INSTITUTO DE FÍSICA
UNIVERSIDAD DE ANTIOQUIA

Teoría de Categorías Aplicada a los Algoritmos de Aprendizaje Supervisados

Trabajo presentado como requisito para optar por el titulo de Físico

Luis Abelardo Papiernik Caicedo

Asesores
Alejandro Sanchez Yali
Ricardo Rueda Robayo

12 de abril de 2023

Resumen

En este trabajo se estudian conceptos de categorías y de categorías monoidales, también se estudian los diagramas de cuerdas que permiten manipular pictóricamente relaciones algebraicas en una categoría monoidal y se termina con la presentación de la aplicación de la teoría de categorías al estudio de los algoritmos de aprendizaje supervisado.

Este trabajo esta dedicado a todos los que me apoyaron, gracias.

Índice general

Introducción

El uso de técnicas de machine learning, en particular el uso de las redes neuronales ha crecido exponencialmente debido a su efectividad en la resolución de problemas reales, su gran efectividad se debe en parte al algoritmo de backpropagation, que permite que el proceso de aprendizaje para una red neuronal sea rápido y efectivo. Debido a este crecimiento exponencial en su uso, se ha hecho difícil manejar su complejidad y entender como áreas de este inmenso campo investigación interactúan entre si, porque en general diferentes áreas del machine learning no comparten un mismo marco conceptual y es difícil usar conceptos que sirvan para todos los contextos, dado que cada uno tiene sus propios significados teóricos, esto es, fundamentos, mejores practicas, medidas de evaluación y a menudo lenguajes y formas de pensar diferente sobre el area en su conjunto.

En contraste, se tiene a la teoría de categorías que nació en la década de 1940, y cuya principal característica es que sus principios de organización se han utilizado para remodelar y reformular problemas dentro de las matemáticas puras, incluida la topología, la teoría de la homotopía y la geometría algebraica. La teoría de categorías tiene luz sobre esos problemas, haciéndolos más fáciles de resolver y abriendo puertas para nuevas vías de investigación. Históricamente, la teoría de categorías ha encontrado una aplicación inmensa dentro de las matemáticas.

Sin embargo, más recientemente, la teoría de categorías ha encontrado aplicaciones en una amplia gama de disciplinas fuera de las matemáticas puras, incluso más allá de los campos estrechamente relacionados a la computación y a la física cuántica. Estas disciplinas incluyen química, neurociencia, biología de sistemas, procesamiento de lenguaje natural, causalidad, sistemas dinámicos y teoría de bases de datos, por nombrar algunos. ¿Y qué tienen todos en común? Eso es mucho de lo que la teoría de categorías aplicada actual busca descubrir. En otras palabras, las técnicas, herramientas e ideas de la teoría de categorías se están utilizando para identificar temas recurrentes en estas diversas disciplinas con el fin de hacerlos más formales.

El objetivo de este trabajo es introducir el formalismo de la Teoría de Categorías para luego aplicarlo de manera efectiva en el estudio de redes neuronales, en específico, de las redes neuronales profundas y sus métodos de entrenamiento, para demostrar que los algoritmos de aprendizaje supervisado tienen estructura composicional, lo que permite un estudio mas estructural de los mismos.

Capítulo 1

Construcciones básicas

En este capítulo se definirán los conceptos de categoría, funtor y transformación natural, se verá cómo construir categorías haciendo uso del principio de dualización y categorías producto. A su vez se estudiaran los objetos y los morfismos especiales de una categoría, con particular atención en las categorías Cat y Fun$(\mathscr{C}, \mathscr{D})$ para $\mathscr{C}$ y $\mathscr{D}$ categorías. Por último se verán las propiedades y las construcciones universales del producto y el coproducto.

1.1. Fundamentos de Conjuntos

se define un **universo**, como un conjunto U con las siguientes propiedades

(i) Si $x \in u$ con $u \in U$, entonces $x \in U$.

(ii) Si $u \in U$ y $v \in U$, entonces $\{u, v\} \in U$, $(u, v) \in U$ y $u \times v \in U$.

(iii) Si $x \in U$, entonces $P(x) \in U$ y $\bigcup x \in U$. con $P(x) = \{v|v \in x\}$ y $\bigcup x = \{y|y \in z \text{ para algún } z \in x\}$.

(iv) $\omega \in U$. Donde $\omega = \{0, 1, 2, \cdots\}$ es el conjunto de todos los ordinales finitos.

(v) Si $f\colon a \to b$ es una función sobreyectiva con $a \in U$ y $b \subset U$, entonces $b \in U$.

estas propiedades aseguran que cualquier operación estándar de conjuntos aplicada sobre elementos de U, siempre producirán elementos de U. En todo este trabajo se usaran 3 universos $U_1 \in U_2 \in U_3$ y se dirá que s es pequeño si $s \in U_1$, b es grande si $b \in U_2$ y que r es muy grande si $r \in U_3$, lo que implica que se podrá manipular conjuntos sin riesgo a caer en contradicciones o en paradojas. Para un tratamiento mas riguroso véase [Mac10, Capitulo I sección 6].

1.2. Categorías

Definición 1.2.1. Una **categoría** $\mathscr{C}$ consta de:

1. Una colección de **objetos** denotado por $\mathrm{Ob}(\mathscr{C})$.
2. Para todo par de objetos $X, Y \in \mathrm{Ob}(\mathscr{C})$, una colección de **morfismos** $\mathrm{Hom}_{\mathscr{C}}(X, Y)$.
3. Para todo $X, Y, Z \in \mathrm{Ob}(\mathscr{C})$ una operación **composición**

$$\circ \colon \mathrm{Hom}_{\mathscr{C}}(Y, Z) \times \mathrm{Hom}_{\mathscr{C}}(X, Y) \to \mathrm{Hom}_{\mathscr{C}}(X, Z),$$

tal que se satisfacen las siguientes propiedades:

(i) **Identidad**: Para todo objeto $X \in \mathrm{Ob}(\mathscr{C})$, existe un morfismo $id_X \in \mathrm{Hom}_{\mathscr{C}}(X, X)$ tal que para todo objeto $Y \in \mathrm{Ob}(\mathscr{C})$ y morfismos $f \in \mathrm{Hom}_{\mathscr{C}}(X, Y)$ y $g \in \mathrm{Hom}_{\mathscr{C}}(Y, X)$, se tiene que $f \circ id_X = f$ y $id_X \circ g = g$.

(ii) **Asociatividad**: Para todo $X, Y, Z, W \in \mathrm{Ob}(\mathscr{C})$ y morfismos $f \in \mathrm{Hom}_{\mathscr{C}}(X, Y)$, $g \in \mathrm{Hom}_{\mathscr{C}}(Y, Z)$ y $h \in \mathrm{Hom}_{\mathscr{C}}(Z, W)$, se tiene que $(h \circ g) \circ f = h \circ (g \circ f)$.

Una categoría $\mathscr{C}$ es llamada **localmente pequeña**, si para cada par de objetos $X, Y \in \mathrm{Ob}(\mathscr{C})$, se tiene que $\mathrm{Hom}_{\mathscr{C}}(X, Y)$ es un conjunto pequeño y $\mathscr{C}$ es llamada **categoría pequeña**, si además $\mathrm{Ob}(\mathscr{C})$ es un conjunto pequeño. En otros casos $\mathscr{C}$ es llamada categoría grande, es decir, cuando $Ob(\mathscr{C})$ es un conjunto grande y $\mathrm{Hom}_{\mathscr{C}}(X, Y)$ para todo $X, Y \in \mathrm{Ob}(\mathscr{C})$ es un conjunto grande; en este trabajo todas las categorías se consideraran grandes a menos que explícitamente se diga lo contrario.

Notación 1.2.2. Si no hay lugar a confusión se escribirá:

1. $X \in \mathscr{C}$ para indicar que X esta en $\mathrm{Ob}(\mathscr{C})$ y en algunas ocasiones se escribirá también «X es un objeto de $\mathscr{C}$».
2. $f \colon X \to Y$ para decir que $f \in \mathrm{Hom}_{\mathscr{C}}(X, Y)$, estos morfismos también son llamados flechas y en algunas ocasiones se escribirá «$f \colon X \to Y$ es un morfismo (o flecha) de $\mathscr{C}$».
3. $\mathrm{Hom}(X, Y)$ en vez de $\mathrm{Hom}_{\mathscr{C}}(X, Y)$.

Dados tres morfismos $f\colon A \to B$, $g\colon B \to C$ y $h\colon A \to C$ se pueden representar mediante el siguiente diagrama

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

y se dice que el diagrama es conmutativo si $h = g \circ f$.

En general, los objetos de $\mathscr{C}$ no son conjuntos y por lo tanto $\mathrm{Hom}(X, Y)$ para dos objetos $X, Y \in \mathscr{C}$ no son funciones entre conjuntos, como se verá más adelante.

Ejemplo 1.2.3. La **categoría de conjuntos**, denotada Set, es aquella que tiene como objetos conjuntos pequeños y para cada par de objetos $X, Y \in \mathrm{Set}$ se toma a $\mathrm{Hom}_{\mathrm{Set}}(X, Y)$ como el conjunto de funciones de X en Y, donde la composición se toma como la composición usual de funciones. Con estas consideraciones claramente Set es una categoría, dado que se cumple la propiedad (i) de la Definición 1.2.1, pues id_X es la función identidad definida por la asignación $x \mapsto x$ para todo $X \in \mathrm{Set}$ y la propiedad (ii) de la Definición 1.2.1, se sigue de la composición de funciones entre conjuntos.

Observacion 1.2.4. Notese que en el ejemplo anterior los únicos objetos de la categoría son los conjuntos pequeños. Se toman solo conjuntos de este tipo, porque de lo contrario Ob(Set) seria un conjunto muy grande, y por tanto se tendría una categoría muy grande la Definición 1.2.1

Continuando con esta clase de ejemplos, se pueden considerar los siguientes.

Ejemplos 1.2.5. En las siguientes construcciones la composición esta definida como la composición usual de funciones, y está bien definida, es decir, la composición de homomorfismos de grupos es homomorfismo, la composición de funciones monótonas para Poset es monótona, entre otras.

1. La categoría Grp que denota a la **categoría de grupos**, toma como objetos a los grupos, y para objetos $G, H \in \mathrm{Grp}$, el conjunto $\mathrm{Hom}_{\mathrm{Grp}}(G, H)$ consta de los homomorfismos de grupos de G en H.

2. La categoría Mon, llamada la **categoría de monoides** tiene como objetos a los monoides y como morfismos entre objetos a los homomorfismos de monoides.

3. Dado un anillo R, la **categoría de módulos** denotada por $R-$mod tiene $R-$modulos como objetos y homomorfismos entre $R-$modulos como morfismos.

4. La categoría Top que denota a la **categoría de espacios topológicos** tiene como objetos espacios topológicos y como morfismos entre objetos a las funciones continuas entre espacios topológicos.

Las categorías cuyos objetos son conjuntos y cuyos morfismos son funciones entre conjuntos son llamadas categorías **concretas**, por ejmplo las categorias definidas anteriormente. En muchas ocasiones resulta útil dado un objeto X de una categoría $\mathscr{C}$ definir una categoría asociada a X, lo que permite visualizar a una categoría como un conjunto de operaciones que se pueden componer, más aún, como una colección de objetos que se relacionan de forma consistente.

Ejemplo 1.2.6. Sea N el conjunto de los números naturales más el cero, esto es, $N = \mathbb{N}\cup\{0\}$. Se define la categoría Nat, como la categoría cuyos objetos son elementos de N y para dos números naturales m y n existe un único morfismo $m \to n$ si n divide a m. La anterior es en realidad una categoría, justificada por la observación anterior, pues obsérvese que si $n \mid m$ entonces necesariamente $n \leq m$, es decir, la condición de divisibilidad define en realidad un orden parcial.

Continuando con ejemplos relacionados al álgebra lineal considere los siguientes

Ejemplo 1.2.7. Sea $\mathbb{K}$ un campo. Se define la categoría $\mathrm{Mat}_{\mathbb{K}}$, como la categoría que tiene como objetos a los números enteros positivos, es decir, $\mathrm{Ob}(\mathrm{Mat}) = \mathbb{N} \cup \{0\}$ y como morfismos tiene a las matrices con entradas en $\mathbb{K}$, específicamente, para objetos m y n una matriz de m filas y n columnas, es un morfismo $n \to m$ y la composición esta dada por la multiplicación.

El ejemplo anterior define en realidad una categoría, pues para un objeto n se define el morfismo identidad como la matriz identidad (que existe) de n filas y n columnas y las propiedades (i) y (ii) en la Definición 1.2.1 se siguen de la asociatividad y reglas de multiplicación de las matrices. Notese además que se tiene el objeto 0, y por tanto, se toman también matrices sin dimensión (o vacías) como morfismos.

Ejemplo 1.2.8. Sea R un conjunto y sea $\sim$ una relación de equivalencia sobre R, la categoría asociada a R toma como objetos los elementos de R y para objetos $a, b \in R$ hay un único morfismo $a \to b$ si $a \sim b$, es decir,

$$\mathrm{Hom}_R(a,b) = \begin{cases} \{\to\}, & \text{si } a \sim b \\ \emptyset, & \text{en otro caso} \end{cases}$$

La composición esta determinada de manera única por la transitividad de la relación $\sim$ y por la unicidad de los morfismos entre dos objetos, además, de forma análoga las propiedades necesarias para que se tenga una categoría se siguen de la unicidad de los morfismos y de la reflexividad de la relación $\sim$, esto ultimo asegura la existencia del morfismo identidad para todo objeto en R y el cumplimiento de la asociatividad en la composición de morfismos, en efecto, al tener morfismos $f\colon a \to b$, $g\colon b \to c$ y $h\colon c \to d$, asociando a izquierda se obtiene que $(h \circ g) \circ f\colon a \to d$ y al asociar por derecha $h \circ (g \circ f)\colon a \to d$, pero por definicion existe solo un morfismo $a \to d$, por tanto, $(h \circ g) \circ f = h \circ (g \circ f)$

En el ejemplo anterior solo se usaron las propiedades de reflexividad y transitividad de la relación $\sim$, es decir, si R es un conjunto ordenado o parcialmente ordenado (se tiene una relación de orden o un preorden respectivamente) también se puede definir la categoría asociada a R de forma análoga.

Ejemplo 1.2.9. Sea G un grupo. La **categoría asociada** a G, denotada por BG, consta de un único objeto denotado por la operación del grupo $*$ y por cada elemento $g \in G$ se define un morfismo $g\colon * \to *$ donde la composición viene determinada por la operación del grupo, para la que se escoge la convención $g \circ h = g * h$ con $g, h \in G$. BG es una categoría, pues las propiedades (i) y (ii) de la Definición 1.2.1 se siguen de las propiedades de grupo. Notese además que $\mathrm{Hom}(*, *) = G$.

Un **grupoide** es una categoría pequeña en donde todo morfismo es un isomorfismo. Se dice que los grupoides son la generalización de los grupos, y esto es debido al siguiente hecho, tener un grupoide G de un solo objeto, es en particular tener un grupo G, consecuencia que se sigue de el Ejemplo 1.2.9.

Ejemplo 1.2.10. Sea $\mathbb{K}$ un campo. La **categoría de espacios de Hilbert**, denotada por $\mathrm{Hilb}_{\mathbb{K}}$, es aquella que tiene como objetos espacios de Hilbert sobre $\mathbb{K}$ y dados dos objetos $H, K \in \mathrm{Ob}(\mathrm{Hilb}_{\mathbb{K}})$, $\mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(H, K)$ consta de todas las transformaciones lineales acotadas (o continuas) entre H y K. Con estas consideraciones, $\mathrm{Hilb}_{\mathbb{K}}$ es efectivamente una categoría, ya que la función identidad de un espacio de Hilbert H es una transformación lineal continua y la composición de transformaciones lineales acotadas es también acotada.

Definición 1.2.11. Una categoría $\mathscr{D}$ es una **subcategoría** de una categoría $\mathscr{C}$ si:

1. Los objetos de $\mathscr{D}$ son también objetos de $\mathscr{C}$.
2. Para todo par de objetos $B, B' \in \mathscr{D}$, $\mathrm{Hom}_{\mathscr{D}}(B, B') \subseteq \mathrm{Hom}_{\mathscr{C}}(B, B')$.

3. La composición y los morfismos identidad de $\mathscr{D}$ son los mismos que en $\mathscr{C}$.

Ejemplos 1.2.12. Tome en consideracion los siguientes ejemplos:

1. La categoría $\mathscr{D}$ que consta de:

   a) La colección de todos los conjuntos de un solo elemento.

   b) Para dos objetos $\{x\}, \{y\} \in \mathscr{D}$, $\mathrm{Hom}_{\mathscr{D}}(\{x\}, \{y\}) = \mathrm{Hom}_{\mathrm{Set}}(\{x\}, \{y\})$

   c) La composición es la composición usual de funciones entre conjuntos.

   Es claro que $\mathscr{D}$ definida como en el ejemplo anterior es en realidad una categoría, más aun, $\mathscr{D}$ es una subcategoría de la categoría Set por definición.

2. Todo espacio de Hilbert es en particular un espacio vectorial, de modo que la categoría Hilb_K es una subcategoría de la categoría Vect_K.

   Al estudiar a los espacios de Hilbert desde el contexto de las categorías, puede llegar a ser inconveniente usar el hecho de que Hilb_K es una subcategoría de Vect_K pues para un objeto $H \in \mathrm{Hilb}_K$ se pierde información que permite definir y medir distancias para llegar a conclusiones geométricas importantes y por tanto puede ser más adecuado considerar a la categoría Hilb_K como una subcategoría de la categoría Ban, que se define como sigue.

3. Sea $\mathbb{K}$ un campo. La **categoría de espacios de Banach**, denotada por Ban_K, es aquella que tiene como objetos espacios de Banach sobre $\mathbb{K}$ y dados dos objetos $H, K \in \mathrm{Ban}_K$, $\mathrm{Hom}_{\mathrm{Ban}_{\mathbb{K}}}(H, K)$ consta de todas las transformaciones lineales acotadas (o continuas) entre H y K. Así, $\mathrm{Ban}_{\mathbb{K}}$ es en efecto una categoría, pues la función identidad de un espacio de Banach H es una transformación lineal continua y la composición de transformaciones lineales acotadas es también acotada, en particular se cumple también que Hilb_K es una subcategoria de Ban_K.

1.3. Funtores

Definición 1.3.1. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías. Un **funtor (covariante)** $F\colon \mathscr{C} \to \mathscr{D}$ de $\mathscr{C}$ en $\mathscr{D}$ consta de lo siguiente:

1. Una asignación $\mathrm{Ob}(\mathscr{C}) \to \mathrm{Ob}(\mathscr{D})$, que envía a cada objeto X de $\mathscr{C}$ en un objeto $F(X)$ de $\mathscr{D}$.

2. Para cada par de objetos $X, Y \in \mathscr{C}$ una asignación $\mathrm{Hom}_{\mathscr{C}}(X,Y) \to \mathrm{Hom}_{\mathscr{D}}(F(X), F(Y))$ que envía al morfismo $f\colon X \to Y$ de $\mathscr{C}$ en el morfismo $F(f)\colon F(X) \to F(Y)$ de $\mathscr{D}$.

Tal que se satisface:

(i) Para cada objeto X de $\mathscr{C}$, se cumple que $F(id_X) = id_{F(X)}$.

(ii) Para cada par de morfismos $f\colon X \to Y$ y $g\colon Y \to Z$ en $\mathscr{C}$, se tiene que $F(g \circ f) = F(g) \circ F(f)$.

Así, si se toma un morfismo $f\colon X \to Y$ que es transformado mediante el funtor a $F(f)\colon F(X) \to F(Y)$, gráficamente se tiene lo siguiente

$$\begin{tikzcd} X \arrow[r] \arrow[d, "f"'] & F(X) \arrow[d, "F(f)"] \\ Y \arrow[r] & F(Y) \end{tikzcd}$$

Observacion 1.3.2. En teoría de categorías es muy común cambiar el sentido de los morfismos, proceso que se puede formalizar por medio de los funtores, es decir, a cada morfismo $f\colon X \to Y$ de $\mathscr{C}$ se asigna un morfismo $F(f)\colon F(Y) \to F(X)$ de $\mathscr{D}$ (y por tanto se hace necesario cambiar el orden de la composición). Proceso que se define formalmente como sigue.

Definición 1.3.3. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías. Un **funtor (contravariante)** F de $\mathscr{C}$ en $\mathscr{D}$ consta de lo siguiente:

1. Una asignación $\mathrm{Ob}(\mathscr{C}) \to \mathrm{Ob}(\mathscr{D})$, que envía a cada objeto X de $\mathscr{C}$ en un objeto $F(X)$ de $\mathscr{D}$.

2. Para cada par de objetos $X, Y \in \mathscr{C}$ una asignación $\mathrm{Hom}_{\mathscr{C}}(X,Y) \to \mathrm{Hom}_{\mathscr{D}}(F(Y), F(X))$ que envía al morfismo $f\colon X \to Y$ de $\mathscr{C}$ en el morfismo $F(f)\colon F(Y) \to F(X)$ de $\mathscr{D}$.

Tal que se satisface:

(i) Para cada objeto X de $\mathscr{C}$, se cumple que $F(id_X) = id_{F(X)}$.

(ii) Para cada par de morfismos $f\colon X \to Y$ y $g\colon Y \to Z$ en $\mathscr{C}$, se tiene que $F(g \circ f) = F(f) \circ F(g)$.

Así, si se toma un morfismo $f\colon X \to Y$ que es transformado mediante el funtor a $F(f)\colon F(Y) \to F(X)$, gráficamente se tiene lo siguiente

$$\begin{tikzcd} X \arrow[r] \arrow[d, "f"'] & F(Y) \arrow[d, "F(f)"] \\ Y \arrow[r] & F(X) \end{tikzcd}$$

Definición 1.3.4. Se define la categoría Cat como la categoría cuyos objetos son categorías pequeñas y cuyos morfismos son funtores.

En la definición anterior para la categoría Cat la composición entre los funtores es como se espera, es decir, si se tienen dos funtores $F\colon \mathscr{C} \to \mathscr{D}$ y $G\colon \mathscr{D} \to \mathscr{E}$, el funtor $H = G \circ F$ asigna a cada objeto $C \in \mathscr{C}$ al objeto $G(F(C)) \in \mathscr{E}$ y a cada morfismo $f\colon X \to Y$ al morfismo $G(F(f))\colon G(F(X)) \to G(F(Y)) \in \mathrm{Hom}_{\mathscr{E}}(G(F(X)), G(F(Y)))$. La funtorialidad de H esta asegurada porque los funtores son estructuras conformadas por funciones entre conjuntos.

Ejemplo 1.3.5. Una clase de ejemplos básicos de funtores esta conformada por los **funtores olvido**, que como su nombre indica, son funtores que *olvidan* parte de la estructura o propiedades extra de la categoría dominio (en general cuando esta es una categoría concreta).

1. El funtor olvido $\mathrm{Grp} \to \mathrm{Set}$ que mapea un grupo G en el conjunto subyacente G y mapea un homomorfismo de grupos $\phi\colon G \to H$ en la función suyacente $\phi\colon G \to H$.

Ejemplo 1.3.6. Sean G y H dos grupos. Una función $f\colon G \to H$ es un homomorfismo de grupos si y solo si el mapeo inducido $\mathrm{B}G \to \mathrm{B}H$ es un funtor. Para comprobar esto basta con ver que la composibilidad del funtor inducido $\mathrm{B}G \to \mathrm{B}H$, se traduce como la conservación de las operaciones de los grupos en la función f.

Una representación lineal sobre un grupo G se define como un homomorfismo de grupos $\phi\colon G \to GL_{\mathbb{K}}(V)$ donde $\mathbb{K}$ es un campo y $GL_{\mathbb{K}}(V)$ representa el conjunto de automorfismos de V[1]. Sea G un grupo y $\mathbb{K}$ un campo. Analogamente, una **representación lineal** de G se puede ver como un funtor $R\colon \mathrm{B}G \to \mathrm{Vect}_{\mathbb{K}}$.

[1]Es totalmente equivalente a la acción del grupo G sobre el espacio vectorial V en caso de dimensión finita.

La definiciones anteriores son en realidad equivalentes, y ambas se hacen en la forma en la que se hace asegura que una representación lineal funcione como la traducción de un grupo (abstracto) en algo más concreto con lo que se pueda operar (en este caso transformaciones lineales, o en algunas ocasiones específicos matrices), comparando ambas perspectivas se obtiene lo siguiente:

1. Cada objeto de $\mathrm{B}G$ necesita ser mapeado en un objeto de $\mathrm{Vect}_{\mathbb{K}}$, dado que en $\mathrm{B}G$ solo hay un objeto se debe seleccionar un espacio vectorial particular de $\mathrm{Vect}_{\mathbb{K}}$, esto es, $R(*) = V$ para $V \in \mathrm{Vect}_{\mathbb{K}}$. En teoría de representaciones se dice que G actúa sobre el espacio V.

2. Para cada morfismo de $\mathrm{B}G$, es decir, para cada elemento $g \in G$ se tiene un morfismo (transformación lineal) $Rg\colon V \to V$ en $\mathrm{Vect}_{\mathbb{K}}$. Por las propiedades de grupos y por funtorialidad de R se sigue que esos morfismos son en realidad isomorfismos.

3. Por funtorialidad la asignación $g \mapsto Rg$ debe mapear la identidad $e \in G$ en el morfismo identidad $id_V \in \mathrm{Hom}_R(V, V)$.

Otra forma de ver esta equivalencia, es recordando que el conjunto de automorfismo de un espacio vectorial, es en realidad un grupo y este grupo se puede ver como una categoría (ver el Ejemplo 1.2.9), por tanto el Ejemplo 1.3.6 garantiza dicha equivalencia.

