducción fuerte. Sea P un predicado que toma valores sobre los números naturales. Para verificar que P es verdadero para todo natural mayor o igual que algún natural k_0 basta realizar lo siguiente:

a) Base Inductiva. Demostrar que $P(k_0)$ es verdadero.

b) Hipótesis de Inducción. Supondremos que n es un natural tal que para todo $k \in \mathbb{N}$ tal que $k_0 \leq k \leq n$ se satisface que $P(k)$ es verdadera.

c) Paso Inductivo. Se demuestra que $P(n+1)$ es verdadera.

Entonces P es verdadera para todo natural mayor o igual que k_0.

Lema (propiedades del producto y la suma) Para cualesquiera naturales m, n y k, se satisfacen los siguientes incisos:

a) $0 + n = n + 0 = n$.

b) $m + n = n + m$.

c) $(m + n) + r = m + (n + r)$.

d) Si $m + n = m + k$, entonces $n = k$.

e) Si $n \neq 0$, entonces $n + m \neq 0$.

f) $n + m = 0$ si y sólo si $n = 0$ y $m = 0$.

g) $n0 = 0n = 0$.

h) $1n = n1 = n$.

i) $mn = nm$.

j) $m(n + k) = mn + mk$.

k) $(nm)k = n(mk)$

l) $mn = 0$ si y sólo si $m = 0$ o $n = 0$.

m) Si $mn = mk$ y $m \neq 0$, entonces $n = k$.

Lema (propiedades del producto y suma) Sean $n_1, \ldots, n_k$ números naturales. Los siguientes enunciados se satisfacen:

a) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i = t$, entonces $\sum_{i=1}^{k} n_i = kt$.

b) Si $j \in \{1, \ldots, k-1\}$, entonces $\sum_{i=1}^{k} n_i = \left(\sum_{i=1}^{j} n_i\right) + \left(\sum_{i=j+1}^{k} n_i\right)$.

c) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i = t$, entonces $\prod_{i=1}^{k} n_i = t^k$.

d) Si $j \in \{1, \ldots, k\}$, entonces $\prod_{i=1}^{k} n_i = \left(\prod_{i=1}^{j} n_i\right) \left(\prod_{i=j+1}^{k} n_i\right)$.

Orden usual. Sean a y b naturales arbitrarios. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $a \leq b$

b) Existe $k \in \mathbb{N}$ tal que $a + k = b$

a) $a \not\leq b$

b) $b < a$

Lema (propiedades de orden) Si m, n, k y l son naturales, entonces los siguientes enunciados se satisfacen.

a) Si $m \leq n$, entonces $m + k \leq n + k$.

b) Si $m \leq n$ y $k \leq l$, entonces $m + k \leq n + l$.

c) Si $m \leq n$, entonces $mk \leq nk$

d) Si $m \leq n$ y $k \leq l$, entonces $mk \leq nl$.

e) $0 \leq n$.

f) $n \leq n + k$

g) Si $n \not\leq m$, entonces $m < n$.

h) Si $n < m$, entonces $n + 1 \leq m$.

i) Si $n \leq m$ y $m \leq n + 1$, entonces $m = n$ o $m = n + 1$.

Lema (propiedades de orden) Sean $n_1, \ldots, n_k$ y $b_1, \ldots, b_k$ números naturales y $t \in \mathbb{N}$. Los siguientes enunciados se satisfacen:

a) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq b_i$, entonces $\sum_{i=1}^{k} n_i \leq \sum_{i=1}^{k} b_i$.

b) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq t$, entonces $\sum_{i=1}^{k} n_i \leq kt$.

c) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq b_i$, entonces $\prod_{i=1}^{k} n_i \leq \prod_{i=1}^{k} b_i$.

d) Si para todo $i \in \{1, \ldots, k\}$ se satisface que $n_i \leq t$, entonces $\prod_{i=1}^{k} n_i \leq t^k$.

Teorema del Buen Orden en $\mathbb{N}$. El orden usual en los naturales es un buen orden.

Teorema (de los conjuntos finitos). Sea A un conjunto arbitrario. A es finito si y sólo si $A = \emptyset$ o existe $n \geq 1$ tal que $|A| = n$.

Lema (igualdad y contención de conjuntos finitos). Sean A un conjunto finito y B un conjunto arbitrario. Los siguientes enunciados se satisfacen.

a) Si $B = A$, entonces B es finito y $|B| = |A|$.

b) Si $B \subseteq A$, entonces B es finito y $|B| \leq |A|$.

c) Si $B \subseteq A$ y $|A| = |B|$, entonces $A = B$.

Teorema (cardinales de operaciones de conjuntos). Sean A y B conjuntos finitos. Los siguientes enunciados se satisfacen:

a) $A \setminus B$ es finito y $|A \setminus B| = |A| - |A \cap B|$.

b) $A \cup B$ es finito y $|A \cup B| = |A| + |B| - |A \cap B|$.

c) $A \times B$ es finto y $|A \times B| = |A| \cdot |B|$.

d) $\mathcal{P}(A)$ es finito y $|\mathcal{P}(A)| = 2^{|A|}$.

Teorema (comparabilidad de cardinales finitos). Sean A y B conjuntos finitos. Los siguientes enunciados se satisfacen.

a) Existe una función inyectiva de A en B si y sólo si $|A| \leq |B|$.

b) Existe una función sobreyectiva de A en B si y sólo si $|B| \leq |A|$.

c) Existe una función biyectiva de A en B si y sólo si $|B| = |A|$.

Teorema (de la cajera). Si A es un conjunto finito y no vacío y $\{A_i : i \in \{1, \ldots, n\}\}$ es una partición de A, entonces:

$$|A| = \sum_{i=1}^{n} |A_i|.$$

Principio de las casillas.

Si A es un conjunto finito y no vacío y $\{A_i : i \in \{1, \ldots, n\}\}$ es una partición de A, entonces los siguientes enunciados se satisfacen:

a) Existe $\alpha \in \{1, \ldots, n\}$ tal que $|A_\alpha| \leq \frac{|A|}{n}$.

b) Existe $\beta \in \{1, \ldots, n\}$ tal que $|A_\beta| \geq \frac{|A|}{n}$.

Principio del palomar. Si A es un conjunto finito con n elementos y $\mathscr{A}$ es una partición de A con $n-1$ elementos, entonces existe $B \in \mathscr{A}$ tal que $2 \leq |B|$.

Sobre el uso de los números combinatorios. En la siguiente tabla se muestra el uso de los números combinatorios de acuerdo al problema que se esté trabajando. En dicho problema debe elegirse ciertos elementos de un conjunto de n objetos.

Uso de símbolos	¿Hay orden?	¿Hay repetición?	Se requiere:	Símbolo:	Cálculo:
Se usan los n objetos	Sí	Sí	Permutación con repetición[1]	$P_n^{a_1 a_2 \ldots a_n}$	$\frac{n!}{a_1! a_2! \ldots a_n!}$
		No	Permutación	P_n	$n!$
	No	Sí	x		1
		No	x		1
Se usan sólo m objetos $(m < n)$	Sí	Sí	Ordenación con repetición	OR_n^m	n^m
		No	Ordenación	O_n^m	$\frac{n!}{(n-m)!}$
	No	Sí	Combinación con repetición	CR_n^m	$\frac{(n+m-1)!}{m!(n-m)!}$
		No	Combinación	$\binom{n}{m}$	$\frac{n!}{m!(n-m)!}$

Lema (propiedades de combinaciones). Sean $\{n, m, k\} \subseteq \mathbb{N}$. Los siguientes enunciados se satisfacen:

a) $\binom{n}{0} = 1$ y $\binom{n}{n} = 1$.

b) $\binom{n}{m} = \binom{n}{n-m}$.

c) $n\binom{n-1}{k-1} = k\binom{n}{k}$.

d) $\binom{n}{m}\binom{n-m}{k} = \binom{n}{k}\binom{n-k}{m}$.

e) $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$.

f) $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

Teorema (Del Binomio). Si $\{a, b\} \subseteq \mathbb{R}$ y $n \in \mathbb{N}$, entonces

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

1. En esta notación, el valor a_i representa cuántas veces se repite el i-ésimo objeto

Capítulo 4

Espacios Vectoriales

4.1. Espacios vectoriales

Un **espacio vectorial sobre** $\mathbb{R}$ consiste en un conjunto no vacío, digamos V, y dos funciones $+ : V \times V \to V$ llamada **suma de vectores** y $\cdot : \mathbb{R} \times V \to V$ llamada **producto por escalares**, que satisfacen:

a) Para todo $x \in V$ y $y \in V$, $x + y = y + x$.

b) Para todo $x \in V$, $y \in V$ y $z \in V$ se cumple que $x + (y + z) = (x + y) + z$.

c) Existe un elemento $x_0 \in V$ tal que para todo $x \in V$ se cumple que $x + x_0 = x$.

d) Para todo $x \in V$ existe $y \in V$ tal que $x + y = x_0$.

e) Para todo $\lambda \in \mathbb{R}$ y $\mu \in \mathbb{R}$ se cumple que $(\lambda\mu)x = \lambda(\mu x)$.

f) Existe $\lambda_0 \in \mathbb{R}$ tal que para todo $x \in V$ se cumple que $\lambda_0 x = x$.

g) Si $\lambda \in \mathbb{R}$ y $x \in V$ y $z \in V$, entonces $\lambda(x + z) = \lambda x + \lambda z$.

h) Para todo $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$ y $x \in V$ se cumple que $(\lambda + \mu)x = \lambda x + \mu x$.

A los elementos del conjunto v se les llama **vectores**, mientras que a los elementos de $\mathbb{R}$ se les llama **escalares**. Es importante mencionar que los vectores y escalares pueden ser objetos de diferentes naturalezas, por lo que es importante hacer distinción de ellos. Con esto en mente, para lo que resta de estas notas denotaremos a los vectores por objetos con una línea sobre ellos, por ejemplo, $\overline{x}$ o $\overline{v}$, mientras que los escalares no contarán con dicha notación y se denotarán por letras griegas, por ejemplo, λ y μ. El siguiente es un lema sencillo de verificar a partir de las propiedades de espacio vectorial.

Lema 4.1.1. *Sea V un espacio vectorial. Los siguientes enunciados se satisfacen:*

a) *Si $\overline{x}$, $\overline{z}$ y $\overline{w}$ son vectores tales que $\overline{x}+\overline{z}=\overline{x}+\overline{w}$, entonces $\overline{z}=\overline{w}$.*

b) *El vector descrito en el punto (c) de la definición de espacio vectorial es único [a].*

c) *El vector descrito en el punto (d) de la definición de espacio vectorial es único y además se satisface que $\overline{x}+(-1)\overline{x}=\overline{0}$ [b].*

d) *Para todo $\overline{x}\in V$, $0\overline{x}=\overline{0}$.*

e) *Para todo $\lambda\in\mathbb{R}$ y todo $\overline{x}\in V$ se cumple que $(-\lambda)\overline{x}=-(\lambda\overline{x})=\lambda(-\overline{x})$.*

f) *Para todo $\lambda\in\mathbb{R}$ se cumple que $\lambda\overline{0}=\overline{0}$.*

[a] A este vector se le conoce como el **vector cero** y se denota por $\overline{0}$

[b] A este vector se le conoce como el **inverso de** $\overline{x}$ y se denota por $-\overline{x}$

Demostración. Sólo realizaremos la demostración de algunos de los incisos de este lema.

(b) El vector descrito en el punto (c) de la definición de espacio vectorial es único.

Supongamos que $\overline{w}$ es un vector tal que para todo $\overline{x}\in V$, $\overline{x}+\overline{w}=\overline{x}$. En tal caso, $\overline{w}+\overline{x}_0=\overline{x}_0$ y $\overline{w}+\overline{x}_0=\overline{w}$, por lo que $\overline{x}_0=w$, concluyendo que $\overline{x}_0$ es único.

(e) Sea $\overline{x}\in V$. Notemos que $\overline{0}x=(1-1)\overline{x}$ y por la propiedad (h) de espacio vectorial, se concluye que $0\overline{x}=\overline{x}+(-1)\overline{x}$, sin embargo, por el inciso (c) de este lema, $0\overline{x}=\overline{0}$.

□

Aunque existen varios espacios vectoriales, en este curso tendremos un interés particular en los espacios $\mathbb{R}^n$, para ello, es necesario definir una suma vectorial y un producto escalar.

Teorema 4.1.2. *El conjunto $\mathbb{R}^n$ es un espacio vectorial con las siguientes operaciones:*

a) $(x_1,\ldots,x_n)+(z_1,\ldots,z_n)=(x_1+z_1,\ldots,x_n+z_n)$.

b) $\lambda(x_1,\ldots,x_n)=(\lambda x_1,\ldots,\lambda x_n)$.

Demostración. Esta demostración se dará por hecho. □

Aunque en estas notas nos enfocaremos a estudiar los espacios $\mathbb{R}^n$ vistos como espacios vectoriales, la mayoría de los resultados aquí descritos son válidos para espacios vectoriales en general.

Si V es un espacio vectorial y W es un subconjunto de V, diremos que W es un **subespacio vectorial de V**, denotado por $W \leq V$, si W es un espacio vectorial con la suma y producto por escalares en V. A grandes rasgos, un subespacio vectorial es un espacio vectorial por sí mismo. Y de acuerdo a la definición de subespacio, si queremos demostrar que un subconjunto W es un subespacio vectorial, entonces debemos mostrar que se satisfacen las ocho propiedades de espacio vectorial, más las propiedades de cerradura de la suma vectorial y el producto escalar, lo cual es bastante talachudo. Sin embargo, hay dos conjuntos particulares para los cuales se puede demostrar que son subespacios vectoriales sin mucha dificultad. El siguiente teorema será de gran utilidad para demostrar que algo es un subespacio vectorial.

Teorema 4.1.3 (Del subespacio). *Sea V un $\mathbb{R}$ espacio vectorial y $W \subseteq V$. W es un subespacio vectorial de V si y sólo si W cumple las siguientes propiedades:*

a) $\overline{0} \in W$.

b) Si $\overline{x} \in W$ *y* $\overline{z} \in W$*, entonces* $\overline{x} + \overline{z} \in W$.

c) Si $\lambda \in \mathbb{R}$ *y* $\overline{x} \in W$*, entonces* $\lambda\overline{x} \in W$.

Con el resultado anterior, es sencillo verificar los siguientes:

Lema 4.1.4. *Si V es un espacio vectorial, entonces $\{\overline{0}\}$ y V son subespacios vectoriales de V.*

Al espacio $\{\overline{0}\}$ se le conocerá como **espacio vectorial trivial**. Un segundo resultado del teorema del subespacio es el siguiente.

Corolario 4.1.5. *Si V es un espacio vectorial y W_1 y W_2 son subespacios vectoriales de $\mathbb{R}$, entonces $W_1 \cap W_2$ es un subespacio vectorial de V.*

Demostración. De acuerdo al teorema del subespacio (teorema 4.1.3) basta verificar lo siguiente.

a) $\overline{0} \in W_1 \cap W_2$.

Para este punto, basta ver que como tanto W_1 como W_2 son subespacios vectoriales, entonces por el punto (a) del teorema del subespacio se tiene que $\overline{0} \in W_1$ y $\overline{0} \in W_2$, por lo cual, $\overline{0} \in W_1 \cap W_2$.

b) Si $\overline{x} \in W_1 \cap W_2$ y $\overline{z} \in W_1 \cap W_2$, entonces $\overline{x} + \overline{z} \in W_1 \cap W_2$.

Como $\overline{x} \in W_1 \cap W_2$, en particular se tiene que $\overline{x} \in W_1$ y $\overline{x} \in W_2$. Análogamente, Como $\overline{z} \in W_1 \cap W_2$, en particular se tiene que $\overline{z} \in W_1$ y $\overline{z} \in W_2$.

Ya que $\overline{x} \in W_1$ y $\overline{z} \in W_1$, se sigue del punto (b) del teorema del subespacio que $\overline{x} + \overline{z} \in W_1$. Análogamente, como $\overline{x} \in W_2$ y $\overline{z} \in W_2$, podemos concluir que $\overline{x} + \overline{z} \in W_1 + W_2$. Por definición de intersección se concluye que $\overline{x} + \overline{z} \in W_1 \cap W_2$.

c) Si $\lambda \in \mathbb{R}$ y $\overline{x} \in W_1 \cap W_2$, entonces $\lambda\overline{x} \in W_1 \cap W_2$.

Como $\overline{x} \in W_1 \cap W_2$, en particular se tiene que $\overline{x} \in W_1$ y $\overline{x} \in W_2$. Ya que $\lambda \in \mathbb{R}$ y $\overline{x} \in W_1$, se tiene del punto (c) del teorema del subespacio que $\lambda\overline{x} \in W_1$. Análogamente, como $\lambda \in \mathbb{R}$ y $\overline{x} \in W_2$, podemos garantizar que $\lambda\overline{x} \in W_1 \cap W_2$.

De los incisos previos y gracias al teorema del subespacio, se tiene que $W_1 \cap W_2$ es un subespacio vectorial de V. □

Corolario 4.1.6. *Sean V un espacio vectorial y W un subespacio de V. Si $\{\overline{w}_1, \ldots, \overline{w}_r\} \subseteq W$ y $c_1, \ldots, c_r \in \mathbb{R}$, entonces $c_1\overline{x}_1 + \cdots + c_r\overline{x}_r \in W$.*

Demostración. Mostraremos por inducción matemática sobre r que si $\{\overline{w}_1, \ldots, \overline{w}_r\} \subseteq W$ y $c_1, \ldots, c_r \in \mathbb{R}$, entonces $c_1\overline{x}_1 + \cdots + c_r\overline{x}_r \in W$.

Base de inducción. Si $\{\overline{w}_1\} \subseteq W$ y $c_1 \in \mathbb{R}$, entonces $c_1\overline{x}_1 \in W$.

Esto es inmediato del punto (c) del teorema del subespacio.

Hipótesis de inducción. Supongamos que k es un natural tal que si $\{\overline{w}_1, \ldots, \overline{w}_k\} \subseteq W$ y $c_1, \ldots, c_k \in \mathbb{R}$, entonces $c_1\overline{x}_1 + \cdots + c_k\overline{x}_k \in W$.

Paso inductivo. Si $\{\overline{w}_1, \ldots, \overline{w}_{k+1}\} \subseteq W$ y $c_1, \ldots, c_{k+1} \in \mathbb{R}$, entonces $c_1\overline{x}_1 + \cdots + c_{k+1}\overline{x}_{k+1} \in W$.

Notemos que por hipótesis de inducción, se tiene que $c_1\overline{x}_1 + \cdots + c_k\overline{x}_k \in W$. Por otro lado, como W es un subespacio vectorial, $c_{k+1} \in \mathbb{R}$ y $\overline{x}_{k+1} \in W$, entonces por el punto (c) del teorema del subespacio podemos concluir que $c_{k+1}\overline{x}_{k+1} \in W$. Así, como $c_1\overline{x}_1 + \cdots + c_k\overline{x}_k \in W$ y $c_{k+1}\overline{x}_{k+1} \in W$, por el punto (b) del teorema del subespacio se concluye que $c_1\overline{x}_1 + \cdots + c_{k+1}\overline{x}_{k+1} \in W$.

Por el principio de inducción matemática podemos garantizar que si W es un subespacio de V, $\{\overline{w}_1, \ldots, \overline{w}_r\} \subseteq W$ y $c_1, \ldots, c_r \in \mathbb{R}$, entonces $c_1\overline{x}_1 + \cdots + c_r\overline{x}_r \in W$. □

4.2. Combinaciones lineales

Uno de los conceptos más importantes de los espacios vectoriales es el de combinación lineal. Dicho concepto es de suma importancia pues todas las definiciones y resultados posteriores serán derivados de éste.

