MATEMÁTICAS

1 de diciembre de 2024

Índice general

Parte I

Álgebra lineal y geometría

Capítulo 1

Grupos

1.1. Definición y ejemplos

Un *grupo* es un conjunto G junto con una operación $\cdot$ que cumple las propiedades:

- Asociativa:
$$g \cdot (g' \cdot g'') = (g \cdot g') \cdot g'' \quad \forall g, g', g'' \in G$$

- Existe un elemento e, al que llamaremos *elemento neutro*, tal que:
$$g \cdot e = g = e \cdot g \quad \forall g \in G$$

- Para cada $g \in G$ existe un elemento, al que denominaremos el *inverso* de g y denotaremos por g^{-1}, tal que
$$g \cdot g^{-1} = e = g^{-1} \cdot g \quad \forall g \in G$$

Si se cumple también la propiedad conmutativa:

$$g \cdot g' = g' \cdot g \quad \forall g, g' \in G$$

diremos que el grupo es *conmutativo* o *abeliano*. En este caso, la operación se denota a menudo por $+$, el elemento neutro por 0 (y se denomina *cero*) y el inverso por $-g$ (y se denomina el *opuesto* de g).

Cuando indicamos la operación por $\cdot$ (notación multiplicativa), el elemento neutro se acostumbra a llamar *unidad* y a escribir 1. Con esta notación multiplicativa, es costumbre suprimir el punto que indica la operación y escribir simplemente gg' para indicar $g \cdot g'$.

Ejemplo 1.1. Los números enteros $\mathbb{Z}$ con la suma forman un grupo conmutativo. Lo mismo vale para los racionales $\mathbb{Q}$ y los reales $\mathbb{R}$. Los números naturales $\mathbb{N} = \{1, 2, \dots\}$ no son un grupo con la suma.

Los números racionales no nulos, $\mathbb{Q} - \{0\}$, con el producto forman un grupo conmutativo. Lo mismo vale para $\mathbb{R} - \{0\}$. Ni $\mathbb{Z} - \{0\}$ ni $\mathbb{N}$ son grupos con el producto.

Ejemplo 1.2. Los números complejos $\mathbb{C}$ con la suma son un grupo conmutativo, $\mathbb{C} - \{0\}$ con el producto es un grupo conmutativo. La circunferencia $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ con el producto es un grupo conmutativo.

Ejemplo 1.3. Todos los ejemplos anteriores son grupos conmutativos. Los ejemplos más sencillos de grupos *no* conmutativos surgen en la geometría al estudiar determinados conjuntos de movimientos. Así, por ejemplo, el conjunto de movimientos del plano que dejan fijo un triángulo equilátero está formado por tres simetrías respecto a ejes que pasan por un vértice y el punto medio del lado opuesto, los giros de $120°$ y $240°$ alrededor del baricentro, y la identidad o giro de $0°$. En estos ejemplos geométricos, la operación es la *composición*: la composición de los movimientos g' y g es el movimiento $g \circ g'$ que resulta de efectuar sucesivamente los movimientos g' y g. (¡Atención al orden!) Esta operación no es conmutativa.

Esos grupos de movimientos aparecerán de manera natural al estudiar la geometría. A continuación, vamos a ocuparnos de otros grupos no conmutativos sencillos: los grupos de permutaciones.

Parte II

Análisis matemático

Capítulo 1

Límites y continuidad

1.1. Introducción

Supongamos al lector ya familiarizado con el concepto de límites tal como es introducido en el cálculo elemental donde es corriente presentar varios tipos de límites. Por ejemplo, el *límite de una sucesión* de números reales $\{x_n\}$, que simbolizamos cuando escribimos

$$\lim_{n\to\infty} x_n = A$$

significa que para cada número $\varepsilon > 0$ existe un entero $N > 0$ tal que

$$|x_n - A| < \varepsilon \quad \text{siempre que } n \geq N$$

Este límite pretende transmitir la idea intuitiva de que x_n puede estar suficientemente próximo a A en el supuesto de que n sea suficientemente grande. También se da el *límite de una función*, indicado por medio de la notación

$$\lim_{x\to p} f(x) = A$$

que significa que para cada $\varepsilon > 0$ existe otro número $\delta > 0$ tal que

$$|f(x) - A| < \varepsilon \quad \text{siempre que } 0 < |x - p| < \delta$$

Esta definición expresa la idea de que $f(x)$ puede conseguirse tan próximo a A como queramos, siempre que x se tome lo suficientemente próximo a p, sin llegar a ser p.