Note que si $\mathscr{C}$ es una categoría localmente pequeña y A es un objeto fijo de $\mathscr{C}$ entonces se define el funtor $\mathrm{Hom}(A, -)\colon \mathscr{C} \to Set$ que envia a X en $\mathrm{Hom}(A, X)$ y para un morfismo $f\colon X \to Y$ se tiene

$$\begin{tikzcd} X \arrow[r] \arrow[d, "f"'] & \mathrm{Hom}(A, X) \arrow[d, "\mathrm{Hom}(A,f)"] \\ Y \arrow[r] & \mathrm{Hom}(A, Y) \end{tikzcd}$$

donde $\mathrm{Hom}(A, f)$ es el morfismo cuya asignacion esta caracterizada por el siguiente diagrama conmutativo

$$\begin{tikzcd}
A \arrow[r, "g"] \arrow[dr, "f\circ g"'] & X \arrow[d, "f"] \\
 & Y
\end{tikzcd}$$

es decir, $\mathrm{Hom}(A,f)(g) = f \circ g$. Asi, $\mathrm{Hom}(A, id_X)(g) = id_X \circ g = g$, igualdad que se cumple para todo $X \in \mathscr{C}$. Ahora, si $f\colon X \to Y$, $g\colon Y \to Z$ y $h\colon A \to X$ son morfismos componibles en $\mathscr{C}$, entonces

$$\begin{tikzcd}
A \arrow[r, "h"] \arrow[d, "(g\circ f)\circ h"'] \arrow[dr, "f\circ h"] & \mathrm{Hom}(A,X) \arrow[d, "f"] \\
Z \arrow[r, "g"'] & Y
\end{tikzcd}$$

es decir, $(g \circ f) \circ h = g \circ (f \circ h)$, por lo tanto

$$\begin{aligned}
\mathrm{Hom}_{\mathscr{C}}(A, g\circ f)(h) &= \mathrm{Hom}_{\mathscr{C}}(A,g)(f\circ h)\\
&= \mathrm{Hom}_{\mathscr{C}}(A,g)(\mathrm{Hom}_{\mathscr{C}}(A,f)(h))\\
&= \{\mathrm{Hom}_{\mathscr{C}}(A,g)\circ \mathrm{Hom}_{\mathscr{C}}(A,f)\}(h)
\end{aligned}$$

Esto muestra que $\mathrm{Hom}(A,-)\colon \mathscr{C} \to Set$ es un funtor covariante.

Observacion 1.3.7. De forma análoga, al fijar un objeto B en una categoría $\mathscr{C}$ se puede definir un funtor contravariante denotado por $\mathrm{Hom}_{\mathscr{C}}(\cdot, B)$ de la siguiente forma. A cada objeto $A \in \mathscr{C}$ se le asigna el objeto $\mathrm{Hom}_{\mathscr{C}}(A,B)$ en la categoría Set, y a cada morfismo $f\colon X \to Y$ se le asigna el morfismo $\mathrm{Hom}_{\mathrm{Set}}(f,B)\colon \mathrm{Hom}_{\mathscr{C}}(Y,B) \to \mathrm{Hom}_{\mathscr{C}}(X,B) \in$ Set que a cada morfismo $g\colon Y \to B \in \mathscr{C}$ le asigna el morfismo $g \circ f \in \mathscr{C}$, esto es, $\mathrm{Hom}(f,B)(g) = g \circ f$.

Definición 1.3.8. Sea $\mathscr{C}$ una categoría. La **categoría opuesta** o **categoría dual** de $\mathscr{C}$, denotada por $\mathscr{C}^{op}$, es aquella en la que:

1. $\mathrm{Ob}(\mathscr{C}^{op}) = \mathrm{Ob}(\mathscr{C})$.
2. Para todo par de objetos $A, B \in \mathscr{C}^{op}$ se tiene $\mathrm{Hom}_{\mathscr{C}^{op}}(B,A) = \mathrm{Hom}_{\mathscr{C}}(A,B)$, esto es, un morfismo $f\colon A \to B \in \mathscr{C}$ corresponde a un morfismo $f^{op}\colon B \to A$ en $\mathscr{C}^{op}$, y las identidades de $\mathscr{C}^{op}$ son las mismás que en $\mathscr{C}$.

3. Si $g\colon C \to B$ y $f\colon B \to A$ son morfismos en $\mathscr{C}$, para los morfismos correspondientes $f^{op}\colon A \to B$ y $g^{op}\colon B \to A$ en $\mathscr{D}$ la composición $g^{op} \circ f^{op}\colon A \to C$ en $\mathscr{D}$ esta dada por $g^{op} \circ^{op} f^{op} = (f \circ g)^{op}$.

Es claro que esta definición es totalmente equivalente a la Definición 1.2.1 salvo la inversión de los morfismos y las composiciones, además, notese que un funtor contravariante F de $\mathscr{C}$ en $\mathscr{D}$ es totalmente equivalente a un funtor covariante F de $\mathscr{C}^{op}$ en $\mathscr{D}$.

1.4. Funtores representables

Los funtores representables son una herramienta que permite extraer información de una categoría arbitraria $\mathscr{C}$ mediante la relación que esta tiene con la categoría de conjuntos Set. La importancia de estos funtores no solo radica en la relación anterior, es también importante pues son usados en la construcción (enunciado) del lemma de Yoneda, que es uno de los resultados más importantes de la teoría de categorías.

Definición 1.4.1. Sea $\mathscr{C}$ una categoría localmente pequeña y $A \in \mathscr{C}$. Se define el **funtor representable** por A, denotado $\mathrm{Hom}_{\mathscr{C}}(A, -)\colon \mathscr{C} \to \mathrm{Set}$ como el funtor que

1. A cada objeto $B \in \mathscr{C}$ le asigna el objeto $\mathrm{Hom}_{\mathscr{C}}(A, B) \in \mathrm{Set}$.

2. A un morfismo $f \in \mathrm{Hom}_{\mathscr{C}}(B, C)$ le asigna $\mathrm{Hom}_{\mathscr{C}}(A, f) \in \mathrm{Hom}_{\mathrm{Set}}(\mathrm{Hom}_{\mathscr{C}}(A, B), \mathrm{Hom}_{\mathscr{C}}(A, C))$, donde $\mathrm{Hom}_{\mathscr{C}}(A, f)(g) = f \circ g$ con $g \in \mathrm{Hom}_{\mathscr{C}}(A, B)$.

El objeto A es llamado **objeto representante**. Nótese que este es un funtor covariante y que dado un objeto B, este funtor lo asigna al objeto $\mathrm{Hom}_{\mathscr{C}}(A, B)$, esto es, el funtor brinda información de todas las formás posibles en las que se puede mapear el objeto A en el objeto B, es decir, ayuda a estudiar la información entre todos los objetos de la categoría $\mathscr{C}$ y el objeto representante A (que en general es escogido a conveniencia, por ejemplo, como el objeto del que más se tenga información).

Definición 1.4.2. Sea $\mathscr{C}$ una categoría localmente pequeña y $B \in \mathscr{C}$. Se define el **funtor co-representable** por B, denotado $\mathrm{Hom}_{\mathscr{C}}(\cdot, B)\colon \mathscr{C} \to \mathrm{Set}$ como el funtor que

1. A cada objeto $A \in \mathscr{C}$ le asigna el objeto $\mathrm{Hom}_{\mathscr{C}}(A, B) \in \mathrm{Set}$.

2. A un morfismo $f \in \mathrm{Hom}_{\mathscr{C}}(A, C)$ le asigna $\mathrm{Hom}_{\mathscr{C}}(f, B) \in \mathrm{Hom}_{\mathrm{Set}}(\mathrm{Hom}_{\mathscr{C}}(C, B), \mathrm{Hom}_{\mathscr{C}}(A, B))$, donde $\mathrm{Hom}_{\mathscr{C}}(f, B)(g) = g \circ f$ con $g \in \mathrm{Hom}_{\mathscr{C}}(C, B)$.

Al igual que en el caso anterior, el objeto A es llamado **objeto representante**. Nótese que este es un funtor contravariante (que se puede obtener desde la versión covariante mediante el principio de dualizacion) y que dado un objeto B, este funtor lo asigna al objeto $\mathrm{Hom}_{\mathscr{C}}(B, A)$, esto es, el funtor brinda información de todas las formás posibles en las que se puede mapear el objeto B en el objeto A, es decir, ayuda a estudiar la información entre todos los objetos de la categoría $\mathscr{C}$ y el objeto representante A, en particular, visualizando a A como un espacio (en un sentido general), entonces los mapeos codifican todas las posibles formás en las que se puede "proyectar" un objeto B en A.

Notese que los funtores representables y co-representables ponen en evidencia un hecho importante, y es que en las categorias, la información importante no esta codificada en los objetos, si no, en los morfismos y en la composición.

Ejemplo 1.4.3. El funtor representable $h_k\colon \mathrm{Mat}_{\mathbb{K}} \to \mathrm{Set}$ asociado a $k \in Mat_{\mathbb{K}}$ (de el Ejemplo 1.2.7), asigna a cada objeto $n \in \mathrm{Mat}_{\mathbb{K}}$ el objeto $\mathrm{Hom}_{\mathrm{Mat}_{\mathbb{K}}}(n, k) \in \mathrm{Set}$, es decir, el conjunto de matrices con k columnas. Dado un morfismo $A\colon n \to m$, lo asigna a la función $\mathrm{Hom}_{\mathrm{Mat}_{\mathbb{K}}}(k, n) \to \mathrm{Hom}_{\mathrm{Mat}_{\mathbb{K}}}(k, m)$ que transforma una matriz B de n filas y k columnas en una matriz C de m filas y k columnas, mediante la multiplicación a izquierda, esto es, $AB = C$.

La construcción del espacio dual (espacio doble dual, ...) de un espacio de Hilbert V define una asignación que es funtorial (un funtor representable contravariante), como se vera en los siguientes ejemplos. En realidad estas construcciones se pueden extender de manera canónica de la categoría $\mathrm{Hilb}_{\mathbb{K}}$ a la categoría $\mathrm{Vect}_{\mathbb{K}}$.

Ejemplo 1.4.4. Considerando la categoría $\mathrm{Hilb}_{\mathbb{K}}$ se define el funtor **espacio dual** como el funtor $\mathrm{ED}\colon \mathrm{Hilb}_{\mathbb{K}} \to \mathrm{Vect}_{\mathbb{K}}$ que

1. A cada objeto (espacio de hilbert) $V \in \mathrm{Hilb}_{\mathbb{K}}$ le asigna el objeto $\mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(V, \mathbb{K})$ en la categoría $\mathrm{Vect}_{\mathbb{K}}$, esto es, el conjunto de transformaciones lineales de V en $\mathbb{K}$.

2. A cada morfismo $T\colon V \to W$ en $\mathrm{Hilb}_{\mathbb{K}}$ le asigna el morfismo $\mathrm{ED}(T)\colon \mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(W, \mathbb{K}) \to \mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(V, \mathbb{K})$ definido por la regla,

$$\mathrm{ED}(T)(g) = g \circ T$$

   para g en $\mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(W, \mathbb{K})$.

Observacion 1.4.5. Sobre el ejemplo anterior cabe notar

1. La funtorialidad de ED se sigue de que la composición entre funciones es lineal.

2. El funtor co-representable $\mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(\cdot, \mathbb{K})$ se puede escribir como la composición entre el funtor ED y el funtor olvido $U\colon \mathrm{Vect}_{\mathbb{K}} \to \mathrm{Set}$, así, $\mathrm{Hom}_{\mathrm{Hilb}_{\mathbb{K}}}(\cdot, \mathbb{K}) = U \circ \mathrm{ED}$. Note que como se menciono anteriormente, el objeto representante (en este caso $\mathbb{K}$) tiene en general características importantes, por ejemplo, todo $\mathbb{K}-$espacio vectorial de dimensión n es isomorfo a $\mathbb{K}^n$.

3. Aplicar el funtor de la definición anterior de forma iterativa define el funtor **espacio n-dual**, donde n es el numero de veces que se aplica el funtor. Es de interés el caso en que $n = 2$, pues como se verá más adelantes, muestra que un espacio vectorial V es *canónicamente* isomorfo a su espacio doble dual V^{**} en el caso de espacios vectoriales de dimensión finita.

1.5. Transformaciones Naturales

Definición 1.5.1. Sean $\mathscr{C}$, $\mathscr{D}$ categorías y F, G funtores de $\mathscr{C}$ en $\mathscr{D}$. Una **transformación natural** $\alpha\colon F \Rightarrow G$ de F en G es una familia de morfismos $\alpha = \{\alpha_X : F(X) \to G(X)\}_{X \in \mathscr{C}}$ denominados **componentes de α en X**, tal que para cada morfismo $f\colon X \to X' \in \mathscr{C}$, se debe cumplir que $G(f) \circ \alpha_X = \alpha_{X'} \circ F(f)$, es decir, el siguiente diagrama conmuta en $\mathscr{D}$

$$\begin{tikzcd} F(X) \arrow[r, "F(f)"] \arrow[d, "\alpha_X"'] & F(X') \arrow[d, "\alpha_{X'}"] \\ G(X) \arrow[r, "G(f)"'] & G(X') \end{tikzcd}$$

Definición 1.5.2. Sean $\mathscr{C}$ una categoría pequeña y $\mathscr{D}$ una categoría. Se define la **categoría** de **funtores** denotada por $\mathrm{Fun}(\mathscr{C}, \mathscr{D})$, como la categoría cuyos objetos son funtores $\mathscr{C} \to \mathscr{D}$ y cuyos morfismos son transformaciones naturales entre estos funtores.

Es inmediato verificar que existe una transformación natural identidad y la composición de transformaciones naturales se hace de la forma natural, es decir, teniendo dos transformaciones naturales $\alpha\colon F \Rightarrow G$ y $\beta\colon G \Rightarrow H$, esto es,

$$\begin{tikzcd}
F(X) \arrow[r, "F(f)"] \arrow[d, "\alpha_X"'] & F(Y) \arrow[d, "\alpha_Y"] \\
G(X) \arrow[r, "G(f)"'] & G(Y)
\end{tikzcd}$$

y

$$\begin{tikzcd}
G(X) \arrow[r, "G(f)"] \arrow[d, "\beta_X"'] & G(Y) \arrow[d, "\beta_Y"] \\
H(X) \arrow[r, "H(f)"'] & H(Y)
\end{tikzcd}$$

se define la composición $\beta \circ \alpha \colon F \Rightarrow H$ con componentes a $(\beta \circ \alpha)_C = \beta_C \circ \alpha_C \colon F(C) \xrightarrow{\alpha_C} G(C) \xrightarrow{\beta_C} HC$, que es natural, pues para $f \colon X \to Y$ se cumple que $\alpha_Y \circ F(f) = G(f) \circ \alpha_X$ y que $\beta_Y \circ G(f) = H(f) \circ \beta_x$ y por tanto

$$\begin{aligned}
H(f) \circ \beta_X \circ \alpha_X &= \beta_Y \circ \alpha_Y \circ F(f) \\
&= \beta_Y \circ (\alpha_Y \circ F(f)) \\
&= \beta_Y \circ (G(f) \circ \alpha_X) \\
&= (\beta_Y \circ G(f)) \circ \alpha_X \\
&= (H(f) \circ \beta_X) \circ \alpha_X
\end{aligned}$$

es decir, el siguiente diagrama conmuta

$$\begin{tikzcd}
F(X) \arrow[r, "\alpha_X"] \arrow[d, "F(f)"] & G(X) \arrow[r, "\beta_X"] \arrow[d, "G(f)"] & H(X) \arrow[d, "H(f)"] \\
F(Y) \arrow[r, "\alpha_Y"'] & G(Y) \arrow[r, "\beta_Y"'] & H(Y)
\end{tikzcd}$$

así la composición de transformaciones naturales definida como anteriormente es también natural.

En teoría de representaciones, además de definir representaciones lineales sobre grupos, se definen también mapeos entre dichas representaciones, llamados mapeos equivariantes. Si se tienen dos representaciones lineales V y W entonces una función $f\colon V \to W$ es equivariante si $f(g \cdot v) = g \cdot f(v)$. De forma análoga se puede tomar la siguiente definición.

Definición 1.5.3. Sea G un grupo y $\mathbb{K}$ un campo. Un **mapeo equivariante** entre dos representaciones lineales $R, S\colon \mathrm{B}G \to \mathrm{Vect}_{\mathbb{K}}$ es una transformación natural $\alpha\colon R \Rightarrow S$.

Para ver la equivalencia entre las distintas definiciones usadas para mapeo equivariante, considere lo siguiente:

1. Para las representaciones lineales R y S se debe hacer la elección de los espacios vectoriales sobre los que actúan, así, se fijan los objetos V y W de $\mathrm{Vect}_{\mathbb{K}}$ respectivamente. Luego, por la Definición 1.5.1, se sabe que por cada objeto de $\mathrm{B}G$ debe haber una componente de la transformación natural α, es decir, se tiene que $\alpha_*\colon V \to W$, pero $\mathrm{B}G$ solo tiene un objeto, por tanto sin perdida de generalidad se puede escribir $\alpha\colon V \to W$.

2. Por ser α una transformation natural, para cada morfismo de $\mathrm{B}G$, es decir, por cada $g \in G$, el siguiente diagrama debe conmutar

$$\begin{tikzcd} V \arrow[r, "R(g)"] \arrow[d, "\alpha"] & V \arrow[d, "\alpha"] \\ W \arrow[r, "S(g)"'] & W \end{tikzcd}$$

   o de forma equivalente se debe cumplir la siguiente igualdad $\alpha \circ R(g) = S(g) \circ \alpha$, en particular, si para todo $v \in V$, $w \in W$ y para todo $g \in \mathrm{B}G$ se denota por $g\cdot$ a la acción de g sobre un elemento de v (y de la misma forma a la acción de g sobre un elemento de W), se obtiene la condición $\alpha(g \cdot v) = g \cdot \alpha(v)$.

Por tanto se concluye que $\alpha\colon R \Rightarrow S$ es una transformación natural si y solo si es un mapeo equivariante entre R y S.

Ejemplo 1.5.4. Fijando un campo $\mathbb{K}$, en el Ejemplo 1.4.4 se construyeron los funtores espacio n-dual, y se afirmo que todo espacio vectorial V de dimensión finita es canónicamente isomorfo a su espacio doble dual V^{**}, es decir, al conjunto de funcionales lineales de la forma $V^* \to \mathbb{K}$. Ahora se vera que en verdad el funtor doble dual es naturalmente isomorfo al funtor identidad. Como punto de partida se tiene lo siguiente,

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d] & W \arrow[d] \\ V^{**} \arrow[r, "f^{**}"'] & W^{**} \end{tikzcd}$$

es decir, falta determinar las componentes de la transformación natural. Para la definición de dichas componentes notese que cada $v \in V$ define un elemento en V^{**} mediante evaluación como sigue: Un elemento en V^{**} es un funcional lineal $V^* \to \mathbb{K}$ y elementos $\omega \in V^*$ se pueden mapear a $\mathbb{K}$ mediante la evaluación de $v \in V$ en ω, así, $\omega \mapsto \omega(v)$. Por tanto, fijando $v \in V$ la función $f_v \colon V^* \to \mathbb{K}$ definida por la asignación $f_v(\omega) = \omega(v)$ ($\omega \in V^*$) es un elemento de V^{**}. Ahora nombrando a la transformación natural con η, una propuesta valida es tomar las componentes $\eta_V \colon V \to V^{**}$ definidas por la asignación $\eta_V(v) = f_v$. Solo resta probar entonces que el siguiente diagrama es conmutativo,

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\eta_V"'] & W \arrow[d, "\eta_W"] \\ V^{**} \arrow[r, "f^{**}"'] & W^{**} \end{tikzcd}$$

es decir, que para todo $v \in V$, $f^{**}(\eta_V(v))$ es igual a $\eta_W(f(v))$, estos son elementos de W^{**}, por tanto, la igualdad se da si y solo si $f^{**}(\eta_V(v))(\omega) = \eta_W(f(v))(\omega)$ para todo $\omega \in W^*$. Por definición de η:

$$\eta_W(f(v))(\omega) = \omega(f(v))$$

Ahora, del otro lado de la igualdad usando la construcción de f^{**} (de el Ejemplo 1.4.4), la definición de η y la construcción de f^* (de el Ejemplo 1.4.4) respectivamente, se obtiene la sucesión de igualdades.

$$f^{**}(\eta_V(v))(\omega) = \eta_V(v)(f^*(\omega)) = (f^*(\omega))(v) = \omega(f(v))$$

Por tanto η es una transformación natural.

1.6. Monomorfismos, Epimorfismos e Isomorfismos

Definición 1.6.1. Un morfismo $f\colon B \to C$ en una categoría $\mathscr{C}$ es un **monomorfismo** (o es **mono**), si para cada par de morfismos $g\colon A \to B$ y $h\colon A \to B$ en $\mathscr{C}$, la igualdad $f \circ g = f \circ h$ implica que $g = h$.

Aplicando el principio de dualización a la definición anterior, se obtiene lo siguiente.

Definición 1.6.2. Un morfismo $f\colon A \to B$ en una categoría $\mathscr{C}$ es un **epimorfismo** (o es **epi**), si para cada par de morfismos $g\colon B \to C$ y $h\colon B \to C$ en $\mathscr{C}$, la igualdad $g \circ f = h \circ f$ implica $g = h$

En la categoría Set los monomorfismos corresponden a las funciones inyectivas, y los epimorfismos (que son la noción dual de los monomorfismos) corresponden a las funciones sobreyectivas, como lo muestra la siguiente proposición.

Proposición 1.6.3. En la categoría de conjuntos un morfismo $f\colon A \to B$ es un monomorfismo, si y solo si la función subyacente es inyectiva.

Demostración. Primero se supondrá un monomorfismo $f\colon A \to B \in \mathscr{C}$ y se vera que es una función inyectiva. Suponga $a, b \in A$ tal que $a \neq b$, se debe mostrar que $f(a) \neq f(b)$, para esto considere $g, f\colon \{*\} \to A$ definidos por las asignaciones $* \mapsto a$ y $* \mapsto b$ respectivamente, así, por hipótesis $g \neq h$ de donde $f \circ g \neq f \circ h$, es decir, $f(a) \neq f(b)$. Por otro lado, veamos que una función inyectiva $f\colon A \to B$ es necesariamente un monomorfismo, así, suponga $f, g\colon C \to A$ tales que $g \neq h$, es decir, existe $C \in \mathscr{C}$ tal que $g(C) \neq h(C)$, luego, por ser f inyectiva $f(g(C)) \neq f(h(C))$, de donde $f \circ g = f \circ h$. □

Y de forma análoga se prueba el siguiente resultado

Proposición 1.6.4. En la categoría de conjuntos un morfismo $f\colon A \to B$ es un epimorfismo, si y solo si la función subyacente es sobreyectiva.

Ejemplo 1.6.5. Como ejemplo considere las monomorfimos y epimorfismos en las siguientes categorías.

1. En la categoría Grp los monomorfimos son los homomorfismos de grupos inyectivos y los epimorfismos son los homomorfismos de grupos sobreyectivos.

2. En la categoría Field los monomorfimos son los homomorfismos de campos inyectivos y los epimorfismos son los homomorfismos de campos sobreyectivos.

3. En la categoría BM asociada a un monoide M los monomorfismos son los elementos de M cancelables a izquierda y los epimorfismos son los elementos de M cancelables a derecha. Por tanto en la categoría BG asociada a un grupo G, todos los morfismos son monomorfismos y epimorfismos (isomorfismo).

4. En la categoría de el Ejemplo 1.2.8 todos los morfismos son monomorfismos y epimorfismos (isomorfismo), esto se sigue de la unicidad de las flechas entre dos objetos dados.

Erróneamente se puede llegar a pensar que cuando se tiene una categoría concreta, los monomorfismos son necesariamente funciones inyectivas y que los epimorfismos son necesariamente funciones sobreyectivas, pero esta correspondencia no se cumple en general, pues se puede usar la estructura adicional de los objetos de forma que se contradiga este hecho. Un ejemplo se obtiene al considerar la categoría Mon, tomando como objetos a el monoide de los números enteros no negativos bajo la suma $(\mathbb{N}, +, 0)$ y al monoide de los números enteros bajo la suma $(\mathbb{Z}, +, 0)$, en este caso el morfismo inclusión $i\colon (\mathbb{N}, +, 0) \to (\mathbb{Z}, +, 0)$ es un epimorfismo, aunque claramente no es una función sobreyectiva.

Para mostrar esto se debe probar que dados los morfismos $h, g\colon (\mathbb{Z}, +, 0) \to (A, *, e)$ tales que $g \circ i = h \circ i$ se cumple que $g = h$. Es claro que la inclusión asegura la igualdad $g|_{\mathbb{N}} = h|_{\mathbb{N}}$, por tanto, para probar que $g = h$ falta mostrar que la igualdad se da también para los valores negativos, para lo que solo se necesita mostrar que $g(-1) = h(-1)$ (dado que $g(-n)$ con $n \in \mathbb{N}$ es $g(-n) = g(-1) + g(-1) \cdots g(-1) = ng(-1)$ y análogamente para h), así

$$\begin{aligned}
g(-1) &= g(-1) * e \\
&= g(-1) * h(0) \\
&= g(-1) * h(1 + (-1)) \\
&= g(-1) * h(1) * h(-1) \\
&= g(-1) * g(1) * h(-1) \\
&= g(0) * h(-1) \\
&= e * h(-1) \\
&= h(-1)
\end{aligned}$$

Por otro lado, en la categoría Cat no es usual tratar con monomorfismos y epimorfismos, es decir, funtores que sean monomorfismos y funtores que sean epimorfismos. En cambio se usan las siguientes caracterizaciones.

Definición 1.6.6. Un funtor $F\colon \mathscr{C} \to \mathscr{D}$ entre dos categorías $\mathscr{C}$ y $\mathscr{D}$ puede ser

1. **Fiel**: Para todo objeto $C, C' \in \mathscr{C}$ y cada par de morfismos $f, f'\colon C \to C'$ de $\mathscr{C}$, se tiene que si $F(f) = F(f')$ en $\mathscr{D}$, entonces $f = f'$ en $\mathscr{C}$.

2. **Pleno**: Para todo objeto $C, C' \in \mathscr{C}$ y cada morfismo $g\colon F(C) \to F(C')$ de $\mathscr{D}$, existe un morfismo $f\colon C \to C'$ de $\mathscr{C}$ tal que $F(f) = g$.

3. **Plenamente fiel**: Es pleno y es fiel.

4. **Esencialmente sobreyectivo**: Para cada objeto $D \in \mathscr{D}$ existe un objeto $C \in \mathscr{C}$ y un isomorfismo $F(C) \to D$ en $\mathscr{D}$.

Observacion 1.6.7. Se sabe que dados dos objetos C, C' en $\mathscr{C}$, el funtor $F\colon \mathscr{C} \to \mathscr{D}$ induce una función

$$\mathrm{Hom}_{\mathscr{C}}(C, C') \to \mathrm{Hom}_{\mathscr{D}}(F(C), F(C'))$$

que permite afirmar lo siguiente: El funtor F es fiel precisamente si la función inducida es inyectiva para todo $C, C' \in \mathscr{C}$, análogamente, el funtor F es pleno si la función inducida es sobreyectiva. Además, notese que el funtor es esencialmente sobreyectivo si es sobreyectivo salvo isomorfismo en objetos.

Definición 1.6.8. Un morfismo $f : A \to B$ en una categoría $\mathscr{C}$ es llamado un **isomorfismo** si existe un morfismo $g : B \to A$ en $\mathscr{C}$, tal que $f \circ g = id_B$ y $g \circ f = id_A$. Dos objetos A, B de $\mathscr{C}$ son llamados **isomorfos** si existe un isomorfismo entre ellos.