Dado un espacio vectorial V, $S \subseteq V$ y $\overline{x} \in V$, decimos que $\boldsymbol{x}$ **es una combinación lineal de** $\boldsymbol{S}$ si existen vectores $\overline{x}_1, \ldots, \overline{x}_n$ en S y escalares $\lambda_1, \ldots \lambda_n$ en $\mathbb{R}$ tales que

$$\overline{x} = \lambda_1 \overline{x}_1 + \lambda_2 \overline{x}_2 + \cdots + \lambda_n \overline{x}_n$$

De manera similar, si $\overline{x}_1, \ldots, \overline{x}_n$ son vectores en V y $\overline{x} \in V$, entonces decimos que $\overline{\boldsymbol{x}}$ **es combinación lineal de** $\overline{\boldsymbol{x}}_{\mathbf{1}}, \ldots, \overline{\boldsymbol{x}}_{\boldsymbol{n}}$ si

$$\overline{x} = \lambda_1 \overline{x}_1 + \lambda_2 \overline{x}_2 + \cdots + \lambda_n \overline{x}_n$$

Dados dos vectores $\overline{x}$ y $\overline{z}$ en V, decimos que $\overline{\boldsymbol{x}}$ **es múltiplo escalar de** $\overline{\boldsymbol{z}}$ si $\overline{x} = \lambda \overline{z}$ para algún $\lambda \in \mathbb{R}$.

Es sencillo verificar que no todos los vectores son combinación lineal de un conjunto de vectores dado, por ejemplo, si consideramos al vector $(0, 0, 1)$ y al conjunto $S = \{(2, 0, 0), (0, 1, 0)\}$ es claro que $(0, 0, 1)$ no es combinación lineal de S. Por otro lado, si $S' = \{(2, 0, 1), (1, 0, 0)\}$, entonces $(0, 0, 1)$ sí es combinación lineal de S', pues $(0, 0, 1) = 1(2, 0, 1) - 2(1, 0, 0)$. El siguiente resultado muestra que el vector cero siempre es combinación lineal de cualquier conjunto no vacío de vectores.

Lema 4.2.1. *Si V es un espacio vectorial y S es un subconjunto no vacío de V, entonces $\overline{0}$ es una combinación lineal de S.*

Demostración. Basta ver que como S es no vacío, si $x \in S$, entonces $\overline{0} = 0\overline{x}$, por lo que $\overline{0}$ es combinación lineal de S. $\square$

El lema 4.2.1 nos indica que el cero vector es combinación lineal de cualquier subconjunto no vacío. En particular, si $\overline{x}_1, \ldots, \overline{x}_n$ son vectores, a la combinación lineal $0\overline{x}_1 + 0\overline{x}_2 + \cdots + 0\overline{x}_n$ se le conoce como **combinación trivial.**

Si S es un subconjunto de un espacio vectorial, denotamos por $\langle S \rangle$ al conjunto de todas las combinaciones lineales de S. Por conveniencia, definiremos $\langle \emptyset \rangle = \{\overline{0}\}$. Además, si el conjunto S es finito, digamos $S = \{\overline{x}_1, \ldots, \overline{x}_n\}$, omitiremos la notación conjuntista y escribiremos simplemente $\langle \overline{x}_1, \ldots, \overline{x}_n \rangle$ en lugar de $\langle \{\overline{x}_1, \ldots, \overline{x}_n\} \rangle$. La primera nota que nos será de ayuda es la definición de pertenencia al conjunto $\langle S \rangle$ y es la siguiente.

Nota 4.2.2. *Sea V un espacio vectorial, $S \subseteq V$ y $\overline{x} \in V$. Son equivalentes.*

a) $\overline{x} \in \langle S \rangle$.

b) *Existe* $\{\overline{v}_1, \ldots, \overline{v}_n\} \subseteq S$ *y* $\{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{R}$ *tales que* $\overline{x} = \alpha_1 \overline{v}_1 + \cdots + \alpha_n \overline{v}_n$.

Con la nota previa, tenemos el siguiente teorema referente al conjunto $\langle S \rangle$.

Teorema 4.2.3. *Sea V un espacio vectorial. Si $S \subseteq V$, entonces $\langle S \rangle$ es un subespacio vectorial de V.*

Demostración. Para verificar que $\langle S \rangle$ es un subespacio vectorial de V, basta verificar los siguientes puntos.

(a) $\overline{0} \in \langle S \rangle$.

En este caso, si $S = \emptyset$, entonces por definición de $\langle S \rangle$ concluimos que $\overline{0} \in \langle S \rangle$. En caso de que S no sea vacío, entonces por el lema 4.2.1 se sabe que $\overline{0}$ es combinación lineal de S y por definición de $\langle S \rangle$, concluimos que $\overline{0} \in \langle S \rangle$.

(b) Si $\overline{x} \in \langle S \rangle$ y $\overline{z} \in \langle S \rangle$, entonces $\overline{x} + \overline{z} \in \langle S \rangle$.

Como $\overline{x} \in \langle S \rangle$, entonces existen $\overline{v}_1, \ldots, \overline{v}_n \in S$ y $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tales que $\overline{x} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n$. De manera análoga, como se sabe que $\overline{z} \in \langle S \rangle$, entonces existen $\overline{w}_1, \ldots, \overline{w}_m \in S$ y $\beta_1, \ldots, \beta_m \in \mathbb{R}$ tales que $\overline{z} = \beta_1\overline{w}_1 + \cdots + \beta_m\overline{w}_m$. Así, podemos concluir que

$$\overline{x} + \overline{z} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n + \beta_1\overline{w}_1 + \cdots + \beta_m\overline{w}_m$$

Siguiéndose de ello que existan $\overline{v}_1, \ldots, \overline{v}_n, \overline{w}_1, \ldots, \overline{w}_m \in S$ y $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in \mathbb{R}$ tal es que

$$\overline{x} + \overline{z} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n + \beta_1\overline{w}_1 + \cdots + \beta_m\overline{w}_m$$

Concluyendo de la definición de $\langle S \rangle$ el que $\overline{x} + \overline{z} \in \langle S \rangle$.

(c) Si $\lambda \in \mathbb{R}$ y $\overline{x} \in \langle S \rangle$, entonces $\lambda\overline{x} \in \langle S \rangle$.

Como $\overline{x} \in \langle S \rangle$, entonces existen $\overline{v}_1, \ldots, \overline{v}_n \in S$ y $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tales que $\overline{x} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n$. En tal caso

$$\lambda\overline{x} = \lambda\alpha_1\overline{v}_1 + \cdots + \lambda\alpha_n\overline{v}_n$$

Es decir, existen $\overline{v}_1, \ldots, \overline{v}_n \in S$ y $\lambda\alpha_1, \ldots, \lambda\alpha_n \in \mathbb{R}$ tales que

$$\lambda\overline{x} = \lambda\alpha_1\overline{v}_1 + \cdots + \lambda\alpha_n\overline{v}_n$$

Concluyendo de la definición de $\langle S \rangle$ el que $\lambda\overline{x} \in \langle S \rangle$.

Por lo antes mostrado y el teorema del subespacio, se concluye que $\langle S \rangle$ es un subespacio vectorial de V. $\square$

Gracias al teorema previo, se sigue que dado cualquier conjunto S de un espacio vectorial V, el conjunto $\langle S \rangle$ es un subespacio vectorial de V, por lo que a $\langle S \rangle$ se le conoce como **subespacio generado por S**.

Como consecuencia de la definición de espacio generado, es sencillo verificar el siguiente resultado.

Lema 4.2.4. *Sea V un espacio vectorial y U y S dos subconjuntos de V. Los siguientes enunciados se satisfacen:*

a) $U \subseteq \langle U \rangle$.

b) *Si* $U \subseteq S$*, entonces* $\langle U \rangle \subseteq \langle S \rangle$.

c) U *es subespacio de* V *si y sólo si* $U = \langle U \rangle$.

Demostración. Ahora mostraremos ambos incisos del lema.

(a) $U \subseteq \langle U \rangle$.

Sea $\overline{x} \in U$, como existe $\overline{x} \in U$ y $1 \in \mathbb{R}$ tal que $\overline{x} = 1\overline{x}$, entonces por definición de espacio generado, se concluye que $\overline{x} \in \langle U \rangle$.

(b) Si $U \subseteq S$, entonces $\langle U \rangle \subseteq \langle S \rangle$.

Sea $\overline{x} \in \langle U \rangle$. Por la nota 4.2.2 podemos garantizar que existen $\overline{v}_1, \ldots, \overline{v}_n \in U$ y $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tales que $\overline{x} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n$.

Por otro lado, como $U \subseteq S$, podemos garantizar que existen $\overline{v}_1, \ldots, \overline{v}_n \in S$ y $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tales que $\overline{x} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n$. De acuerdo a la definición de subespacio generado, podemos concluir que $\overline{x} \in \langle S \rangle$.

(c) U es subespacio de V si y sólo si $U = \langle U \rangle$.

Para la parte necesaria, tenemos que como $U = \langle U \rangle$ y $\langle U \rangle$ es un subespacio de V, entonces U es un subespacio de V.

Ahora procederemos a mostrar la parte suficiente, es decir, si U es un subespacio de V, entonces $U = \langle U \rangle$.

Por el inciso (a) de este lema, se sabe que $U \subseteq \langle U \rangle$, por lo que sólo mostraremos que $\langle U \rangle \subseteq U$.

Sea $\overline{x} \in \langle U \rangle$. Por definición, existen $\overline{v}_1, \ldots, \overline{v}_n \in U$ y $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tales que $\overline{x} = \alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n$. Como U es un subespacio vectorial, se tiene del corolario 4.1.6 que $\alpha_1\overline{v}_1 + \cdots + \alpha_n\overline{v}_n \in U$ y con ello, $\overline{x} \in U$. De lo antes mostrado, $U = \langle U \rangle$.

□

Una consecuencia de la definición de espacio generado es que podemos definir ciertos tipos de conjuntos de vectores que serán muy relevantes. Si V es un espacio vectorial, W es un subespacio de V y $S \subseteq V$, diremos que $\boldsymbol{S}$ **es un conjunto de generadores de** $\boldsymbol{W}$ o que $\boldsymbol{S}$ **es un conjunto generador de** $\boldsymbol{W}$ si $\langle S \rangle = W$.

Para fines prácticos, el siguiente teorema es una forma sencilla de verificar que algo es un conjunto de generadores de un subespacio vectorial dado.

Teorema 4.2.5. *Sean V un espacio vectorial, $B \subseteq V$ y W un subespacio vectorial de V. B es un conjunto de generadores de W si y sólo si se cumple que:*

a) $B \subseteq W$.

b) $W \subseteq \langle B \rangle$.

Demostración. Primero mostraremos la parte suficiente del teorema: si B es un conjunto de generadores de W, es decir, $W = \langle B \rangle$ entonces se cumple que (i) $B \subseteq W$ y (ii) $W \subseteq \langle B \rangle$.

Como por hipótesis $W = \langle B \rangle$, en particular se satisface (ii), es decir, $W \subseteq \langle B \rangle$. Además, por le lema 4.2.4 (a) se sigue que $B \subseteq \langle B \rangle$ y nuevamente por hipótesis podemos concluir (i), es decir, $B \subseteq W$.

Ahora mostraremos la parte suficiente del enunciado, es decir, $W = \langle B \rangle$ entonces se cumple que (i) $B \subseteq W$ y (ii) $W \subseteq \langle B \rangle$, entonces B es un conjunto de generadores de W, es decir, $W = \langle B \rangle$.

Para ver que $\langle B \rangle \subseteq W$, basta ver que por (i) se tiene que $B \subseteq W$, en tal caso, por le lema 4.2.4 (b) se tiene que $\langle B \rangle \subseteq \langle W \rangle$, sin embargo, por ser W un subespacio vectorial de V, entonces por el lema 4.2.4 (c) se concluye que $W = \langle W \rangle$ y así, $\langle B \rangle \subseteq W$.

Por otro lado, notemos que por hipótesis (ii) se tiene que $W \subseteq \langle B \rangle$, concluyendo que $W = \langle B \rangle$, es decir, B es un conjunto de generadores de W. □

Gracias al teorema previo, tenemos la siguiente nota para conjuntos generadores finitos.

Nota 4.2.6. *Sean V un espacio vectorial, W un subespacio de V y $\{\overline{x}_1, \ldots, \overline{x}_n\} \subseteq V$. Los siguientes son equivalentes.*

a) El conjunto $\{\overline{x}_1, \ldots, \overline{x}_n\}$ es un conjunto de generadores de W.

b) $\{\overline{x}_1, \ldots, \overline{x}_n\} \subseteq W$ y para todo $\overline{z} \in W$, existen $\{c_1, \ldots, c_n\} \subseteq \mathbb{R}$ tales que $\overline{z} = c_1\overline{x}_1 + \cdots + c_n\overline{x}_n$.

El siguiente lema nos presenta condiciones para garantizar que, dado un conjunto de generadores, es posible eliminar algunos de los vectores de dicho conjunto sin perder la propiedad de ser generador.

Lema 4.2.7. *Sea V un espacio vectorial y $B \subseteq V$ un conjunto finito no vacío. Si existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\} \rangle$, entonces $\langle B \rangle = \langle B \setminus \{\overline{x}\} \rangle$.*

Demostración. Para simplificar esta prueba, supongamos que $B = \{\overline{x}_1, \ldots, \overline{x}_n\}$ y supongamos que $\overline{x}_n \in B$ es tal que $\overline{x}_n \in \langle B \setminus \{\overline{x}_n\}\rangle$. De la definición de espacio generado, existen $c_1, \ldots, c_{n-1} \in \mathbb{R}$ tales que

$$\overline{x}_n = c_1\overline{x}_1 + \cdots + c_{n-1}\overline{x}_{n-1} \tag{4.1}$$

Por otro lado, mostraremos por doble contención que $\langle B\rangle = \langle B \setminus \{\overline{x}\}\rangle$. Es claro que como $B \setminus \{\overline{x}\} \subseteq B$ se tenga que $\langle B \setminus \{\overline{x}\}\rangle \subseteq \langle B\rangle$, por lo que resta mostrar que $\langle B\rangle = \langle B \setminus \{\overline{x}\}\rangle$.

Sea $\overline{w} \in \langle B\rangle$. Se sigue de la definición de espacio generado que existan $a_1, \ldots, a_n \in \mathbb{R}$ tales que se satisface que $\overline{w} = a_1\overline{x}_1 + \cdots + a_{n-1}\overline{x}_{n-1} + a_n\overline{x}_n$. Con ello, y de acuerdo a la ecuación 4.1 se concluye que:

$$\begin{aligned}
\overline{w} &= a_1\overline{x}_1 + \cdots + a_{n-1}\overline{x}_{n-1} + a_n\overline{x}_n \\
&= a_1\overline{x}_1 + \cdots + a_{n-1}\overline{x}_{n-1} + a_n(c_1\overline{x}_1 + \cdots + c_{n-1}\overline{x}_{n-1}) \\
&= a_1\overline{x}_1 + \cdots + a_{n-1}\overline{x}_{n-1} + a_nc_1\overline{x}_1 + \cdots + a_nc_{n-1}\overline{x}_{n-1} \\
&= (a_1 + a_nc_1)\overline{x}_1 + \cdots + (a_{n-1} + a_nc_{n-1})\overline{x}_{n-1}
\end{aligned}$$

Y por la definición de espacio generado, se sigue que $\overline{w} \in \langle B \setminus \{\overline{x}_n\}\rangle$, concluyendo la prueba. □

Un espacio vectorial es **finitamente generado** si existe $S \subseteq V$, con S finito, tal que $\langle S\rangle = V$. En lo consecuente, consideraremos solamente subespacios finitamente generados, aunque muchos de los resultados próximos también son válidos en cierta medida para los espacios vectoriales que no son finitamente generados. El último resultado de esta sección muestra que los espacios $\mathbb{R}^n$ son finitamente generados. Para ello definiremos una colección de vectores en $\mathbb{R}^n$ que serán de importancia. Para cada $i \in \{1, \ldots, n\}$, sea $\overline{e}_i$ el vector de $\mathbb{R}^n$ en el cual la i-ésima coordenada es igual a 1 y el resto de las coordenadas son iguales a 0. Al vector $\overline{e}_i$ se le conoce como el **i-ésimo vector canónico de $\mathbb{R}^n$**. Con esto, tenemos el siguiente resultado.

Teorema 4.2.8. *El conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es un conjunto de generadores de $\mathbb{R}^n$. En particular, el espacio vectorial $\mathbb{R}^n$ es finitamente generado.*

Demostración. Primero veremos que el conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es un conjunto de generadores de $\mathbb{R}^n$. Sea $\overline{x} = (x_1, \ldots, x_n)$ un vector arbitrario en $\mathbb{R}^n$. Es sencillo verificar que $\overline{x} = x_1\overline{e}_1 + \ldots x_n\overline{x}_n$, por lo que el conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es un conjunto de generadores de $\mathbb{R}^n$. Con ello, en particular se sigue que $\mathbb{R}^n$ es finitamente generado. □

4.3. Dependencia e independencia lineal

Sea V un espacio vectorial y B un subconjunto finito de V. Decimos que **B es linealmente independiente** si la única combinación lineal de B igual al vector cero es la trivial. Si un conjunto B no es linealmente independiente, entonces diremos que $\boldsymbol{B}$ **es linealmente dependiente**. La siguiente nota es útil para realizar demostraciones de independencia lineal mediante el uso de la definición.

Nota 4.3.1. *Sean V un espacio vectorial y $B = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k\}$ un subconjunto de V. Son equivalentes:*

a) *B es linealmente independiente.*

b) *Si $\lambda_1\overline{x}_1 + \lambda_2\overline{x}_2 + \cdots + \lambda_k\overline{x}_k = \overline{0}$, entonces $\lambda_i = 0$ para todo $i \in \{1, \ldots, k\}$.*

Igualmente, podemos establecer una nota para la definición de dependencia lineal.

Nota 4.3.2. *Sean V un espacio vectorial y $B = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k\}$ un subconjunto de V. Son equivalentes:*

a) *B es linealmente dependiente.*

b) *Existen $\lambda_1, \lambda_2, \ldots, \lambda_k$ en $\mathbb{R}$ tales que $\lambda_1\overline{x}_1 + \lambda_2\overline{x}_2 + \cdots + \lambda_k\overline{x}_k = \overline{0}$ y $\lambda_j \neq 0$ para algún $j \in \{1, \ldots, k\}$.*

Un primero resultado sobre independencia lineal es el siguiente.

Lema 4.3.3. *Sean V un espacio vectorial y $B \subseteq V$ linealmente independiente y finito. Si $\overline{x} \in V \setminus \langle B \rangle$, entonces $B \cup \{\overline{x}\}$ es linealmente independiente.*

Demostración. Para simplificar la notación de esta demostración, supongamos que $B = \{\overline{x}_1, \ldots, \overline{x}_n\}$. La demostración la realizaremos por contradicción, es decir, supondremos que $B \cup \{\overline{x}\}$ es linealmente dependiente. Por definición, se sigue que existen $\lambda_0, \lambda_1, \ldots, \lambda_n$ en $\mathbb{R}$ tales que $\overline{0} = \lambda_0\overline{x} + \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$ y $\lambda_j \neq 0$ para algún $j \in \{0, \ldots, n\}$.