Las aplicaciones del cálculo a los problemas geométricos y físicos del espacio tridimensional y a las funciones de varias variables nos obligan a extender estos conceptos a $\mathbb{R}^n$. Es tan necesario como fácil dar un paso más e introducir los límites en el marco más general de los espacios métricos. Esto simplifica la teoría puesto que elimina restricciones innecesarias y al mismo tiempo cubre casi todos los aspectos necesarios del análisis.

Primeramente discutiremos los límites de las sucesiones de puntos de un espacio métrico y después discutiremos los límites de funciones y el concepto de continuidad.

1.2. Sucesiones convergentes en un espacio métrico

Definición 1.1. Una sucesión $\{x_n\}$ de puntos de un espacio métrico (S, d) es convergente si existe un punto p de S que satisfaga la siguiente propiedad:

Para todo $\varepsilon > 0$ existe un entero N tal que

$$d(x_n, p) < \varepsilon \quad \text{siempre que } n \geq N$$

Diremos también que $\{x_n\}$ converge hacia p y escribiremos $x_n \to p$ cuando $n \to \infty$, o simplemente $x_n \to p$. Si no existe un tal número p de S, se dice que la sucesión $\{x_n\}$ es divergente.

La definición de convergencia implica que

$$x_n \to p \quad \text{si, y sólo si, } d(x_n, p) \to 0$$

La convergencia de la sucesión $\{d(x_n, p)\}$ hacia 0 se realiza en el espacio euclídeo $\mathbb{R}$.

Ejemplo 1.1. En el espacio euclídeo $\mathbb{R}$, una sucesión $\{x_n\}$ se llama *creciente* si $x_n \leq x_{n+1}$ para todo n. Si una sucesión creciente está acotada superiormente (esto es, si $x_n \leq M$ para un $M > 0$ y para todo n), entonces $\{x_n\}$ converge hacia el supremo de su recorrido, $\sup\{x_1, x_2, \dots\}$. Análogamente, $\{x_n\}$ se llama *decreciente* si $x_{n+1} \leq x_n$ para todo n. Cada sucesión decreciente acotada inferiormente converge hacia el ínfimo de su recorrido. Por ejemplo, $\{1/n\}$ converge a 0.

Ejemplo 1.2. Si $\{a_n\}$ y $\{b_n\}$ son sucesiones reales que convergen hacia 0, entonces $\{a_n + b_n\}$ también converge hacia 0. Si $0 \leq c_n \leq a_n$ para todo n y si $\{a_n\}$ converge hacia 0, entonces $\{c_n\}$ también converge hacia 0. Estas propiedades elementales de las sucesiones $\mathbb{R}$ pueden ser útiles para simplificar algunas de las demostraciones concernientes a límites de un espacio métrico general.

Ejemplo 1.3. En el plano complejo $\mathbb{C}$, sea

$$z_n = 1 + \frac{1}{n^2} + \left(2 - \frac{1}{n}\right) i$$

Entonces $\{z_n\}$ converge hacia $1 + 2i$ puesto que

$$d(z_n, 1 + 2i)^2 = |z_n - (1 + 2i)|^2 = \left|\frac{1}{n^2} - \frac{1}{n} i\right|^2 = \frac{1}{n^4} + \frac{1}{n^2} \to 0 \text{ cuando } n \to \infty$$

luego $d(z_n, 1 + 2i) \to 0$.

Teorema 1.1. *Una sucesión $\{x_n\}$ en un espacio métrico (S, d) puede converger hace un punto de S, a lo sumo.*

Demostración. Supongamos que $x_n \to p$ y que $x_n \to q$. Probaremos que $p = q$. En virtud de la desigualdad triangular se tiene

$$0 \leq d(p, q) \leq d(p, x_n) + d(x_n, q)$$

Como $d(p, x_n) \to 0$ y $d(x_n, q) \to 0$ se tiene que $d(p, q) = 0$, luego $p = q$. $\square$

Si una sucesión $\{x_n\}$ converge, el único punto hacia el que converge se llama *límite* de la sucesión y se designa por medio de $\lim x_n$ o por medio de $\lim_{n\to\infty} x_n$.

Parte III

Cálculo vectorial

Capítulo 1

Espacios lineales

1.1. Introducción

A lo largo de la Matemática se encuentran muchos ejemplos de objetos matemáticos que pueden sumarse unos con otros y multiplicarse por números reales. Ante todo, los números reales son objetos de tal naturaleza. Otros ejemplos son las funciones vectoriales, los números complejos, las series y los vectores en el espacio n-dimensional. En este capítulo tratamos un concepto matemático general, llamado *espacio lineal*, que incluye todos esos ejemplos y muchos otros como casos particulares.