En la categoría Set, una función $f\colon A \to B$ se llama biyectiva (isomorfismo) si es inyectiva (monomorfismo) y sobreyectiva (epimorfismo), entonces es natural pensar en que en general en una categoría $\mathscr{C}$ arbitraria, si un morfismo $f\colon A \to B$ es monomorfismo y epimorfismo, entonces se tiene que f es isomorfismo, pero esto es falso, la afirmación correcta es como sigue.

Proposición 1.6.9. Sea $f\colon A \to B$ un isomorfismo en una categoría $\mathscr{C}$, entonces f es monomorfismo y epimorfismo.

Demostración. Sea $f\colon A \to B$ un isomorfismo en una categoría $\mathscr{C}$, por tanto existe un morfismo $f^{-1}\colon B \to A$ tal que $f \circ f^{-1} = id_B$ y $f^{-1} \circ f = id_A$, así, al tener morfismos $g, h\colon D \to A$ tales que $f \circ g = f \circ h$ se sigue inmediatamente que $g = h$, es decir, f es un monomorfismo. De forma análoga, al tener morfismos $g, h\colon A \to D$ tales que $g \circ f = h \circ f$ se sigue inmediatamente que $g = h$, esto es, f es un epimorfismo. $\square$

Al considerar la categoría Cat y fijar las categorías $\mathscr{C}$ y $\mathscr{D}$ un morfismo (un fuctor) $F\colon \mathscr{C} \to \mathscr{D}$ es un isomorfismo si existe otro morfismo $G\colon \mathscr{D} \to \mathscr{C}$ tal que $F \circ G = id_{\mathscr{D}}$ y $G \circ F = id_{\mathscr{C}}$. Es valido afirmar que la condición anterior es muy restrictiva en el contexto de las categorías, dado que la primera igualdad (de forma análoga la segunda igualdad) implica que dado un objeto $D \in \mathscr{D}$ (lo mismo para los morfismos) se debe cumplir $(F \circ G)(D) = F(G(D)) = id_{\mathscr{D}}(D)$, restrictiva en el sentido de que se puede cambiar la igualdad por un isomorfismo (y que la igualdad implica un isomorfismo, pero no al reves). Hablar de un isomorfismo va más de acuerdo con la filosofía en el estudio de las categorías, dado que para verificar si dos estructuras u objetos son iguales, en general se necesita revisar la estructura interna de ambos objetos, cuando se comparan estos mediante la búsqueda de un isomorfismo, en esencia se esta estudiando como se relacionan los objetos entre si en la categoría. Por esta observación, en teoría de categorías no es muy usual hablar de categorías isomorfas si no de categorías equivalentes, que se define como sigue.

Definición 1.6.10. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías. Una **equivalencia de categorías** entre $\mathscr{C}$ y $\mathscr{D}$ consiste de un par de funtores $F\colon \mathscr{C} \to \mathscr{D}$ y $G\colon \mathscr{D} \to \mathscr{C}$ e isomorfismos naturales $\eta\colon G \circ F \to id_{\mathscr{C}}$ y $\epsilon\colon F \circ G \to id_D$. En este caso, se llama a G **pseudo inversa** de F (y viceversa).

En general dado un funtor $F\colon \mathscr{C} \to \mathscr{D}$ es difícil encontrar un funtor $G\colon \mathscr{D} \to \mathscr{C}$ que sea el pseudo inverso del primero. El siguiente es un resultado importante, porque permite encontrar una equivalencia de categorías, solo mediante la inspección de las propiedades de un solo funtor.

Teorema 1.6.11. Un funtor $F\colon \mathscr{C} \to \mathscr{D}$ define una equivalencia de categorías si y solo si es plenamente fiel y esencialmente sobreyectivo.

Por ejemplo, al fijar un campo $\mathbb{K}$ existe una equivalencia entre la categoria de los espacios vectoriales finito dimensionales sobre $\mathbb{K}$ denotado por $\mathrm{Vect}_{\mathbb{K}}^{\leq\infty}$ y la subcategoria de Mat de matrices con entradas en $\mathbb{K}$ denotada por $\mathrm{Mat}_{\mathbb{K}}$. Dicha equivalencia esta justificada por el funtor $F\colon \mathrm{Vect}_{\mathbb{K}}^{\leq\infty} \to \mathrm{Mat}_{\mathbb{K}}$, que a cada espacio vectorial V le asigna el numero $dim_{\mathbb{K}}(V)$ y para cada transformacion lineal $T\colon V \to W$ le asigna la matrix asociada a la transformacion $[T]_{dim_{\mathbb{K}}(W),dim_{\mathbb{K}}(V)}$. El funtor F es plenamente fiel, pues, las transformaciones lineales entre espacios vectoriales sobre un campo $\mathbb{K}$ estan en biyeccion con las matrices con entradas en $\mathbb{K}$. El funtor F es esencialmente sobreyecivo, pues dado cualquier numero n, se tiene que $F(\mathbb{K}^n) = dim(\mathbb{K}^n) = n$.

1.7. Objetos iniciales, objetos Terminales y objeto Cero

Así como hay morfismos especiales, en las categorías también existen objetos que exhiben algunas propiedades especiales, que por ejemplo permiten hacer algunas conexiones directas con la estructura interna de los objetos en una categoría (cuando la categoría es concreta).

Definición 1.7.1. Un objeto $\emptyset$ en una categoría $\mathscr{C}$ es llamado objeto **inicial** si para todo objeto A existe exactamente un morfismo $f\colon \emptyset \to A$.

Ejemplo 1.7.2. En la categoría Set un objeto inicial es el conjunto $\emptyset$. Dado que por vacuidad, para todo conjunto A la función vacía es la única que va del conjunto $\emptyset$ a A.

Ejemplo 1.7.3. Al considerar la categoría de el Ejemplo 1.2.6 se tiene que el único objeto inicial es 0, pues se cumple que para todo $n \in \mathrm{Nat}$, $n \mid 0$, esto es, existe un único morfismo $0 \to n$.

Ejemplo 1.7.4. En la categoría de anillos con unidad un objeto inicial es el anillo de los números enteros $\mathbb{Z}$. En efecto, dado que por los homomorfismos preservar la unidad, cualquier homomorfismo de $\mathbb{Z}$ en un anillo arbitrario R, $f\colon \mathbb{Z} \to R$, queda completamente determinado sabiendo que $f(1_{\mathbb{Z}}) = 1_R$ y dicha asignación define un homomorfismo de anillos.

Definición 1.7.5. Un objeto 1 en una categoría $\mathscr{C}$ es llamado objeto **terminal** si para todo objeto A existe exactamente un morfismo $f\colon A \to 1$.

Ejemplo 1.7.6. Nuevamente considerando la categoría Set cualquier conjunto de un solo elemento es un objeto terminal, pues un morfismo de un conjunto arbitrario A en 1 debe mapear todos los elementos de A al único elemento en 1 y dicho morfismo existe.

Ejemplo 1.7.7. En la categoría de el Ejemplo 1.2.6 1 es el unico objeto terminal, pues se cumple que para todo $n \in \mathrm{Nat}$, $1 \mid n$, esto es, existe un único morfismo $n \to 1$.

No toda categoría tiene un objeto inicial o un objeto terminal y en algunos casos un objeto pueden ser inicial y final, a este se le llama objeto **cero** denotado por 0, como ejemplos considere los siguientes.

Ejemplo 1.7.8. Cuando se considera $(R, \leq)$ como una categoría, un objeto terminal 1 (inicial $\emptyset$ respectivamente) es todo aquel para el que se cumple que $r \leq 1$ para todo $r \in R$ ($0 \leq r$ para todo $r \in R$ respectivamente). Por ejemplo, tomando como caso particular a la categoría $(\mathbb{R}, \leq)$, como $\mathbb{R}$ no tiene elemento máximo ni elemento mínimo, la categoría asociada no tiene ni objetos iniciales ni objetos finales, más aun, no existen tales objetos en cualquier subcategoría de esta cuyo conjunto de objetos forme un intervalo abierto, por la misma razón.

Ejemplo 1.7.9. En la categoría Field no puede existir un objeto terminal 1 (o inicial $\emptyset$) pues no puede existir un homomorfismo de campos entre dos campos con característica diferente. Un hecho diferente seria si se considera la subcategoría Field_p[2] $\subset$ Field, en este caso un objeto inicial es el campo $\mathbb{Z}_p$ por la misma razón de el Ejemplo 1.7.4, esto es, un morfismo de campos $f\colon \mathbb{Z}_p \to \mathbb{K}$ queda completamente determinado por la asignación $f(1_{\mathbb{Z}_p}) = 1_{\mathbb{K}}$.

En el siguiente ejemplo se vera como los objetos terminales sirven para construir un análogo categórico a lo que son los elementos de un conjunto en la categoría Set.

Ejemplo 1.7.10. Sea S un conjunto. A cada elemento $x \in S$ se le puede asociar un morfismo $x\colon 1 \to S$, de tal forma que si se tiene una función $f\colon S \to T$ donde T es un conjunto arbitrario entonces $f(x)$ (evaluar f en x) va a ser el único elemento en T que es la imagen de la composición $f \circ x$.
Así, a todo morfismo de un objeto terminal en un objeto S se le llama elemento global de S o constante de S.

Observacion 1.7.11. Es importante notar que la construcción anterior es funtorial, más aun, la construcción anterior define el funtor representable $Hom_{\text{Set}}(1, \cdot)$ y por tanto se tiene una relación directa con el lema de Yoneda (ver la Sección 1.9), relación que se hará evidente al final del capitulo.

En todos los ejemplos anteriores tanto los objetos iniciales como los objetos terminales son únicos salvo isomorfismo, este hecho no es una coincidencia, en realidad esta unicidad se cumple para cualquier categoría, como se establece en la siguiente proposición.

Proposición 1.7.12. Sea $\mathscr{C}$ una categoría, si existe un objeto inicial $\emptyset$ este es único salvo isomorfismo.

[2]En esta se tienen como objetos campos de característica p con p primo, y los morfismos son homomorfismos de campos

Demostración. Sean $\emptyset$ y $\emptyset'$ dos objetos iniciales en una categoría $\mathscr{C}$, entonces existen morfismos $f\colon \phi \to \phi'$ y $g\colon \phi' \to \phi$, así se tiene el siguiente diagrama

$$\begin{tikzcd}
\emptyset \arrow[r, "f"] \arrow[dr, dotted, "g\circ f"'] & \emptyset' \arrow[d, "g"] \arrow[dr, dotted, "f\circ g"] & \\
& \emptyset \arrow[r, "f"'] & \emptyset'
\end{tikzcd}$$

entonces por unicidad $f \circ g = id_{\emptyset'}$ y $g \circ f = id_{\emptyset}$, esto es, $\emptyset$ y $\emptyset'$ son isomorfos. □

Por principio de dualidad se sigue que los objetos terminales también son únicos salvo isomorfismo.

1.8. Propiedades universales

Muchas definiciones categóricas sobre una categoría $\mathscr{C}$, se dan en términos de objetos terminales de un subconjunto de objetos de la categoría $\mathscr{C}$, que comparten propiedades en común (y en general forman una categoría), junto con morfismos que acompañan a esta clase de objetos. Esto describe un patrón para construcciones universales[3], como ilustración considere la generalización del producto cartesiano en la categoría Set a otras categorías.

Se tomara el caso del producto de dos conjuntos, pero este puede ser extendido fácilmente a n conjuntos o a una colección no finita de conjuntos. Se debe notar entonces que el producto cartesiano $A \times B$ de dos conjuntos A y B esta equipado de forma canónica con dos morfismos $\pi_1\colon A \times B \to A$ y $\pi_2\colon A \times B \to B$ llamados morfismos proyección, y estos están tan estrechamente ligados al producto cartesiano, que lo correcto es considerar al producto cartesiano de A y B como la tupla $(A \times B, \pi_1, \pi_2)$.

Definición 1.8.1. Sea $\mathscr{C}$ una categoría. Un **producto** de dos objetos $A, B \in \mathscr{C}$ es un objeto denotado por $A \times B$ junto con dos morfismos $\pi_1\colon A \times B \to A$ y $\pi_2\colon A \times B \to B$, llamados morfismos proyección, tal que para todo objeto $C \in \mathscr{C}$ y morfismos $f_1\colon C \to A$ y $f_2\colon C \to B$ existe un único morfismo $h\colon C \to A \times B$ tal que se cumple $\pi_1 \circ h = f_1$ y $\pi_2 \circ h = f_2$, esto es, el siguiente diagrama conmuta en $\mathscr{C}$

[3]Poner referencia a los limites.

$$\begin{tikzcd}
 & C \arrow[ld, "f_1"'] \arrow[d, dashed, "\exists! h"] \arrow[rd, "f_2"] & \\
A & A \times B \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B
\end{tikzcd}$$

Notación 1.8.2. Por abuso de notación, a veces se omitirán los morfismos proyección y se escribirá $A \times B$ para denotar al producto de A y B.

De acuerdo a lo dicho anteriormente se vera como en una categoría $\mathscr{C}$ el producto de dos objetos $A, B \in \mathscr{C}$ es un objeto terminal de cierta categoría $\mathscr{D}$. Para esto tome a $\mathscr{D}$ como la categoría tal que

$$\mathrm{Ob}(\mathscr{D}) = \{(C, f_1 \colon C \to A, f_2 \colon C \to B) \mid C \in \mathscr{C}\}$$

y para dos objetos $X_1 = (C, f_1, f_2)$ y $X_2 = (D, g_1, g_2)$ en $\mathscr{D}$ se tiene que,

$$\mathrm{Hom}_{\mathscr{D}}(X_1, X_2) = \{h \colon C \to D \mid f_1 = g \circ h \text{ y } f_2 = g_2 \circ h\}$$

en efecto, por construcción un producto de A y B (objetos de C) es precisamente un objeto terminal de $\mathscr{D}$, más aun, cualquier objeto terminal de $\mathscr{D}$ es un producto de A y B, es decir, el producto de dos objetos es único salvo isomorfismo único. Por construcción, en la categoría Set, el producto de A y B es precisamente el producto cartesiano.

Ejemplo 1.8.3. Considerando algunas de las categorías de ejemplos 1.2.5 se tiene lo siguiente.

1. En la categoría Grp, el producto de dos grupos G y H es precisamente $(G \times H, \pi_1, \pi_2)$, donde $G \times H$ ha sido dotada con la operación definida por la regla $(g, h) *_{G \times H} (g', h') = (g *_G g', h *_H h')$ con $(g, h), (g', h') \in G \times H$.

2. En la categoría $R-$mod el producto de dos objetos M y N esta dado por el producto de $R-$modulos con los homomorfismos proyección.

3. En la categoría Top, el producto de dos espacios topológicos X y Y es $(X \times Y, \pi_1, \pi_2)$ donde el conjunto $X \times Y$ ha sido dotado con la topología producto, y π_1, π_2 corresponden a las funciones proyección.

Ejemplo 1.8.4. Un ejemplo poco convencional del producto, es el producto de la categoría definida en el Ejemplo 1.2.6. Aplicando la construcción para dos números $p, q \in \mathrm{Nat}$, se obtiene el siguiente diagrama

$$\begin{tikzcd} & k \arrow[dl, "f_1"'] \arrow[d, dashed, "\exists! h"] \arrow[dr, "f_2"] & \\ p & d \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & q \end{tikzcd}$$

Es decir, el objeto producto d es divisible por p y q. Y siempre que p y q dividan a un numero k, se tiene que $d \mid k$, más aun, que $d \leq k$. En otras palabras, el producto de dos números p y q resulta ser el mínimo común múltiplo de dichos números.

Ejemplo 1.8.5. Un caso más concreto de producto surge al considerar a las categorias $\mathrm{Vect}_{\mathbb{K}}^{<\infty}$ y a la categoria $\mathrm{Mat}_{\mathbb{K}}$. Aplicando la construccion del producto para dos espacios vectoriales V y W finito dimensionales, se obtiene lo siguiente

$$\begin{tikzcd} & Z \arrow[dl, "T_1"'] \arrow[d, dashed, "\exists! h"] \arrow[dr, "T_2"] & \\ V & V \times W \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & W \end{tikzcd}$$

que mediante el funtor que da la equivalencia entre $\mathrm{Vect}_{\mathbb{K}}^{<\infty}$ y $\mathrm{Mat}_{\mathbb{K}}$ se transforma en

$$\begin{tikzcd} & dim_{\mathbb{K}}(Z) \arrow[dl, "{[T_1]}"'] \arrow[d, dashed, "\exists! {[h]}"] \arrow[dr, "{[T_2]}"] & \\ dim_{\mathbb{K}}(V) & dim_{\mathbb{K}}(V \times W) \arrow[l, "{[A]}"] \arrow[r, "{[B]}"'] & dim_{\mathbb{K}}(W) \end{tikzcd}$$

Recordando que $dim_{\mathbb{K}}(V \times W) = dim_{\mathbb{K}}(V) + dim_{\mathbb{K}}(W)$, A es la matrix de dimensiones $dim_{\mathbb{K}}(V) \times (dim_{\mathbb{K}}(V) + dim_{\mathbb{K}}(W))$ asociada a la transformacion lineal proyeccion en la primera componente y B es la matriz de dimensiones $dim_{\mathbb{K}}(W) \times (dim_{\mathbb{K}}(V) + dim_{\mathbb{K}}(W))$ asociada a la transformacion lineal proyeccion en la segunda componente, esto es,

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{bmatrix} B = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Observacion 1.8.6. Aunque los funtores preservan diagramás conmutativos, en general los productos no se preservan. Notese además, que al fijar objetos Z y $\{X_i\}_{i\in I}$ en $\mathscr{C}$ se tiene el siguiente isomorfismo,

$$\mathrm{Hom}_{\mathscr{C}}(Z, \prod_{i\in I} X_i) \cong \prod_{i\in I} \mathrm{Hom}_{\mathscr{C}}(Z, X_i)$$

lo que se prueba considerando el diagrama asociado

$$\begin{tikzcd} & Z \arrow[dl, "f_j"'] \arrow[d, dashed, "\exists! h"] \\ X_j & \prod_{i\in I} X_i \arrow[l, "\pi_j"] \end{tikzcd}$$

así, se define el mapeo

$$r\colon \mathrm{Hom}_{\mathscr{C}}(Z, \prod_{i\in I} X_i) \to \prod_{i\in I} \mathrm{Hom}_{\mathscr{C}}(Z, X_i)$$

que envia a h en $\prod_{i\in I} \pi_i \circ h$ y el mapeo

$$s\colon \prod_{i\in I} \mathrm{Hom}_{\mathscr{C}}(Z, X_i) \to \mathrm{Hom}_{\mathscr{C}}(Z, \prod_{i\in I} X_i)$$

que envia a $\prod_{i\in I} f_i$ en h, por construcción, se cumple que $s \circ r = id_{\mathrm{Hom}_{\mathscr{C}}(Z, \prod_{i\in I} X_i)}$ y que $r\circ s = id_{\prod_{i\in I} \mathrm{Hom}_{\mathscr{C}}(Z, X_i)}$, es decir, en efecto se tiene el isomorfismo $\mathrm{Hom}_{\mathscr{C}}(Z, \prod_{i\in I} X_i) \cong \prod_{i\in I} \mathrm{Hom}_{\mathscr{C}}(Z, X_i)$.

Como se vio anteriormente, al aplicar el principio de dualización a una construcción se obtiene la construcción dual, en este caso, al aplicar dicho principio al del producto, se obtiene el coproducto. En otras palabras, para una categoría $\mathscr{C}$ se define el **coproducto** de dos objetos $A, B \in \mathscr{C}$ como un producto de dos objetos A, B en la categoría opuesta $\mathscr{C}^{op}$.

Ejemplo 1.8.7. Nuevamente tomando algunas de las categorías de ejemplos 1.2.5, se tiene lo siguiente.

1. Para la categoría Set, el coproducto de dos conjuntos A y B corresponde a la unión disjunta, junto con los morfismos inclusión.

2. Para la categoría Grp el coproducto de dos grupos G y H es el grupo libre formado con G y H. De forma análoga para la categoría de monoides, es decir, dados dos monoides M y N, el coproducto de M y N es el monoide libre formado por M y N.

3. En la categoría Top, el coproducto de dos espacios topológicos X y Y esta definido por la unión disjunta $X \sqcup Y$, donde $X \sqcup Y$ ha sido dotada con la topología disjunta.

Ejemplo 1.8.8. Aplicando el principio de dualización al ejemplo el Ejemplo 1.8.4 se obtiene que el coproducto de dos números p y q en la categoría Nat es el máximo común divisor de p y q.

1.9. Lema de Yoneda

Lema 1.9.1 (Yoneda)**.** Sea $\mathscr{C}$ una categoría localmente pequeña. Sea X un objeto de $\mathscr{C}$ y $F\colon \mathscr{C}^{op} \to Set$ un funtor, considere además el mapeo

$$\mathrm{Hom}_{[\mathscr{C}^{op},Set]}(\mathrm{Hom}_{\mathscr{C}}(-,X),F) \to FX$$

que a cada transformación natural $\alpha\colon \mathrm{Hom}_{\mathscr{C}}(-,X) \Rightarrow F$ le asigna el elemento $\alpha_X(id_X) \in FX$. Se afirma que dicha asignación es una biyección, más aun, que es natural en X y en F.

Demostración. Para demostrar el teorema se deben probar dos cosas, que el mapeo es en realidad biyectivo y que el mapeo es natural en X y en F.

biyectividad. Primero se mostrara que la asignación $\alpha \to \alpha_X(id_X)$ es inyectiva. Así, sean $\alpha, \beta\colon \mathrm{Hom}_{\mathscr{C}}(-,X) \Rightarrow F$ transformaciones naturales y supóngase que $\alpha_X(id_X) = \beta_X(id_X)$, se vera que esto implica $\alpha = \beta$, esto es, para todo objeto $Y \in \mathscr{C}$ las componentes $\alpha_Y, \beta_Y\colon \mathrm{Hom}_{\mathscr{C}}(Y,X) \to FY$ son iguales, así, tómese un objeto $Y \in \mathscr{C}$, para todo morfismo $f \in \mathrm{Hom}_{\mathscr{C}}(Y,X)$ se tiene por naturalidad de α que el siguiente diagrama conmuta

$$\begin{tikzcd}
\mathrm{Hom}_{\mathscr{C}}(X,X) \arrow[r, "\mathrm{Hom}_{\mathscr{C}}(X,f)"] \arrow[d, "\alpha_X"'] & \mathrm{Hom}_{\mathscr{C}}(Y,X) \arrow[d, "\alpha_Y"] \\
F(X) \arrow[r, "F(f)"'] & F(Y)
\end{tikzcd}$$

es decir, si se empieza con la identidad $id_X \in \mathrm{Hom}_{\mathscr{C}}(X,X)$ se tiene la siguiente relación

$$F(f)(\alpha_X(id_X)) = \alpha_Y(id_X \circ f) = \alpha_Y(f) \tag{1.1}$$

y por tanto

$$\alpha_Y(f) = F(f)(\alpha_X(id_X)) = F(f)(\beta_X(id_X)) = \beta_Y(f)$$

esto se cumple para todo todo morfismo f en $\mathscr{C}$ y todo objeto Y en $\mathscr{C}$, por tanto $\alpha = \beta$, es decir, la asignación es inyectiva.

Ahora se vera que la asignación es sobreyectiva, es decir, para todo $p \in F(X)$ debe existir una transformación natural $\alpha\colon \mathrm{Hom}_{\mathscr{C}}(-,X) \Rightarrow F$ tal que $\alpha_X(id_X) = p$. Se definirá dicha transformación, como la transformación que tiene por componentes $\alpha_Y\colon \mathrm{Hom}_{\mathscr{C}}(Y,X) \Rightarrow F(Y)$ que mapea a $f \in \mathrm{Hom}_{\mathscr{C}}(Y,X)$ en $\{F(f)\}(p)$ para todo $Y \in \mathscr{C}$. Se afirma que α es una transformación natural, es decir, para todo morfismo $g\colon Y \to Z$ en $\mathscr{C}$ el siguiente diagrama conmuta

$$\begin{tikzcd}
\mathrm{Hom}_{\mathscr{C}}(Z,X) \arrow[r, "\mathrm{Hom}_{\mathscr{C}}(X,g)"] \arrow[d, "\alpha_Z"'] & \mathrm{Hom}_{\mathscr{C}}(Y,X) \arrow[d, "\alpha_Y"] \\
F(Z) \arrow[r, "F(f)"'] & F(Y)
\end{tikzcd}$$

en otras palabras, al tomar $h \in \mathrm{Hom}_{\mathscr{C}}(Z,X)$ se debe cumplir que $F(g)(\alpha_Z(h)) = \alpha_Y(h \circ g)$. Considerando que F es un funtor contravariante se obtiene la siguiente sucesión de igualdades

$$F(g)(\alpha_Z(h)) = F(g)(F(h)(p)) = F(h \circ g)(p) = \alpha_Y(h \circ g)$$

Por tanto α es una transformación natural, cuya componente en X mapea a id_X en

$$\alpha_X(id_X) = F(id_X)(p) = id_{FX}(p) = p$$

Y por tanto la asignación $\alpha \to \alpha_X(id_X)$ es sobreyectiva, más aun, es una biyección.

naturalidad en F y X. Primero se probara la naturalidad en X. Dado un morfismo $h\colon X \to Y$ de $\mathscr{C}$, este induce una transformación natural $h_*\colon \mathrm{Hom}_{\mathscr{C}}(-,X) \Rightarrow \mathrm{Hom}_{\mathscr{C}}(-,Y)$ definida por la composición de h a izquierda, es decir, para todo objeto A la componente $(h_*)_A\colon \mathrm{Hom}_{\mathscr{C}}(A,X) \Rightarrow \mathrm{Hom}_{\mathscr{C}}(A,Y)$ asigna a morfismos $g \in \mathrm{Hom}_{\mathscr{C}}(A,X)$ en morfismos $h \circ g \in \mathrm{Hom}_{\mathscr{C}}(A,Y)$. Ahora, la naturalidad de X significa que el siguiente diagrama conmuta

$$\begin{tikzcd}
\mathrm{Hom}_{[\mathscr{C}^{op},\mathrm{Set}]}(\mathrm{Hom}_{\mathscr{C}}(-,Y),F) \arrow[r, "-\circ h_*"] \arrow[d, "\cong"'] & \mathrm{Hom}_{[\mathscr{C}^{op},\mathrm{Set}]}(\mathrm{Hom}_{\mathscr{C}}(-,X),F) \arrow[d, "\cong"] \\
F(Y) \arrow[r, "F(h)"'] & F(X)
\end{tikzcd}$$

es decir, para cada transformación natural $\alpha\colon \mathrm{Hom}_{\mathscr{C}}(-,Y) \Rightarrow$ se tiene que $F(h)(\alpha_Y(id_Y))$ es igual a $\alpha_X \circ h_*(id_X)$, lo que es cierto dado que

$$\alpha_X(h_*(id_X)) = \alpha_X(id_X \circ h) = \alpha_X(h)$$

y por la ecuación 1.1 $\alpha_X(h)$ es igual a $F(h)(\alpha_Y(id_Y))$.