Si $\lambda_0 \neq 0$, un simple despeje mostraría que:

$$\overline{x} = \left(\frac{\lambda_1}{-\lambda_0}\right)\overline{x}_1 + \left(\frac{\lambda_2}{-\lambda_0}\right)\overline{x}_2 + \cdots + \left(\frac{\lambda_n}{-\lambda_0}\right)\overline{x}_n$$

Así $\overline{x} \in \langle B \rangle$, lo cual no es posible por hipótesis sobre $\overline{x}$. Con ello, $\lambda_0 = 0$ y en tal caso, $\overline{0} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$ y $\lambda_j \neq 0$ para algún $j \in \{1, \ldots, n\}$, lo cual no es posible pues B es linealmente independiente. Así concluimos que $B \cup \{\overline{x}\}$ es linealmente independiente. $\square$

Los conceptos de dependencia e independencia lineal son cruciales en muchos aspectos de espacios vectoriales, por lo que será importante saber verificar cuándo un conjunto es o no linealmente independiente y en algunas ocasiones el uso de las definiciones no necesariamente serán las mejores opciones. Por ello tendremos diversos resultados que nos auxiliaran a determinar cuándo un conjunto es linealmente dependiente o linealmente independiente.

Teorema 4.3.4. *Sean V un espacio vectorial y B un subconjunto finito de V. B es linealmente dependiente si y sólo si existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$.*

Demostración. Para simplificar esta prueba, supongamos que $B = \{\overline{x}_1, \ldots, \overline{x}_k\}$.

Primero mostraremos la parte suficiente del teorema, es decir, si B es linealmente dependiente, entonces existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$.

Para esto, por ser B linealmente dependiente, existen $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ tales que $\overline{0} = \lambda_1\overline{x}_1 + \cdots + \lambda_k\overline{x}_k$ y $\lambda_j \neq 0$ para algún $j \in \{1, \ldots, k\}$. Supongamos, sin pérdida de generalidad, que $j = k$. Con ello, es sencillo verificar que:

$$\overline{x}_k = \left(\frac{\lambda_1}{-\lambda_k}\right)\overline{x}_1 + \left(\frac{\lambda_2}{-\lambda_k}\right)\overline{x}_2 + \cdots + \left(\frac{\lambda_{k-1}}{-\lambda_k}\right)\overline{x}_{k-1}$$

Gracias a lo anterior podemos garantizar de la definición de espacio generado que $\overline{x}_k \in \langle B \setminus \{\overline{x}_k\}\rangle$ probando así lo deseado.

Ahora mostraremos la parte suficiente del enunciado, es decir, si existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$, entonces B es linealmente dependiente.

Para ello supongamos sin pérdida de generalidad que $\overline{x} = \overline{x}_k$, es decir, $\overline{x}_k \in B$ y $\overline{x}_k \in \langle B \setminus \{\overline{x}_k\}\rangle$ Como $\overline{x}_k \in \langle B \setminus \{\overline{x}_k\}\rangle$, entonces existen $\mu_1, \ldots, \mu_{k-1} \in \mathbb{R}$ tales que $\overline{x}_k = \mu_1\overline{x}_1 + \cdots + \mu_{k-1}\overline{x}_{k-1}$ y en particular:

$$\overline{0} = \mu_1\overline{x}_1 + \cdots + \mu_{k-1}\overline{x}_{k-1} + (-1)\overline{x}_k$$

Así, existe una combinación lineal de $\overline{x}_1, \ldots, \overline{x}_k$ igual al vector cero y al menos uno de sus coeficientes es diferente de 0, a saber, el coeficiente de $\overline{x}_k$. Con ello, B es linealmente dependiente. $\square$

Como consecuencia del teorema previo, obtendremos los siguientes corolarios.

Corolario 4.3.5. *Sean V un espacio vectorial y B un subconjunto finito de V. B es linealmente independiente si y sólo para todo $\overline{x} \in B$, $\overline{x} \notin \langle B \setminus \{\overline{x}\}\rangle$.*

Demostración. Este corolario es inmediato del teorema 4.3.4. $\square$

Corolario 4.3.6. *Sean V un espacio vectorial y $B \subseteq V$. Si $\overline{0} \in B$, entonces B es linealmente dependiente.*

Demostración. Como $\overline{0} \in B$ y $\overline{0} \in \langle B \setminus \{\overline{0}\}\rangle$, entonces por el teorema 4.3.4 se concluye que B es linealmente dependiente. $\square$

Corolario 4.3.7. *Sean V un espacio vectorial y B un subconjunto finito de V. Si existen dos vectores distintos en B, digamos $\overline{x}$ y $\overline{z}$, tales que $\overline{x} = c\overline{z}$ para alguna $c \in \mathbb{R}$, entonces B es un conjunto linealmente dependiente.*

Demostración. Para este punto, notemos que como $\overline{x} = c\overline{z}$ donde $\overline{z} \in B \setminus \{\overline{x}\}$ y $c \in \mathbb{R}$, entonces $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$ y por el teorema 4.3.4, B es linealmente dependiente. □

Corolario 4.3.8. *Sean V un espacio vectorial y B y B' subconjuntos finitos de V tales que $B \subseteq B'$. Si B es linealmente dependiente, entonces B' es linealmente dependiente.*

Demostración. Como B es linealmente dependiente, entonces por el teorema 4.3.4 existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$. Notemos que $\overline{x} \in B'$ y además, por el lema 4.2.4 (b), $\langle B \setminus \{\overline{x}\}\rangle \subseteq \langle B' \setminus \{\overline{x}\}\rangle$, concluyendo que existe $\overline{x} \in B'$ tal que $\overline{x} \in \langle B' \setminus \{\overline{x}\}\rangle$. Por el teorema 4.3.4 se concluye que B' es linealmente dependiente. □

El último corolario de esta sección es el siguiente.

Corolario 4.3.9. *Sean V un espacio vectorial y B y B' subconjuntos finitos de V tales que $B \subseteq B'$. Si B' es linealmente independiente, entonces B es linealmente independiente.*

Demostración. Este corolario es la contraposición del corolario 4.3.8. □

4.4. Bases de espacios vectoriales

Sea V un espacio vectorial y $B \subseteq V$ finito. Decimos que **B es una base de V** si B es linealmente independiente y es un conjunto generador de V. Debido a la importancia de las bases, es natural que existan diversas caracterizaciones de dichos conjuntos. En particular tenemos el siguiente teorema.

Teorema 4.4.1 (De la expresión única)**.** *Sean V un espacio vectorial no trivial y $B = \{\overline{x}_1, \ldots, \overline{x}_n\}$ un conjunto de vectores de V. Son equivalentes:*

a) *B es una base de V.*

b) *Para todo $\overline{w} \in V$ existen escalares únicos $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$.*

Demostración. Primero mostraremos que (a) implica (b), es decir, si B es una base de V, entonces Para todo $\overline{w} \in V$ existen escalares únicos $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$.

Sea $\overline{w} \in V$. Notemos que por ser B una base de V, entonces B es un conjunto de generadores y con ello, existen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$.

Ahora procederemos a mostrar que dichos escalares son únicos. Para ello, supongamos que $\mu_1, \ldots, \mu_n \in \mathbb{R}$ son tales que $\overline{w} = \mu_1\overline{x}_1 + \cdots + \mu_n\overline{x}_n$. En tal caso, se tiene que $\lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n = \mu_1\overline{x}_1 + \cdots + \mu_n\overline{x}_n$ y con un simple despeje se puede concluir que:

$$\overline{0} = (\lambda_1 - \mu_1)\overline{x}_1 + (\lambda_2 - \mu_2)\overline{x}_2 + \cdots + (\lambda_n - \mu_n)\overline{x}_n$$

Así, como B es una base de V, en particular es linealmente independiente, por lo que para toda $i \in \{1, \ldots, n\}$ se satisface que $\lambda_i - \mu_i = 0$ y así, $\lambda_i = \mu_i$, concluyendo que dichos escalares son únicos.

Ahora procederemos a mostrar que (b) implica (a), es decir, si para todo $\overline{w} \in V$ existen escalares únicos $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$, entonces B es una base de V.

Primero veremos que B es un conjunto de generadores de V. Para ello consideremos $\overline{w} \in V$. Notemos que por hipótesis, existen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$, concluyendo que B es un conjunto de generadores de V.

Ahora para mostrar que B es linealmente independiente, consideremos una combinación lineal de B igual al vector cero, digamos $\overline{0} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$. Notemos que por hipótesis dichos escalares son únicos. Por otro lado, como la combinación trivial también es una combinación del vector cero, se concluye que para toda $i \in \{1, \ldots, n\}$, $\lambda_i = 0$, mostrando así que B es linealmente independiente. Por lo anterior, B es una base de V. $\square$

Los siguientes resultados muestran que es posible encontrar bases a partir de conjuntos de generadores finitos y también a partir de conjuntos linealmente independientes.

Teorema 4.4.2. *Sean V un espacio vectorial no trivial y $B \subseteq V$. Si B es un conjunto finito de generadores de V, entonces existe una base de V, digamos B', tal que $B' \subseteq B$.*
En particular, si V es finitamente generado, entones todo conjunto finito de generadores de V contiene una base finita de V.

Demostración. Mostraremos por inducción sobre $|B|$ que si $|B| = n$ y B es un conjunto de generadores de V, entonces existe $B' \subseteq B$ tal que B' es base de V.

Base de inducción. Si $|B| = 1$ y B es un conjunto de generadores de V, entonces existe una base de V, digamos B', tal que $B' \subseteq B$.

En este caso, como B tiene un elemento, es un conjunto de generadores de V y V es no trivial, entonces el vector en B debe ser distinto del vector cero, por lo que B es linealmente independiente. Así, B es una base de G.

Hipótesis de inducción. Supongamos que k es un natural tal que si $|S| = k$ y S es un conjunto generador de V, entonces existe una base de V, digamos S', tal que $S' \subseteq S$.

Paso inductivo. Si $|B| = k+1$ y B es un conjunto generador de V, entonces existe una base de V, digamos B', tal que $B' \subseteq B$.

Para esta demostración basta ver que si B es linealmente independiente, entonces B es la base buscada. Para lo que resta de la demostración, supondremos que B es linealmente dependiente. En tal caso, por el teorema 4.3.4 existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\}\rangle$. En tal caso, por el lema 4.2.7 se concluye que $\langle B\rangle = \langle B \setminus \{\overline{x}\}\rangle$. Sin embargo, como B es un conjunto de generadores, entonces $B \setminus \{\overline{x}\}$ es un conjunto de generadores con k elementos, siguiéndose de la hipótesis de inducción que exista $B' \subseteq B \setminus \{\overline{x}\}$ tal que B' es una base de V. Así, $B' \subseteq B$ es la base deseada.

Por el principio de inducción matemática se muestra que si B es un conjunto generador finito de V, entonces existe $B' \subseteq B$ tal que B' es base de V. □

Un resultado análogo para los conjuntos linealmente independientes.

Teorema 4.4.3. *Sean V un espacio vectorial no trivial, finitamente generado y $B \subseteq V$. Si B es un conjunto finito y linealmente independiente de V, entonces existe una base de V, digamos B', tal que $B \subseteq B'$*

Demostración. Sea G_0 un conjunto de generadores finito de V. Mostraremos por inducción sobre $|G_0 \setminus B|$ que si B es un conjunto linealmente independiente y $|G_0 \setminus B| = n$, entonces existe una base de V, digamos B', tal que $B \subseteq B'$.

Base de inducción. Si B es un conjunto finito, linealmente independiente y $|G_0 \setminus B| = 0$, entonces existe una base de V, digamos B', tal que $B \subseteq B'$.

En este caso, $G_0 \setminus B = \emptyset$, por lo que $G_0 \subseteq B$. Así, por el lema 4.2.4 (a) se sigue que $\langle G_0\rangle \subseteq \langle B\rangle$. Como G_0 es un conjunto de generadores, entonces $V = \langle G_0\rangle$, concluyendo que $V = \langle B\rangle$. Así, B es un conjunto de generadores de V, por lo que B es la base deseada.

Hipótesis de inducción. Supongamos que k es un natural tal que si S es un conjunto finito, linealmente independiente y $|G_0 \setminus S| = k$, entonces existe una base de V, digamos S', tal que $S \subseteq S'$

Paso inductivo. Si B es un conjunto finito, linealmente independiente y $|G_0 \setminus B| = k+1$, entonces existe una base de V, digamos B', tal que $B \subseteq B'$

Para esta demostración verificaremos dos posibles casos.

Caso 1. $G_0 \subseteq \langle B \rangle$.

En este caso, por el lema 4.2.4 (b) se tiene que $\langle G_0 \rangle \subseteq \langle \, \langle B \rangle \, \rangle$. Sin embargo, como $\langle B \rangle$ es un subespacio vectorial, se tiene que $\langle \, \langle B \rangle \, \rangle = \langle B \rangle$ (lema 4.2.4 (c)), con lo cual, $\langle G_0 \rangle \subseteq \langle B \rangle$ y por ser G_0 un conjunto de generadores, entonces $V = \langle G_0 \rangle$, concluyendo que $V = \langle B \rangle$, es decir, B es un conjunto de generadores y así, B es la base deseada.

Caso 2. $G_0 \not\subseteq \langle B \rangle$.

En este caso consideremos $\overline{x} \in G_0$ tal que $\overline{x} \notin \langle B \rangle$. Primero notemos que gracias al lema 4.2.4 (a), $\overline{x} \notin B$. Además de lo anterior, por el lema 4.3.3 se sigue que $S = B \cup \{\overline{x}\}$ es un conjunto linealmente independiente de G. Ahora notemos que:

$$\begin{aligned}
|G_0 \setminus S| &= |G_0| - |G_0 \cap S| && \text{Cardinal de la diferencia} \\
&= |G_0| - |G_0 \cap (B \cup \{\overline{x}\})| && \text{Definición de } S \\
&= |G_0| - |(G_0 \cap B) \cup (G_0 \cap \{\overline{x}\})| && \text{Lema 1.4.7 (b)} \\
&= |G_0| - (|G_0 \cap B| + |G_0 \cap \{\overline{x}\}| - |(G_0 \cap B) \cap (G_0 \cap \{\overline{x}\}|) && \text{Cardinal de la unión} \\
&= |G_0| - (|G_0 \cap B| + |G_0 \cap \{\overline{x}\}|) && \overline{x} \notin B \\
&= |G_0| - (|G_0 \cap B| + 1) && \overline{x} \in G_0 \\
&= |G_0 \setminus B| - 1 && \text{Cardinal de la diferencia} \\
&= k
\end{aligned}$$

Con lo anterior, S es un conjunto linealmente independiente de V tal que $|G_0 \setminus S| = k$, por lo que por hipótesis de inducción sobre S se tiene que existe una base de V, digamos B', tal que $S \subseteq B'$. Como $B \subseteq S$, entonces $B \subseteq B'$, por lo que B' es la base deseada.

Por el principio de inducción matemática podemos concluir que si B es un conjunto linealmente independiente de V, entonces existe una base de V, digamos B', tal que $B \subseteq B'$. □

Una primera consecuencia del resultado anterior es el siguiente, el cual recibe categoría de teorema por su importancia.

Teorema 4.4.4 (Existencia de Bases). *Todo espacio vectorial no trivial y finitamente generado tiene al menos una base.*

Demostración. Sea V un espacio vectorial finitamente generado no trivial. Como V es no trivial, podemos considerar $\overline{x} \in V$ tal que $\overline{x} \neq \overline{0}$. Así, $B = \{\overline{x}\}$ es un conjunto linealmente independiente de V, siguiéndose del teorema 4.4.3 que exista una base de V, digamos B', tal que $B \subseteq B'$. En particular, B' es una base de V. □

A simple vista el teorema anterior podría parecer no muy relevante, sin embargo quiero mencionar algunas cosas de dicho resultado antes de continuar. En primer lugar, el teorema 4.4.4 también es verdadero para los espacios vectoriales que no son finitamente generados, es decir:

Teorema 4.4.5 (Existencia de Bases). *Todo espacio vectorial no trivial tiene al menos una base.*

Este último teorema, aunque sencillo de enunciar, puede ser considerado uno de los teoremas más importantes de Álgebra Lineal, por no decir de toda la Matemática en general. Su relevancia radica no sólo en el gran poder, tanto teórico como práctico, que albergan las bases en los espacios vectoriales y cuya existencia es garantizada por el teorema 4.6.1, sino también por la belleza de su demostración, misma que no veremos en estas notas, y su equivalencia al Axioma de elección, estableciendo un vínculo con diversas ramas de las matemáticas como lo son la Teoría de Conjuntos, la Topología y la Teoría de Ordenes. En una nota personal, este es uno de mis teoremas favoritos de toda la Matemática.

A pesar del resultado anterior, es sencillo verificar que algunos espacios vectoriales tienen una base, por ejemplo, los espacios $\mathbb{R}^n$.

Lema 4.4.6. *El conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es una base de $\mathbb{R}^n$.*

Demostración. Para verificar que $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es linealmente independiente en $\mathbb{R}^n$, consideremos $c_1, \ldots, c_n \in \mathbb{R}$ tales que $\overline{0} = c_1\overline{e}_1 + \cdots + c_n\overline{e}_n$. Se sigue de la definición de los elementos en $\{\overline{e}_1, \ldots, \overline{e}_n\}$ que para cada $i \in \{1, \ldots, n\}$, $c_i = 0$, por lo que $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es linealmente independiente.

Ahora, para mostrar que $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es un conjunto de generadores de $\mathbb{R}^n$, basta considerar $(x_1, \ldots, x_n) \in \mathbb{R}^n$ arbitrario. Nuevamente, por la definición de los elementos en $\{\overline{e}_1, \ldots, \overline{e}_n\}$ se tiene que $(x_1, \ldots, x_n) = x_1\overline{e}_1 + \cdots + x_n\overline{e}_n\}$, por lo que $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es un conjunto de generadores, concluyendo que dicho conjunto es una base de $\mathbb{R}^n$. $\square$

Al conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ se le suele llamar **base canónica de $\mathbb{R}^n$** y la usaremos en algunas ocasiones consecuentes.

Continuando con las notas, tenemos la siguiente equivalencia de las bases, mostrando que son conjuntos óptimos en cuestiones de independencia lineal y de conjuntos de generadores.

Teorema 4.4.7. *Sean V un espacio vectorial y $B \subseteq V$. Son equivalentes:*

a) B es una base de V.

b) B es linealmente independiente maximal, es decir, B es linealmente independiente y si B' es un conjunto linealmente independiente y $B \subseteq B'$, entonces $B = B'$.

c) B es generador minimal de V, es decir, B es un conjunto de generadores de V y si B' es un conjunto de generadores de V tal que $B' \subseteq B$, entonces $B = B'$.

Demostración. Para la prueba de este enunciado, mostraremos que (a) implica (b), (b) implica (c) y por último mostraremos que (c) implica (a).

(a) implica (b). Si B es un base de V, entonces (i) B es linealmente independiente y (ii) si B' es un conjunto linealmente independiente y $B \subseteq B'$, entonces $B = B'$.

Primero notemos que por ser B una base, trivialmente se satisface (i), por lo que procederemos a demostrar (ii). Procederemos por contradicción al suponer que B' es un conjunto linealmente independiente tal que $B \subseteq B'$ y $B \neq B'$. Así, podemos considerar $\overline{x} \in B'$ tal que $\overline{x} \notin B$. Notemos que como B es una base, en particular es un conjunto de generadores de V, por lo que $\overline{x} \in \langle B \rangle$, en tal caso, por el teorema 4.3.4 se sigue que $B \cup \{\overline{x}\}$ es linealmente dependiente. Sin embargo, es sencillo ver que $B \cup \{\overline{x}\} \subseteq B'$, por lo que por el corolario 4.3.8, se sigue que B' es linealmente dependiente, lo cual no es posible por hipótesis sobre B'. Así podemos concluir que $B = B'$.