Por ultimo, la naturalidad en F significa que para cada funtor $G\colon C^{op} \to \mathrm{Set}$ y cada transformación natural $\beta\colon F \to G$ el siguiente diagrama conmuta

$$\begin{tikzcd}
\mathrm{Hom}_{[\mathscr{C}^{op},\mathrm{Set}]}(\mathrm{Hom}_{\mathscr{C}}(-,X),F) \arrow[r, "\beta\circ -"] \arrow[d, "\cong"'] & \mathrm{Hom}_{[\mathscr{C}^{op},\mathrm{Set}]}(\mathrm{Hom}_{\mathscr{C}}(-,X),G) \arrow[d, "\cong"] \\
F(X) \arrow[r, "\beta_X"'] & G(X)
\end{tikzcd}$$

es decir, para cada transformación natural $\alpha\colon \mathrm{Hom}_{\mathscr{C}}(-,X) \Rightarrow F$ se tiene que $\beta_X(\alpha_X(id_X))$ es igual a $(\beta \circ \alpha)_X(id_X)$. Pero note que esta es precisamente la definición de la composición de transformaciones naturales, por tanto, esta igualdad esta asegurada.

Por lo tanto la asignación $\alpha \to \alpha_X(id_X)$ es una biyección, más aun, es natural en F y en X. $\square$

En muchas ocasiones resulta útil la versión dual del lema de Yoneda, que es como sigue.

Lema 1.9.2 (Dual del lema de Yoneda)**.** Sea $\mathscr{C}$ una categoría localmente pequeña. Sea X un objeto de $\mathscr{C}$ y $F\colon \mathscr{C} \to Set$ un funtor, considere además el mapeo

$$\mathrm{Hom}_{[\mathscr{C},Set]}(\mathrm{Hom}_{\mathscr{C}}(X,-),F) \to FX$$

que a cada transformación natural $\alpha\colon \mathrm{Hom}_{\mathscr{C}}(X,-) \Rightarrow F$ le asigna el elemento $\alpha_X(id_X) \in FX$. Se afirma que dicha asignación es una biyección, más aun, que es natural en X y en F.

La mejor forma de entender los resultados anteriores, es pensar a que pregunta quiere responder. El lema de Yoneda esta dando una caracterización de todas las transformaciones naturales que existe entre un funtor representable $\mathrm{Hom}_{\mathscr{C}}(X,-)$ (co-representable) y un funtor $F\colon \mathscr{C} \to \mathrm{Set}$ covariante (contravariante).

Para considerar un caso concreto, tómese la categoría $\mathrm{Mat}_{\mathbb{K}}$. En el Ejemplo 1.4.3 se vio que un funtor representable h_k es básicamente el conjunto de todas las matrices de k columnas, así, al considerar una transformación natural entre dos funtores representables h_k y h_j se puede pensar en que es una trasnformacion que cambia el numero de columnas en las matrices de forma compatible, en el siguiente sentido, el siguiente diagrama debe conmutar

$$\begin{tikzcd}
h_k(n) \arrow[r, "h_k(A)"] \arrow[d, "\alpha_n"'] & h_k(m) \arrow[d, "\alpha_m"] \\
h_j(n) \arrow[r, "h_j(A)"'] & h_j(m)
\end{tikzcd}$$

De acuerdo a esto es natural entonces preguntarse ¿Cuales son todas las posibles transformaciones naturales $\alpha\colon h_k \Rightarrow h_j$? La versión dual del lemma de Yoneda da la respuesta a esto mediante la biyección

$$\mathrm{Hom}_{[\mathrm{Mat},Set]}(h_k,h_j) \to h_j(k)$$

más específicamente, el lema de Yoneda dice que toda transformación natural $\alpha\colon h_k \Rightarrow h_j$ esta determinada por una única matriz de k filas y j columnas y esa matriz es obtenida mediante la aplicación de la transformación natural a la matriz identidad $(\alpha(I_k))$ y por tanto la operación columna aplicada sobre una matrix C es determinada por la multiplicación a derecha $(\alpha(C) = C\alpha(I_k))$.

Capítulo 2

Categorías Monoidales

En este capitulo se estudiaran las categorías monoidales no unitales, categorías monoidales, functores monoidales laxos y functores monoidales. Este capitulo esta inspirado en [Lur21].

2.1. Categorías monoidales no unitales

Definición 2.1.1. Un **monoide** sin **unidad** es un conjunto M equipado con un mapeo $m\colon M \times M \to M$ que asigna a $(x, y) \in M \times M$ en $xy \in M$, tal que satisface la ley de asociatividad, esto es, para todo $x, y, z \in M$ se cumple que $x(yz) = (xy)z$.

Ejemplos 2.1.2.

1. Cualquier grupo G es en particular un monoide sin unidad, por tanto, dado un grupo G la categoría $\mathrm{B}G$ determina un monoide sin unidad. Más aun, toda categoría con un solo objeto determina un monoide sin unidad.

2. El conjunto de los números naturales $\mathbb{N}$ con la suma usual es un monoide sin unidad.

3. El conjunto de los números enteros $\mathbb{Z}$ con la multiplicacion usual determina un monoide sin unidad.

A continuacion se pretende extender la definición anterior al contexto de las categorías. Este cambio tiene varias implicaciones sobre el resto de la definición, una de ellas es respecto a la función $m\colon M \times M \to M$, pues al tener una categoría $\mathscr{C}$ se debe considerar un funtor $\otimes : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$; la otra implicación es acerca de el cumplimiento de la ley de asociatividad y como se traduce esta al ambiente de las categorías, Una de las posibles extensiones que se puede hacer es la siguiente.

Definición 2.1.3. Sea $\mathscr{C}$ una categoría. Una **estructura monoidal estricta sin unidad** sobre $\mathscr{C}$ es un funtor

$$\otimes\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C} \qquad (X, Y) \mapsto X \otimes Y$$

estrictamente asociativo en el siguiente sentido:

(i) Para todo $X, Y, Z \in \mathscr{C}$, se tiene la *igualdad* $X \otimes (Y \otimes Z) = (X \otimes Y) \otimes Z$.

(ii) Para morfismos $f\colon X \to X'$, $g\colon Y \to Y'$ y $h\colon Z \to Z'$ se tiene la igualdad $f \otimes (g \otimes h) = (f \otimes g) \otimes h$ de morfismos en $\mathscr{C}$ con dominio $X \otimes (Y \otimes Z) = (X \otimes Y) \otimes Z$ y codominio $X' \otimes (Y' \otimes Z') = (X' \otimes Y') \otimes Z'$.

Una **categoría monoidal sin unidad estricta** es un par $(\mathscr{C}, \otimes)$, donde $\mathscr{C}$ es una categoría y $\otimes\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ es una estructura monoidal sin unidad estricta sobre $\mathscr{C}$, pero a veces se abusara de la notación y se escribirá $\mathscr{C}$ para denotar a $(\mathscr{C}, \otimes)$.

La definición anterior implica que el siguiente diagrama conmuta

$$\begin{tikzcd}
X \otimes (Y \otimes Z) \arrow[r, "f\otimes(g\otimes h)"] \arrow[d, "id_{X\otimes(Y\otimes Z)}"'] & X' \otimes (Y' \otimes Z') \arrow[d, "id_{X'\otimes(Y'\otimes Z')}"] \\
(X \otimes Y) \otimes Z \arrow[r, "(f\otimes g)\otimes h"'] & (X' \otimes Y') \otimes Z'
\end{tikzcd}$$

De la definición anterior cabe resaltar que la propiedad (I) es algo restrictiva, pues cuando se quiere probar la igualdad entre dos objetos, por ejemplo de dos conjuntos A y B con propiedades adicionales, se hace necesario operar con la estructura interna de dichos conjuntos o conocer las condiciones que se cumplen. Como se vio en el Capitulo 1 en la teoría de categorías mas que la estructura interna de los objetos, interesa la forma en la que los objetos se relacionan entre si mediante morfismos, es entonces necesario encontrar una manera mas adecuada de formular una *ley de asociatividad*, por ejemplo, una en la que el producto tensorial de espacios vectoriales dote a $\mathrm{Vect}_{\mathbb{K}}$ con estructura monoidal estricta sin unidad, mas específicamente.

Ejemplo 2.1.4. Sea $\mathbb{K}$ un campo y U, V, W espacios vectoriales sobre $\mathbb{K}$. Un mapeo $b\colon U \times V \to W$ es llamado $\mathbb{K}$-bilineal si satisface las siguientes identidades

$$\begin{aligned} b(u+u', v) &= b(u,v) + b(u', v) \\ b(u, v+v') &= b(u,v) + b(u, v') \\ b(\lambda u, v) &= \lambda b(u,v) = b(u, \lambda v) \end{aligned}$$

Para todo $u \in U$, $v \in V$, $w \in W$ y $\lambda \in \mathbb{K}$. Y se dice que un mapeo $\mathbb{K}$-bilineal $b\colon U \times V \to W$ es universal si para cualquier $\mathbb{K}$-espacio vectorial W' y mapeo $\mathbb{K}$-bilineal $t\colon U \times V \to W'$ existe un único morfismo $h\colon W \to W'$ que hace que el siguiente diagrama conmute.

$$\begin{tikzcd} & W' \arrow[dl, "t"'] \arrow[d, dashed, "\exists! h"] \\ U \times V & W \arrow[l, "b"] \end{tikzcd}$$

O en otras palabras, la composición con b induce la biyección

$$\{\text{mapeos } \mathbb{K}\text{-lineales } W \to W'\} \cong \{\text{mapeos } \mathbb{K}\text{-bilineales } U \times V \to W\}$$

Cuando lo anterior se cumple, W esta determinado por U y V salvo isomorfismo y es llamado producto tensorial de U y V denotado por $U \otimes_{\mathbb{K}} V$. Esta construccion induce un funtor $\otimes_{\mathbb{K}}\colon \mathrm{Vect}_{\mathbb{K}} \times \mathrm{Vect}_{\mathbb{K}} \to \mathrm{Vect}_{\mathbb{K}}$ llamado funtor producto tensorial.

Es entonces claro que se necesita una nocion de asociatividad que tenga en cuenta en que los objetos son unicos salvo isomorfismo, dado que la igualdad estricta en realidad esta comparando las imágenes de dos functores, a saber, las imágenes de los functores

$$\mathscr{C} \times \mathscr{C} \times \mathscr{C} \to \mathscr{C} \qquad (X, Y, Z) \mapsto X \otimes (Y \otimes Z)$$

y

$$\mathscr{C} \times \mathscr{C} \times \mathscr{C} \to \mathscr{C} \qquad (X, Y, Z) \mapsto (X \otimes Y) \otimes Z$$

Entonces, una mejor formulación de la ley de asociatividad se puede dar en termino de transformaciones naturales, como sigue.

Definición 2.1.5. Sea $\mathscr{C}$ una categoría. Una **estructura monoidal sin unidad** sobre $\mathscr{C}$ consiste de lo siguiente:

1. Un funtor $\otimes\colon \mathscr{C}\times\mathscr{C}\to\mathscr{C}$ llamado producto tensorial.
2. Un isomorfismo natural $\alpha\colon F\Rightarrow G$ con componentes de la forma $\alpha_{X,Y,Z}\colon (X\otimes Y)\otimes Z\to X\otimes(Y\otimes Z)$ (llamadas restricciones de asociatividad de $\mathscr{C}$) donde $F\colon \mathscr{C}\times\mathscr{C}\times\mathscr{C}\to\mathscr{C}$ mapea (X,Y,Z) en $(X\otimes Y)\otimes Z$ y $G\colon \mathscr{C}\times\mathscr{C}\times\mathscr{C}\to\mathscr{C}$ mapea (X,Y,Z) en $X\otimes(Y\otimes Z)$,

tal que se cumple la **Identidad pentagonal**, es decir, el siguiente diagrama conmuta para todos los objetos y morfismos (compatibles) involucrados:

$$\begin{tikzcd}
 & (W\otimes X)\otimes(Y\otimes Z) \arrow[dr, "\alpha_{W,X,Y\otimes Z}"] & \\
((W\otimes X)\otimes Y)\otimes Z \arrow[ur, "\alpha_{W\otimes X,Y,Z}"] \arrow[d, "\alpha_{W,X,Y}\otimes id_Z"'] & & (W\otimes(X\otimes(Y\otimes Z))) \\
(W\otimes(X\otimes Y))\otimes Z \arrow[rr, "\alpha_{W,X\otimes Y,Z}"'] & & W\otimes((X\otimes Y)\otimes Z) \arrow[u, "id_W\otimes\alpha_{X,Y,Z}"']
\end{tikzcd}$$

Una **categoría monoidal sin unidad** es una tupla $(\mathscr{C},\otimes,\alpha)$, donde $\mathscr{C}$ es una categoría y $(\otimes,\alpha)$ es una estructura monoidal sin unidad sobre $\mathscr{C}$. A veces se abusara de la notación y se escribirá $\mathscr{C}$ para denotar a $(\mathscr{C},\otimes,\alpha)$.

Observacion 2.1.6. En la definición anterior la condición de conmutatividad del diagrama es conocida como condición pentagonal o condición de coherencia y esta garantiza que dos transformaciones naturales (construidas a partir de la condición de asociatividad α) entre dos functores (construidos usando el funtor producto tensorial $\otimes$) son iguales, o en otras palabras, asegura que la asociatividad se cumpla para casos en los que se tengan el producto tensorial de mas de tres objetos.

La Definición 2.1.5 es una generalización de la Definición 2.1.3, pues tomando a las restricciones de asociatividad como los morfismos identidad se recupera el caso de la asociatividad estricta. Al considerar nuevamente el Ejemplo 2.1.4 se nota como La Definición 2.1.5 no encapsula toda la información de dicho ejemplo, a saber, se cumple que $\mathbb{K}\otimes V=V=V\otimes\mathbb{K}$. Es decir, se necesita de una noción de elemento identidad.

2.2. Categorías monoidales

En esta sección se pretende extender las construcciones hechas anteriormente, dotando con un objeto unidad, ciertos morfismos y transformaciones naturales a estas categorías monoidales no unitales, antes se revisara la definición de monoide y las propiedades (o condiciones) que cumple el elemento identidad.

Definición 2.2.1. Un monoide sin unidad M es llamado **monoide con unidad** (o simplemente **monoide**), si existe un elemento $e \in M$ tal que para todo $x \in M$, se cumple que $ex = x = xe$. en este caso, el elemento e es llamado elemento unidad de M.

Ejemplo 2.2.2. Sea $\mathscr{C}$ una categoría y X un objeto de $\mathscr{C}$. Un **endomorfismo** de X es un morfismo de X en X en la categoría $\mathscr{C}$, se denota entonces por $\mathrm{End}_{\mathscr{C}}(X) = \mathrm{Hom}_{\mathscr{C}}(X, X)$ al conjunto de todos los endomorfismos de X. Note que la operación composición de la categoría $\mathscr{C}$ induce un mapeo

$$\mathrm{End}_{\mathscr{C}}(X) \times \mathrm{End}_{\mathscr{C}}(X) \to \mathrm{End}_{\mathscr{C}}(X) \qquad (f, g) \mapsto f \circ g$$

y este mapeo dota a $\mathrm{End}_{\mathscr{C}}(X)$ con estructura de monoide, donde el elemento identidad es id_X.

Al igual que en la sección anterior se tenían varias formas de extender la ley de asociatividad de los monoides no unitales, en esta sección se tienen varias formas de generalizar la noción de elemento unidad de los monoides. Una posible generalizacion es la siguiente.

Definición 2.2.3. Sea $\mathscr{C}$ una categoría. Una **estructura monoidal estricta** sobre $\mathscr{C}$ es una estructura monoidal estricta sin unidad $\otimes\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ para la que existe un objeto $1 \in \mathscr{C}$ que satisface la siguiente propiedad:
Para cada objeto $X \in \mathscr{C}$, se tiene que

$$X \otimes 1 = X = 1 \otimes X$$

mas aun, para cada morfismo $f\colon X \to X'$ de $\mathscr{C}$, se tiene que

$$f \otimes id_1 = f = id_1 \otimes f$$

como un morfismo de X en X'.

Una **categoría monoidal estricta** es un par $(\mathscr{C}, \otimes)$ donde $\mathscr{C}$ es una categoría y $\otimes\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ es una estructura monoidal estricta sobre $\mathscr{C}$.

Al igual que en los casos anteriores, se tiene una construcción muy restrictiva, en el mismo sentido que en el caso previo. Así, en pro a extender la definición del elemento unidad en un monoide (de una forma mas fiel) a un objeto que haga las veces de unidad en una categoría monoidal sin unidad, es necesario el siguiente resultado, donde se caracteriza de forma mas precisa (mediante condiciones) al elemento unidad de un monoide.

Proposición 2.2.4. Sea M un monoide sin unidad. Entonces un elemento $e \in M$ es una unidad si y solo si las siguientes condiciones se satisfacen.

(i) El elemento e es idempotente, esto es, se cumple que $ee = e$.

(ii) El elemento e es cancelable a izquierda, esto es, el mapeo $x \mapsto ex$ es un monomorfismo de M en si mismo.

(iii) El elemento e es cancelable a derecha, esto es, el mapeo $x \mapsto xe$ es un monomorfismo de M en si mismo.

Demostración. Suponiendo que e es una unidad en el monoide M es inmediato concluir que se cumplen las propiedades (I), (II) y (III), en efecto, $ee = e$ y las propiedades (I) y (II) se siguen porque las funciones inducidas son el morfismo identidad, que es monomorfismo. Por otro lado, al suponer (I), (II) y (III) y aplicar (I) a la expresion xe para todo $x \in M$ se obtiene que $xe = xee$ y por ser la funcion asociada monomorfismo se obtiene que $x = xe$, es decir, e actua como unidad a derecha, de forma analoga se concluye que e actua tambien como unidad a izquierda, asi, e es una unidad en M. □

En analogia con este resultado, se puede decir de la unidad en una categoria monoidal $\mathscr{C}$ para cada item anterior, lo siguiente:

1. Dado que la igualdad involucra a un solo elemento de M, al extender esta condición a una categoría monoidal sin unidad $\mathscr{C}$, basta con considerar un objeto 1 que cumpla $1 \otimes 1 \cong 1$.

2. En este caso, la condición involucra un mapeo de M, por tanto al considerar el análogo en la categoría Cat se debe tomar el funtor que mapea a un objeto C (de alguna categoría $\mathscr{C}$) en el objeto $1 \otimes C$ (de la misma categoría $\mathscr{C}$) y se debe exigir que el funtor sea plenamente fiel.

3. En analogía al caso anterior el funtor que mapea al objeto C (de alguna categoría $\mathscr{C}$) en el objeto $C \otimes 1$ (de la misma categoría $\mathscr{C}$) debe ser también plenamente fiel.

De forma mas rigurosa la definicion de la unidad es como sigue.

Definición 2.2.5. Sea $\mathscr{C}$ una categoría monoidal sin unidad. Una unidad de $\mathscr{C}$ es un par $(1, v)$, donde 1 es un objeto de $\mathscr{C}$ y $v\colon 1 \otimes 1 \xrightarrow{\sim} 1$ es un isomorfismo, tal que se satisface lo siguiente: Los functores

$$\mathscr{C} \to \mathscr{C} \qquad C \mapsto 1 \otimes C$$

$$\mathscr{C} \to \mathscr{C} \qquad C \mapsto C \otimes 1$$

son plenamente fiel.

Por la metodología que se esta usando, en la que se esta haciendo una analogía entre los monoides en la categoría Set y las categorías monoidales (que se quieren construir), es natural pensar que así como la unidad en un monoide esta determinado de manera univoca, también lo debería estar la unidad en una categoría monoidal sin unidad, esto es, $(1, v)$ debería estar determinado salvo isomorfismo, y en efecto, según el siguiente resultado.

Lema 2.2.6 (Unicidad de la unidad)**.** Sea $\mathscr{C}$ una categoría monoidal sin unidad equipada con unidades $(1, v)$ y $(1', v')$, entonces existe un isomorfismo único $u\colon 1 \xrightarrow{\sim} 1'$ para el que el siguiente diagrama conmuta.

$$\begin{tikzcd}
1 \otimes 1 \arrow[r, "v"] \arrow[d, "u \otimes u"'] & 1 \arrow[d, "u"] \\
1' \otimes 1' \arrow[r, "v'"'] & 1'
\end{tikzcd}$$

Con las consideraciones anteriores, se puede definir finalmente lo que es una categoría monoidal.

Definición 2.2.7. Sea $\mathscr{C}$ una categoría. Una **estructura monoidal** sobre $\mathscr{C}$ es una estructura monoidal sin unidad $(\otimes, \alpha)$ sobre $\mathscr{C}$ junto con la elección de una unidad $(1, v)$. Una **categoría monoidal** es una categoría $\mathscr{C}$ junto con una estructura monoidal $(\otimes, \alpha, 1, v)$ sobre $\mathscr{C}$. Se hace referencia a 1 como la unidad de $\mathscr{C}$ y a el isomorfismo $v\colon 1 \otimes 1 \to 1$ como la restricción de la unidad de $\mathscr{C}$.

Notación 2.2.8. Algunas veces se abusara de la notación y se escribirá $\mathscr{C}$ para denotar a la categoría monoidal $(\mathscr{C}, \otimes, \alpha, 1, v)$.

Observacion 2.2.9. La definición anterior tiene una forma alterna, en esa se define una **categoría monoidal** como una categoría monoidal sin unidad $\mathscr{C}$ que admite una unidad. Que es única por el Lema 2.2.6

Ejemplo 2.2.10. Sea $\mathscr{C}$ una categoría que admite productos finitos, es decir, para objetos arbitrarios X y Y en $\mathscr{C}$ existe el producto $X \times Y$. Entonces $\mathscr{C}$ puede ser dotado con una estructura monoidal como sigue.

El producto de dos objetos X y Y en $\mathscr{C}$ es en realidad la tupla $(X \times Y, \pi^1_{X,Y} \colon X \times Y \to X, \pi^2_{X,Y} \colon X \times Y \to Y)$. Esto define el funtor producto $\mathscr{C} \times \mathscr{C} \to \mathscr{C}$ que mapea al objeto $(X, Y) \in \mathscr{C} \times \mathscr{C}$ en el objeto $X \times Y \in \mathscr{C}$. Y dados dos morfismos $f \colon X \to Y$ y $g \colon X' \to Y'$ estos son mapeados a $f \times g \colon X \times X' \to Y \times Y'$ donde $f \times g$ es el único morfismo para el que el siguiente diagrama conmuta.

$$\begin{tikzcd}
X \arrow[d, "f"'] & X \times Y \arrow[l, "\pi^1_{X,Y}"'] \arrow[r, "\pi^2_{X,Y}"] \arrow[d, dotted, "f \times g"] & Y \arrow[d, "g"] \\
X' & X' \times Y' \arrow[l, "\pi^1_{X',Y'}"] \arrow[r, "\pi^2_{X',Y'}"'] & X
\end{tikzcd}$$

Las restricciones de asociatividad $\alpha_{X,Y,Z} \colon X \times (Y \times Z) \xrightarrow{\sim} (X \times Y) \times Z$ para todo $X, Y, Z \in \mathscr{C}$ estan garantizadas por la existencia de los productos, mas especificamente, $\alpha_{X,Y,Z}$ es el unico morfismo que hace que el siguiente diagrama conmute.

$$\begin{tikzcd}
& X \times (Y \times Z) \arrow[rr, "\alpha_{X,Y,Z}"] \arrow[ddl, "\pi^1_{X,Y \times Z}"'] \arrow[drr] & & (X \times Y) \times Z \arrow[dll] \arrow[ddr, "\pi^2_{X \times Y, Z}"] & \\
& X \times Y \arrow[dl, "\pi^1_{X,Y}"] \arrow[dr, "\pi^2_{X,Y}"] & & Y \times Z \arrow[dl, "\pi^1_{Y,Z}"'] \arrow[dr, "\pi^2_{Y,Z}"'] & \\
X & & Y & & Z
\end{tikzcd}$$

En este punto se ha dotado a $\mathscr{C}$ con una estructura monoidal sin unidad. Si $\mathscr{C}$ tiene un objeto terminal 1, entonces puede ser extendida a una categoria monoidal, donde la unidad es el objeto terminal 1 y la restriccion de la unidad esta dada por el unico morfismo (isomorfismo) de 1×1 en 1.

Usualmente cuando se define una categoría monoidal no se requiere de la existencia de una unidad y restricción de la unidad (en el sentido de la definición 2.2.5), en cambio se exigen isomorfismos naturales llamados unitor a izquierda y unitor a derecha, esto es, transformaciones naturales con componentes de la forma $\lambda_X \colon 1 \otimes X \xrightarrow{\sim} X$ y $\rho_X \colon X \otimes 1 \xrightarrow{\sim} X$ respectivamente. Pero se vera que ambos resultados son en realidad equivalentes, para esto considere primero la siguiente proposición.

Proposición 2.2.11. Sea $F \colon \mathscr{C} \to \mathscr{D}$ un funtor plenamente fiel. Sean C, C' objetos en $\mathscr{C}$ y sea $\phi \colon F(C) \to F(C')$ un isomorfismo en $\mathscr{D}$. Entonces existe un único isomorfismo $\phi' \colon C \to C'$ tal que $F(\phi') = \phi$.

Demostración. Sea $\phi\colon F(C) \to F(C')$ un isomorfismo con inversa $\psi\colon F(C') \to F(C)$, como F es plenamente fiel deben existir unicos $\phi'\colon C \to C'$ y $\psi'\colon C' \to C$ en $\mathscr{C}$ tales que $\phi = F(\phi')$ y $\psi = F(\psi')$, es necesario probar entonces que $\phi' \circ \psi' = id_{C'}$ y que $\psi' \circ \phi' = id_C$. Se tiene que $\psi \circ \phi = id_{F(C)}$ que es equivalente a $F(\psi') \circ F(\phi') = id_{F(C)}$ que por functorialidad se transforma en $F(\psi' \circ \phi') = F(id_C)$, pero como F es plenamente fiel, se sigue que $\psi' \circ \phi' = id_C$, de forma analoga se concluye que $\phi' \circ \psi' = id_{C'}$. Por tanto ϕ' es el unico isomorfismo tal que $F(\phi') = \phi$. $\square$

Construcción 2.2.12 (Restricciones de la unidad a izquierda y derecha)**.** Sea $\mathscr{C} = (\mathscr{C}, \otimes, \alpha, 1, v)$ una categoría monoidal. Para todo $X \in \mathscr{C}$, se tienen los isomorfismos naturales

$$1 \otimes (1 \otimes X) \xrightarrow{\alpha_{1,1,X}} (1 \otimes 1) \otimes X \xrightarrow{v \otimes id_X} 1 \otimes X$$

Como el funtor $Y \mapsto 1 \otimes Y$ es plenamente fiel, por la Proposición 2.2.11, existe un único isomorfismo $\lambda_X\colon 1 \otimes X \xrightarrow{\sim} X$ que hace que el siguiente diagrama conmute

$$\begin{tikzcd}
1 \otimes (1 \otimes X) \arrow[rr, "\alpha_{1,1,X}", "\sim"'] \arrow[dr, "\sim", "id_1 \otimes \lambda_X"'] & & (1 \otimes 1) \otimes X \arrow[dl, "\sim"', "v \otimes id_X"] \\
& 1 \otimes X &
\end{tikzcd}$$

Se hace referencia a λ_X como la **restricción de la unidad a izquierda**. De forma análoga se concluye que existe un único isomorfismo $\rho_X\colon X \otimes 1 \xrightarrow{\sim} X$ que hace que el siguiente diagrama conmute

$$\begin{tikzcd}
X \otimes (1 \otimes 1) \arrow[rr, "\alpha_{X,1,1}", "\sim"'] \arrow[dr, "\sim", "id_X \otimes v"'] & & (X \otimes 1) \otimes 1 \arrow[dl, "\sim"', "\rho_X \otimes id_1"] \\
& X \otimes 1 &
\end{tikzcd}$$

Se hace referencia a ρ_X como la **restricción de la unidad a derecha**.

Note que por cada $X \in \mathscr{C}$ se tiene un morfismo λ_X y un morfismo ρ_X, es decir, las restricciones de la unidad a derecha y a izquierda son en realidad transformaciones naturales.

Teorema 2.2.13. Sea $\mathscr{C}$ una categoría monoidal con objeto unidad 1, entonces las restricciones de la unidad a izquierda y derecha $\lambda_1, \rho_1\colon 1 \otimes 1 \xrightarrow{\sim} 1$ son iguales a la restricción de la unidad $v\colon 1 \otimes 1 \xrightarrow{\sim} 1$.

Este teorema establece la equivalencia mencionada anteriormente, en cuanto a la definición de unidad en una categoría monoidal adoptada en este trabajo y las definiciones adoptadas en otros escritos.

2.3. Functores monoidales no unitales

Recordando que si M y M' son monoides no unitales. Se dice que una función $f\colon M \to M'$ es un **homomorfismo de monoides no unitales** si para todo par de elementos $x, y \in M$, se cumple que $f(xy) = f(x)f(y)$. Si M y M' son monoides, se dice que f es un **homomorfismo de monoides**, si es un homomorfismo de monoides no unitales que mapea a el elemento unidad $e \in M$ en el elemento unidad $e' \in M'$.

Dadas dos categorías monoidales (no unitales) $\mathscr{C}$ y $\mathscr{D}$, para construir el análogo de la definicion anterior, se necesita un funtor $F\colon \mathscr{C} \to \mathscr{D}$ que se comporte bien con las estructuras monoidales (no unitales) de $\mathscr{C}$ y $\mathscr{D}$, es decir, en el que las imágenes de los functores que mapean $(X, Y) \mapsto F(X) \otimes F(Y)$ y $(X, Y) \mapsto F(X \otimes Y)$ se puedan comparar con una condición análoga a la de la definición, en la versión mas simplificada la condición es de igualdad estricta. Además, la condición de comparación debe ser compatible con los morfismos en la categoría, y mas aun, con las restricciones de asociatividad de ambas categorías.

Definición 2.3.1. San $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales no unitales. Un **funtor monoidal estricto sin unidad** de $\mathscr{C}$ en $\mathscr{D}$ es un funtor $F\colon \mathscr{C} \to \mathscr{D}$ con las siguientes propiedades:

1. El siguiente diagrama de functores es estrictamente conmutativo

$$\begin{tikzcd}
\mathscr{C} \times \mathscr{C} \arrow[r, "\otimes_{\mathscr{C}}"] \arrow[d, "F \times F"'] & \mathscr{C} \arrow[d, "F"] \\
\mathscr{D} \times \mathscr{D} \arrow[r, "\otimes_{\mathscr{D}}"'] & \mathscr{D}
\end{tikzcd}$$

   es decir, para cualquier par de objetos $X, Y \in \mathscr{C}$ se tiene la igualdad $F(X) \otimes_{\mathscr{D}} F(Y) = F(X \otimes_{\mathscr{C}} Y)$ de objetos en $\mathscr{D}$

2. Para todo $X, Y, Z \in \mathscr{C}$ el funtor F mapea la restricción de asociatividad $\alpha_{X,Y,Z}\colon X \otimes_{\mathscr{C}} (Y \otimes_{\mathscr{C}} Z) \to (X \otimes_{\mathscr{C}} Y) \otimes_{\mathscr{C}} Z$ (de la estructura monoidal de $\mathscr{C}$) en la restricción de asociatividad $\alpha_{F(X),F(Y),F(Z)}\colon F(X) \otimes_{\mathscr{D}} (F(Y) \otimes_{\mathscr{D}} F(Z)) \to (F(X) \otimes_{\mathscr{D}} F(Y)) \otimes_{\mathscr{D}} F(Z)$.

Ejemplo 2.3.2. Sea $\mathscr{C}$ una categoría monoidal sin unidad, entonces el funtor identidad $id_{\mathscr{C}}$ es un funtor monoidal estricto sin unidad de $\mathscr{C}$ en $\mathscr{C}$.

Por la misma razón por la que se definió un concepto un poco mas general que la definición 2.1.3 se debe buscar una definición mas amplia que permita mas posibilidades en cuanto a la forma en la que se estudian las categorías monoidales, es decir, en donde los objetos se puedan igualar salvo isomorfismo.

Definición 2.3.3. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales no unitales y $F : \mathscr{C} \to \mathscr{D}$ un funtor de $\mathscr{C}$ en $\mathscr{D}$. Una **estructura monoidal laxa sin unidad** sobre F es una colección de morfismos $\mu = \{\mu_{X,Y} \colon F(X) \otimes_{\mathscr{D}} F(Y) \to F(X \otimes_{\mathscr{C}} Y)\}_{X,Y \in \mathscr{C}}$ que satisfacen las siguientes propiedades

1. μ es una transformación natural, es decir, para cada $X, Y \in \mathscr{C}$, $X', Y' \in \mathscr{D}$ y morfismos $f \colon X \to X'$ y $g \colon Y \to Y'$ el siguiente diagrama conmuta

$$\begin{tikzcd}
F(X) \otimes_{\mathscr{D}} F(Y) \arrow[r, "\mu_{X,Y}"] \arrow[d, "F(f)\otimes_{\mathscr{D}} F(g)"'] & F(X \otimes_{\mathscr{C}} Y) \arrow[d, "F(f \otimes_{\mathscr{C}} g)"] \\
F(X') \otimes_{\mathscr{D}} F(Y') \arrow[r, "\mu_{X',Y'}"] & F(X' \otimes_{\mathscr{C}} Y')
\end{tikzcd}$$

2. Los morfismos $\mu_{X,Y}$ son compatibles con las restricciones de asociatividad de $\mathscr{C}$ y $\mathscr{D}$, es decir, para todo $X, Y, Z \in \mathscr{C}$ el siguiente diagrama conmuta en $\mathscr{D}$

$$\begin{tikzcd}
F(X) \otimes_{\mathscr{D}} (F(Y) \otimes_{\mathscr{D}} F(Z)) \arrow[r, "\alpha^{\mathscr{D}}_{F(X),F(Y),F(Z)}"] \arrow[d, "id_{F(X)} \otimes \mu_{Y,Z}"'] & (F(X) \otimes_{\mathscr{D}} F(Y)) \otimes_{\mathscr{D}} F(Z) \arrow[d, "\mu_{X,Y} \otimes_{\mathscr{D}} id_{F(Z)}"] \\
F(X) \otimes_{\mathscr{C}} F(Y \otimes_{\mathscr{D}} Z) \arrow[d, "\mu_{X, Y \otimes_{\mathscr{C}} Z}"'] & F(X \otimes_{\mathscr{D}} Y) \otimes_{\mathscr{C}} F(Z) \arrow[d, "\mu_{X \otimes_{\mathscr{C}} Y, Z}"] \\
F(X \otimes_{\mathscr{C}} (Y \otimes_{\mathscr{C}} Z)) \arrow[r, "F(\alpha^{\mathscr{C}}_{X,Y,Z})"'] & F((X \otimes_{\mathscr{C}} Y) \otimes_{\mathscr{C}} Z)
\end{tikzcd}$$

   donde $\alpha^{\mathscr{C}}$ y $\alpha^{\mathscr{D}}$ denotan las restricciones de asociacitivdad en la respectiva categoría monoidal.

Un **funtor monoidal laxo sin unidad** de $\mathscr{C}$ en $\mathscr{D}$ es un par (F, μ) donde $F \colon \mathscr{C} \to \mathscr{D}$ es un funtor y $\mu = \{\mu_{X,Y}\}_{X,Y \in \mathscr{C}}$ es una estructura monoidal laza sin unidad sobre F. Se hace referencia a los morfismos $\{\mu_{X,Y}\}_{X,Y \in \mathscr{C}}$ como restricciones tensoriales de F.

Básicamente esta definición busca mantener la compatibilidad de la estructura monoidal entre dos categorías $\mathscr{C}$ y $\mathscr{D}$.

1. Con la condición (1) se asegura que los functores producto tensorial de ambas categorías sean equivalentes, es decir, se preserva el producto tensorial.

2. La condición (2) asegura que el funtor F conserve las restricciones de asociatividad de las estructuras monoidales de ambas categorías.

Definición 2.3.4. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales no unitales y $F : \mathscr{C} \to \mathscr{D}$ un funtor de $\mathscr{C}$ en $\mathscr{D}$. Una **estructura monoidal sin unidad** sobre F es una estructura monoidal laxa sin unidad $\mu = \{\mu_{X,Y}\}_{X,Y\in\mathscr{C}}$ sobre F, tal que cada restricción de asociatividad $\mu_{X,Y} \colon F(X) \otimes_{\mathscr{C}} F(Y) \to F(X \otimes_{\mathscr{D}} Y)$ es un isomorfismo. Un **funtor monoidal sin unidad** de $\mathscr{C}$ en $\mathscr{D}$ es un par (F, μ), donde $F \colon \mathscr{C} \to \mathscr{D}$ es un funtor y μ es una estructura monoidal sin unidad sobre F.

Ejemplo 2.3.5. Sean $\mathscr{C}$ y $\mathscr{D}$ cagegoria monoidales no unitales y $F \colon \mathscr{C} \to \mathscr{D}$ un funtor monoidal estricto sin unidad, entonces es claro que F admite una estructura monoidal sin unidad $\{\mu_{X,Y}\}_{X,Y\in\mathscr{C}}$, pues se puede tomar a $\mu_{X,Y}$ como el morfismo identidad de $F(X) \otimes_{\mathscr{D}} F(Y)$ en $F(X \otimes_{\mathscr{C}} Y)$.

Ejemplo 2.3.6. Sean M y M' monoides no unitales considerados como categorías monoidales no unitales con solo morfismos identidad, entonces, functores monoidales laxos no unitales de M en M' pueden ser identificados con homomorfismos de monoides de M en M'.

Así como se tienen equivalencias de categorías en la categoría Cat, o así como se tiene la categoría $\mathrm{Fun}(\mathscr{C}, \mathscr{D})$ para dos categorías $\mathscr{C}$ y $\mathscr{D}$, también se puede pensar en tener equivalencias de categorías monoidales, o la categoría $\mathrm{Func}[\mathscr{C}, \mathscr{D}]$ para categorías monoidales $\mathscr{C}$ y $\mathscr{D}$. Para esto se debe definir primero lo que es una transformación natural monoidal.

Definición 2.3.7. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales no unitales y $F, F' \colon \mathscr{C} \to \mathscr{D}$ functores equipados con estructuras monoiedales laxas no unitales μ y μ' respectivamente. Se dice que una transformación natural de functores $\gamma \colon F \to F'$ es **monoidal sin unidad**, si para todo $X, Y \in \mathscr{C}$, el siguiente diagrama conmuta

$$\begin{tikzcd}
F(X) \otimes_{\mathscr{D}} F(Y) \arrow[r, "\mu_{X,Y}"] \arrow[d, "\gamma(X)\otimes_{\mathscr{D}}\gamma(Y)"'] & F(X \otimes_{\mathscr{C}} Y) \arrow[d, "\gamma(X\otimes_{\mathscr{C}} Y)"] \\
F'(X) \otimes_{\mathscr{D}} F'(Y) \arrow[r, "\mu'_{X,Y}"] & F'(X \otimes_{\mathscr{C}} Y)
\end{tikzcd}$$

2.4. Functores monoidales

Dadas las consideraciones de las secciones anteriores, para la generalización de un homomorfismo de monoides, solo hace falta dar una noción de conservación de la unidad en los functores monoidales no unitales, para ello, primero se definirá lo que se entiendo por unidad de un funtor monoidal laxo.

Definición 2.4.1. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales con objetos unidad $1_{\mathscr{C}}$ y $1_{\mathscr{D}}$ respectivamente y $F\colon \mathscr{C} \to \mathscr{D}$ un funtor monoidal laxo sin unidad con restricciones tensoriales $\mu = \{\mu_{X,Y}\}_{X,Y\in\mathscr{C}}$, sea también un morfismo $\epsilon\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$ en $\mathscr{D}$. Se dice que ϵ es una unidad a izquierda para F, si para todo X en $\mathscr{C}$, la restricción de la unidad a izquierda $\lambda_{F(X)}\colon 1_{\mathscr{D}} \otimes F(X) \xrightarrow{\sim} F(X)$ en $\mathscr{D}$ es igual a la composición

$$1_{\mathscr{D}} \otimes_{\mathscr{D}} F(X) \xrightarrow{\epsilon\otimes_{\mathscr{D}} id_{F(X)}} F(1_{\mathscr{C}}) \otimes_{\mathscr{D}} F(X) \xrightarrow{\mu_{1_{\mathscr{C}},X}} F(1_{\mathscr{C}} \otimes_{\mathscr{C}} X) \xrightarrow{F(\lambda_X)} F(X)$$

donde $\lambda_X\colon 1_{\mathscr{C}} \otimes_{\mathscr{C}} X \xrightarrow{\sim} X$ es la restricción de la unidad a izquierda en $\mathscr{C}$. Se dice que ϵ es una unidad a derecha para F, si para todo X en $\mathscr{C}$, la restricción de la unidad a derecha $\rho_{F(X)}\colon F(X) \otimes_{\mathscr{D}} 1_{\mathscr{D}} \xrightarrow{\sim} F(X)$ en $\mathscr{D}$ es igual a la composición

$$F(X) \otimes_{\mathscr{D}} 1_{\mathscr{D}} \xrightarrow{id_{F(X)}\otimes_{\mathscr{D}}\epsilon} F(X) \otimes_{\mathscr{D}} F(1_{\mathscr{C}}) \xrightarrow{\mu_{X,1_{\mathscr{C}}}} F(X \otimes_{\mathscr{C}} 1_{\mathscr{C}}) \xrightarrow{F(\rho_X)} F(X)$$

Se dice que ϵ es una unidad para F, si es unidad a izquierda y unidad a derecha de F.

En caso de que un funtor monidal laxo sin unidad posea unidades a izquierda y derecha, estas unidades en realidad son la misma y son unicas, como se establece en los siguientes resultados.

Proposición 2.4.2. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales con objetos unidades $1_{\mathscr{C}}$ y $1_{\mathscr{D}}$ respectivamente y sea $F\colon \mathscr{C} \to \mathscr{D}$ un funtor monoidal laxo sin unidad. Suponga que F admite una unidad a izquierda $\epsilon_L\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$ y una unidad a derecha $\epsilon_R\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$, entonces $\epsilon_L = \epsilon_R$.

Corolario 2.4.3. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales y sean $F\colon \mathscr{C} \to \mathscr{D}$ un fuctor monoidal laxo sin unidad. Entonces F admite una unidad $\epsilon\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$ si y solo si admite una unidad a izquierda y una unidad a derecha. En este caso, la unidad ϵ es única.

Notese que a diferencia de las categorías monoidales, la unidad no es unica salvo isomorfismo, en este caso la unidad es unica (bajo igualdad estricta).

Definición 2.4.4. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales y sea $F\colon \mathscr{C} \to \mathscr{D}$ un funtor. Una estructura monoidal laxa sobre F es una estructura monoidal laxo sin unidad $\mu = \{\mu_{X,Y} = X, Y \in \mathscr{C}\}$ para la que existe una unidad $\epsilon\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$. Un funtor monoidal laxo de $\mathscr{C}$ en $\mathscr{D}$ es un par (F, μ), donde $F\colon \mathscr{C} \to \mathscr{D}$ es un funtor y μ es una estructura monoidal laxa sobre F. En este caso se hace referencia a el morfismo $\epsilon\colon 1_{\mathscr{D}} \to F(1_{\mathscr{C}})$, como la unidad de F.

Para definir equivalencias entre categorías monoidales (ente otras construcciones mencionadas en la sección anterior), se necesita definir primero que es una transformación natural monoiedal de functores monoidales laxos, así, considere:

Definición 2.4.5. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales y sean $F, F'\colon \mathscr{C} \to \mathscr{D}$ functores monoidales laxos de $\mathscr{C}$ en $\mathscr{D}$. Se dice que una transformación natural $\gamma\colon F \to F'$ es **monoidal** si se satisfacen las siguientes condiciones:

1. La transformación natural γ es monoidal sin unidad en el sentido de la definición 2.3.7.

2. La unidad de F' es igual a la composición

$$1_{\mathscr{D}} \xrightarrow{\epsilon} F(1_{\mathscr{C}}) \xrightarrow{\gamma(1_{\mathscr{C}})} F'(1_{\mathscr{C}})$$

   donde ϵ es la unidad de F.

Definición 2.4.6. Sean $\mathscr{C}$ y $\mathscr{D}$ categorías monoidales y $F\colon \mathscr{C} \to \mathscr{D}$ un funtor. Una estructura monoidal sobre F es una estructura monoidal laxa sin unidad $\mu = \{\mu_{X,Y}\}_{X,Y\in\mathscr{C}}$, tal que se satisfacen las siguientes propiedades:

1. Para todo $X, Y \in \mathscr{C}$, la restricción tensorial $\mu_{X,Y}\colon F(X) \otimes F(Y) \to F(X \otimes Y)$ es un isomorfismo en $\mathscr{D}$.

2. Existe un isomorfismo $\epsilon\colon 1_{\mathscr{D}} \xrightarrow{\sim} F(1_{\mathscr{C}})$, en el sentido de la definición 2.4.1

Un **funtor monoidal** de $\mathscr{C}$ en $\mathscr{D}$ es un par (F, μ), donde F es un funtor de $\mathscr{C}$ en $\mathscr{D}$ y μ es una estructura monoidal sobre F.

Capítulo 3

Diagramas de Cuerda

En este capítulo se estudiarán los diagramas de cuerda, los cuales permitirán un mejor manejo de relaciones algebraicas expresadas en términos de morfismos y objetos de una categoría monoidal mediante métodos geométricos.

3.1. Estrictificación

Las categorías monoidales usualmente son usadas en áreas donde se tienen procesos (y estados entre estos procesos) que se pueden componer tanto vertical (la composición usual) como horizontalmente (lo que se define como producto tensorial), por ejemplo, en la descripción de reacciones químicas, en los diagramas de Feynman, en lógica, entre otras. En estos casos concretos, en general resulta mucho más conveniente trabajar con categorías estrictas, que permite expresar mediante igualación relaciones entre expresiones formadas con objetos, morfismos, funtores y transformaciones naturales que involucran a categorías monoidales. El siguiente resultado muestra que en una categoría monoidal arbitraria M se puede cambiar la comparación por isomorfismos $\cong$ por la comparación con igualdad $=$.

Teorema 3.1.1. Toda categoría monoidal M es monoidalmente equivalente a una categoría monoidal estricta S.

Demostración. Sea $(M, \otimes_M, \alpha_M, 1_M, \upsilon_M)$ una categoría monoidal arbitraria. Para demostrar el teorema se debe encontrar una categoría monoidal estricta $(S, \otimes_S, 1_S)$, junto con funtores monoidales $F\colon S \to M$ y $G\colon M \to S$ con identidades $\epsilon\colon 1_M \xrightarrow{\sim} F(1_S)$, $\epsilon'\colon 1_S \xrightarrow{\sim} G(1_M)$ y restricciones tensoriales $\mu = \{\mu_{A,B}\}_{A,B\in M}$, $\mu' = \{\mu'_{A,B}\}_{A,B\in S}$ respectivamente, tales que existan isomorfismos naturales monoidales $\eta\colon id_S \Rightarrow G \circ F$

y $\gamma\colon F \circ G \Rightarrow id_M$.

Para esto se partirá desde la construcción de S considerando solo los objetos, por tanto, para construir el funtor F solo es necesario especificar la acción sobre los objetos (porque S aun no cuenta con más estructura), y a medida que S se vaya completando a una categoría monoidal, F se ira completando a un funtor monoidal.

Así, considere el conjunto de objetos de S como el conjunto de sucesiones finitas de objetos en M, incluyendo la sucesión vacía denotada por la lista vacia $[\,]$. Luego, se define el mapeo del funtor F sobre los objetos, como la asignación

$$F([A_1, A_2, \ldots, A_n]) = A_1 \otimes_M (A_2 \otimes_M (A_3 \otimes_M \ldots \otimes_M (A_{n-1} \otimes_M A_n)))$$

en donde, por definición, $F([\,]) = 1_M$. Finalmente se usa esta asignación para definir los morfismos de S, esto es, para dos objetos $A, B \in S$ se define

$$\mathrm{Hom}_S(A, B) = \mathrm{Hom}_M(F(A), F(B))$$

de donde es claro que la composición en S es determinada por la composición en M vía el funtor F. Con todas estas consideraciones, por definición, S es una categoría y F es un funtor. Para dotar de estructura monoidal a S considere el producto tensorial $\otimes_S\colon S \times S \to S$ definido por:

1. La concatenación, esto es, dados dos objetos $[A_1, A_2, \ldots A_n]$ y $[B_1, B_2, \ldots B_m]$ en S

$$[A_1, A_2, \ldots A_n] \otimes_S [B_1, B_2, \ldots B_m] = [A_1, A_2, \ldots A_n, B_1, B_2, \ldots B_m].$$

2. Dados dos morfismos $f\colon A \to C$ y $g\colon B \to D$ en S, donde, por definición $f \in \mathrm{Hom}_M(F(A), F(C))$ y $g \in \mathrm{Hom}_M(F(B), F(D))$, se quiere definir su producto tensorial

$$f \otimes_S g \in \mathrm{Hom}_S(A \otimes_S B, C \otimes_S D) = \mathrm{Hom}_M(F(A \otimes_S B), F(C \otimes_S D))$$

Se tomara ese producto como la siguiente composición:

$$F(A \otimes_S B) \to F(A) \otimes_M F(B) \to F(C) \otimes_M F(D) \to F(C \otimes_S D)$$

donde el primer morfismo $F(A \otimes_S B) \to F(A) \otimes_M F(B)$ esta dado por la única[1] composición de asociadores de M que cambia los paréntesis de la manera

[1]Unicidad garantizada por el teorema de coherencia para categorías monoidales aplicado a M.

necesaria y de forma análoga para el tercer morfismo $F(C)\otimes_M F(D) \to F(C\otimes_S D)$. El morfismo del medio $F(A) \otimes_M F(B) \to F(C) \otimes_M F(D)$ se define como el producto tensorial de f y g en M. Por construcción, la funtorialidad de $\otimes_S$ se sigue del teorema de coherencia y de la funtorialidad de $\otimes_M$.

Ahora se toma como unidad tensorial 1_S a la sucesión vacía $[\,]$, y las restricciones de asociatividad, y las restricciones de la unidad (unitor a izquierda y unitor a derecha) se toman como las respectivas transformaciones identidad. Así, se satisfacen la identidad triangular y la identidad pentagonal, de manera que $(S, \otimes_S, 1_S)$ es en efecto una categoría monoidal, más aun, es una categoría monoidal estricta.

Lo siguiente, es dotar a F con una estructura monoidal, lo cual se consigue tomando las siguientes consideraciones:

1. Se debe notar que el funtor F actúa como el mapeo identidad sobre los morfismos y por tanto existe el morfismo identidad $1_M \to F(1_S)$, así, se toma el morfismo $\epsilon\colon 1_M \to F(1_S)$ como el morfismo identidad de 1_M.

2. Las restricciones funtoriales $\mu_{A,B}\colon F(A) \otimes_M F(B) \to F(A \otimes_S B)$ están determinadas por las restricciones de asociatividad de M que cambian los paréntesis de la manera necesaria

La naturalidad de μ se sigue de manera directa de la unicidad garantizada por el teorema de coherencia junto con la funtorialidad de $\otimes_M$. Es fácil verificar que las condiciones de asociatividad y unitalidad se cumplen, de manera que (F, μ, ϵ) es un funtor monoidal laxo, más aun, es un funtor monoidal, dado que tanto ϵ como μ son isomorfismos.

Ahora, se debe definir un funtor monoidal $G\colon M \to S$, el cual se construye como sigue:

1. Sobre objetos G mapea a $A \in M$ en $[A] \in S$ y sobre morfismos actúa como un mapeo identidad.

2. El morfismo $\epsilon'\colon 1_S \to G(1_M)$ se toma como el mapeo identidad de 1_M en M.

3. Las restricciones tensoriales $\mu'_{M,N}\colon G(M)\otimes_S G(N) \to G(M\otimes_M N)$ están dados por la identidad de $M \otimes_M N$.

Tanto en el segundo como en el tercer numeral, se debe recordar la definición de $\mathrm{Hom}_S(A, B)$ y que F actúa como la identidad sobre los morfismos. Esto garantiza que (G, μ', ϵ') es un funtor monoidal laxo, más aun, es un funtor monoidal, dado que tanto ϵ como μ son isomorfismos.