(b) implica (c). Si B es un conjunto de vectores tal que (i) si B' es un conjunto linealmente independiente y $B \subseteq B'$, entonces $B = B'$ y (ii) B es linealmente independiente, entonces (I) B es un conjunto de generadores de V y (II) si B' es un conjunto de generadores de V tal que $B' \subseteq B$, entonces $B = B'$.

Primero mostraremos que B es un conjunto de generadores de V. Para ello consideremos $\overline{x} \in V$. Si $\overline{x} \in B$, trivialmente se satisface que $\overline{x} \in \langle B \rangle$, por lo que supondremos que $\overline{x} \notin B$. Notemos que por (i) se tiene que $B \cup \{\overline{x}\}$ no es linealmente independiente, por lo que por la contraposición del teorema 4.3.3 se sigue que $\overline{x} \in \langle B \rangle$, concluyendo que B es un conjunto de generadores de V.

Ahora sólo resta mostrar (II). Procederemos por contradicción al suponer que B' es un conjunto de generadores de V tal que $B' \subseteq B$ y $B' \neq B$. Consideremos $\overline{x} \in B$ tal que $\overline{x} \notin B'$. Notemos que por ser B' un conjunto de generadores se sigue que $\overline{x} \in \langle B' \rangle$, pero como $B' \subseteq B \setminus \{\overline{x}\}$, entonces $\langle B' \rangle \subseteq \langle B \setminus \{\overline{x}\} \rangle$, en particular, $\overline{x} \in \langle B \setminus \{\overline{x}\} \rangle$ y de acuerdo al teorema 4.3.4 podemos garantizar que B es linealmente dependiente, lo cual no es posible por (ii). Así se concluye que si B' es un conjunto de generadores de V tal que $B' \subseteq B$, entonces $B = B'$.

(c) implica (a). Si B es un conjunto de vectores tales que (i) si B' es un conjunto de generadores de V tal que $B' \subseteq B$, entonces $B = B'$ y (ii) B es un conjunto de generadores de V, entonces B es una base de V.

Dado que B es un conjunto de generadores, entonces basta mostrar que B es un conjunto linealmente independiente. Procediendo por contradicción, supongamos que B es linealmente dependiente, por lo que de acuerdo al teorema 4.3.4, existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\} \rangle$. Por el lema 4.2.7 se sigue que $\langle B \rangle = \langle B \setminus \{\overline{x}\} \rangle$, por lo que $B \setminus \{\overline{x}\}$ es un conjunto de generadores de V, lo cual no es posible por (i). Así concluimos que B es un conjunto linealmente independiente y por (ii), B es una base de V.

□

El ultimo resultado de esta sección es el siguiente.

Lema 4.4.8 (Del reemplazo). *Sean V un espacio vectorial y B_0 una base finita de V. Si S es un conjunto finito y linealmente independiente y $|S| \leq |B_0|$, entonces existe $S' \subseteq B_0$ tal que $|S'| = |B_0| - |S|$ y $\langle S \cup S' \rangle = V$.*

Demostración. Mostraremos mediante inducción matemática sobre $|S|$ que: Si $|S| = n$, S es un conjunto linealmente independiente y $|S| \leq |B_0|$, entonces existe $S' \subseteq B_0$ tal que $|S'| = |B_0| - |S|$ y $\langle S \cup S' \rangle = V$.

Base de Inducción. Si $|S| = 0$, S es un conjunto linealmente independiente y $|S| \leq |B_0|$, entonces existe $S' \subseteq B_0$ tal que $|S'| = |B_0| - |S|$ y $\langle S \cup S' \rangle = V$.

En este caso, $S = \emptyset$ y como B_0 es una base de V, entonces $S' = B_0$ satisface las condiciones deseadas.

Hipótesis de Inducción. Supongamos que k es un natural tal que si $|T| = k$, T es un conjunto linealmente independiente y $|T| \leq |B_0|$, entonces existe $T' \subseteq B_0$ tal que $|T'| = |B_0| - |T|$ y $\langle T \cup T' \rangle = V$.

Paso Inductivo. Si $|S| = k + 1$, S es un conjunto linealmente independiente y $|S| \leq |B_0|$, entonces existe $S' \subseteq B_0$ tal que $|S'| = |B_0| - |S|$ y $\langle S \cup S' \rangle = V$.

Para simplificar notación, supongamos que $S = \{\overline{s}_1, \ldots, \overline{s}_{k+1}\}$ y $|B_0| = n$. De acuerdo al corolario 4.3.9, el conjunto $T = \{\overline{s}_1, \ldots, \overline{s}_k\}$ es linealmente independiente, por lo que, de acuerdo a la hipótesis de inducción, existe $T' \subseteq B_0$ tal que $|T'| = |B_0| - |T|$ y $\langle T \cup T' \rangle = V$. Supongamos que $T' = \{\overline{x}_1, \ldots, \overline{x}_{n-k}\}$.

Como $\langle T \cup T' \rangle = V$, en particular se tiene que $\overline{s}_{k+1}$ es combinación lineal de $T \cup T'$, por lo que de acuerdo a la definición de espacio generado, existen $c_1, \ldots, c_k, b_1, \ldots, b_{n-k} \in \mathbb{R}$ tales que:

$$\overline{s}_{k+1} = c_1\overline{s}_1 + \cdots + c_k\overline{s}_k + b_1\overline{x}_1 + \cdots + b_{n-k}\overline{x}_{n-k} \tag{4.2}$$

Notemos que si para todo $i \in \{1, \ldots, n-k\}$ se cumple que $b_i = 0$, entonces por la igualdad 4.2 se tiene que $\overline{s}_{k+1} = c_1\overline{s}_1 + \cdots + c_k\overline{s}_k$, siguiéndose de ello que $\overline{s}_{k+1} \in \langle S \setminus \{\overline{s}_{k+1}\}\rangle$, lo cual no es posible pues S es linealmente independiente. Así, podremos suponer sin perdida de generalidad que $b_{n-k} \neq 0$.

Con lo anterior, propondremos $S' = \{\overline{x}_1, \ldots, \overline{x}_{n-k-1}\}$ y mostraremos que S_1 es el conjunto deseado. Primero notemos que por elección de S' se tiene que $S' \subseteq B_0$ y que $|S'| = |B_0| - |S|$, por lo que solo resta mostrar que $\langle S \cup S_1\rangle = V$.

Notemos que de un simple despeje de la ecuación 4.2 se tiene que:

$$\overline{x}_{n-k} = \left(\frac{c_1}{-b_{n-k}}\right)\overline{s}_1+\cdots+\left(\frac{c_k}{-b_{n-k}}\right)\overline{s}_k+\left(\frac{-1}{-b_{n-k}}\right)\overline{s}_{k+1}+\left(\frac{b_1}{-b_{n-k}}\right)\overline{x}_1+\cdots+\left(\frac{-b_{n-k-1}}{b_{n-k}}\right)\overline{x}_{n-k-1} \tag{4.3}$$

Ahora procederemos a mostrar que $V = \langle S \cup S'\rangle$. Como $\langle T \cup T'\rangle$ es un conjunto de generadores, bastará verificar que $\langle T \cup T'\rangle \subseteq \langle S \cup S'\rangle$. Para ello, consideremos $\overline{w} \in \langle T \cup T'\rangle$. En tal caso existen $a_1, \ldots, a_k, b_1, \ldots, b_{n-k} \in \mathbb{R}$ tales que

$$\overline{w} = a_1\overline{s}_1 + \cdots + a_k\overline{s}_k + b_1\overline{x}_1 + \cdots + b_{n-k-1}\overline{x}_{n-k-1} + b_{n-k}\overline{x}_{n-k}$$

Para simplificar notación, denotaremos por $\overline{v}$ al vector $a_1\overline{s}_1 + \cdots + a_k\overline{s}_k$, por $\overline{s}$ al vector $b_{n-k}\overline{x}_{n-k}$ y por $\overline{u}$ al vector $b_1\overline{x}_1 + \cdots + b_{n-k-1}\overline{x}_{n-k-1}$. Así, se tiene que $\overline{w} = \overline{v} + \overline{u} + \overline{s}$. Notemos que $\overline{v} \in \langle S\rangle$ y, en particular, $\overline{v} \in \langle S \cup S'\rangle$. Del mismo modo, $\overline{u} \in \langle S'\rangle$, por lo que $\overline{u} \in \langle S \cup S'\rangle$. Por último, gracias a la ecuación 4.3 se tiene que $\overline{s} \in \langle S \cup S'\rangle$. Ya que $\langle S \cup S'\rangle$ es un subespacio vectorial, en particular es cerrado bajo la suma vectorial, por lo que $\overline{w} \in \langle S \cup S'\rangle$. Con lo anterior, $\langle T \cup T'\rangle \subseteq \langle S \cup S'\rangle$. Como $V = \langle T \cup T'\rangle$, se concluye que $V = \langle S \cup S'\rangle$, mostrando que S' es el conjunto deseado.

Por el principio de inducción matemática se tiene que si S es un conjunto finito y linealmente independiente y $|S| \leq |B_0|$, entonces existe $S' \subseteq B_0$ tal que $|S'| = |B_0| - |S|$ y $\langle S \cup S'\rangle = V$. □

4.5. Dimensión de espacios vectoriales

A partir del teorema del remplazo será posible empezar a deducir cicertas propiedades de finitud para los conjuntos linealmente indpendientes, los conjuntos generadores y las bases.

Lema 4.5.1. *Sean V un espacio vectorial y B una base finita de V. Si S es un conjunto finito, linealmente independiente y $|S| = |B|$, entonces S es una base de V.*

Demostración. Sea S un conjunto linealmente independiente y finito tal que $|S| = |B|$. Como B es una base finita de V, entonces por el teorema del remplazo (teorema 4.1.3) existe $S' \subseteq B$ tal que $|S'| = |B| - |S|$ y

$\langle S \cup S' \rangle = V$. Notemos que debido a que $|S| = |B|$, entonces S' es vacío, por lo que $S \cup S' = S$ y con ello, $\langle S \rangle = V$. Así, S es linealmente independiente y generador, por lo que S es una base de V. □

Corolario 4.5.2. *Sean V un espacio vectorial y B una base finita de V. Si $S \subseteq V$ es linealmente independiente, entonces S es finito y $|S| \leq |B|$.*

Demostración. Procederemos por contradicción al suponer que S es infinito o $|B| < |S|$. En cualquiera de los casos anteriores es posible encontrar $S' \subseteq S$ tal que $|S| = |B|$. Como S es linealmente independiente, entonces S' es linealmente independiente, y como $|S'| = |B|$, por el lema 4.5.1 se concluye que S' es una base de V. Sin embargo, $S' \subseteq S$, $S' \neq S$ y S es linealmente independiente, lo cual no es posible pues S' es linealmente independiente maximal. Así podemos concluir que $|S| \leq |B|$. □

Derivado de los resultados anteriores tenemos el siguiente teorema, el cual tiene gran relevancia en toda el Álgebra Lineal. A pesar de ser un resultado directo de lo antes mostrado, tendrá categoría de teorema por su utilidad.

Teorema 4.5.3. *(De la Dimensión) Sea V un espacio vectorial finitamente generado. Si B y B' son bases de V, entonces $|B| = |B'|$.*

Demostración. Primero veamos que como V es finitamente generado, entonces existe una base finita de V, digamos B_0 (teorema 4.4.2). Así, como B_0 es una base finita y B es un conjunto linealmente independiente, se tiene del corolario 4.5.2 que B es finito y $|B| \leq |B_0|$. Del mismo modo, como B es una base finita y B_0 es un conjunto linealmente independiente, entonces $|B_0| \leq |B|$, concluyendo que $|B_0| = |B|$. Un razonamiento análogo muestra que $|B'| = |B_0|$, por lo que $|B| = |B'|$. □

El teorema anterior establece que en un espacio vectorial finitamente generado todas las bases son finitas y tienen exactamente el mismo número de elementos. Dicha propiedad será importante e incluso se le asocia una definición. Dado un espacio vectorial finitamente generado V, la **dimensión de $\boldsymbol{V}$**, denotada por $dim(V)$, es el número de elementos de cualquier base de V. Si V es trivial, diremos que su dimensión es 0.

Es sencillo calcular la dimensión de los espacios $\mathbb{R}^n$

Lema 4.5.4. *Si $n \geq 1$, entonces $dim(\mathbb{R}^n) = n$.*

Demostración. Como el conjunto $\{\overline{e}_1, \ldots, \overline{e}_n\}$ es una base de $\mathbb{R}^n$ y tiene n elementos, entonces se sigue de la definición de dimensión que $dim(\mathbb{R}^n) = n$. □

Además de lo anterior, es posible mostrar ciertas cotas para los conjuntos generadores y los conjuntos linealmente independientes.

Lema 4.5.5. *Sea V un espacio vectorial finitamente generado y $B \subseteq V$ finito. Los siguientes enunciados se satisfacen.*

a) Si B es un conjunto de generadores de V, entonces $dim(V) \leq |B|$.

b) Si B es un conjunto linealmente independiente de V, entonces $|B| \leq dim(V)$.

Demostración. Ahora mostraremos ambos incisos de este resultado.

(a) Si B es un conjunto de generadores de V, entonces $dim(V) \leq |B|$.

Como B es un conjunto finito de generadores de V, entonces por el teorema 4.4.2 se concluye que existe una base de V, digamos B', tal que $B' \subseteq B$. Así, $|B'| \leq |B|$ y por definición de dimensión, $dim(V) \leq |B|$.

(b) Si B es un conjunto linealmente independiente de V, entonces $|B| \leq dim(V)$.

Como B es un conjunto finito y linealmente independiente, entonces por el teorema 4.4.3 se concluye que existe una base de B, digamos B', tal que $B \subseteq B'$. Así, $|B| \leq |B'|$ y por definición de dimensión, $|B| \leq dim(V)$.

□

Como consecuencia del resultado anterior, tenemos el siguiente corolario.

Corolario 4.5.6. *Sea V un espacio vectorial finitamente generado y $B \subseteq V$. Los siguientes enunciados se satisfacen.*

a) Si $|B| < dim(V)$, entonces B no es un conjunto generador de V.

b) Si $|B| > dim(V)$, entonces B no es un conjunto linealmente independiente.

c) Si $|B| \neq dim(V)$, entonces B no es base de V.

Demostración. El inciso (a) es la contraposición del inciso (a) del lema 4.5.5. Análogamente, el inciso (b) es la contraposición del inciso (b) del lema 4.5.5. Por último, el inciso (c) se sigue de la definición de dimensión. □

Un segundo resultado que relaciona la dimensión con los conjuntos generadores y conjuntos linealmente independientes es el siguiente.

Teorema 4.5.7. *Sean V un espacio vectorial finitamente generado y $B \subseteq V$ finito. Los siguientes enunciados se satisfacen.*

a) Si B es un conjunto de generadores de V y $|B| = dim(V)$, entonces B es una base de V.

b) Si B es linealmente independiente y $|B| = dim(V)$, entonces B es una base de V.

Demostración. Ahora veremos la demostración de ambos incisos.

(b) Si B es un conjunto de generadores de V y $|B| = dim(V)$, entonces B es una base de V.

Como B es un conjunto finito de generadores de V, entonces por el teorema 4.4.2 se concluye que existe una base de V, digamos B', tal que $B' \subseteq B$. Como por hipótesis, $|B| = dim(V)$, se seguiría que $|B| = |B'|$ y debido a que $B' \subseteq B$, se tiene que $B = B'$. Así, B es una base de V.

(a) Si B es linealmente independiente y $|B| = dim(V)$, entonces B es una base de V.

Como B es un conjunto finito y linealmente independiente, entonces por el teorema 4.4.3 se concluye que existe una base de B, digamos B', tal que $B \subseteq B'$. Como por hipótesis, $|B| = dim(V)$, se seguiría que $|B| = |B'|$ y debido a que $B \subseteq B'$, se tiene que $B = B'$. Así, B es una base de V.

□

El último resultado es referente a la dimensión de subespacios vectoriales.

Lema 4.5.8. *Sean V un espacio vectorial finitamente generado y W un subespacio de V. Los siguientes enunciados se satisfacen:*

a) $dim(W) \leq dim(V)$.

b) Si $dim(W) = dim(V)$, entonces $V = W$.

Demostración. Ahora mostraremos ambos incisos de este resultado.

(a) $dim(W) \leq dim(V)$.

Notemos que si $W = \{\overline{0}\}$, entonces $dim(W) = 0$ y el enunciado se satisface, por lo que supondremos que W es no trivial. Con ello, sea B una base de W. Como B es un conjunto linealmente independiente, entonces por el lema 4.5.5 (b) se sigue que $|B| \leq dim(V)$ y y por la definición de dimensión y elección de B, concluimos que $dim(W) \leq dim(V)$.

(b) Si $dim(W) = dim(V)$, entonces $V = W$.

Notemos que si $W = \{\bar{0}\}$, entonces $dim(W) = 0$ por hipótesis se concluiría que $V = \{\bar{0}\}$ con lo cual, $W = V$. Ahora supondremos que W es no trivial. Con ello, sea B una base de W. Notemos que B es un conjunto linealmente independiente. Por otro lado, como $dim(W) = dim(V)$, se tiene que $|B| = dim(W)$ por el lema 4.5.7 (b) se concluye que B es una base de V. Con ello, por ser B un conjunto de generadores de W y un conjunto de generadores de V, entonces $V = W$.

$\square$

4.6. Apéndice del Capítulo 4

En este apéndice se resumen los conceptos y los resultados más relevantes de este capítulo. Es importante mencionar que varios de ellos son verdaderos para espacios vectoriales en general y no solamente los espacios $\mathbb{R}^n$.

Subespacio vectorial (por teorema). Sean V un espacio vectorial y $W \subseteq V$ en V. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) W es un subespacio vectorial de V.

b) W satisface las siguientes propiedades:

- b.1) $\overline{0} \in W$.
- b.2) Si $\overline{x} \in W$ y $\overline{z} \in W$, entonces $\overline{x} + \overline{z} \in W$.
- b.3) Si $\lambda \in \mathbb{R}$ y $\overline{x} \in W$, entonces $\lambda\overline{x} \in W$.

Combinación lineal. Sean V un espacio vectorial, $W \subseteq V$, y $\overline{x} \in V$. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $\overline{x}$ es combinación lineal de W

b) Existe $\{\overline{z}_1, \ldots, \overline{z}_n\} \subseteq W$ y existe $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{R}$ tales que $\overline{x} = \lambda_1\overline{z}_1 + \cdots + \lambda_n\overline{z}_n$

a) $\overline{x}$ no es combinación lineal de W.

b) Para todo $\{\overline{z}_1, \ldots, \overline{z}_n\} \subseteq W$ y todo $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{R}$, se satisface que $\overline{x} \neq \lambda_1\overline{z}_1 + \cdots + \lambda_n\overline{z}_n$.