Por ultimo resta verificar la existencia de isomorfismos naturales monoidales $\eta\colon id_S \to G \circ F$ y $\gamma\colon F \circ G \to id_M$. Pero, $F \circ G$ es el funtor identidad (que es monoidal) por construcción, luego, se puede tomar a γ como la isomorfismo natural monoidal identidad. Falta definir η. Para esto se describirá explícitamente el funtor monoidal $(G \circ F, \epsilon'', \mu'')$:

1. El funtor subyacente $G \circ F\colon S \to S$ sobre objetos esta dado por la regla

   $$G(F([A_1, A_2, \ldots, A_n])) = [A_1 \otimes_M (A_2 \otimes_M (A_3 \otimes_M \ldots \otimes_M (A_{n-1} \otimes_M A_n)))]$$

   y en morfismos es el morfismo identidad.

2. El morfismo $\epsilon''\colon 1_S \to GF(1_S)$ es el mapeo identidad de 1_M.

3. La transformación natural $\mu''_{A,B}\colon G(F(A)) \otimes_S G(F(B)) \to G(F(A \otimes_S B))$ para objetos $A, B \in S$ esta determinada por la única composición de las restricciones de asociatividad de M que cambia los paréntesis de manera necesaria. Por tanto la naturalidad de μ'' es directa.

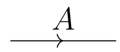
Así, definiendo las componentes η_A de la transformación natural η como el mapeo identidad para cada $A \in S$, se obtiene el isomorfismo natural monoidal $\eta\colon id_S \to G \circ F$ requerido. $\square$

3.2. Diagramas de cuerda para categorías

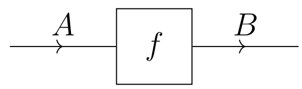
En el primer capítulo, se mostró una forma de presentar información de una categoría mediante figuras llamadas diagramas conmutativos. En esta sección se pretenden empezar a desarrollar los diagramas de cuerda. Para categorías que no tienen estructura adicional (no son monoidales, o abelianas, ...), en este caso los diagramas de cuerda son totalmente equivalentes a los vistos en el capítulo I.

Una formalización completa y rigurosa de los diagramas de cuerda se encuentra en [JS91] y un tratamiento más moderno esta en [Sel10].

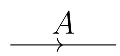
Así, dada una categoría $\mathscr{C}$, las diferentes estructuras que se pueden formar usando elementos de la categoría se representan con los siguientes diagramas de cuerda. Un objeto A en $\mathscr{C}$ es representado por el diagrama de cuerda



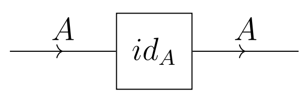
y un morfismo $f\colon A \to B$ se representa mediante el diagrama de cuerda



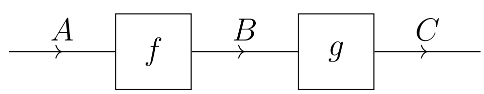
Esta forma de representar a objetos y morfismos, viene de la idea de considerar a los morfismos como cajas negras, como un mecanismo que recibe un objeto de entrada y lo transforma en un objeto de salida, sin saber exactamente que tipo de transformación realiza, en parte porque no se tiene información de la estructura interna de los objetos de la categoría a la que pertenecen los morfismos. Así, el morfismo identidad $id_A\colon A \to A$ para un objeto A, se representa con el siguiente diagrama de cuerda



En este caso se usa la misma representación del objeto A para el morfismo id_A, dado que el morfismo identidad no transforma el objeto de entrada A en ningún otro objeto B, por tanto, hacer lo siguiente seria redundante



Dados dos morfismos $f\colon A \to B$ y $g\colon B \to C$, la composición se representa mediante el diagrama de cuerda

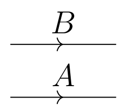


Teorema 3.2.1 (Teorema de coherencia)**.** Una expresión bien formada entre dos términos de morfismos en el lenguaje de categorías se sigue de los axiomas categóricos si y solo si se cumple en el lenguaje gráfico salvo isomorfismo de diagramas (de cuerda).

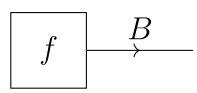
El teorema anterior básicamente establece que hacer cálculos de forma gráfica mediante los diagramas de cuerda es equivalente a hacerlos de forma algebraica. Claramente se prefiere una aproximación sobre la otra cuando sea conveniente, pero en este punto no sirve de mucho, porque la categoría no tiene estructura adicional. En las siguientes secciones se introducirá los diagramas asociados para representar la información de la categoría que se tiene junto con su respectivo diagrama de cuerda.

3.3. Diagramas de cuerdas para categorías monoidales

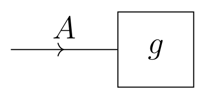
Al tomar una categoría monoidal $\mathscr{C}$, además de considerar los diagramas anteriores, se deben introducir nuevas reglas para representar objetos y morfismos que están relacionados de forma directa con la estructura monoidal de $\mathscr{C}$, por ejemplo, se debe especificar como representar objetos de la forma $A \otimes B \in \mathscr{C}$ y morfismos de la forma $f\colon A_1 \otimes \cdots \otimes A_n \to B_1 \otimes \cdots \otimes B_m$. Así, dados dos objetos A y B el producto tensorial $A \otimes B$ se representa mediante el diagrama



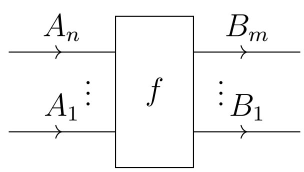
Al elemento identidad 1 de la categoría monoidal $\mathscr{C}$ no se le asigna ninguna representación, es decir, se representa sin dibujar nada, por tanto, para el morfismo $f\colon 1 \to B$ se tiene el siguiente diagrama



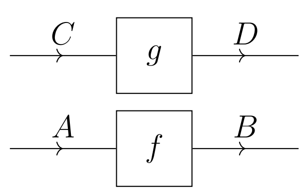
y para el morfismo $g\colon A \to 1$



Para representar un morfismo de la forma $f\colon A_1 \otimes \cdots \otimes A_n \to B_1 \otimes \cdots \otimes B_m$, se usa el siguiente diagrama



Y por ultimo, cuando se tiene el producto tensorial $f \otimes g$ entre dos morfismos $f\colon A \to B$ y $g\colon C \to D$, se usa el diagrama



Teorema 3.3.1 (Teorema de coherencia)**.** Una ecuación bien formada entre términos de morfismos en el lenguaje de categorías monoidales, se sigue de los axiomas de las categorías monoidales si y solo si se cumple en el lenguaje gráfico salvo isotopia planar.

3.4. Diagramas de cuerda para categorías monoidales simétricas

Definición 3.4.1. Un **trenzado** sobre una categoría monoidal $\mathscr{C}$ es una transformación natural $c = \{c_{A,B} \colon A \otimes B \to B \otimes A\}_{A,B \in \mathscr{C}}$, tal que los siguientes siguientes diagramas son conmutativos:

$$\begin{tikzcd}
B \otimes (A \otimes C) \arrow[r, "\alpha_{B,A,C}"] \arrow[d, "id_B \otimes \mu_{A,C}"'] & (B \otimes A) \otimes C \arrow[d, "\mu_{X,A} \otimes id_C"] \\
B \otimes (A \otimes C) \arrow[d, "\mu_{X,A\otimes C}"'] & (X \otimes A) \otimes C \arrow[d, "\mu_{X \otimes A,C}"] \\
X \otimes (A \otimes C) \arrow[r, "(\alpha^{\mathscr{C}}_{X,A,C})"'] & (X \otimes A) \otimes C
\end{tikzcd}$$

$$\begin{tikzcd}
B \otimes (A \otimes C) \arrow[r, "\alpha_{B,A,C}"] \arrow[d, "id_B \otimes \mu_{A,C}"'] & (B \otimes A) \otimes C \arrow[d, "\mu_{X,A} \otimes id_C"] \\
B \otimes (A \otimes C) \arrow[d, "\mu_{X,A\otimes C}"'] & (X \otimes A) \otimes C \arrow[d, "\mu_{X \otimes A,C}"] \\
(X \otimes (A \otimes C)) \arrow[r, "(\alpha^{\mathscr{C}}_{X,A,C})"'] & ((X \otimes A) \otimes C)
\end{tikzcd}$$

Una categoría monoidal trenzada es un par $(\mathscr{C}, c)$ donde $\mathscr{C}$ es una categoría monoidal y c es un trenzado sobre $\mathscr{C}$.

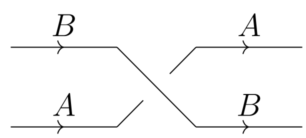
Definición 3.4.2. Un funtor $F \colon \mathscr{C} \to \mathscr{D}$ es un **funtor monoidal trenzado** entre dos categorías monoidales trenzadas $(\mathscr{C}, c)$ y $(\mathscr{D}, d)$, si se satisface lo siguiente:

1. F es un funtor monoidal con restricción de asociatividad μ.

2. Si es compatible con el trenzado de $\mathscr{C}$ y $\mathscr{D}$, esto es, el siguiente diagrama conmuta

$$\begin{tikzcd}
F(A) \otimes_{\mathscr{D}} F(B) \arrow[r, "\mu_{A,B}"] \arrow[d, "c_{F(A),F(B)}"'] & F(A \otimes_{\mathscr{C}} B) \arrow[d, "F(c_{A,B})"] \\
F'(B) \otimes_{\mathscr{D}} F'(A) \arrow[r, "\mu_{B,A}"] & F'(B \otimes_{\mathscr{C}} A)
\end{tikzcd}$$

Para representar gráficamente una categoría trenzada se deben tomar las consideraciones de las secciones anteriores y el morfismo $c_{A,B} \colon A \otimes B \to B \otimes A$ es equivalente a

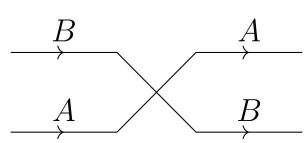


Teorema 3.4.3. Una ecuación bien formada entre términos de morfismos en el lenguaje de categorías monoidales trenzadas, se sigue de los axiomas de las categorías monoidales trenzadas si y solo si se cumple en el lenguaje gráfico salvo isotopia en 3 dimensiones.

Definición 3.4.4. Una categoría monoidal simétrica, es una categoría monoidal trenzada $(\mathscr{C}, c)$ donde las componentes $c_{A,B}$ son su misma inversa, esto es, se cumple que $c_{A,B} = c^{-1}_{B,A}$. En este caso el trenzado es llamado simétrico.

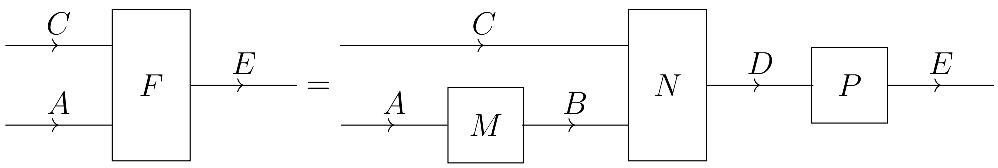
Definición 3.4.5. Un **funtor monoidal simétrico** es un funtor monoidal trenzado cuyo dominio y codominio son categorías monoidales simétricas.

Y extendiendo los diagramas de cuerda de las categorías monoidales trenzadas a las categorías simétricas, se usa el siguiente diagrama



Teorema 3.4.6. Una ecuación bien formada entre términos de morfismos en el lenguaje de categorías monoidales simétricas, se sigue de los axiomas de las categorías monoidales simétricas si y solo si se cumple en el lenguaje gráfico salvo isomorfismo de diagramas.

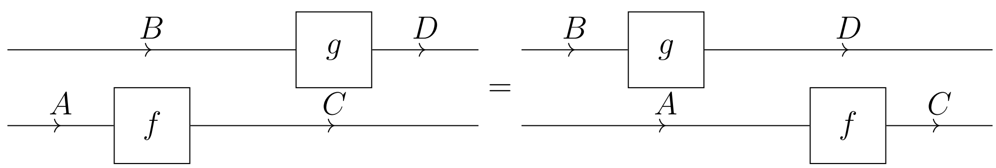
Ahora, de acuerdo a todo lo anterior, se verán algunos casos de como usar diagramas de cuerda para obtener conclusiones de una categoría monoidal $\mathscr{C}$. Así, si se tienen los morfismos $M\colon A \to B$, $N\colon B \otimes C \to D$, $P\colon D \to E$ y $F\colon A \otimes C \to E$ para objetos A, B, C, D y E de una categoría monoidal $\mathscr{C}$, tales que $F = P \circ N \circ (M \otimes id_C)$, entonces, en términos de diagramas de cuerda



Por otro lado, en una categoría monoidal $\mathscr{C}$ por la bifuntorialidad de el producto tensorial $\otimes\colon \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ se cumple que

$$(id_C \otimes g) \circ (f \otimes id_B) = (f \otimes id_D) \circ (id_A \otimes g)$$

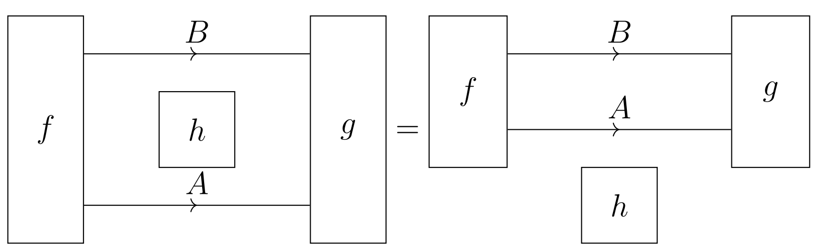
que es fácilmente verificable usando diagramas de cuerda, pues esta igualdad se transforma en



Por ultimo, probar la siguiente igualdad

$$g \circ ((\rho_A \circ (id_A \otimes h) \circ \rho_A^{-1}) \otimes id_B) \circ f = g \circ ((\lambda_A \circ (h \otimes id_A)) \circ \lambda_A^{-1}) \circ f$$

para morfismos $f\colon I \to A \otimes B$, $g\colon A \otimes B \to I$ y $h\colon I \to I$, es equivalente a mostrar que se cumple la siguiente igualdad



Capítulo 4

Teoría de Categorías Aplicada a los Algoritmos de Aprendizaje Supervisados

En este capitulo se aplicara el lenguaje de la teoría de categorías al estudio de los algoritmos de aprendizaje supervisados. Se definirá la categoría **Learn** cuyos objetos son conjuntos y cuyos morfismos son aprendices, también se definirá la categoría **Para**, cuyos objetos son espacios euclidianos finitos y cuyos morfismos son funciones parametrizadas, para probar que el gradiente descendiente y el *backpropagation* inducen un functor de la categoría **Para** en la categoría **Learn** cuando se fija una función de error y una taza de aprendizaje. Luego, se definirá la categoría **NNet** cuyos objetos son números naturales y cuyos morfismos son redes neuronales y por ultimo se demostrara que existe un funtor de la categoría **NNet** en la categoría **Learn**. Esto asegurará que los algoritmos de aprendizaje supervisados son composicionales, más aun, se verá que estas categorías son además, categorías monoidales. Una referencia fundamental para este tema se encuentra en [FST19].

4.1. La categoría Learn

En esta sección se verá que los algoritmos de aprendizaje supervisados forman una categoría. Habitualmente, estos algoritmos de aprendizaje supervisado se describen como algoritmos que buscan la mejor aproximación a una función ideal $f\colon A \to B$, en un conjunto de funciones de tipo $A \to B$ parametrizadas por un espacio de parámetros P, lo que se hace tomando ejemplos $(a, f(a)) \in A \times B$ y actualizando los parámetros de acuerdo a alguna regla. Esta idea se hace precisa con la siguiente definición.

Definición 4.1.1. Sean A y B conjuntos. Un **algoritmo de aprendizaje supervisado** o simplemente un **aprendiz** de tipo $A \to B$, es una tupla (P, I, U, r) donde P es un conjunto y I, U y r son funciones de tipos:

$$\begin{aligned} I &\colon P \times A \to B, \\ U &\colon P \times A \times B \to P, \\ r &\colon P \times A \times B \to A. \end{aligned}$$

En la definición anterior, a P se le llama **espacio de parámetros**. El mapeo I es denominado **función de implementación**, U es llamada **función de actualización** y r es llamada **función de petición** o **función de solicitud**.

Ejemplo 4.1.2. En este ejemplo se escribirá el algoritmo de aprendizaje supervisado que aprende a predecir una función lineal en el marco de la definición anterior, es decir, se definirá como un aprendiz denotado por L. En este caso se tiene una función $f\colon \mathbb{R} \to \mathbb{R}$ a la cual se le quiere encontrar la mejor aproximación lineal, es decir, el aprendiz L es de tipo $\mathbb{R} \to \mathbb{R}$ y el espacio de parámetros es $P = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ y por lo tanto la función de implementación $I\colon \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ esta definida por la asignación $I(m, b, x) = mx + b$. Se quiere ajustar los valores de m y b buscando que $I(m, b, x) \approx f(x)$, por tanto, la función de actualización $U\colon \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ se define por la asignación

$$U(m, b, x, y) = (m - \alpha(mx + b - y) \cdot x, b - \alpha(mx + b - y))$$

donde $\alpha \in \mathbb{R}$ es un valor fijo. Notese que la Definición 4.1.1 no exige que las funciones I, U y r cumplan alguna condición, por tanto se puede escoger a $r\colon \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ de forma arbitraria[1] como la función con la asignación $r(m, b, x, y) = x$.

Según el ejemplo anterior, a priori, puede parecer que el requerimiento de una función de petición r en la construcción del aprendiz (P, I, U, r) es innecesaria, dado que su escogencia puede ser arbitraria. Pero antes de responder a el porque se hace necesaria una función de petición r, vale la pena mostrar como la Definición 4.1.1 generaliza la definición usual de lo que es un algoritmo de aprendizaje supervisado.

Los algoritmos de aprendizaje, en particular los algoritmos de aprendizaje supervisado encuentran su formalización en la teoría de la medida. Así, sean X, Y y Z espacios

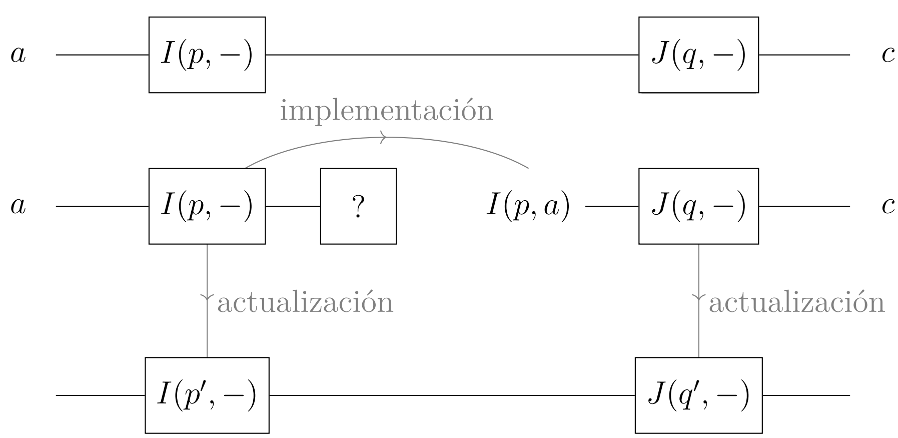
[1]Más adelante se vera como hacer una escogencia adecuada.

medibles donde Z es un subconjunto de $X \times Y$, esto es, Z expresa una relación entre los datos de X y Y. En una tarea de aprendizaje el objetivo es poder describir la relación Z a partir de una muestra de datos $S = (z^{(i)})_{i=1}^{m} \in Z^m$ y una función de perdida $L\colon M(X,Y) \times Z \to \mathbb{R}$, donde $M(X,Y)$ es el conjunto de funciones medibles de X en Y, esto implica hacer la elección de un conjunto de hipótesis $F \subset M(X,Y)$ y la construcción de un algoritmo de aprendizaje, es decir, un mapeo $A\colon \bigcup_{m\in\mathbb{N}} Z^m \to F$ que a partir de la muestra de datos S permita encontrar un modelo $f_s = A(s) \in F$ con *buen desempeño* en S y con *capacidad de generalizar* para los datos desconocidos $z \in Z$. El buen desempeño es medido con la función de perdida L y el valor especifico de la función de perdida $L(f_s, z)$ restringida a S. En este contexto generalizar significa que el desempeño del modelo encontrado sobre los datos observados sea similar al desempeño sobre los datos no observados.

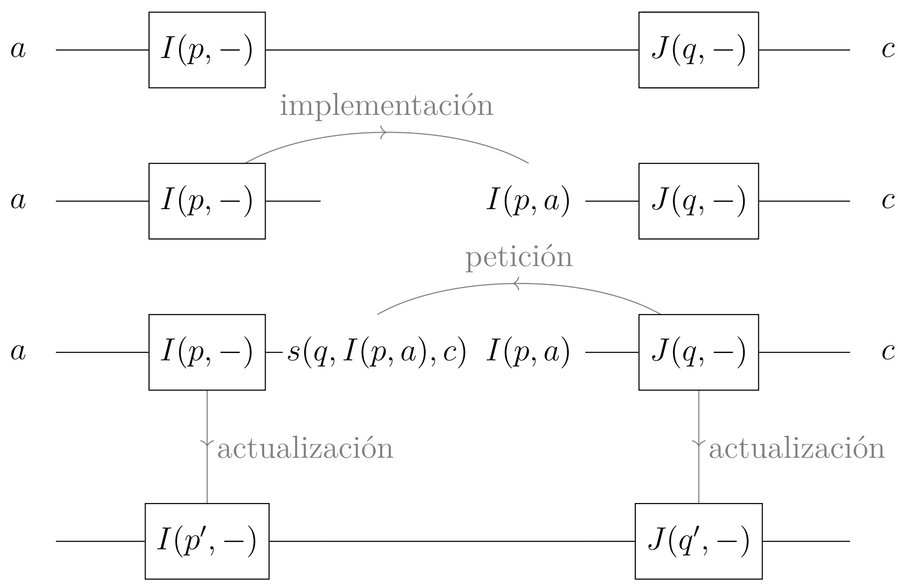
Así, se sigue de forma directa que la Definición 4.1.1 generaliza esta noción de algoritmo de aprendizaje, en efecto, tomando los elementos de la discusión anterior para formar el aprendiz $(\bigcup_{m\in\mathbb{N}} Z^m, f_s, A, r)$ donde $r\colon \bigcup_{m\in\mathbb{N}} Z^m \times X \times Y \to X$. Además, dado un aprendiz (P, I, U, R) de tipo $A \to B$ no se le exige que los conjuntos P, A y B sean medibles (no se le están aplicando restricciones y por tanto contienen al conjunto de los espacios medibles), más aun, no se tiene una función de perdida (no explícitamente) que restrinja la escogencia de la función de implementación I y la escogencia de la función de actualización U para asegurar un buen desempeño de el algoritmo de aprendizaje, es decir, según la discusión anterior, la función de actualización U no siempre mejora el desempeño de la predicción f_s. De acuerdo a esto, sin perdida de generalidad en este trabajo se tomara a un algoritmo de aprendizaje supervisado como un aprendiz.

Concretizando, un aprendiz (P, I, U, r) de tipo $A \to B$ tiene como objetivo encontrar la mejor aproximación posible a una función $f\colon A \to B$, para hacer esto el supervisor provee una lista de ejemplos $(a, b) \in A \times B$, donde cada ejemplo es una estimación de los valores tomados por f, es decir, se tiene que $b \approx f(a)$ (en particular puede ser que $b = f(a)$). El supervisor define también un espacio de funciones sobre las que el algoritmo debe buscar, esto se formaliza mediante la escogencia de un espacio de parámetros P y una función $I\colon P \times A \to B$. Así, la función I implementa un parámetro $p \in P$ como una función $I(p, -)\colon A \to B$. Entonces, dado un ejemplo $(a, b) \in A \times B$, el algoritmo toma la aproximación hipotética actual de f determinada por p, es decir, $I(p, -)$ y trata de mejorarla a una función $I(p', -)$ con el nuevo valor estimado del parámetro p'. Esto ultimo se formaliza con la función de actualización U, pues dado un ejemplo $(a, b) \in A \times B$ y el parámetro actual $p \in P$, U produce un *mejor* (idealmente, pero no necesariamente) parámetro $p' = U(p, a, b) \in P$, es decir, la idea es que la función actualizada $I(p', -)\colon A \to B$ estima el valor que toma la función en a mejor de lo que lo hace la función $I(p, -)$.

Con el objetivo de mostrar el porque la función de petición r para un aprendiz (P, I, U, r) es necesaria, se debe recordar que el objetivo es definir una categoría **`Learn`** con aprendices como morfismos, es decir, el objetivo es dotar a los aprendices con una estructura composicional. Así, si se tienen dos aprendices (P, I, U) y (Q, J, V) de tipo $A \to B$ y $B \to C$ respectivamente (donde se han eliminado las funciones de petición para hacer claro el porque se necesita), se quiere obtener un aprendiz (N, K, W) de tipo $A \to C$. Para el espacio de parámetros de el nuevo aprendiz se toma $N = P \times Q$ y por tanto las funciones K y W tienen los tipos $K\colon P \times Q \times A \to C$ y $W\colon P \times Q \times A \times C \to P \times Q$, y se define K con la asignación $K(p, q, a) = J(q, I(p, a))$. El problema surge cuando se quiere definir de forma explicita quien es W, para dotar a los aprendices con una estructura composicional, la nueva funcion de actualización W debe usar las funciones de actualización U y V, es decir, de los parámetros $(p, q, a, c) \in P \times Q \times A \times C$ que recibe la función de actualización W se deben producir ejemplos $(a', b') \in A \times B$ y $(b'', c'') \in B \times C$ para U y V respectivamente. Dado un ejemplo (a, c) para W es simple producir un ejemplo (b'', c'') para V, en efecto, se toma el par $(I(p, a), c) \in B \times C$. Pero no existe una forma natural de producir un par $(a', b') \in A \times B$ para U, porque la elección de b' debe codificar información de c y de J, pero no se ha especificado un mecanismo para esto. Gráficamente se tiene lo siguiente.

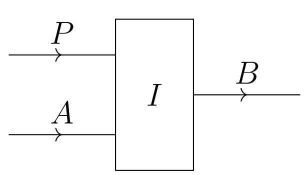


Así, para completar el marco composicióńal en los aprendices, se hace necesario que el segundo aprendiz pueda pasar al primer aprendiz elementos de B que contengan información de c y de J, es decir, se necesita una función $s\colon Q \times B \times C \to B$ que dado el parámetro actual $q \in Q$ y un ejemplo (b'', c'') retorne $b' := s(q, b'', c'')$, con el fin de producir el ejemplo (a, b') para la función U. Esta función es denominada precisamente función de petición. En el ejemplo anterior (y en general) se puede entender como que el primer aprendiz le pregunta al segundo aprendiz por el mejor valor de $b' \in B$ que puede usar U para el proceso de actualización del parámetro p y por tanto de la función I. En la sección de gradiente descendiente y *backpropagation* estas funciones de petición tomaran más sentido, pues se verá que ellas encapsulan la funcionalidad del *backpropagation*. Ahora gráficamente se tiene lo siguiente.

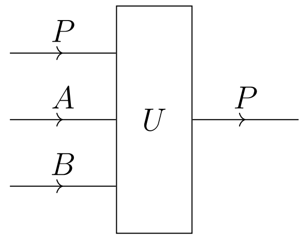


4.1.1. Diagramas de cuerdas

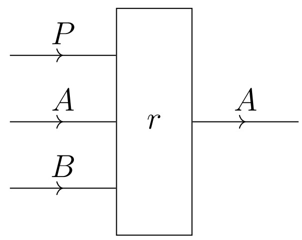
Para probar los siguientes resultados, resulta conveniente usar los diagramas de cuerda en la categoría Set, así, se puede representar a la función de implementación I como:



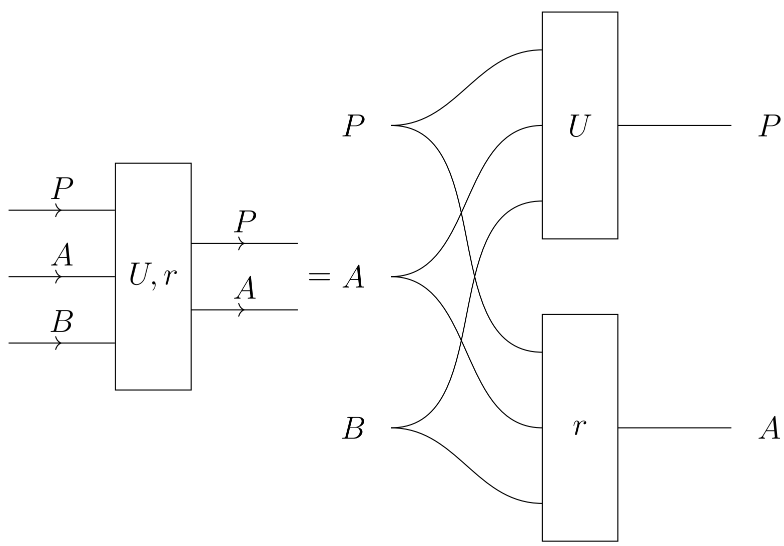
la función de actualización U como:



y la función de petición r como:



Pero es conveniente combinar U y r para representarlas en un solo diagrama vía la siguiente equivalencia



Definición 4.1.3. Se dirá que dos aprendices (P, I, U, r) y (P', I', U', r') de tipo $A \to B$ son equivalentes, si existe una función biyectiva $f \colon P \to P'$ tal que las

siguientes condiciones se cumplen para todo $p \in P$, $a \in A$ y $b \in B$:

$$\begin{aligned} I'(f(p), a) &= I(p, a), \\ U'(f(p), a, b) &= f(U(p, a, b)), \\ r(f(p), a, b) &= r(p, a, b). \end{aligned}$$

Proposición 4.1.4. La relación definida en el ejemplo anterior define una relación de equivalencia.

Demostración. Se probará que la relación cumple con ser reflexiva, simétrica y transitiva.

1. Reflexividad. Al tomar un aprendiz arbitrario (P, I, U, r) de tipo $A \to B$, esta propiedad se sigue de manera directa al tomar la función $id_P \colon P \to P$.

2. Simetría. Sean (P, I, U, r) y (Q, J, V, s) aprendices arbitrarios de tipo $A \to B$ tal que $(P, I, U, r) \sim (Q, J, V, s)$, es decir, existe una función biyectiva $f \colon P \to Q$ tal que para todo $p \in P$, $a \in A$ y $b \in B$

$$\begin{aligned} J(f(p), a) &= I(p, a) \\ V(f(p), a, b) &= f(U(p, a, b)) \\ s(f(p), a, b) &= r(p, a, b) \end{aligned} \tag{4.1}$$

como f es biyectiva, existe $f^{-1} \colon Q \to P$ tal que $f \circ f^{-1} = id_Q$ y $f^{-1} \circ f = id_P$, además, para todo $p \in P$ existe un $q \in Q$ tal que $f(p) = q$ y $f^{-1}(q) = p$. Así, de las expresiones en 4.1 se sigue que

$$\begin{aligned} J(q, a) &= I(f^{-1}(q), a) \\ f^{-1}(V(q, a, b)) &= U(f^{-1}(q), a, b) \\ s(q, a, b) &= r(f^{-1}(q), a, b) \end{aligned} \tag{4.2}$$

es decir, $(Q, J, V, s) \sim (P, I, U, r)$ y por tanto la relación es simétrica.

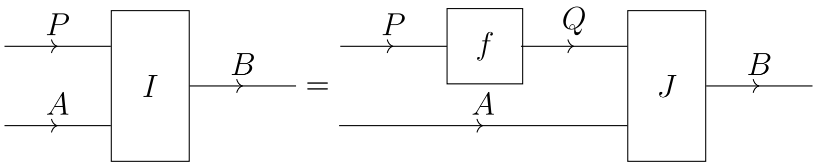
3. Transitividad. Sean (P, I, U, r), (Q, J, V, s) y (N, K, W, t) aprendices arbitrarios de tipo $A \to B$ tales que $(P, I, U, r) \sim (Q, J, V, s)$ y $(Q, J, V, s) \sim (N, K, W, t)$, es decir, existen biyecciones $f \colon P \to Q$ y $g \colon Q \to N$ tales que se cumplen las siguientes conjuntos de ecuaciones

$$\begin{aligned} J(f(p), a) &= I(p, a) \\ V(f(p), a, b) &= f(U(p, a, b)) \\ s(f(p), a, b) &= r(p, a, b) \end{aligned} \tag{4.3}$$

y,

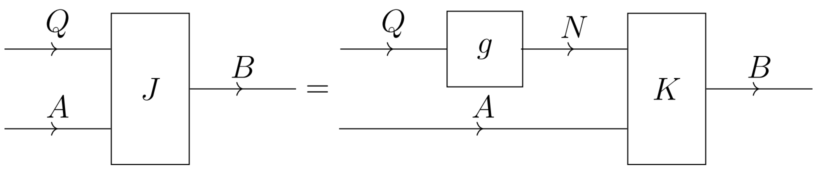
$$\begin{aligned} K(g(q), a) &= J(q, a) \\ W(g(q), a, b) &= g(V(q, a, b)) \\ t(g(q), a, b) &= s(q, a, b) \end{aligned} \tag{4.4}$$

De las ecuaciones 4.3 y 4.4 se obtienen los siguientes diagramas



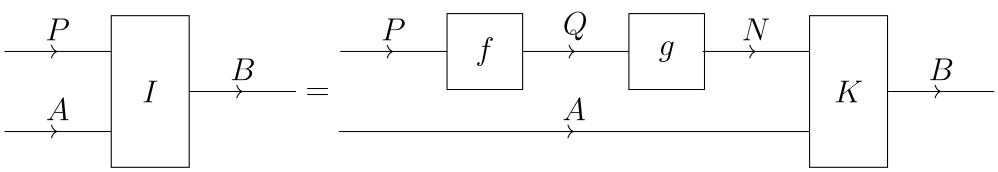
(4.5)

y,

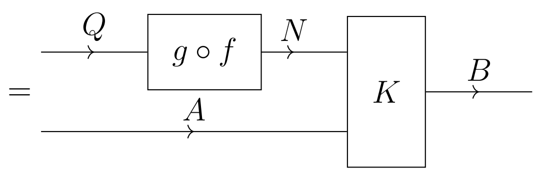


(4.6)

ahora, combinando 4.5 y 4.6



(4.7)



(4.8)

es decir, $K(g(f(p)), a) = I(p, a)$ y de forma análoga se sigue que $W(g(f(p)), a, b) = g(f(U(p, a, b)))$ y que $t(g(f(p)), a, b) = r(p, a, b)$, así, la relación es transitiva.

Por tanto, la relación definida en el ejemplo anterior es una relación de equivalencia. $\square$

Teorema 4.1.5. Existe una categoría monoidal simétrica **Learn** cuyos objetos son conjuntos y cuyos morfismos son clases de equivalencia de aprendices.

Demostración. Para probar el teorema se debe mostrar de forma explicita cual es la composición de aprendices y para cada objeto $A \in$ **Learn** cual es el morfismo identidad, esto asegura que **Learn** es en efecto una categoría. Y para probar que **Learn** tiene estructura monoidal simétrica se debe especificar el producto tensorial, y el trenzado de este producto tensorial que la hace simétrica.

1. (Composición) Supóngase que se tienen dos aprendices (P, I, U, r) y (Q, J, V, s) de tipos $A \to B$ y $B \to C$, esquemáticamente

$$A \xrightarrow{(P,I,U,r)} B \xrightarrow{(Q,J,V,s)} C$$

el aprendiz composición de tipo $A \to C$ es denotado por $(P \times C, I * J, U * V, r * s)$, donde la función implementación $I * J$ esta definida por la asignación

$$(I * J)(p, q, a) := J(q, I(p, a))$$

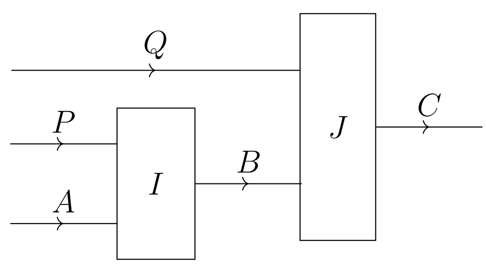
la función de actualización es

$$(U * V)(p, q, a, c) := (U(p, a, s(q, I(p, a), c)), V(q, I(p, a), c))$$

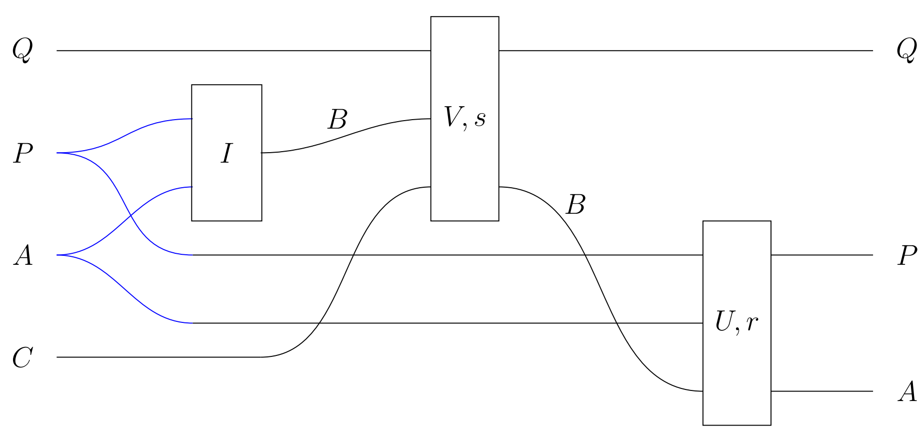
y la función de petición esta dada por

$$(r * s)(p, q, a, c) = r(p, a, s(q, I(p, a), c))$$

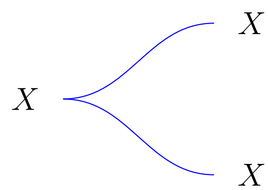
Pictóricamente se tiene, para la composición de las funciones de implementación



y para las composiciones de las funciones de actualización y las funciones de petición.

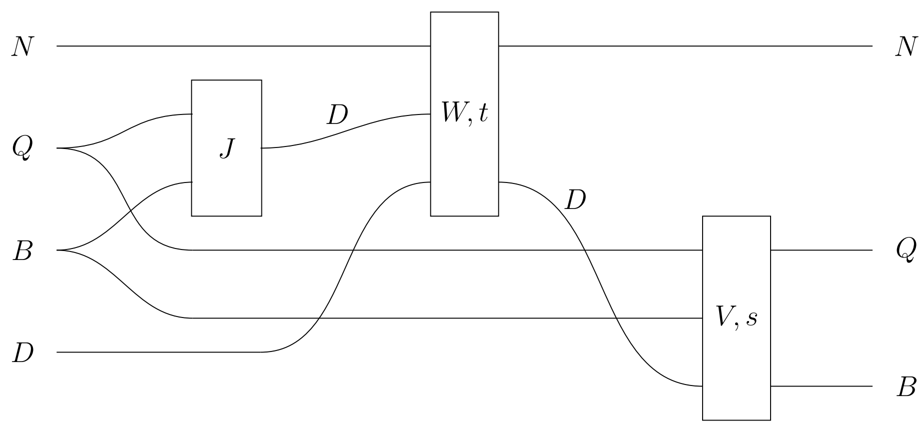


Donde el diagrama

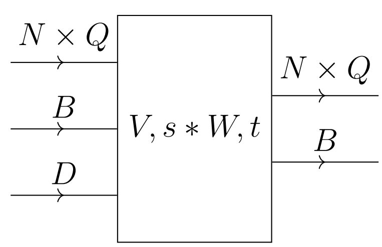


representa al mapeo diagonal $X \to X \times X$ que asigna $x \mapsto (x, x)$. Ahora se verá que esta composición es asociativa. La asociatividad de la composición en las funciones de implementación es fácil de verificar, en efecto, solo se debe verificar que la representación de diagramas de cuerda para $I * (J * K)$ y $(I * J) * K$ coinciden, por tanto solo se hará la verificación para las funciones de actualización y para las funciones de petición.

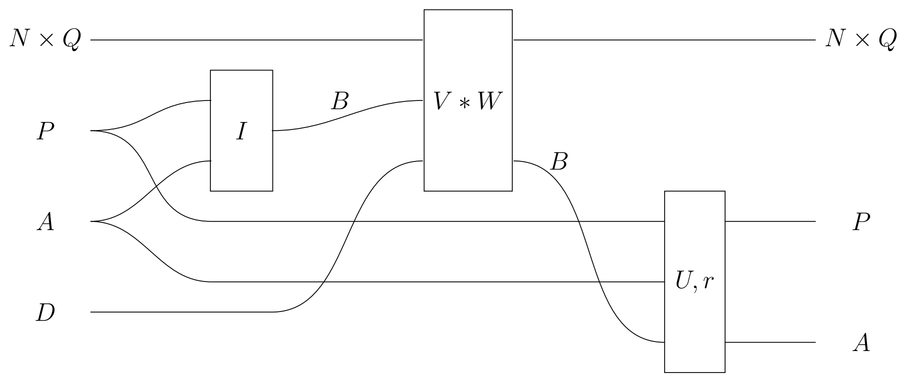
Para probar la asociatividad se tomaran 3 aprendices (P, I, U, r), (Q, J, V, s) y (N, K, W, t) de tipo $A \to B$, $B \to C$ y $C \to D$ respectivamente, se quiere ver que $U * (V * W) = (U * V) * W$ y que $r * (s * t) = (r * s) * t$. Considerando la expresión $U * (V * W)$, para $V * W$ de acuerdo a la definición se tiene



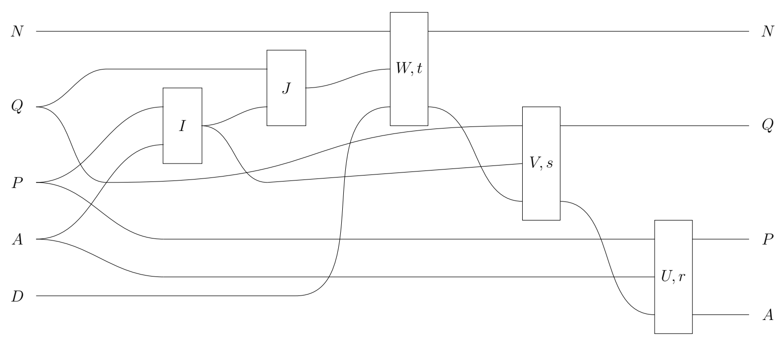
y simplificando el diagrama



ahora, aplicando la definición de la composición a U y $(V * W)$ para obtener $U * (V * W)$

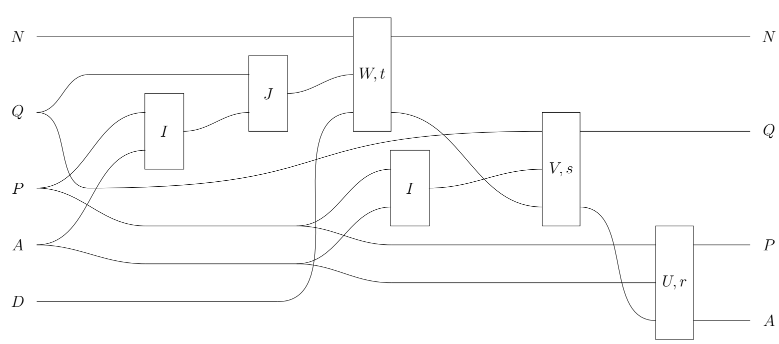


dado que $P \times (Q \times N) \cong (P \times Q) \times N \cong P \times Q \times N$ se puede expandir el diagrama anterior para obtener



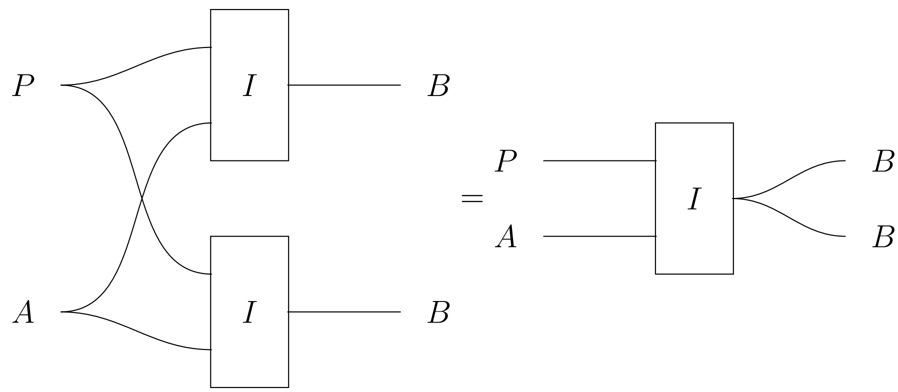
(4.9)

de forma análoga se llega a que el diagrama de cuerda asociado a $(U * V) * W$ es



(4.10)

ahora, para representar $(I(p,a), I(p,a))\colon P \times A \to B \times B$ se tiene la siguiente equivalencia



por tanto, los diagramas 4.9 y 4.10 son equivalentes, así que, se cumple la asociatividad para las funciones de actualización y para las funciones de petición.

2. (Morfismo identidad) Para cada objeto A, se tiene el mapeo identidad

$$(\mathbb{R}^0, id, !, \pi_2)\colon A \to A$$

donde en realidad $id = \pi^2_{\mathbb{R}^0 \times A}$ es la proyección en la segunda componente, pero como $\mathbb{R}^0 \times A \cong A$ se abusará de la notación y se escribirá id ignorando a $\mathbb{R}^0$, pues $\pi^2_{\mathbb{R}^0 \times A}$ actúa como la función identidad en la segunda componente. ! es el único morfismo que existe de $\mathbb{R}^0 \times A \times A$ en $\mathbb{R}^0$ y $\pi_2\colon \mathbb{R}^0 \times A \times A \to A$ es la proyección en el segundo factor (ignorando a $\mathbb{R}^0$ como en $\pi^2_{\mathbb{R}^0 \times A}$). Para demostrar que $(\mathbb{R}^0, id, !, \pi_2)$ es en realidad la identidad tómese un aprendiz (P, I, U, r), así, se tiene que la composición $(P, I, U, r) * (\mathbb{R}^0, id, !, \pi_2)$ tiene como espacio de parámetros a $P \times \mathbb{R}^0 \cong P$, como función de implementación tiene a $(I * id)(r, p, a) = id(r, I(p, a)) = I(p, a)$, como función de actualización tiene a $(U*!)(r, p, a, b) = U(p, a, \pi_2(r, I(a, b), b)) = U(p, a, b)$ y por último, la función de petición está dada por $(r * \pi_2)(r, p, a, b) = r(p, a, \pi_2(r, I(p, a), b)) = r(p, a, b)$. Análogamente se prueba que también es identidad a izquierda para los aprendices con tipos adecuados.

3. (Producto monoidal) Hasta el momento se ha probado que **`Learn`** es una categoría, ahora note lo siguiente, existe un funtor de la categoría Set en la categoría **`Learn`**, donde este funtor es el mapeo identidad sobre los objetos y para conjuntos A y B la función $f\colon A \to B$ es mapeada en la función trivialmente parametrizada $\bar{f}\colon \mathbb{R}^0 \times A \to B$ es decir, el funtor mapea como la

identidad sobre los morfismos (porque $\mathbb{R}^0 \times B \cong B$ para todo conjunto B). Recuerde que (Set, $\times$) es una categoría monoidal simétrica y por tanto **`Learn`** puede heredar la estructura monoidal simétrica si se define el producto tensorial como en Set (producto cartesiano para los objetos, que son conjuntos y para los learnes se define el producto tensorial componente a componente como el producto tensorial de las funciones en la categoría Set), dado que un funtor preserva diagramas conmutativos, las restricciones de asociatividad α de Set, y las restricciones de la unidad ρ y λ de Set, son heredadas por Learn vía el funtor que las transforma en $\bar{\alpha}$, $\bar{\rho}$ y $\bar{\lambda}$ respectivamente.

De acuerdo a lo anterior, **`Learn`** tiene estructura monoidal si se define el producto tensorial sobre objetos A y B como el producto cartesiano $A \times B$. Y para morfismos $(P, I, U, r)\colon A \to B$ y $(Q, J, V, s)\colon B \to C$ se define el producto como $(P \times Q, I \parallel J, U \parallel V, r \parallel s)$ donde la función de implementación es

$$(I \parallel J)(p, q, a, c) := (I(p, a), J(q, c))$$

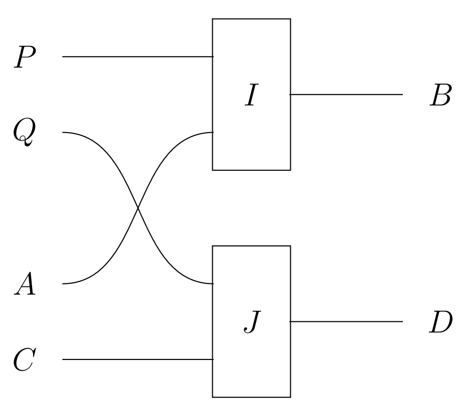
la función de actualización es

$$(U \parallel V)(p, q, a, c, d, b) := (U(p, a, b), J(q, c, d))$$

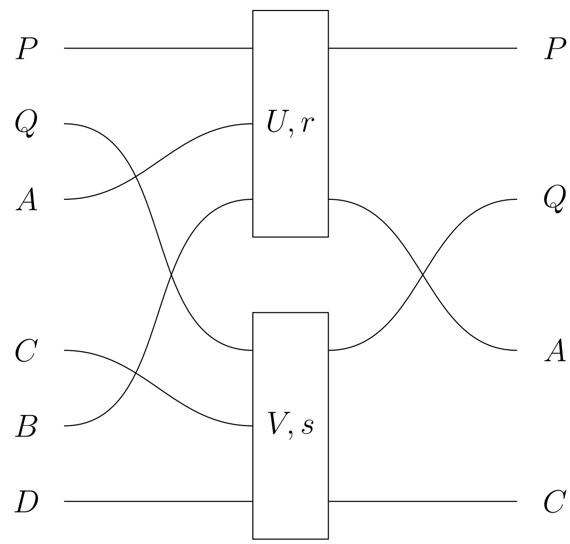
y la función de petición es

$$(r \parallel s)(p, q, a, c, b, d) := (r(p, a, b), s(q, c, d))$$

Nótese que el producto sobre los morfismos está actuando como el producto cartesiano de funciones componente a componente. Gráficamente el producto tensorial sobre las funciones de implementación es como sigue



y el producto tensorial sobre las funciones de actualización y de petición luce como sigue



4. (Trenzado) Para asegurar que la estructura monoidal de **Learn** es simétrica, se debe definir el trenzado de **Learn** como el trenzado de Set (y asi las condiciones de conmutatividad son heredadas por **Learn**, via el funtor que existe entre Set y **Learn**), específicamente el trenzado simétrico de tipo $A \times B \to B \times A$ esta dado por $(\mathbb{R}^0, \sigma, !, \sigma \circ \pi)$ donde $\sigma \colon A \times B \to B \times A$ es la función que permuta el producto, asignando (a, b) en (b, a) y $\pi \colon \mathbb{R}^0 \times (A \times B) \times (B \times A) \to B \times A$ es nuevamente la proyección en el segundo factor (ignorando a $\mathbb{R}^0$)

Por tanto **Learn** es una categoría monoidal simétrica. □

4.2. El descenso del gradiente y el *backpropagation*

En esta sección se verá que el gradiente descendiente y el *backpropagation* definen un funtor monoidal simétrico de la categoría monoidal simétrica Para (de funciones parametrizadas diferenciables) a la categoría monoidal simétrica **Learn**.

Definición 4.2.1. Sean A y B espacios euclidianos. Se define una **función parametrizada diferenciables** de tipo $A \to B$ como un par (P, I), donde P es un espacio euclidiano y $I \colon P \times A \to B$ es una función diferenciable.

Definición 4.2.2. Se dice que dos funciones parametrizadas diferenciables (P, I) y (P', I') de tipo $A \to B$ son equivalentes, si existe una función biyectiva $f \colon P \to P'$ tal que para todo $p \in P$ y $a \in A$ se tiene que $I'(f(p), a) = I(p, a)$.

Proposición 4.2.3. La relación de la definición anterior, es una relación de equivalencia.

Demostración. Note que este es un caso particular de la Proposición 4.1.4 si se ignoran las funciones de actualización y de petición. $\square$

Observacion 4.2.4. Se abusará de la notación y se escribirá $I\colon P \times A \to B$ para representar a la función parametrizable diferenciable (P, I).

Definición 4.2.5. Se define la categoría **`Para`** como la categoría cuyos objetos son espacios euclidianos de dimensión finita y cuyos morfismos $\mathbb{R}^n \to \mathbb{R}^m$ para dos objetos $\mathbb{R}^n$ y $\mathbb{R}^m$, son clases de equivalencias de funciones parametrizadas diferenciables de tipo $\mathbb{R}^n \to \mathbb{R}^m$. Donde

1. La composición de dos morfismos $(P, I)\colon \mathbb{R}^n \to \mathbb{R}^m$ y $(Q, J)\colon \mathbb{R}^m \to \mathbb{R}^l$ está dada por $(P \times Q, I * J)$ con

$$(I * J)(p, q, a) := J(q, I(p, a))$$

2. Para cada objeto $\mathbb{R}^n$, se tiene el mapeo identidad

$$(\mathbb{R}^0, id)\colon \mathbb{R}^n \to \mathbb{R}^n$$

   donde se ignora a $\mathbb{R}^0$

Proposición 4.2.6. **`Para`** es una categoría, más aun, es una categoría monoidal simétrica.