Espacio generado. Sean V un espacio vectorial, $B \subseteq V$, y $\overline{x} \in V$. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $\overline{x} \in \langle B \rangle$

b) Existe $\{\overline{z}_1, \ldots, \overline{z}_n\} \subseteq B$ y existe $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{R}$ tales que $\overline{x} = \lambda_1\overline{z}_1 + \cdots + \lambda_n\overline{z}_n$

a) $\overline{x} \notin \langle B \rangle$

b) Para todo $\{\overline{z}_1, \ldots, \overline{z}_n\} \subseteq B$ y todo $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{R}$ se satisface que $\overline{x} \neq \lambda_1\overline{z}_1 + \cdots + \lambda_n\overline{z}_n$

Espacio generado. Sean V un espacio vectorial, $S \subseteq V$ con $S \neq \emptyset$ y $\overline{x} \in V$. Los incisos (a) y (b) de los siguientes recuadros son equivalentes:

a) $\overline{x} \in \langle S \rangle$.

b) Existen $\{\overline{v}_1, \ldots, \overline{v}_n\} \subseteq S$ y $\{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{R}$ tales que $\overline{x} = \alpha_1 \overline{v}_1 + \cdots + \alpha_n \overline{v}_n$.

Sea V un espacio vectorial. Si $S \subseteq V$, entonces $\langle S \rangle$ es un subespacio vectorial de V.

Conjunto de generadores. Sean V un espacio vectorial, W un subespacio de V y $\{\overline{x}_1, \ldots, \overline{x}_n\} \subseteq V$. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) El conjunto $\{\overline{x}_1, \ldots, \overline{x}_n\}$ es un conjunto de generadores de W.

b) $\{\overline{x}_1, \ldots, \overline{x}_n\} \subseteq W$ y para todo $\overline{z} \in W$, existen $c_1, \ldots, c_n \in \mathbb{R}$ tales que $\overline{z} = c_1 \overline{x}_1 + \cdots + c_n \overline{x}_n$.

Sea V un espacio vectorial y $B \subseteq V$ un conjunto finito no vacío. Si existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\} \rangle$, entonces $\langle B \rangle = \langle B \setminus \{\overline{x}\} \rangle$.

Conjunto linealmente dependiente. Sean V un espacio vectorial y $B = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k\}$ un subconjunto de V. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) B es linealmente dependiente.

b) Existe $\{\lambda_1, \lambda_2, \ldots, \lambda_k\} \subseteq \mathbb{R}$ tales que $\lambda_1 \overline{x}_1 + \lambda_2 \overline{x}_2 + \cdots + \lambda_k \overline{x}_k = \overline{0}$ y $\lambda_j \neq 0$ para algún $j \in \{1, \ldots, k\}$.

Conjunto linealmente independiente. Sean V un espacio vectorial y $B = \{\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k\}$ un subconjunto de V. Los incisos (a) y (b) del siguiente recuadro son equivalentes:

a) B es linealmente independiente.

b) Si $\lambda_1 \overline{x}_1 + \lambda_2 \overline{x}_2 + \cdots + \lambda_k \overline{x}_k = \overline{0}$, entonces $\lambda_i = 0$ para todo $i \in \{1, \ldots, k\}$.

Los siguientes son resultados importantes respecto a conjuntos linealmente independientes y linealmente dependientes.

Sean V un espacio vectorial y $B \subseteq V$ linealmente independiente y finito. Si $\overline{x} \in V \setminus \langle B \rangle$, entonces $B \cup \{\overline{x}\}$ es linealmente independiente.

Sean V un espacio vectorial y B un subconjunto finito de V. B es linealmente dependiente si y sólo si existe $\overline{x} \in B$ tal que $\overline{x} \in \langle B \setminus \{\overline{x}\} \rangle$.

Sean V un espacio vectorial y B un subconjunto finito de V. B es linealmente independiente si y sólo para todo $\overline{x} \in B$, $\overline{x} \notin \langle B \setminus \{\overline{x}\} \rangle$.

Sean V un espacio vectorial y $B \subseteq V$. Si $\overline{0} \in B$, entonces B es linealmente dependiente.

Sean V un espacio vectorial y B un subconjunto finito de V. Si existen dos vectores distintos en B, digamos $\overline{x}$ y $\overline{z}$, tales que $\overline{x} = c\overline{z}$ para alguna $c \in \mathbb{R}$, entonces B es un conjunto linealmente dependiente.

Sean V un espacio vectorial y B y B' subconjuntos finitos de V tales que $B \subseteq B'$. Si B es linealmente dependiente, entonces B' es linealmente dependiente.

Sean V un espacio vectorial y B y B' subconjuntos finitos de V tales que $B \subseteq B'$. Si B' es linealmente independiente, entonces B es linealmente independiente.

Base. Sean V un espacio vectorial y B un subconjunto de V. Los incisos (a) y (b) del siguiente recuadro son equivalentes.

a) B es una base de V.

b) B es un conjuntos de generadores de V y B es linealmente independiente.

Los siguientes son resultados importantes respecto a las bases.

Teorema 4.6.1 (Existencia de Bases). *Todo espacio vectorial no trivial tiene al menos una base.*

Teorema 4.6.2 (De la expresión única). *Sean V un espacio vectorial no trivial y $B = \{\overline{x}_1, \ldots, \overline{x}_n\}$ un conjunto de vectores de V. Son equivalentes:*

a) *B es una base de V.*

b) *Para todo $\overline{w} \in V$ existen escalares únicos $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\overline{w} = \lambda_1\overline{x}_1 + \cdots + \lambda_n\overline{x}_n$.*

Sean V un espacio vectorial no trivial y $B \subseteq V$. Si B es un conjunto finito de generadores de V, entonces existe una base de V, digamos B', tal que $B' \subseteq B$.
En particular, si V es finitamente generado, entones todo conjunto finito de generadores de V contiene una base finita de V.

Sean V un espacio vectorial no trivial, finitamente generado y $B \subseteq V$. Si B es un conjunto finito y linealmente independiente de V, entonces existe una base de V, digamos B', tal que $B \subseteq B'$

Sean V un espacio vectorial y $B \subseteq V$. Son equivalentes:

a) B es una base de V.

b) B es linealmente independiente maximal, es decir, B es linealmente independiente y si B' es un conjunto linealmente independiente y $B \subseteq B'$, entonces $B = B'$.

c) B es generador minimal de V, es decir, B es un conjunto de generadores de V y si B' es un conjunto de generadores de V tal que $B' \subseteq B$, entonces $B = B'$.

Dimensión. Sean V un espacio vectorial finitamente generado, $S \subseteq V$ y $n \in \mathbb{N}$. Los incisos del siguiente recuadro son equivalentes:

a) $dim(V) = n$.

b) Si S es una base de V, entonces $|S| = n$

Los siguientes son resultados importantes referentes a dimensión de espacios vectoriales.

Sea V un espacio vectorial finitamente generado y $B \subseteq V$ finito. Los siguientes enunciados se satisfacen.

a) Si B es un conjunto de generadores de V, entonces $dim(V) \leq |B|$.

b) Si B es un conjunto linealmente independiente de V, entonces $|B| \leq dim(V)$.

Sea V un espacio vectorial finitamente generado y $B \subseteq V$. Los siguientes enunciados se satisfacen.

a) Si $|B| < dim(V)$, entonces B no es un conjunto generador de V.

b) Si $|B| > dim(V)$, entonces B no es un conjunto linealmente independiente.

c) Si $|B| \neq dim(V)$, entonces B no es base de V.

Sean V un espacio vectorial finitamente generado y $B \subseteq V$ finito. Los siguientes enunciados se satisfacen.

a) Si B es un conjunto de generadores de V y $|B| = dim(V)$, entonces B es una base de V.

b) Si B es linealmente independiente y $|B| = dim(V)$, entonces B es una base de V.

Sean V un espacio vectorial finitamente generado y W un subespacio de V. Los siguientes enunciados se satisfacen:

a) $dim(W) \leq dim(V)$.

b) Si $dim(W) = dim(V)$, entonces $V = W$.

Capítulo 5

Matrices y Determinantes

5.1. Definiciones básicas para matrices

Una **matriz real de orden $\boldsymbol{n \times m}$**, digamos A, es una función $A : \{1, \dots, n\} \times \{1, \dots, m\} \to \mathbb{R}$. Esto se puede interpretar como un arreglo de $n \cdot m$ números reales acomodados en n renglones y m columnas. Ejemplos de matrices de orden $n \times m$ se muestran a continuación.

$$A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \\ 5 & 6 \\ 3 & 4 \end{pmatrix} \qquad A' = \begin{pmatrix} 12 & 25 & -12 & 4 \\ 0 & 1 & 10 & 21 \\ 0 & 0 & 8 & 4 \\ 0 & 0 & 0 & 9 \end{pmatrix} \qquad A^* = \begin{pmatrix} 3 & -3 & 0 & 7 & 0 \end{pmatrix}$$

A los valores de la imagen de la matriz se les llama coloquialmente como **entradas de la matriz** y estas quedan determinadas por su posición en la matriz. Esto es, cada entrada de la matriz A se puede representar por un símbolo genérico acompañado de dos subíndices, por ejemplo: a_{ij}, donde i indica el número del renglón en el que se encuentra el valor a_{ij} y j representa la columna en la que se encuentra dicho valor[1]. Como ejemplo de lo anterior, la entrada a_{32} de la matriz A mostrada anteriormente es el valor que está en el tercer renglón y la segunda columna, es decir, $a_{32} = 6$. Otro ejemplo, el valor a'_{44} de la matriz A' mostrada previamente, es el ubicado en el cuarto renglón y cuarta columna de la matriz, es decir, el valor 9.

Cuando tengamos una matriz A de $n \times m$ y $n \geq 10$ o $m \geq 10$, es conveniente usar una coma en el subíndice para evitar ambigüedades. Por ejemplo, si se tiene una matriz B de 15×15, el valor b_{113} es ambiguo en su notación, pues puede hacer referencia a dos valores distintos y por ello es importante la coma: el valor $b_{1,13}$ hace referencia a la entrada en el primer renglón y la 13-ava columna, mientras que el valor $b_{11,3}$ hace referencia al

[1]Por lo general, a las matrices se les representa con letras mayúsculas y a sus entradas por la misma letra, pero en minúscula.

valor en el 11-avo renglón y tercera columna.

Uno de los primeros problemas al trabajar con matrices es la dificultad de su notación. Una matriz de 3×3 es, en realidad, un arreglo de 9 números reales, mientras que una de 4×5 es una de 20 número reales, por lo que escribir 20 números para cada una de las propiedades que veremos a continuación será un problema. Derivado de ello, utilizaremos una notación auxiliar. Supongamos que tenemos la siguiente matriz:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Para simplificar la notación anterior, podemos escribir simplemente la matriz A por:

$$A = (a_{ij})_{n\times m}$$

Además de lo anterior, si A es una matriz, podemos hacer referencia a la entrada ij de dicha matriz por $(A)_{ij}$, por ejemplo, teniendo en consideración la matriz B mostrada a continuación.

$$B = \begin{pmatrix} -1 & 4 & -2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 8 & -2 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

$$(B)_{22} = 1$$
$$(B)_{34} = 2$$
$$(B)_{41} = 0$$

Notemos que si $A = (a_{ij})_{n\times m}$ y $B = (b_{ij})_{n\times m}$ son dos matrices, entonces éstas son **iguales** si para toda $i \in \{1, \ldots, n\}$ y toda $j \in \{1, \ldots, m\}$, se satisface que $a_{ij} = b_{ij}$.

Al conjunto de todas las matrices reales de orden $n \times m$ se le denota por $M_{n\times m}[\mathbb{R}]$ es decir:

$$M_{n\times m}[\mathbb{R}] = \{(a_{ij})_{n\times m} : a_{ij} \in \mathbb{R} \text{ para toda } i \in \{1, \ldots, n\} \text{ y toda } j \in \{1, \ldots, m\}\}.$$

Por ejemplo, el conjunto $M_{2\times 3}[\mathbb{R}]$ tiene por elementos a todas las matrices con entradas reales con 2 renglones y 3 columnas; mientras que $M_{15\times 20}[\mathbb{R}]$ denota el conjunto de todas las matrices con 15 renglones y 20 columnas.

Otro par de definiciones que necesitaremos antes de continuar con la teoría de matrices son las siguientes. Dada una matriz, digamos $A = (a_{ij})_{n\times m}$, el **$k-$ésimo vector columna de** A, denotado por $C_k(A)$, es el vector $(a_{1k}, a_{2k}, \ldots, a_{nk})$. Del mismo modo, el **$k-$ésimo vector renglón de** A, denotado por $R_k(A)$, es el vector $(a_{k1}, a_{k2}, \ldots, a_{km})$. En muchas ocasiones nos referiremos a los *renglones* o *columnas* de una matriz como los vectores definidos anteriormente. A continuación se da un ejemplo de las notaciones antes mencionadas.

$$B = \begin{pmatrix} -1 & 4 & -2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 8 & -2 \end{pmatrix}$$

$$R_3(B) = (0, 0, 8, -2)$$

$$C_3(B) = (-2, 0, 8)$$

$$R_1(B) = (-1, 4, -2, 0)$$

5.2. Matrices particulares

Dentro de todas las matrices existentes hay algunas que son más requeridas que otras, ya sea por sus propiedades específicas o por la simpleza de su manejo. En esta sexión exploraremos algunas de ellas. Dada una matriz $A \in M_{n\times m}[\mathbb{R}]$, diremos que **$A$ es una matriz cuadrada** si $n = m$. Es decir, una matriz es cuadrada cuando tiene el mismo número de renglones que de columnas.

Dada una matriz cuadrada, digamos $A = (a_{ij})_{n\times m}$, la **diagonal de A** es el vector $(a_{11}, a_{22}, \ldots, a_{nn})$. Con lo anterior, decimos que una matriz es **diagonal** si las entradas de dicha matriz que no estén en su diagonal son iguales a 0. A la matriz diagonal de $n \times n$ en las que todas las entradas de la diagonal son igual a 1 se le llama **matriz identidad de orden n** y se denota por Id_n. A continuación se muestra el ejemplo de una matriz cuadrada de 2×2, una matriz diagonal de 3×3 y la matriz identidad de orden 4.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad Id_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Una matriz cuadrada, digamos $B = (b_{ij})_{n\times n}$, es una **matriz triangular superior** si se cumple que $b_{ij} = 0$ si $j < i$. Intuitivamente hablando, en una matriz triangular superior, todos los valores que estén por debajo de la diagonal deben ser iguales a 0, por lo que si dicha matriz tiene valores no nulos, estos deben estar por encima de la diagonal. De manera análoga, diremos que B es **triangular inferior** si $b_{ij} = 0$ si $i < j$, e intuitivamente hablando, en una matriz triangular inferior, si esta tiene valores no nulos, estos deben estar por debajo de la diagonal

Ejemplo de matrices triangulares son las siguientes:

$$A = \begin{pmatrix} 1 & 3 & 3 & 5 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 9 & 0 & 0 \\ 1 & 2 & 3 & 5 & 0 \\ 3 & 4 & 8 & 2 & 1 \end{pmatrix}$$

Sin embargo no todas las matrices son triangulares. Por ejemplo, una condición para que una matriz sea triangular es que sea una matriz cuadrada, por lo que existen muchas matrices que no son cuadradas y por ello no son triangulares. En tal caso es necesario poder establecer un tipo de matriz que preserve varias propiedades de las matrices triangulares pero que no estén restringidas al hecho de ser matrices cuadradas.

Dada una matriz $A = (a_{ij})_{n\times m}$, diremos que A es una **matriz escalonada por renglones** si satisface las siguientes condiciones:

(i) Si $R_k(B) = \overline{0}$ y $l \geq k$, entonces $R_l(B) = \overline{0}$.

(ii) Si a_{ik} es el primer elemento distinto de cero de renglón i y si a_{rj} es el primer elemento distinto de cero del renglón k con $r > i$, entonces $k < j$.

Estas matrices parecen ser un poco más difíciles de interpretar intuitivamente hablando, por lo que vamos a desmenuzar su información con cuidado. Supongamos que A es una matriz que satisface los puntos (i) y (ii) arriba mencionados. Primero veamos el punto (i) de la definición. Este inciso indica que si la matriz A tiene un renglón en el que todos sus valores son iguales a 0, digamos, el k-ésimo renglón, entonces todos los renglones por debajo de él también tienen todos sus valores iguales a cero. Se debería de ver de la siguiente forma:

$$A = \left(\begin{array}{ccccc} a_{11} & \cdots & a_{1i} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,i} & \cdots & a_{k-1},m \\ \hline 0 & \cdots & 0 & \cdots & 0 \\ \hline 0 & \ddots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \cdots & 0 \end{array}\right)$$

Ahora vamos con el punto (ii). Para este, primero hay que tomar un renglón arbitrario de A. Ahora, nos fijamos en el primer valor de ese renglón que es distinto de 0, digamos que dicho valor es a_{rs}. El punto (ii) indica que cualquier valor que esté por debajo y/o a la izquierda de a_{rs} debe ser igual a 0. Por lo que la matriz debería verse algo así:

$$A = \left(\begin{array}{ccc|ccc}
a_{11} & \cdots & a_{1,s} & a_{1,s+1} & \cdots & a_{1m} \\
a_{21} & \cdots & a_{2s} & a_{2,s+1} & \cdots & a_{2m} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\cline{1-3}
0 & \cdots & a_{rs} & a_{r,s+1} & \cdots & a_{rm} \\
0 & \ddots & 0 & a_{r+1,s+1} & \ddots & a_{r+1,m} \\
0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nm}
\end{array}\right)$$

Con esto en mente, la idea intuitiva de las matrices escalonadas es que los valores no nulos de los primero renglones formen una *escalera* en la matriz. Las matrices A_1 y A_2 mostradas a continuación son escalonadas, mientras que la matriz B no es escalonada.

$$A_1 = \begin{pmatrix} 1 & 0 & 2 & 2 & -2 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 3 & 1 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 & 6 & -6 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 4 \\ 1 & 0 & 0 \end{pmatrix}$$

Un primer resultado respecto de las matrices escalonadas es el siguiente.

Teorema 5.2.1. *Si $A \in M_{n\times m}[\mathbb{R}]$ es una matriz escalonada y $R_1(A), \ldots, R_k(A)$ son los renglones distintos de $\overline{0}$ de A, entonces $\{R_1(A), \ldots, R_t(A)\}$ es un conjunto linealmente independiente.*

Las matrices escalonadas, al igual que las matrices triangulares, son realmente útiles al momento de realizar cuentas y operaciones en matrices, sin mencionar que todas las matrices triangulares son un caso particular de las matrices escalonadas. A pesar de lo anterior, es posible *mejorar* la estructura de una matriz mediante el uso del siguiente concepto. Si $B = (b_{ij})_{n\times m}$ es una matriz, diremos que B es una **matriz escalonada reducida** si B es escalonada y además satisface que:

(iii) Si b_{kl} es el primer valor distinto de cero del renglón k, entonces $b_{kl} = 1$.

(iv) Si a_{ik} es el primer elemento distinto de cero de un renglón i, entonces para todo $j \in \{1, \ldots, n\}$ tal que $i \neq j$, se satisface que $a_{jk} = 0$.

Por último, considerando una matriz $D \in M_{n\times m}[\mathbb{R}]$, definimos la **matriz transpuesta de D**, denotada por D^t, como la matriz en $M_{m\times n}[\mathbb{R}]$ tal que:

$$(D^t)_{ij} = d_{ji}$$

El concepto de matriz transpuesta tampoco es muy complicado una vez que es interpretado. Técnicamente la matriz transpuesta invierte el orden de las columnas y renglones: el renglón 1 cambia por columna 1, el renglón 2 cambia por columna 2, etc.

$$D = \begin{pmatrix} 2 & 0 & -1 & 5 \\ 0 & 3 & 0 & -1 \end{pmatrix} \qquad D^t = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ -1 & 0 \\ 5 & -1 \end{pmatrix}$$

Lema 5.2.2. *Si $A \in M_{n\times k}[\mathbb{R}]$ y $B \in M_{k\times m}[\mathbb{R}]$, entonces*

(i) $(A^t)^t$

(ii) $(A \cdot B)^t = B^t \cdot A^t$

Si $A \in M_{n\times m}[\mathbb{R}]$, denotamos por A_{ij} a la matriz en $M_{(n-1)\times(m-1)}[\mathbb{R}]$ obtenida a partir de A al eliminar el renglón i y la columna j de A[2]. Por ejemplo, tenemos las siguientes matrices.

$$E = \begin{pmatrix} -3 & 6 & 0 & 10 & 1 \\ 0 & -1 & 4 & -4 & 4 \\ 2 & 2 & 2 & 0 & 2 \\ -1 & 6 & 5 & 7 & -8 \end{pmatrix} \qquad E_{24} = \begin{pmatrix} 3 & 6 & 0 & 1 \\ 2 & 2 & 2 & 2 \\ -1 & 6 & 5 & -8 \end{pmatrix} \qquad E_{15} = \begin{pmatrix} 0 & -1 & 4 & -4 \\ 2 & 2 & 2 & 0 \\ -1 & 6 & 5 & 7 \end{pmatrix}$$

5.3. Operaciones con matrices

El primer resultado importante que tendremos es que $M_{n\times m}[\mathbb{R}]$ puede tener estructura de un espacio vectorial con las siguientes operaciones:

Sean $A = (a_{ij})_{n\times m}$ y $B = (b_{ij})_{n\times m}$ dos matrices reales y $\lambda \in \mathbb{R}$. Definimos la **suma de matrices**, denotada por $A + B$, como la matriz

$$A + B = (a_{ij} + b_{ij})_{n\times m}$$

Notemos que la suma de matrices es simplemente el sumar los valores de las coordenadas respectivas de cada matriz, por lo que su operación no debería ser mucho problema para ser interpretado. Por otro lado, definimos el **producto de una matriz por un escalar**, denotado por λA, como la matriz:

$$\lambda A = (\lambda a_{ij})_{n\times m}.$$

[2]No confundir la notación de $(A)_{ij}$ con la notación A_{ij}: la primera denota un número real y la segunda denota una matriz.

De forma similar a la suma de marices, el producto de una matriz por un escalar es simplemente multiplicar cada valor de la matriz por el escalar en cuestión, por lo que tampoco representa mucha dificultad su comprensión.

Teorema 5.3.1. *El conjunto de matrices reales de orden $n \times m$ con la suma y producto por escalares es un espacio vectorial.*

El teorema anterior tiene diversas repercusiones en lo que se refiere a la teoría de matrices, pues muchos de los resultados generales de espacios vectoriales son ciertos considerando a las matrices como vectores, por ejemplo, la existencia de bases, caracterización de subespacios vectoriales, teoremas de dimensión, etc.

Además de la suma y el producto por escalares, es posible definir un producto entre matrices, pero previo a ello conviene introducir un producto entre vectores, conocido como el **producto punto de $\mathbb{R}^n$.** Si $(x_1, \ldots, x_n)$ y $(z_1, \ldots, z_n)$ son dos vectores de $\mathbb{R}^n$, el producto punto de dichos vectores, denotado por $(x_1, \ldots, x_n) \cdot (z_1, \ldots, z_n)$, se define como:

$$(x_1, \ldots, x_n) \cdot (z_1, \ldots, z_n) = x_1 z_1 + \cdots + x_n z_n$$

Gracias a lo anterior, podremos definir de manera más sencilla el producto de dos matrices. Consideremos las matrices $A = (a_{ij})_{n \times k}$ y $B = (b_{ij})_{k \times m}$. El **producto de A por B**, denotado por AB, es la matriz:

$$AB = (c_{ij})_{n \times m}$$

Donde para toda $i \in \{1, \ldots, n\}$ y toda $j \in \{1, \ldots, m\}$ se tiene que:

$$c_{ij} = R_i(A) \cdot C_j(B).$$

Con esto, la entrada ij de la matriz producto AB no es otra cosa que el producto punto del i-ésimo renglón de la matriz A con la j-ésima columna de la matriz B[3].

Notemos primero que el producto de matrices AB sólo se puede realizar cuando el número de columnas de la matriz A tiene el mismo número de renglones de la matriz B, en otro caso dicho producto no está definido. Como segunda observación, nótese que si A es una matriz de $n \times m$ y B es una matriz de $m \times k$, entonces su producto AB es una matriz de $n \times k$. Ahora veamos un ejemplo del producto de dos matrices. Consideremos las matrices:

[3] Aunque nosotros usaremos esta definición para el producto de matrices, es posible que también se encuentre en la literatura que la entrada ij de la matriz producto es

$$c_{ij} = \sum_{t=i}^{n} a_{it} b_{tj}.$$

La cual es, justamente, el producto punto del i-ésimo renglón de A con la j-ésima columna de B.

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}$$

Como la matriz A es de 2×3 y la matriz B es de 3×2, entonces su producto debe ser una matriz de 2×2, digamos:

$$AB = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

Ahora, para determinar los valores c_{ij}, basta realizar el producto punto de el i-ésimo renglón de A con la j-ésima columna de B, es decir:

$c_{11} = R_1(A) \cdot C_1(B) = (0, 0, -1) \cdot (1, 0, 2) = (0)(1) + (0)(0) + (-1)(2) = -2$

$c_{12} = R_1(A) \cdot C_2(B) = (0, 0, -1) \cdot (2, 1, 0) = (0)(2) + (0)(1) + (-1)(0) = 0$

$c_{21} = R_2(A) \cdot C_1(B) = (1, 2, 0) \cdot (1, 0, 2) = (1)(1) + (2)(0) + (0)(2) = 1$

$c_{22} = R_2(A) \cdot C_2(B) = (1, 2, 0) \cdot (2, 1, 0) = (1)(2) + (2)(1) + (0)(0) = 4$

Y de ello, se tiene que:

$$AB = \begin{pmatrix} -2 & 0 \\ 1 & 4 \end{pmatrix}$$

Los siguientes resultados son básicos en lo que refiere a estar primeras tres operaciones de matrices.

Teorema 5.3.2. *Sean A una matriz de orden $n \times k$, B y C matrices de orden $k \times m$, D una matriz de orden $m \times p$ y $\lambda \in \mathbb{R}$. Los siguientes se satisfacen:*

a) $(A \cdot B) \cdot D = A \cdot (B \cdot D)$.

b) $A \cdot (B + C) = A \cdot B + A \cdot C$.

c) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$.

d) $Id_n A = A$ *y* $AId_n = A$.

Es importante mencionar que el producto de matrices no es conmutativo. En primer lugar, recordemos que una condición para que dos matrices A y B puedan multiplicarse, es que número de columnas de A es igual al número de renglones de B, en tal caso, es posible realizar el producto $A \cdot B$. Sin embargo, esto no necesariamente implica que el número de columnas de B sea el número de renglones de A, por lo que $B \cdot A$ ni siquiera está definido. Aunado a eso, aun en las matrices cuadrada la conmutatividad no necesariamente es cierta

Nota 5.3.3. *Si A y B son matrices de orden $n \times n$, entonces no necesariamente es cierto que $A \cdot B = B \cdot A$.*

Las últimas operaciones con matrices que revisaremos son llamadas **operaciones elementales por renglones** [4]:

1. Intercambiar dos renglones de una matriz.

2. Multiplicar todos los valores de un renglón por un mismo número real distinto de cero.

3. Sumar a un renglón un múltiplo escalar de otro renglón.

Si dadas dos matrices A y B se tiene que B se puede obtener mediante la matriz A al realizar operaciones elementales por renglones a la matriz A, entonces diremos que $\boldsymbol{A}$ **y** $\boldsymbol{B}$ **son equivalentes**. Lo anterior lo denotamos simplemente por $A \sim B$. Una matriz $A \in M_{n\times n}[\mathbb{R}]$ es **elemental** si A es obtenida al realizar una operación elemental sobre Id_n.

5.4. El rango de una matriz

El concepto de rango está más relacionado con una propiedad que no se estudiará en este curso: las matrices son una forma simplificada de ver un cierto tipo de funciones entre espacios vectoriales, a saber, las transformaciones lineales. La imagen de una transformación lineal es un subespacio vectorial del codominio, por lo que es posible calcular su dimensión. A pesar de que en el capítulo anterior verificamos que el conocer la dimensión de un espacio es verdaderamente útil, no exploramos la importancia de conocer la dimensión de la imagen de una transformación lineal. Es tan importante que dicho número recibe un nombre y se llama justamente *rango de la matriz*. Sin embargo, dado que este es un curso de Álgebra Superior I y no de álgebra Lineal, nos conformaremos con una de las equivalencias de dicha definición. Esta equivalencia será más fácil de tratar y será suficiente para obtener todos los resultados que necesitamos. Dada una matriz A, el **rango de una matriz**, denotado por $rang(A)$, es la dimensión del espacio generado por sus vectores columna.

La definición anterior ya nos habla de términos más familiares pues conocemos los conceptos de dimension, espacio generado y vectores columna.

Tomemos por ejemplo la siguiente matriz:

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

[4]También existen las **operaciones elementales por columnas** y se definen de manera análoga.

El calcular el rango de la matriz A se reduce a calcular la dimensión del subespacio generado por sus columnas, es decir: $\langle (2,1,0),(4,0,9),),(0,0,0),(7,-3,-2)\rangle$. Como ya trabajamos esto en el capítulo pasado, calcular el rango de una matriz no debería ser difícil, pero ¿a quién le interesa conocer la dimensión del espacio generado por sus columnas? Pues esta última pregunta tiene buenas respuestas, pero nosotros sólo arañaremos un poco de éstas cuando lleguemos a la parte de encontrar soluciones de sistemas de ecuaciones.

Pues bien, si es tan importante encontrar el mentado rango, sería bueno que tuviéramos algunos resultados que nos faciliten más la vida con este concepto. Y sí los hay. Primero veremos que en realidad el espacio generado por las columnas y el espacio generado por los renglones es, de hecho, el mismo.

Teorema 5.4.1. *Si* $A \in M_{n\times m}[\mathbb{R}]$, *entonces* $dim(\langle R_1(A),\ldots,R_m(A)\rangle) = dim(\langle C_1(A),\ldots,C_n(A)\rangle)$

Sin embargo, el resultado anterior no pareciera muy útil al momento de encontrar el rango, sólo dice que es lo mismo calcular la dimensión del espacio generado por las columnas que la dimensión del espacio generado por los renglones de la matriz[5].

El siguiente lema es importante para poder calcular el rango de una matriz.

Teorema 5.4.2. *Si* A *y* B *son matrices en* $M_{n\times m}[\mathbb{R}]$ *y* B *es obtenida a partir de* A *mediante operaciones elementales, entonces* $rang(A) = rang(B)$.

Esto es muy útil a la hora de calcular el rengo de una matriz si no queremos dedicarnos a encontrar bases de espacios generados. Técnicamente no importa qué operación elemental apliquemos sobre la matriz, ésta no alterará su rango. Si además de lo anterior tenemos en consideración que:

Teorema 5.4.3. *Si* $A \in M_{n\times k}[\mathbb{R}]$ *es una matriz escalonada, entonces el rango de* A *es igual al número de renglones de* A *distintos de* $\overline{0}$.

Entonces tenemos una muy buena forma de calcular el rango de una matriz. Por ejemplo, consideremos la siguiente matriz:

$$D = \begin{pmatrix} 2 & 3 & 1 & 7 & 0 & 4 \\ 1 & 3 & 1 & 3 & 0 & 2 \\ 4 & 6 & 2 & 14 & -2 & 9 \\ 4 & 6 & 2 & 14 & 0 & 8 \end{pmatrix}$$

Calcular su rango es calcular una base para el espacio generado por sus columnas (o sus renglones), lo cual

[5]El Teorema 5.4.1 no es un teorema que se deba descartar tan fácilmente: no importa quién sea la matriz, el la dimensión del espacio generado por sus renglones y la dimensión del espacio generado por sus columnas es siempre el mismo.

puede ser un poco engorroso. En todo caso, podemos llevar a la matriz D a una forma escalonada:

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

Por el Teorema 5.4.2 se concluye que el rango de D y de F son exactamente el mismo. Por último, el Teorema 5.4.3 indica que el rango de F es 3, pues tiene tres renglones distintos de cero, con lo cual, el rango de D es 3.

Teorema 5.4.4. *Si $A \in M_{n\times m}[\mathbb{R}]$ entonces $rang(A) = rang(A^t)$.*

5.5. Invertibilidad de matrices

Dentro de esta sección revisaremos el concepto de matriz invertible, pero para ello es necesario conocer un poco de teoría previa respecto al producto de matrices.

Lema 5.5.1. *Si $A \in M_{n\times m}$, $B \in M_{m\times k}$ y $C \in M_{k\times l}[\mathbb{R}]$, entonces $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.*

Lema 5.5.2. *Sean $A \in M_{n\times m}[\mathbb{R}]$ y B y C matrices en $M_{n\times k}[\mathbb{R}]$ y $\alpha \in \mathbb{R}$.*

(i) $A \cdot (B + C) = (A \cdot B) + (A \cdot B)$.

(ii) $\alpha(A \cdot B) = (\alpha A) \cdot B = A \cdot (\alpha B)$.

(iii) $Id_m \cdot A = A$

(iv) $A \cdot Id_m = A$.

Consideremos una matriz $A \in M_{n\times m}[\mathbb{R}]$ y $B \in M_{m\times n}[\mathbb{R}]$. Diremos que **$B$ es un inverso derecho de A** si se cumple que $A \cdot B = Id_{n\times n}$. Del mismo modo, si $C \in M_{l\times m}[\mathbb{R}]$, diremos que **$C$ es un inverso izquierdo de A** si $C \cdot A = Id_n$.

Teorema 5.5.3. *Si $A \in M_{n\times n}[\mathbb{R}]$ es un matriz tal que B y C son matrices inversas izquierda y derecha de A respectivamente, entonces $B = C$.*

Si $A \in M_{n\times n}[\mathbb{R}]$, diremos que **$A$ es invertible** si A tiene un inverso derecho e izquierdo. De acuerdo al Teorema 5.5.3 dicho inverso debe ser único, por lo que se le denota por A^{-1} y recibe el nombre de **matriz inversa de A**. Los siguientes son propiedades básicas y fácilmente verificables sobre matrices invertibles.

Teorema 5.5.4. *Si A y B son matrices invertibles, entonces A^{-1} y $A \cdot B$ son invertibles. Más aun, $(A^{-1})^{-1} = A$ y $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.*

Corolario 5.5.5. *Si $A_1, \ldots A_k$ son matrices en $M_{n\times n}[\mathbb{R}]$ y todas son invertibles, entonces $A_1 \cdot A_2 \cdot \cdots \cdot A_k$ es invertible.*

El Corolario 5.5.5, aunque no lo parezca, será importante para poder encontrar la matriz inversa (si es que existe) de una matriz dada.

Corolario 5.5.6. *Sean A, B y C matrices en $M_{n\times n}[\mathbb{R}]$ tales que $A = B \cdot C$ y C es invertible. A es invertible si y sólo si B es invertible.*

Lema 5.5.7. *Toda matriz elemental es invertible y su inversa es una matriz elemental.*

Teorema 5.5.8. *Si $A \in M_{n\times n}[\mathbb{R}]$ y $E_1, \ldots, E_k$ son matrices elementales en $M_{n\times n}[\mathbb{R}]$ tales que*

$$E_1 \cdot E_2 \cdot \cdots \cdot E_k \cdot A = Id_n$$

Entonces A es invertible y $A^{-1} = E_1 \cdot E_2 \cdot \cdots \cdot E_k$.

El Teorema 5.5.8 será el que nos otorgará una forma justificada de encontrar la matriz inversa si es que ésta existe. Para ello tenemos el siguiente algoritmo:

Algoritmo para encontrar la matriz inversa.

1. Consideramos una matriz cuadrada A y mediante operaciones elementales sobre A intentamos obtener la matriz identidad. Tengamos cuidado de anotar en orden qué operaciones elementales hemos realizado a A.

2. Si es posible obtener a la matriz identidad a partir de A, procedemos al siguiente paso. En caso contrario, la matriz A no es invertible.

3. Aplicamos a la matriz identidad las operaciones elementales que aplicamos previamente a la matriz A y en el mismo orden que fueron aplicadas. La matriz obtenida de este proceso es la matriz inversa de A.

Vamos el algoritmo anterior con un ejemplo sencillo. Consideremos la siguiente matriz A

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

Lo primero que hacemos es utilizar operaciones elementales para obtener la matriz identidad a partir de A:

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 3 & 3 \end{pmatrix} \xrightarrow[R_3 - R_1]{} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow[R_3 - R_2]{} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[R_1 - R_2]{}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_2 - R_3]{} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_1 - R_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ahora aplicamos las mismas operaciones elementales a la matriz identidad:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_3 - R_1]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \xrightarrow[R_3 - R_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \xrightarrow[R_1 - R_2]{}$$

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \xrightarrow[R_2 - R_3]{} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix} \xrightarrow[R_1 - R_2]{} \begin{pmatrix} 0 & -3 & 1 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Y con lo anterior tenemos que:

$$A^{-1} = \begin{pmatrix} 0 & -3 & 1 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

5.6. Determinantes

En esta sección nos limitaremos a calcular el determinante de una matriz sin involucrarnos plenamente en su definición formal. La definición que nosotros daremos será recursiva, por lo que el calcular el determinante de una matriz cuadrada implica un proceso *relativamente* largo en sus cálculos.

Si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ es una matriz real, definimos el **determinante de A**, denotado por $det(A)$ o $|A|$, como

$$det(A) = a_{11}a_{22} - a_{21}a_{12}$$

Si $B = (b_{ij})_{n\times n} \in M_{n\times n}[\mathbb{R}]$ es una matriz cuadrada real, para un valor fijo $k \in \{1,\dots,n\}$ definimos el **determinante de B**, denotado por $det(B)$ o $|B|$ como

$$det(B) = \sum_{i=1}^{n}(-1)^{i+k}a_{ik}det(A_{ik}).$$

Al valor $(-1)^{l+k}det(A_{lk})$ se le llama el **cofactor de a_{lk}**.

Como se podrá notar, la definición de determinante[6] parece bastante rebuscada y de difícil comprensión.

En este punto es importante aclarar algunas cosas, la notación de una matriz está delimitada por paréntesis redondos, mientras que dicha notación se suele cambiar por dos barras rectas cuando se calcula su determinante. Es decir, si tenemos la matriz

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

Usamos la siguiente notación (cuando sea necesaria) para indicar su determinante:

$$\begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix}$$

Es importante poder distinguir entre ambas notaciones: la primera indica una matriz y la segunda un número real, en este caso, $det(A) = -5$.