Demostración. Al dotar a **`Para`** con el producto tensorial definido sobre dos objetos $\mathbb{R}^n$ y $\mathbb{R}^m$ como $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ y sobre dos morfismos $(P, I)\colon \mathbb{R}^n \to \mathbb{R}^m$ y $(Q, J)\colon \mathbb{R}^l \to \mathbb{R}^k$ como $(P \times Q, I \parallel J)$, donde

$$(I \parallel J)(p, q, a, c) = (I(p, a), J(q, c))$$

Y definir el trenzado de tipo $\mathbb{R}^n \parallel \mathbb{R}^m \to \mathbb{R}^m \parallel \mathbb{R}^n$ denotado por $(\mathbb{R}^0, \sigma)$ donde sigma esta definida por la asignación $\sigma(a, b) = (b, a)$, es claro que esta proposición es un caso particular de el Teorema 4.1.5, así la prueba se realiza de forma análoga ignorando las consideraciones que se hacen en el Teorema 4.1.5 sobre las funciones de actualización y las funciones de petición. $\square$

Teorema 4.2.7. Sea $\epsilon > 0$ un numero real fijo, $\alpha\colon \mathbb{N} \to \mathbb{R}_{>0}$ y $e\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ una función diferenciable tal que $\frac{\partial e}{\partial x}(x_0, -)\colon \mathbb{R} \to \mathbb{R}$ es invertible para todo $x_0 \in \mathbb{R}$. Entonces se puede definir un funtor fiel, inyectivo sobre objetos, y con estructura monoidal simétrica

$$L_{\epsilon,e}\colon \mathbf{Para} \to \mathbf{Learn}$$

que mapea cada función parametrizada $I\colon P \times A \to B$ en el aprendiz (P, I, U_I, r_I) donde

$$U_I(p, a, b) := p - \epsilon \nabla_p E_I(p, a, b)$$

y

$$r_I(p, a, b) := f_a(\frac{1}{\alpha_B} \nabla_a E_I(p, a, b))$$

con $E_I(p, a, b) := \alpha_B \sum_j e(I_j(p, a), b_j)$, α_B un valor específico de la función α y f_a la aplicación componente a componente de la inversa de $\frac{\partial e}{\partial x}(a_i, -)$ para cada i. Donde I_j denota la componente j de $I\colon \mathbb{R}^n \to \mathbb{R}^m$.

Demostración. Notese que el funtor $L_{\epsilon,e}$ toma una funcion parametrizada (P, I) y le adjunta dos componenetes U_I y r_I para obtener un aprendiz (P, I, U_I, r_I), es decir, $L_{\epsilon,e}$ inyectivo sobre objetos y sobre morfismos. Por tanto, si $L_{\epsilon,e}$ es funtorial, entonces adicionalmente es fiel e inyectiva sobre objetos.

Para probar la funtorialidad se mostrara que la composiciones de funciones parametrizadas es mapeada en la composición de los learnes asociados y que la identidad de un conjunto A es mapeada en el aprendiz identidad de A en la categoría **Learn**. Así, tomando las funciones parametrizadas $I\colon P \times \mathbb{R}^n \to \mathbb{R}^m$ y $J\colon Q \times \mathbb{R}^m \to \mathbb{R}^l$, se tiene para.

1. (funciones de implementación). Se sigue de inmediato, dado que las funciones de implementación asociadas a I y J son las mismas funciones en la categoría **Para** y en la categoría **Learn**, es decir, se hereda la asociatividad.

2. (funciones de actualización). Por definición la composición para las funciones de actualización U_I y U_J asociadas a I y a J está definida por

$$(U_I * U_J)(p, q, a, c) = (U_I(p, a, r_J(q, I(p, a), c)), U_J(q, I(p, a), c))$$

que reemplazando la definición de U_I para la función parametrizada I se obtiene

$$(U_I * U_J)(p, q, a, c) = (p - \epsilon\nabla_p E_I(p, a, r_J(q, I(p, a), c)), q - \epsilon\nabla_q E_J(q, I(p, a), c))$$

mientras que la función de actualización de la composición $I * J$ esta dada por

$$U_{I*J}(p,q,a,c) = (p - \epsilon \nabla_p E_{I*J}(p,q,a,c), q - \epsilon \nabla_q E_{I*J}(p,q,a,c))$$

así, para demostrar que $U_I * U_J = U_{I*J}$ se debe demostrar que las siguientes igualdades se cumplen

$$\nabla_p E_I(p,a,r_J(q,I(p,a),c)) = \nabla_p E_{I*J}(p,q,a,c) \tag{4.11}$$

$$\nabla_q E_J(q,I(p,a),c)) = \nabla_q E_{I*J}(p,q,a,c)) \tag{4.12}$$

tomando la ecuación 4.11 y reemplazando E_I en el termino de la izquierda de acuerdo a la definición se obtiene

$$\nabla_p E_I(p,a,r_J(q,I(p,a),c)) = \nabla_p \alpha_B \sum_i e(I_i(p,a), r_J(q,I(p,a),c)_i)$$

pero $\nabla_p = \frac{\partial}{\partial p} = \frac{\partial I}{\partial p}\frac{\partial}{\partial I}$, por tanto se llega a

$$(\alpha_B \sum_i \frac{\partial e}{\partial x}(I_i(p,a), r_J(q,I(p,a),c)_i)\frac{\partial I_i}{\partial p_l}(p,a))_l$$

ahora, reemplazando la definición de r_J teniendo en cuenta que $I(p,a) = b$ para algún $b \in B$, se tiene

$$(\alpha_B \sum_i \frac{\partial e}{\partial x}(I_i(p,a), f_{I(p,a)}(\frac{1}{\alpha_B}\nabla_b E_J(q,I(p,a),c))_i)\frac{\partial I_i}{\partial p_l}(p,a))_l$$

pero recordando que f es la aplicación de la inversa de $\frac{\partial e}{\partial}(x,-)$ componente a componente, se llega a

$$(\alpha_B \sum_i \frac{1}{\alpha_B}(\nabla_b E_J(q,I(p,a),c))_i \frac{\partial I_i}{\partial p_l}(p,a))_l$$

en la que reemplazando nuevamente la definición de E_J implica

$$(\alpha_B \sum_i \frac{1}{\alpha_B}(\nabla_b \alpha_C \sum_j e(J_j(q,I(p,a)),c_j))_i \frac{\partial I_i}{\partial p_l}(p,a))_l$$

ahora, recordando nuevamente que $\nabla_b = \frac{\partial}{\partial b} = \frac{\partial J}{\partial b}\frac{\partial}{\partial J}$ se llega a

$$(\sum_i \alpha_C \sum_j \frac{\partial e}{\partial x}(J_j(q,I(p,a)),c_j)\frac{\partial J_j}{\partial b_i}(q,I(p,a))\frac{\partial I_i}{\partial p_l}(p,a))_l$$

pero nótese que por regla de la cadena se tiene la siguiente igualdad

$$\sum_i \frac{\partial J_j}{\partial b_i}(q, I(p,a))\frac{\partial I_i}{\partial p_l}(p,a) = \frac{\partial J_j}{\partial p_l}(q, I(p,a))$$

y por tanto se obtiene

$$(\alpha_C \sum_j \frac{\partial e}{\partial x}(J_j(q, I(p,a)), c_j)\frac{\partial J_j}{\partial p_l}(q, I(p,a)))_l$$

tomando nuevamente la definición de ∇_p, se llega a

$$\nabla_p \alpha_C \sum_j e(J_j(q, I(p,a)), c_j)$$

finalmente, usando la definición de E_{I*J} se tiene que

$$\nabla_p E_I(p, a, r_J(q, I(p,a), c)) = \nabla_p E_{I*J}(p, q, a, c)$$

Ahora, para la ecuación 4.12 se sigue de manera directa al usar la definición de E_I y la definición de $I * J$, así

$$\begin{aligned} E_J(q, I(p,a), c) &= \alpha \sum_i e(J(q, I(p,a))_i, c_i) \\ &= E_{I*J}(p, q, a, c) \end{aligned}$$

y por tanto $U_I * U_J = U_{I*J}$

3. (funciones de petición) En este caso se debe probar que

$$r_I(p, a, r_J(q, I(p,a), c)) = r_{I*J}(p, q, a, c) \tag{4.13}$$

reemplazando por las definiciones de r_I y r_{I*J} se obtiene que

$$\begin{aligned} r_I(p, a, r_J(q, I(p,a), c)) &= f_a(\frac{1}{\alpha_B}\nabla_a E_I(p, a, r_J(q, I(p,a), c))) \\ &= f_a(\frac{1}{\alpha_B}\nabla_a E_{I*J}(p, q, a, c)) \\ &= r_{I*J}(p, q, a, c) \end{aligned}$$

es decir, se necesita probar la igualdad

$$f_a(\frac{1}{\alpha_B}\nabla_a E_I(p,a,r_J(q,I(p,a),c))) = f_a(\frac{1}{\alpha_B}\nabla_a E_{I*J}(p,q,a,c))$$

más específicamente, se debe probar que

$$\nabla_a E_I(p,a,r_J(q,I(p,a),c)) = \nabla_a E_{I*J}(p,q,a,c) \tag{4.14}$$

pero nótese que este es el mismo caso que el de la ecuación 4.11, solo que se ha cambiado a p por a, y por lo tanto, de forma análoga se sigue que la igualdad 4.14 se cumple.

Esto demuestra que el funtor respeta la regla de composición de ambas categorías. Ahora se vera que el morfismo identidad para un objeto A en la categoría **Para** es mapeado al morfismo identidad del objeto A en la categoría **Learn**. La identidad de A en **Para** es la función $id_A\colon \mathbb{R}^0 \times A \to A$, como el conjunto de parámetros es trivial, la función de implementación de A (en **Learn**), esto es, la imagen de id_A bajo la aplicación de $L_{\epsilon,e}$ es la identidad en A, además, como el conjunto de parámetros es trivial, la función de actualización es única y por tanto denotada por !. Para la función de petición, de acuerdo a la definición se tiene lo siguiente

$$\begin{aligned} r_{id_A}(0,a,b) &= f_a(\frac{1}{\alpha_B}\nabla_a \sum_i e(a_i,b_i)) \\ &= (f_a(\frac{\partial e}{\partial a_i}(a_i,b_i)))_i \\ &= b \end{aligned}$$

y por tanto, la identidad id_A es mapeada exactamente en la identidad $(\mathbb{R}^0, id_A, !, \pi_2)$ de A en la categoría **Learn**.

Lo anterior demuestra que $L_{\epsilon,e}$ es un funtor. Ahora se vera que $L_{\epsilon,e}$ tiene estructura monoidal, es decir, que dadas dos funciones parametrizadas $I\colon P\times A \to B$ y $K\colon N\times C \to \mathscr{D}$ el aprendiz definido por el producto tensorial $I \parallel J$ vía el funtor $L_{\epsilon,e}$ es igual al aprendiz definido por el producto tensorial de los aprendices definidos por I y J vía el funtor $L_{\epsilon,e}$, salvo el isomorfismo $\mathbb{R}^n \times \mathbb{R}^m \cong \mathbb{R}^{n+m}$. Así, supóngase los aprendices (P, I, U_I, r_I) y (N, K, U_K, r_K) definidos por las funciones parametrizadas $I\colon P\times A \to B$ y $K\colon N\times C\to D$ mediante la aplicación del funtor $L_{\epsilon,e}$ cuyo producto tensorial es $I \parallel K\colon (P\times N)\times(A\times C) \to B\times D$. Note que $E_{I\parallel J}(p,q,a,c,b,d) = E_I(p,a,b) + E_J(q,c,d)$, dado que para obtener el valor de la función de implementación $I \parallel J$

se calcula por separado el valor de I y de J (el producto tensorial es el producto cartesiano), por tanto, los errores se calculan también por separado y el error total es la suma de los errores. Ahora, aplicando este hecho y las definiciones se tiene que

$$\begin{aligned} U_{I\|J}(p,q,a,c,b,d) &= (p - \epsilon\nabla_p E_{I\|J}(p,q,a,c,b,d), q - \epsilon\nabla_q E_{I\|J}(p,q,a,c,b,d)) \\ &= (p - \epsilon\nabla_p E_I(p,a,b), q - \epsilon\nabla_q E_I(q,c,d)) \\ &= (U_I(p,a,b), U_J(q,c,d)) \end{aligned}$$

y por último, para las funciones de petición se tiene

$$\begin{aligned} r_{I\|J}(p,q,a,c,b,d) &= f_{(a,c)}(\frac{1}{\alpha_{B\times D}}\nabla_{(a,c)}E_{I\|J}(p,q,a,c,b,d)) \\ &= f_{(a,c)}(\frac{1}{\alpha_{B\times D}}\nabla_{(a,c)}\alpha_{B\times D}(\sum_i e(I(p,a)_i, b_i) \\ &+ \sum_j e(J(q,c)_j, d_j)) \\ &= (f_a(\frac{1}{\alpha_B}\nabla_a E_I(p,a,b)), f_c(\frac{1}{\alpha_D}\nabla_c E_J(q,c,d))) \\ &= (r_I(p,b,a), r_J(q,c,d)) \end{aligned}$$

Así, la imagen del tensor es el tensor de la imagen y por lo tanto $L_{\epsilon,e}$ es un funtor fiel, inyectivo sobre objetos con estructura monoidal simétrica. $\square$

Corolario 4.2.8. Sea $\epsilon > 0$ un numero real fijo y $e\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ una función diferenciable tal que $\frac{\partial e}{\partial x}(x_0, -)\colon \mathbb{R}\to\mathbb{R}$ es invertible para todo $x_0\in\mathbb{R}$. Entonces se puede definir un funtor fiel, inyectivo sobre objetos, y con estructura monoidal simétrica

$$L_{\epsilon,e}\colon \mathbf{Para} \to \mathbf{Learn}$$

que mapea cada función parametrizada $I\colon P\times A\to B$ en el aprendiz (P, I, U_I, r_I) donde

$$U_I(p,a,b) := p - \epsilon\nabla_p E_I(p,a,b)$$

y

$$r_I(p,a,b) := f_a(\nabla_a E_I(p,a,b))$$

con $E_I(p,a,b) := \sum_j e(I_j(p,a), b_j)$ y f_a la aplicación componente a componente de la inversa de $\frac{\partial e}{\partial x}(a_i, -)$ para cada i.

Demostración. Se sigue de el Teorema 4.2.7 al tomar a α como una función de valor constante r, para algún $r\in\mathbb{R}_{>0}$. $\square$

Ejemplo 4.2.9. Como se menciono en el Ejemplo 4.1.2, existe una forma adecuada de escoger la función de petición (mas aun, de escoger la función de actualización) y esta esta dada por el teorema anterior. Así, al tomar la función $e\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ como el error cuadrático $e(x,y) := \frac{1}{2}(x-y)^2$ el error total estará dado por

$$E_I(p,a,b) = \frac{1}{2}\sum_j (I_j(p,a)-b_j)^2 = \frac{1}{2}\|I(p,q)-b\|$$

Luego, $\frac{\partial e}{\partial x}(x_0,-)(y) = \frac{\partial e}{\partial x}(x_0,y) = x_0 - y$ y esta función es su propia inversa, por lo tanto $f_{x_0}(y) := x_0 - y$. Luego, al fijar una taza de aprendizaje $\epsilon > 0$, las funciones de actualización y de petición para cualquier función parametrizada $I\colon \mathbb{R}^n\times\mathbb{R}\to\mathbb{R}$ quedan determinadas por

$$\begin{aligned} U_I(p,a,b) &:= p - \epsilon\nabla_p E_I(p,a,b) \\ &= (p_k - \epsilon\sum_j (I_j(p,a)-b_j)\frac{\partial I_j}{\partial p_k}(p,a))_k \end{aligned}$$

y

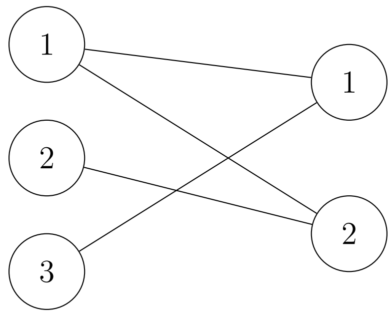
$$\begin{aligned} r_I(p,a,b) &:= f_a(\nabla_a E_I(p,a,b)) \\ &= (a_i - \sum_j (I_j(p,a)-b_j)\frac{\partial I_j}{\partial a_i}(p,a))_i \end{aligned}$$

Observacion 4.2.10. Como consecuencia del teorema se hace claro que en general, para un aprendiz (P,I,U,r), la escogencia de la función de actualización y la escogencia de la función de petición no es arbitraria, en efecto, el teorema muestra que la función de actualización U encapsula la funcionalidad del gradiente descendiente y que la función de petición r encapsula la funcionalidad del *backpropagation*.

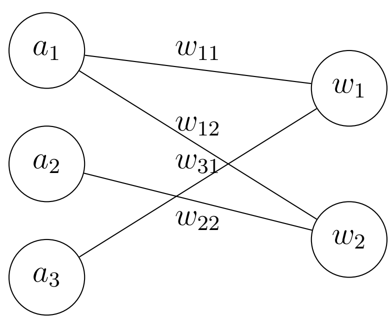
4.3. La categoría NNet

Definición 4.3.1. Sean n y m en $\mathbb{N}$. Una **red neuronal** de 1 **capa** de tipo (n,m) es un subconjunto $C\subseteq [n]\times[m]$, donde $[n]=\{1,\ldots,n\}$ y $[m]=\{1,\ldots,m\}$. Una **red neuronal** de k **capas** de tipo (n,m) es una secuencia de redes neuronales de una capa de tipos $(n_0,n_1),(n_1,n_2),\ldots,(n_{k-1},n_k)$, donde $n_0=n$ y $n_k=m$. Una **red neuronal** de tipo (n,m), es una red neuronal de k capas para algún k.

En una red neuronal de 1 capa $C\subseteq[n]\times[m]$, n y m representan el numero de nodos (neuronas) en cada lado de la red (también en redes neuronales de k capas) y el subconjunto C representa las conexiones, es decir, si $(i,j) \in C$ quiere decir que el nodo en la posición i del lado izquierdo de la red esta conectado con el nodo en la posición j del lado derecho de la red, como se ilustra en la siguiente figura.



En este caso se tiene una red neuronal de 1 capa de tipo $(3,2)$ donde $C = \{(1,1),(1,2),(2,2),(3,1)\}$. Nótese además, que si se fija una función adicional $\sigma \colon \mathbb{R} \to \mathbb{R}$ (a la que más adelante se le llamara función de activación) entonces una red neuronal de 1 capa define una función parametrizada $I \colon \mathbb{R}^{|C|+n_2} \times \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ donde $\mathbb{R}^{|C|}$ encapsula la información de unos números a los que se les llamara pesos y $\mathbb{R}^{n_2}$ encapsula la información de unos números a los que se les llamara umbrales. Así, si se toma la red neuronal mencionada en el párrafo anterior se tiene que la gráfica



define la función parametrizada $I \colon \mathbb{R}^6 \times \mathbb{R}^3 \to \mathbb{R}^2$ definida por la asignación

$$I(w_{11}, w_{12}, w_{22}, w_{31}, w_1, w_2, a_1, a_2, a_3) = (\sigma(w_{11}a_1 + w_{31}a_3 + w_1), \sigma(w_{12}a_1 + w_{22}a_2 + w_2))$$

En realidad, no solo las redes neuronales de 1 capa definen una función parametrizada, esta construcción puede ser extendida a redes neuronales con un numero arbitrario de capas. Pero antes de mostrar cómo, se vera que las redes neuronales forman una categoría.

Definición 4.3.2. La categoría **NNet** de redes neuronales, tiene como objetos números naturales y como morfismos $n \to m$ a las redes neuronales de tipo (n, m)

para n y m dos números naturales. La composición de morfismos esta dada por la concatenación de redes neuronales y el morfismo identidad para un objeto n es la red neuronal de 0 capas.

Dado que la composición es la concatenación (al igual que se hizo en la demostración del teorema de estrictificación de las categorías monoidales), se sigue de inmediato que esta composición es asociativa y por lo tanto **NNet** es una categoría. En el siguiente resultado se mostrara como definir una función parametrizada (de manera funtorial) para cualquier red neuronal.

Teorema 4.3.3. Dada una función diferenciable $\sigma \colon \mathbb{R} \to \mathbb{R}$, se tiene un funtor

$$I^{\sigma} \colon \mathbf{NNet} \to \mathbf{Para}$$

tal que sobre objetos I^{σ} mapea a el numero natural n en el espacio euclidiano de dimensión $\mathbb{R}$. Sobre morfismos cada red neuronal de 1 capa $C \colon m \to n$ es mapeada a la función parametrizada $I^{\sigma}_C \colon \mathbb{R}^{|C|+n} \times \mathbb{R}^m \to \mathbb{R}^n$, definida por la asignación

$$((w_{ji}, w_j), x_i)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \mapsto (\sigma(\sum_i (w_{ji} x_i + w_j)))_{1 \leq j \leq n}$$

σ es llamada **función de activación**, los w_{ji} son llamados **pesos** donde $(i, j) \in C$ y los w_j son llamados **umbrales**, donde $j \in [n]$

Demostración. Dado que una red neuronal arbitraria de tipo (m, n) se puede escribir como la concatenación de múltiples redes neuronales de 1 capa, por definición la composición de redes neuronales es mapeada en la composición de funciones parametrizadas, pero vale la pena mencionar que dado que todas las funciones involucrada son diferenciables, también lo es la imagen de una red neuronal bajo el funtor I^{σ} y por tanto, también lo son las composiciones. Nótese además, que el 0 es mapeado en $\mathbb{R}^0$ y que la red neuronal de 0 capas, es la composición (de capas) vacía lo que implica que I^{σ} mapea identidades en identidades. $\square$

Observacion 4.3.4. La funtorialidad de I^{σ} implica que para una red neuronal de k capas $N = C_1, \ldots, C_k$ la imagen bajo el funtor, es la composición de la imagen de cada capa, así

$$I^{\sigma}_N = I^{\sigma}_{C_1} * \cdots * I^{\sigma}_{C_k}$$

Del resultado anterior es valido afirmar entonces, que las redes neuronales son un lenguaje combinatorial y composicional que permite especificar funciones diferenciables parametrizadas. Ademas, nótese que la composición de I^{σ} con $L_{\epsilon,e}$ definen un funtor de **NNet** en **Learn**, es decir, se tiene el siguiente diagrama conmutativo

$$
\begin{tikzcd}
\texttt{NNet} \arrow[rr] \arrow[dr, "I^{\sigma}"'] & & \texttt{Learn} \\
& \texttt{Para} \arrow[ur, "L_{\epsilon,e}"'] &
\end{tikzcd}
$$

esto muestra que dada una función de activación σ, una función de error e y una taza de aprendizaje ϵ una red neuronal define un algoritmo de aprendizaje supervisado.

Conclusiones

Para resumir, en este trabajo se estudiaron conceptos de categorías y categorías monoidales, luego, para facilitar la manipulación de expresiones formadas con objetos, morfismos, funtores y transformaciones naturales que involucran a categorías monoidales, se estudiaron diagramas de cuerda, para finalmente desarrollar un marco algebraico que permitiera describir la composición de los algoritmos de aprendizaje supervisado. De lo que se concluye lo siguiente:

1. Existe una categoría monoidal simétrica **Learn** que tiene como morfismos a los algoritmos de aprendizaje supervisado.

2. Las funciones parametrizadas $I \colon \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^m$ inducen un algoritmo de aprendizaje supervisado $(\mathbb{R}^k, I, U_I, r_I)$ vía el funtor $L_{\epsilon,e}$, dada una función de error $e \colon \mathbb{R} \to \mathbb{R}$ y una taza de aprendizaje $\epsilon \in \mathbb{R}_{>0}$, es decir, el gradiente descendente y el *backpropagation* inducen un funtor de la categoría **Para** en la categoría **Learn**, donde U_I encapsula la información del gradiente descendente y r_I encapsula la información del *backpropagation*.

3. Las redes neuronales son útiles porque proveen de un lenguaje combinatorial para especificar algoritmos de aprendizaje supervisado vía la composición de funtores $L_{\epsilon,e} \circ I^{\sigma}$.

4. La categoría **Learn** contiene muchos más morfismos que la imagen de la categoría **Para** bajo el funtor $L_{\epsilon,e}$, en efecto, en la categoría **Learn** las funciones de actualización U y r de un learner (P, I, U, r) de tipo $A \to B$ no necesitan tener una forma especifica (ser definidas usando derivadas, por ejemplo), y por tanto, esto muestra que se pueden incluir elementos más generales que las redes neuronales en algoritmos de machine learning. General en el sentido en que los espacios A y B no necesitan ser euclidianos, o que partes del algoritmo aprendan funciones que están restringidas por simetrías, como periodicidad, o equivalentemente, que estén definidas sobre un toroide.

5. Dado que la categoría **Learn** desconoce totalmente la estructura de la categoría **Para** y la estructura del funtor $L_{\epsilon,e}$, en **Learn** se pueden construir algoritmos de aprendizaje supervisado que varían las nociones de errores a lo largo de una red neuronal.

Bibliografía

[Awo10] Steve Awodey. *Category theory*. Second. Vol. 52. Oxford Logic Guides. Oxford University Press, Oxford, 2010, págs. xvi+311.

[FST19] Brendan Fong, David Spivak y Rémy Tuyéras. «Backprop as functor: a compositional perspective on supervised learning». En: *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, [Piscataway], NJ, 2019, [13 pp.]

[JS91] André Joyal y Ross Street. «The geometry of tensor calculus, I». En: *Advances in Mathematics* 88.1 (jul. de 1991), págs. 55-112. DOI: 10.1016/0001-8708(91)90003-p. URL: https://doi.org/10.1016/0001-8708(91)90003-p.

[Lur21] Jacob Lurie. *Kerodon*. 2021. URL: https://kerodon.net.

[Mac10] Saunders Mac Lane. *Categories for the working mathematician*. eng. 2nd. ed., Softcover version of original hardcover edition 1998. Graduate texts in mathematics 5. New York, NY: Springer, 2010. ISBN: 9781441931238.

[Per21] Paolo Perrone. *Notes on Category Theory with examples from basic mathematics*. 2021. arXiv: 1912.10642 [math.CT].

[Sel10] P. Selinger. «A Survey of Graphical Languages for Monoidal Categories». En: *New Structures for Physics*. Springer Berlin Heidelberg, 2010, págs. 289-355. DOI: 10.1007/978-3-642-12821-9_4. URL: https://doi.org/10.1007%2F978-3-642-12821-9_4.