Otro punto importante a tener en cuenta (y que es un resultado bastante largo) es que no importa quién sea el valor de k, el determinante siempre es el mismo:

Teorema 5.6.1. *Si $A = (a_{ij})_{n\times n} \in M_{n\times n}[\mathbb{R}]$ y $s, t \in \{1,\dots,n\}$ son fijos, entonces:*

$$\sum_{i=1}^{n}(-1)^{i+s}a_{is}det(A_{is}) = \sum_{i=1}^{n}(-1)^{i+t}a_{it}det(A_{it}).$$

Ahora vamos a calcular el determinante de una matriz de 3×3. Consideremos la matriz:

$$B = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 4 \end{pmatrix}$$

Lo primero que tenemos que hacer es *fijar* un valor para k, es decir, escogeremos una columna de B (y gracias al Teorema 5.6.1) que sea de nuestro agrado. En mi caso usaré la columna 2, es decir $k = 2$. En tal caso,

[6] En muchas ocasiones la suma antes expuesta se conoce más como **la expansión del determinante de B a través de la columna k** y es propiamente una forma de *calcular* un determinante que su definición en sí.

utilizando la expansión por columnas tenemos que:

$$\begin{aligned}
det(B) = \sum_{i=1}^{3}(-1)^{i+2}a_{i2}det(B_{i2}) &= (-1)^{1+2}a_{12}det(B_{12}) + (-1)^{2+2}a_{22}det(B_{22}) + (-1)^{3+2}a_{32}det(B_{32}) \\
&= (-1)^{3}(-1)det\begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix} + (-1)^{4}(0)det\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} + (-1)^{5}(1)det\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \\
&= (0)(-4-0) + (2)(4-0) + (-1)(0+2) \\
&= 6
\end{aligned}$$

La anterior fue el cálculo de un determinante de una matriz de 3×3 y aunque son cuentas, en realidad no parecen muchas...pero para calcular el determinante de una matriz de 4×4 implica que debemos calcular cuatro determinantes de matrices de 3×3 y éstas a su vez implican el calcular tres determinantes de 3×3 por cada una. Al final de cuentas, entre más grande sea la matriz, más cálculos necesitamos para obtener su determinante.

Un primer resultado importante es que el determinante no cambia si cambiamos la columna que hemos fijado desde el inicio, es decir, pudimos haber escogido cualquier otra columna de la matriz anterior y su determinante no cambia. De hecho podremos decir un poco más: si en lugar de fijar una columna fijamos un renglón, entonces el determinante es exactamente el mismo. Ello se deduce del siguiente resultado:

Teorema 5.6.2. *Si $A \in M_{n\times n}[\mathbb{R}]$, entonces $det(A) = det(A^t)$*

Y como corolario tenemos que la expansión de un determinante se puede realizar indistintamente de si fijamos renglones o si fijamos columnas:

Corolario 5.6.3. *Si $A = (a_{ij})_{n\times n} \in M_{n\times n}[\mathbb{R}]$ y $k, t \in \{1, \ldots, n\}$ son fijos, entonces:*

$$\sum_{i=1}^{n}(-1)^{i+k}a_{ik}det(A_{ik}) = \sum_{i=1} n(-1)^{t+i}a_{ti}det(A_{ti}).$$

Con lo anterior ya podemos encontrar rápidamente el determinante de algunas matrices muy particulares. Por ejemplo:

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

Estas son matrices de 4×4. Calcular su determinante mediante la expansión por columnas implica calcular el determinante de cuatro matrices de 3×3 y cada una de ellas requiere el calcular tres determinantes de 2×2. Toda una lata. Pero si es verdad que podemos escoger cualquier columna o cualquier renglón para calcular el determinante ¿por qué no escoger la segunda columna en el caso de la matriz A y el tercer renglón en caso de la matriz B? la elección de dicha columna y renglón respectivamente nos indicarían que en realidad no hay que hacer cálculo alguno, pues el determinante de ambas matrices es cero.

Corolario 5.6.4. *Si una matriz cuadrada tiene un renglón o una columna nula, entonces su determinante es* 0.

El problema es que pocas matrices tienen un renglón o una columna igual a cero, por lo que el Corolario 5.6.4 no parece muy útil. Es aquí donde entran diversos de nuestros resultados importantes. El primero de ellos es el hacer notar que las operaciones elementales *casi* no alteran el determinante:

Teorema 5.6.5. *Sea* $A \in M_{n\times n}[\mathbb{R}]$. *Los siguientes se satisfacen:*

(i) Si B *es obtenida a partir de* A *al multiplicar un renglón de* A *por un escalar no nulo, digamos* λ*, entonces* $det(B) = \lambda det(A)$.

(ii) Si B *se obtiene a partir de* A *al intercambiar dos renglones de* A*, entonces* $det(B) = -det(A)$.

(iii) Si B *se obtiene a partir de* A *al sumar a un renglón de* A *otro renglón de* A*, entonces* $det(B) = det(A)$.

El teorema anterior podría facilitarnos mucho las cuentas y ayudarnos a calcular el determinante de muchas otras matrices que las consideradas en el Corolario 5.6.4. Por ejemplo, si una matriz tiene dos renglones iguales, podemos usar uno de los dos renglones para hacer que el otro sea cero y de acuerdo al punto (iii) ésto no altera el determinante. Lo anterior se puede refinar si tenemos el siguiente resultado:

Corolario 5.6.6. *Sea* $A \in M_{n\times n}[\mathbb{R}]$. *En cualquiera de los siguientes casos el determinante de* A *es 0.*

(i) A *tiene un renglón (o columna) igual a cero.*

(ii) A *tiene dos renglones (o columnas) iguales.*

(iii) A *tiene un renglón (columna) que es múltiplo escalar de otro renglón (columna).*

(iv) A *tiene un renglón (columna) que es múltiplo escalar de los otros renglones (columnas).*

Y aunque lo anterior indica cómo saber si el determinante de una matriz es cero, no todas las matrices satisfacen algunas de las condiciones anteriores. Pues bien, para las matrices restantes tenemos el siguiente:

Teorema 5.6.7. *Si* $A \in M_{n\times n}[\mathbb{R}]$ *es una matriz triangular, entonces* $det(A) = \prod_{i=1}^{n} a_{ii}$.

Y con este teorema, junto al Teorema 5.6.5 ya no debemos preocuparnos por las expansiones por columnas y por renglones. Sólo hay que tomar una matriz cuadrada, llevarla mediante operaciones elementales a una matriz triangular y con ella podemos calcular el determinante. Tomemos el siguiente ejemplo:

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & -1 & 4 & 1 \\ -1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \end{pmatrix}$$

Utilizaremos operaciones elementales para obtener una forma triangular de la matriz A.

$$\begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & -1 & 4 & 1 \\ -1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \end{pmatrix} \xrightarrow{R_3 + R_1} \underbrace{\begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & -1 & 4 & 1 \\ 0 & 0 & 2 & 3 \\ \frac{1}{2} & 0 & 1 & 0 \end{pmatrix}}_{B_1} \xrightarrow{(-2)R_4} \underbrace{\begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & -1 & 4 & 1 \\ 0 & 0 & 2 & 3 \\ -1 & 0 & -2 & 0 \end{pmatrix}}_{B_2} \xrightarrow{R_4 + R_1} \underbrace{\begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & -1 & 4 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 3 \end{pmatrix}}_{B_3}$$

Dado que la matriz B_3 es un matriz triangular, entonces su determinante es igual al producto de los elementos de su diagonal, por lo que $det(B_3) = (1)(-1)(2)(3) = -6$. Por otro lado, como obtuvimos B_3 a partir de B_2 al sumar dos renglones, de acuerdo al Teorema 5.6.5 se sigue que $det(B_3) = det(B_2)$, por lo que $det(B_2) = -6$. Del mismo modo, como obtuvimos a B_2 a partir de B_1 al multiplicar un escalar por uno de sus renglones, entonces $(-2)det(B_1) = det(B_2)$, por lo que $det(B_1) = 3$. Y por último, como obtuvimos a B_1 a partir de A al sumar dos renglones, entonces $det(A) = det(B_1)$, concluyendo que $det(A) = 3$.

Y para terminar, derivado de lo anterior se tiene que:

Corolario 5.6.8. *Sea $A \in M_{n\times n}[\mathbb{R}]$. Si $rang(A) < n$, entonces $det(A) = 0$.*

Y como consecuencia de lo anterior:

Corolario 5.6.9. *Una matriz $A \in M_{n\times n}[\mathbb{R}]$ es invertible si y sólo si $det(A) \neq 0$.*

Capítulo 6

Sistemas de Ecuaciones Lineales

6.1. Sistemas de ecuaciones

Una **ecuación lineal con n variables** es una ecuación de la forma

$$a_1x_1 + \cdots + a_nx_n = b. \tag{6.1}$$

A los términos x_i se les llaman **variables** y a los términos a_i se les llama **coeficientes**, mientras que al término b se le llama **término independiente**.

En general, los coeficientes y el término independiente suelen ser valores algebraicos de algún domino entero, pero para los resultados de estas notas consideraremos que dichos términos son números reales.

El determinar cuándo existen valores reales para los que la igualdad 6.1 se satisface será el tema central de este capítulo y se auxiliará de los dos capítulos anteriores, por lo que es necesario tenerlos presentes.

Cabe mencionar que si tenemos una ecuación con n variables como la mostrada en 6.1 y un vector $\bar{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$, cada que digamos que **evaluamos la ecuación en el vector $\bar{t}$** esto quiere decir que para todo $k \in \{1, \ldots, n\}$ sustituiremos el valor t_k en la variable x_k. Notemos que el evaluar un vector cualquiera en una ecuación no necesariamente se respeta la igualdad, pero cuando esto último sucede, entonces diremos que el vector es **solución de la ecuación.**

Un **sistema de m ecuaciones con n variables** es un sistema de la forma

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1. \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases} \tag{6.2}$$

Una **solución del sistema** es un vector en $\mathbb{R}^n$, digamos $(s_1, \ldots, s_n)$, tal que para toda $k \in \{1, \ldots, m\}$ se satisface que

$$a_{k,1}s_1 + \cdots + a_{k,n}s_k = b_k.$$

Al vector $(b_1, \ldots, b_m)$ se le llamará **vector de términos independientes**.

Dado que no todo sistema tiene solución o puede tener varias soluciones, es preferible hablar del **conjunto de soluciones del sistema**. Para términos de estas notas, a dicho conjunto lo denotaremos por $\mathcal{S}_\eta$ y está determinado por:

$$\mathcal{S}_\eta = \{\overline{x} \in \mathbb{R}^n : \overline{x} \text{ es solución del sistema } \eta\}.$$

Así, un sistema de ecuaciones η es **consistente** si $\mathcal{S}_\eta$ es no vacío y es **inconsistente** en otro caso.

Nota 6.1.1. *Notemos que es lo mismo decir que un sistema es consistente si éste tiene al menos una solución y es inconsistente si no tiene solución.*

Cuando un sistema de ecuaciones como en 6.2 es tal que para toda $i \in \{1, \ldots, m\}$ se tiene que $b_i = 0$, entonces diremos que es un **sistema homogéneo**. En caso contrario, diremos que es un **sistema no homogéneo**.

Dado que nuestro primer acercamiento es determinar si un sistema tiene o no solución, notemos que todos los sistemas homogéneos tienen al menos una solución: el vector cero. A dicha solución se le conoce como **solución trivial**.

Lema 6.1.2. *Si η es un sistema homogéneo con n variables, entonces $\overline{0} \in \mathcal{S}_\eta$.*

La importancia de los sistemas homogéneos será más evidente cuando se busque el conjunto de soluciones de sistemas no homogéneos al final de este capítulo.

6.2. Sistemas equivalentes.

Con lo anterior, el propósito de todo este capítulo será encontrar el conjunto de soluciones de cualquier sistema de ecuaciones. La tarea anterior no es nada sencilla y dependerá mucho de cómo sea el sistema de ecuaciones, porque no es lo mismo calcular el conjunto de soluciones de un sistema de una ecuación con una variable que el de un sistema de diez ecuaciones con siete variables.

Para ello podremos ver una propiedad importante, existen sistemas de ecuaciones cuyo conjunto de soluciones es el mismo, pongamos por ejemplo los sistemas η y η' tienen el mismo conjunto de soluciones.

$$\eta = \begin{cases} 2x_1 + 3x_2 + x_3 + 7x_4 + 4x_6 = 0 \\ x_1 + 3x_2 + x_3 + 3x_4 + 2_6 = 0 \\ 4x_1 + 6x_2 + 2x_3 + 14x_4 - 2x_5 + 9x_6 = 0 \\ 4x_1 + 6x_2 + 2x_3 + 14x_4 + 8x_6 = 0 \end{cases}$$

$$\eta' = \begin{cases} x_1 + 4x_4 + 2x_6 = 0 \\ 3x_2 - x_3 - x_4 = 0 \\ -2x_5 + x_6 = 0 \end{cases}$$

La propiedad anterior será verdaderamente importante. Dados dos sistemas de ecuaciones, digamos η y ϕ, decimos que los sistemas son **equivalentes** si tienen el mismo conjunto de soluciones, es decir, $\mathcal{S}_\eta = \mathcal{S}_\phi$.

Esto podría ser de gran ayuda, pues si tenemos un sistema de cinco ecuaciones con cinco variables y sabemos que es equivalente a otro sistema con sólo dos ecuaciones y dos variables, entonces será mucho más sencillo calcular el conjunto de soluciones del segundo sistema y que al final de cuentas, es el mismo que el del primero. Más aun, tenemos un resultado bastante interesante respecto a lo anterior.

Teorema 6.2.1. *Sean η y ϕ dos sistemas de n ecuaciones. Si definimos la relación $\sim$ en el conjunto de todos los sistemas de n ecuaciones dada por $\eta \sim \phi$ si y sólo si $\mathcal{S}_\eta = \mathcal{S}_\phi$, entonces $\sim$ es una relación de equivalencia.*

Aunque el teorema anterior ya habla sobre una posible forma de catalogar los sistemas de ecuaciones, tenemos otro resultado que en la práctica puede ser de mayor utilidad.

Teorema 6.2.2. *Si η es un sistema de ecuaciones y ϕ es el sistema obtenido a partir de η al realizar cualquiera de las siguientes operaciones:*

i) Intercambiar de lugar dos ecuaciones de η.

ii) Multiplicar una ecuación de η por una constante no nula.

iii) Sumarle a una ecuación de η cualquier otra ecuación de η.

Entonces η y ϕ son sistemas de ecuaciones equivalentes.

A las operaciones (i), (ii) y (iii) del teorema 6.2.2 les llamaremos **operaciones elementales de sistemas de ecuaciones** o simplemente **operaciones elementales**[1].

[1] Esto cuando no haya ambigüedad sobre qué estamos operando: noten que estas operaciones son similares a las operaciones elementales por renglones en matrices.

La importancia del teorema 6.2.2 radica en poder simplificar sistemas de ecuaciones para obtener, posiblemente, una forma de determinar si éste tiene solución y en caso de que la tenga, cuál es el conjunto de soluciones. Sin embargo, existe una forma más simplificada para manejar los sistemas de ecuaciones y es utilizando matrices. Es deseable hacer notar que la notación hasta ahora introducida y el teorema 6.2.2 son muy similares al de matrices que ya habíamos introducido previamente.

Para continuar con el curso, consideremos el sistema

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1. \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases} \tag{6.3}$$

A este sistema le asociaremos dos matrices. La primera de ellas, llamada la **matriz de coeficientes de η** es una matriz de $m \times n$, la denotaremos por $\mathcal{C}_\eta$ y se define por la matriz:

$$\mathcal{C}_\eta = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

La segunda matriz que asociaremos al sistema η, llamada **matriz aumentada de η** o **matriz extendida de η** es una matriz de $m \times (n+1)$, la denotaremos por $\mathcal{A}_\eta$ y se define por

$$\mathcal{A}_\eta = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{pmatrix}$$

El uso de matrices en lugar de un sistema de ecuaciones tiene una doble finalidad: por un lado, simplifica considerablemente la notación del sistema y lo plantea en términos de una estructura matemática que ya hemos trabajado anteriormente. Por el otro, las propiedades de matrices que usamos en el capítulo anterior tienen una repercusión importante para los sistemas de ecuaciones asociados.

Nota 6.2.3. *Dado un sistema de m ecuaciones con n variables, a éste se le puede asociar una matriz de $m \times (n+1)$, a saber, su matriz aumentada. Del mismo modo, si $n \geq 1$, entonces a una matriz de $m \times (n+1)$, digamos A, se le puede asociar un sistema de m ecuaciones con n variables.*

El teorema 6.2.2 tiene un similar respecto a matrices como se muestra a continuación.

Teorema 6.2.4. *Sean η y ϕ dos sistemas de ecuaciones y $\mathcal{A}_\eta$ y $\mathcal{A}_\phi$ sus matrices extendidas respectivamente. Son equivalentes:*

i) η y ϕ son sistemas de ecuaciones equivalentes.

ii) $\mathcal{A}_\eta$ y $\mathcal{A}_\phi$ son matrices equivalentes.

El resultado anterior será bastante útil y podría ser el teorema principal de esta sección pues combina dos resultados importantes: es similar al resultado del teorema 6.2.2 pero utiliza la notación simplificada de las matrices, cosa que no sucedía con el mencionado teorema.

Notemos que el manejo de este nuevo resultado será importante para los siguientes dos capítulos.

Nota 6.2.5. *Una consecuencia del teorema 6.2.4 es que, dentro de todas las posibilidades de matrices equivalentes, existe una que será de particular utilidad. Ya sabemos que dado un sistema de ecuaciones η es posible asociarle una matriz $\mathcal{A}_\eta$ y ésta podemos operarla bajo cualquier operación elemental. En tal caso, ¿qué pasaría si B es la matriz escalonada reducida de $\mathcal{A}_\eta$? Por el teorema 6.2.4 el sistema de ecuaciones asociada a B es equivalente al sistema η, sin embargo, el sistema de ecuaciones asociado a B es considerablemente más simple que el de η.*

6.3. Criterios de existencia de soluciones.

Ahora que ya tenemos algunas herramientas en lo que refiere a sistemas de ecuaciones, nos centraremos en determinar cuándo un sistema tiene o no solución. Cabe señalar que en esta sección solamente nos centraremos en determinar *si un sistema dado tiene solución*, pero no en *quién es el conjunto de soluciones.*

El lema 6.1.2 ya nos muestra un primer criterio de existencia de soluciones:

Lema 6.3.1. *Todo sistema de ecuaciones homogéneo tiene una solución. Más aun, la solución trivial es solución de un sistema si y sólo si el sistema es homogéneo.*

De lo anterior, se puede preguntar si es posible sabe cuántas soluciones hay y al menos en un sistema homogéneo hay una primera aproximación a ello. Previo a ello, notemos el siguiente resultado:

Lema 6.3.2. *Si η es un sistema homogéneo con m ecuaciones y n variables y $\mathcal{S}_\eta$ es su conjunto de soluciones, entonces $\mathcal{S}_\eta$ es un subespacio vectorial de $\mathbb{R}^n$.*

Demostración.

Supongamos que

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0. \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = 0 \end{cases}$$

Y que $\mathcal{S}_\eta$ es el conjunto de soluciones de η. Para verificar que $\mathcal{S}_\eta$ es un subespacio vectorial de $\mathbb{R}^n$ basta verificar que:

i) $\overline{0} \in \mathcal{S}_\eta$.

Este punto es inmediato del lema 6.1.2.

ii) Si $(s_1, \ldots, s_n) \in \mathcal{S}_\eta$ y $\lambda \in \mathbb{R}$, entonces $\lambda(s_1, \ldots, s_n) \in \mathcal{S}_\eta$

Para demostrar que $\lambda(s_1, \ldots, s_n) \in \mathcal{S}_\eta$ sólo basta ver que dicho vector es solución de η, para ello, consideremos cualquier ecuación de η y evaluamos los valores del vector $(\lambda s_1, \ldots, \lambda s_n)$.

$$\begin{aligned} a_{i,1}(\lambda s_1) + a_{i,2}(\lambda s_2) + \cdots + a_{i,n}(\lambda s_n) &= \lambda(a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n) \\ &= \lambda(0) \\ &= 0 \end{aligned}$$

Con lo anterior, $(\lambda s_1, \ldots, \lambda s_n)$ es solución del sistema η por lo que $(\lambda s_1, \ldots, \lambda s_n) \in \mathcal{S}_\eta$.

iii) Si $(s_1, \ldots, s_n) \in \mathcal{S}_\eta$ y $(t_1, \ldots, t_n) \in \mathcal{S}_\eta$, entonces $(s_1, \ldots, s_n) + (t_1, \ldots, t_n) \in \mathcal{S}_\eta$.

Al igual que el punto anterior, para verificar que $(s_1, \ldots, s_n) + (t_1, \ldots, t_n) \in \mathcal{S}_\eta$ basta ver que $(s_1 + t_1, \ldots, s_n + t_n)$ es solución de η. Para ello, consideremos cualquier ecuación de η y evaluamos los valores de $(s_1 + t_1, \ldots, s_n + t_n)$.

$$\begin{aligned} a_{i,1}(s_1 + t_1) + a_{i,2}(s_2 + t_2) + \cdots + a_{i,n}(s_n + t_n) &= a_{i,1}s_1 + a_{i,1}t_1 + \cdots + a_{i,n}s_n + a_{i,n}t_n \\ &= (a_{i,1}s_1 + \cdots + a_{i,n}s_n) + (a_{i,1}t_1 + \cdots + a_{i,n}t_n) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Con lo anterior, $\mathcal{S}_\eta$ es un subespacio vectorial. $\square$

El lema anterior tiene diversas implicaciones, algunas de ellas realmente importantes. La primera es que si η es un sistema homogéneo, entonces $\mathcal{S}_\eta$ tiene una estructura de espacio vectorial, por lo que podemos hablar de bases y dimensión, las bases de $\mathcal{S}_\eta$ nos ayudarán a poder encontrar *todas* las soluciones de η, mientras que la dimensión de $\mathcal{S}_\eta$ nos indicará *cuántas* soluciones existen. Por otro lado, derivado del lema 6.3.2 es frecuente llamarle a $\mathcal{S}_\eta$ el **espacio de soluciones** en lugar de conjunto de soluciones, pero sólo en el caso de sistemas homogéneos.

Teorema 6.3.3. *Sean η un sistema de ecuaciones, $\overline{b}$ el vector de términos independientes de η y $\overline{c}_1, \ldots, \overline{c}_n$ las columnas de $\mathcal{C}_\eta$. Son equivalentes:*

i) El sistema η es consistente.

ii) $\overline{b} \in \langle \overline{c}_1, \ldots, \overline{c}_n \rangle$.

Demostración.

Supongamos que

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1. \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

Con lo anterior, tenemos que $\overline{b} = (b_1, \ldots, b_n)$ y que $\overline{c}_k = (a_{1,k}, a_{2,k}, \ldots, a_{m,k})$. Mostraremos primero que:

i) Si el sistema η es consistente, entonces $\overline{b} \in \langle \overline{c}_1, \ldots, \overline{c}_n \rangle$.

Para ello, basta mostrar que existen escalares $\lambda_1, \ldots, \lambda_n$ tales que $\overline{b} = \lambda_1\overline{c}_1 + \cdots + \lambda_n\overline{c}_n$. La hipótesis de la condicional indica que existe al menos una solución del sistema η, digamos $\overline{s} = (s_1, \ldots, s_n)$. Afirmamos que los escalares s_i son los buscados. Para ello notemos que:

$$\begin{aligned} s_1\overline{c}_1 + \cdots + s_n\overline{c}_n &= s_1(a_{1,1}, a_{2,1}, \ldots, a_{m,1}) + s_2(a_{1,2}, a_{2,2}, \ldots, a_{m,2}) + \cdots + s_n(a_{1,n}, a_{2,n}, \ldots, a_{m,n}) \\ &= (s_1a_{1,1}, s_1a_{2,1}, \ldots, s_1a_{m,1}) + (s_2a_{1,2}, s_2a_{2,2}, \ldots, s_2a_{m,2}) + \cdots + (s_na_{1,n}, s_na_{2,n}, \ldots, s_na_{m,n}) \\ &= (s_1a_{1,1} + \cdots + s_na_{1,n}, s_1a_{2,1} + \cdots + s_na_{2,n}, \ldots, s_1a_{m,1} + \cdots + s_na_{m,n}) \\ &= (b_1, b_2, \ldots, b_n) \\ &= \overline{b} \end{aligned}$$

Con lo anterior, $\overline{b}$ se escribe como combinación lineal de los vectores $\overline{c}_1, \ldots, \overline{c}_m$, por lo que $\overline{b} \in \langle \overline{c}_1, \ldots, \overline{c}_n \rangle$.

Ahora sólo resta demostrar la segunda implicación del enunciado, es decir:

ii) Si $\overline{b} \in \langle \overline{c}_1, \ldots, \overline{c}_n \rangle$, entonces η es un sistema consistente.

Para ver que η es un sistema consistente, basta verificar que existe un vector $\overline{s} \in \mathbb{R}^n$ tal que para cada ecuación de η se satisface que

$$a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n = b_n$$

Sabemos por hipótesis que $\overline{b} \in \langle \overline{c}_1, \ldots, \overline{c}_n \rangle$, por lo que existen escalares $\lambda_1, \ldots, \lambda_n$ tales que $\overline{b} = \lambda_1\overline{c}_1 + \cdots + \lambda_n\overline{c}_n$. Afirmamos que el vector $(\lambda_1, \ldots, \lambda_n)$ es solución de η. Notemos que

$$\begin{aligned} \lambda_1\overline{c}_1 + \cdots + \lambda_n\overline{c}_n &= \lambda_1(a_{1,1}, \ldots, a_{m,1}) + \cdots + \lambda_n(a_{1,n}, \ldots, a_{m,n}) \\ &= (\lambda_1a_{1,1}, \ldots, \lambda_1a_{m,1}) + \cdots + (\lambda_na_{1,n}, \ldots, \lambda_na_{m,n}) \\ &= (\lambda_1a_{1,1} + \cdots + \lambda_na_{1,n}, \ldots, \lambda_1a_{m,1} + \cdots + \lambda_na_{m,n}) \end{aligned}$$

Con lo anterior se concluye que $(b_1, \ldots, b_n) = (\lambda_1 a_{1,1} + \cdots + \lambda_n a_{1,n}, \ldots, \lambda_1 a_{m,1} + \cdots + \lambda_n a_{m,n})$ y de esto se sigue que para toda $i \in \{1, \ldots, n\}$ se cumple que

$$b_i = \lambda_1 a_{i,1} + \lambda_2 a_{i,2} + \cdots + \lambda_n a_{i,n}$$

Por lo que $(\lambda_1, \ldots, \lambda_n)$ es solución de η y así éste es consistente. $\square$

Teorema 6.3.4. *Sea η un sistema homogéneo de m ecuaciones con n variables y $\mathcal{C}_\eta$ su matriz aumentada. Si $\mathcal{S}_\eta$ es el espacio de soluciones de η, entonces $dim(S_\eta) = n - rank(\mathcal{C}_\eta)$.*

Demostración.

Corolario 6.3.5. *Sea η un sistema homogéneo de m ecuaciones con n variables. Si $m < n$ entonces el sistema tiene una ecuación no trivial. En este caso, η tiene una infinidad de soluciones.*

Demostración.

Dado que $rank(\mathcal{C}_\eta) \leq m$, entonces $rank(\mathcal{C}_\eta) < n$. En tal caso, del teorema 6.3.4 se concuye que $dim(S_\eta)$ es un subespacio vectorial no trivial por lo que η tiene una solución no trivial. Además, de acuerdo al ejercicio ??????? se sigue que η tiene una infinidad de soluciones. $\square$

Aunque los resultados anteriores ofrecen una cantidad considerable de información respecto a los sistemas homogéneos, el siguiente resultado muestra que también es posible determinar en sistemas no homogéneos.

Teorema 6.3.6. *Sean η un sistema de ecuaciones, $\mathcal{A}_\eta$ y $\mathcal{C}_\eta$ son la matriz asociada aumentada y la matriz de coeficientes de η respectivamente. Son equivalentes:*

ii) El sistema η es consistente.

ii) $rank(\mathcal{C}_\eta) = rank(\mathcal{A}_\eta)$

Demostración.

Sean η un sistema de ecuaciones, $\bar{b}$ el vector de términos independientes de η, $\mathcal{A}_\eta$ la matriz aumentada de η, $\mathcal{C}_\eta$ la matriz de coeficientes de η y consideremos $\bar{c}_1, \ldots, \bar{c}_n$ las columnas de $\mathcal{C}_\eta$. Además de lo anterior, consideremos $W = \langle \bar{b}, \bar{c}_1, \ldots, \bar{c}_n \rangle$ y $U = \langle \bar{c}_1, \ldots, \bar{c}_n \rangle$.

Primero demostraremos que (i) implica (ii), es decir:

Si el sistema η es consistente, entonces $rank(\mathcal{C}_\eta) = rank(\mathcal{A}_\eta)$.

Sabemos del teorema 6.3.3 que $\bar{b} \in U$, por lo que $\langle \bar{b}, \bar{c}_1, \ldots, \bar{c}_n \rangle = \langle \bar{c}_1, \ldots, \bar{c}_n \rangle$, es decir, $W = U$. En particular, $dim(W) = dim(U)$ y de la definición de rango de una matriz, se sigue que $rank(\mathcal{A}_\eta) = rank(\mathcal{C}_\eta)$.

Ahora procederemos a mostrar que (ii) implica (i), es decir:

Si $rank(\mathcal{A}_\eta) = rank(\mathcal{C}_\eta)$, entonces η es consistente.

Para ello notemos del lema 4.2.4 tenemos que $U \subseteq W$ y por hipótesis, $rank(\mathcal{A}_\eta) = rank(\mathcal{C}_\eta)$, es decir, $dim(U) = dim(W)$. En tal caso, por el lema 4.5.8 concluimos que $U = W$. En particular, como $\overline{b} \in W$ se concluye que $\overline{b} \in U$ y por el teorema 6.3.3 concluimos que η es consistente. □

Nota 6.3.7. *El teorema 6.3.6 nos ofrece una forma sencilla y práctica de verificar cuándo un sistema de ecuaciones tiene solución. Si η es un sistema de ecuaciones y $\mathcal{A}_\eta$ y $\mathcal{C}_\eta$ son su matriz aumentada y su matriz de coeficientes respectivamente, basta llevar ambas a matrices escalonadas reducidas y si ambas tienen eñ mismo número de renglones no nulos, entonces el sistema η tiene una solución. En cualquier otro caso, el sistema η es inconsistente.*

6.4. Conjunto de soluciones

Los resultados de la sección anterior son de gran ayuda al determinar cuándo un sistema de ecuaciones tiene solución y no estar buscando inútilmente un objeto que no existe. Sin embargo, ninguno de los resultados hasta ahora vistos nos dice cómo encontrar una solución dado un sistema de ecuaciones en general. Para ello es esta última sección.

Hasta ahora ya contamos con suficientes resultados para poder encontrar un algoritmo que determine todas las posibles soluciones de un sistema homogéneo.

Algoritmo para encontrar el conjunto de soluciones de un sistema de ecuaciones.

i) Consideremos un sistema de ecuaciones η, digamos,

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0 \\ \qquad\qquad\qquad\qquad \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = 0 \end{cases}$$

Además, supongamos que $\mathcal{A}_\eta$ es la matriz aumentada del sistema y B es la matriz escalonada reducida de $\mathcal{S}_\eta$.

ii) Encontramos el sistema asociado a B, supongamos que dicho sistema es ϕ

$$\phi = \begin{cases} x_{k_1} + \sum_{i=1}^{\alpha_{k_1}} b_i x_{\beta_i} = 0. \\ \qquad\qquad\qquad \vdots \\ x_{k_t} + \sum_{i=1}^{\alpha_{k_t}} b_i x_{\beta_i} = 0 \end{cases}$$

iii) Por el teorema 6.2.4 se sigue que $\mathcal{S}_\eta = \mathcal{S}_\phi$, por lo que bastará encontrar $\mathcal{S}_\phi$.

iv) A las variables $x_{k_1}, \ldots, x_{k_t}$ les llamaremos **variables dependientes** y a los otros términos se les llamaran **variables independientes.** Dado que la matriz B es una matriz escalonada reducida, ninguna variable independiente es también una variable dependiente.

v) Se despejan las variables dependientes del sistema ϕ.

vi) El conjunto $\mathcal{S}_\phi$ tiene vectores de la forma $(w_1, \ldots, w_n)$ donde los índices de las variables dependientes respetan las ecuaciones obtenidas en (v).

vii) De acuerdo al punto (vi), es posible dar valores a las variables independientes para encontrar vectores que pertenezcan a $\mathcal{S}_\eta$. Podemos encontrar valores de tal forma que los vectores obtenidos sean linealmente independientes.

viii) Con los vectores antes descritos es posible describir a todo el espacio de soluciones.

Aunque el procedimiento antes mencionado es sencillo de verificar, la descripción del espacio de soluciones sólo es válido si el sistema es homogéneo, pero no dice qué sucede en otro caso. Sin embargo, podremos verificar una manera sencilla de calcular una solución particular para el caso no homogéneo.

Algoritmo para encontrar una solución particular de un sistema no homogéneo.

i) Consideremos un sistema de ecuaciones η, digamos,

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

Además, supongamos que $\mathcal{A}_\eta$ es la matriz aumentada del sistema y B es la matriz escalonada reducida de $\mathcal{S}_\eta$.

ii) Encontramos el sistema asociado a B, supongamos que dicho sistema es ϕ

$$\phi = \begin{cases} x_{k_1} + \sum_{i=1}^{\alpha_{k_1}} c_i x_{\beta_i} = d_1. \\ \vdots \\ x_{k_t} + \sum_{i=1}^{\alpha_{k_t}} c_i x_{\beta_i} = d_t \end{cases}$$

iii) Por el teorema 6.2.4 se sigue que $\mathcal{S}_\eta = \mathcal{S}_\phi$, por lo que bastará encontrar una solución para ϕ.

iv) Se despejan las variables dependientes del sistema ϕ.

v) De acuerdo al punto (iv), es posible dar valores a las variables independientes para encontrar vectores que pertenezcan a $\mathcal{S}_\eta$.

Aunque el resultado anterior nos habla de una sola solución a un sistema no homogéneo, ésta será suficiente para describir todo el conjunto de soluciones de dicho sistema, como se demuestra en el siguiente teorema.

Teorema 6.4.1. *Sean η un sistema no homogéneo consistente y ϕ su sistema homogéneo asociado. Si $\overline{s}_0$ es una solución de η, entonces*

$$\mathcal{S}_\eta = \{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s} \text{ donde } \overline{s} \in \mathcal{S}_\phi\}.$$

Demostración.

Supongamos que

$$\eta = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0. \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = 0 \end{cases}$$

Y además $\overline{s}_0 = (s_1, \ldots, s_n)$. Para demostrar el teorema basta realizar una doble contención.

i) $\mathcal{S}_\eta \subseteq \{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s}$ donde $\overline{s} \in \mathcal{S}_\phi\}$.

Sea $\overline{z} \in \mathcal{S}_\eta$. Notemos primero que el vector $\overline{w} = \overline{z} - \overline{s}_0$ es una solución de ϕ. Para ello tomemos cualquier ecuación de ϕ y evaluemos $\overline{z} - \overline{s}_0$ en dicha ecuación y verifiquemos que se preserva la igualdad.

$$\begin{aligned} a_{i,1}(z_1 - s_1) + a_{i,2}(z_2 - s_2) + \cdots + a_{i,n}(z_n - s_n) &= a_{i,1}z_1 - a_{i,1}s_1 + a_{i,2}z_2 - a_{i,2}s_2 + \cdots + a_{i,n}z_n + a_{i,n}s_n \\ &= (a_{i,1}z_1 + \cdots + a_{i,n}z_n) - (a_{i,1}z_1 + \cdots + a_{i,n}s_n) \\ &= b_i - b_i \\ &= 0 \end{aligned}$$

De lo anterior se concluye que $\overline{w} \in \mathcal{S}_\phi$. Es sencillo notar que $\overline{z} = \overline{s}_0 + \overline{w}$, por lo que $\overline{z} \in \{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s}$ donde $\overline{s} \in \mathcal{S}_\phi\}$. Mostrando así la primera contención.

ii) $\{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s}$ donde $\overline{s} \in \mathcal{S}_\phi\} \subseteq \mathcal{S}_\eta$.

Consideremos un vector $\overline{z} \in \{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s}$ donde $\overline{s} \in \mathcal{S}_\phi\}$, es decir, $\overline{z} = \overline{s}_0 + \overline{w}$, donde $\overline{w} = (w_1, \ldots, w_n) \in \mathcal{S}_\phi$. Para ver que $\overline{z} \in \mathcal{S}_\eta$ y concluir esta contención, basta evaluar el vector $\overline{z}$ en una ecuación arbitraria de η y verificar que la ecuación se preserva.

$$\begin{aligned} a_{i,1}z_1 + a_{i,2}z_2 + \cdots + a_{i,n}z_n &= a_{i,1}(s_1 + w_1) + a_{i,2}(s_2 + w_2) + \cdots + a_{i,n}(s_n + w_n) \\ &= a_{i,1}s_1 + a_{i,1}w_1 + \cdots + a_{i,n}s_n + a_{i,n}w_n \\ &= (a_{i,1}s_1 + \cdots + a_{i,n}s_n) + (a_{i,1}w_1 + \cdots + a_{i,n}w_n) \\ &= b_i + 0 \\ &= b_i \end{aligned}$$

Con lo anterior, $\overline{z}$ es una solución del sistema η, por lo que $\overline{z} \in \mathcal{S}_\eta$, mostrando la segunda contención.

De los puntos (i) y (ii) se concluye que $\mathcal{S}_\eta = \{\overline{x} \in \mathbb{R}^n : \overline{x} = \overline{s}_0 + \overline{s}$ donde $\overline{s} \in \mathcal{S}_\phi\}$. □

Ejemplo. Consideremos el siguiente sistema no homogéneo.

$$\eta = \begin{cases} x_1 + x_3 - 2x_5 = -1 \\ x_2 + x_5 = 2 \\ 2x_1 + 2x_3 + x_4 - x_6 = 2 \\ x_1 + x_2 + x_3 + x_4 - x_6 = 2 \\ x_1 + x_2 + x_3 + 2x_4 + x_5 - 2x_6 = 5 \end{cases}$$

i) Consideramos el sistema homogéneo asociado a η:

$$\eta_0 = \begin{cases} x_1 + x_3 - 2x_5 = 0 \\ x_2 + x_5 = 0 \\ 2x_1 + 2x_3 + x_4 - x_6 = 0 \\ x_1 + x_2 + x_3 + x_4 - x_6 = 0 \\ x_1 + x_2 + x_3 + 2x_4 + x_5 - 2x_6 = 0 \end{cases}$$

ii) Encontramos la matriz aumentada de η_0, digamos $\mathcal{A}_{\eta_0}$ y la llevamos a su matriz escalonada reducida, digamos $\mathcal{A}_{\phi_0}$.

$$\mathcal{A}_{\eta_0} = \begin{pmatrix} 1 & 0 & 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 2 \\ 2 & 0 & 2 & 1 & 0 & -1 & 2 \\ 1 & 1 & 1 & 1 & 0 & -6 & 2 \\ 1 & 1 & 1 & 2 & 1 & -2 & 5 \end{pmatrix}$$