Brevemente, un espacio lineal es un conjunto de elementos de naturaleza cualquiera sobre el que pueden realizarse ciertas operaciones llamadas *adición* y *multiplicación por números*. Al definir un espacio lineal no especifacmos la naturaleza de los elementos ni decimos cómo se realizan las operaciones entre ellos. En cambio, exigimos que las operaciones tengan ciertas propiedades que tomamos como axiomas de un espacio lineal. Vamos ahora a hacer con detalle una descripción de esos axiomas.

1.2. Definición de espacio lineal

Sea V un conjunto no vacío de objetos, llamados *elementos*. El conjunto V se llama espacio lineal si satisface los diez axiomas siguientes que se enuncian en tres grupos.

Axiomas de clausura

AXIOMA 1. CLAUSURA RESPECTO A LA ADICIÓN. *A todo par de elementos x e y de V corresponde un elemento único de V llamado suma de x e y, designado por $x + y$.*

AXIOMA 2. CLAUSURA RESPECTO A LA MULTIPLICACIÓN POR NÚMEROS REALES. *A todo $x$de V y todo número real a corresponde un elemento de V llamado producto de a por x, designado por ax.*

Axiomas para la adición

AXIOMA 3. LEY CONMUTATIVA. *Para todo* x *y todo* y *de* V*, tenemos*

$$x + y = y + x$$

AXIOMA 4. LEY ASOCIATIVA. *Cualesquiera que sean* x, y, z *de* V*, tenemos*

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

AXIOMA 5. EXISTENCIA DE ELEMENTO CERO. *Existe un elemento en* V*, designado por el símbolo* 0*, tal que*

$$x + 0 = x, \quad \textit{para todo } x \textit{ de } V$$

AXIOMA 6. EXISTENCIA DE OPUESTOS. *Para todo* x *de* V*, el elemento* $(-1)x$ *tiene la propiedad*

$$x + (-1)x = 0$$

Axiomas para la multiplicación por números

AXIOMA 7. LEY ASOCIATIVA. *Para todo* x *de* V*, y todo par de números* a *y* b*, tenemos*

$$a(bx) = (ab)x$$

AXIOMA 8. LEY DISTRIBUTIVA PARA LA ADICIÓN EN V. *Para todo* x *y todo* y *de* V *y todo número real* a*, tenemos*

$$a(x + y) = ax + ay$$

AXIOMA 9. LEY DISTRIBUTIVA PARA LA ADICIÓN DE NÚMEROS. *Para todo* x *de* V *y todo par de números reales* a *y* b*, tenemos*

$$(a + b)x = ax + bx$$

AXIOMA 10. EXISTENCIA DE ELEMENTO IDENTIDAD. *Para todo* x *de* V*, tenemos*

$$1x = x$$

Los espacios lineales así definidos, se llaman, a veces, espacios lineales *reales* para resaltar el hecho de que se multiplican los elementos de V por números reales. Si en los axiomas 2, 7, 8 y 9 se reemplaza *número real* por *número complejo*, la estructura que resulta se llama *espacio lineal complejo*. Algunas veces un espacio lineal se llama también *espacio vectorial lineal* o simplemente *espacio vectorial*; los números utilizados como multiplicadores se llaman *escalares*. Un espacio lineal real tiene números reales como escalares; un espacio lineal complejo tiene como escalares números complejos. Si bien consideraremos principalmente ejemplos de espacios lineales reales, todos los teoremas son válidos para espacios lineales complejos. Cuando digamos espacio lineal sin más, se sobrentenderá que el espacio puede ser real o complejo.

1.3. Ejemplos de espacios lineales

Si precisamos el conjunto V y decimos cómo se suman sus elementos y cómo se multiplican por números, obtenemos un ejemplo concreto de espacio lineal. El lector fácilmente puede comprobar que cada uno de los ejemplos siguientes satisface todos los axiomas para un espacio lineal real.

Ejemplo 1.1. Sea $V = \mathbb{R}$, el conjunto de todos los números reales, y sean $x + y$ y ax la adición y la multiplicación ordinarias de números reales.

Ejemplo 1.2. Sea $V = \mathbb{C}$ el conjunto de todos los números complejos, definimos $x + y$ como la adición ordinaria de números complejos y ax como la multiplicación del número complejo x por el número real a. Aunque los elementos de V sean números complejos, éste es un espacio lineal real porque los escalares son reales.

Capítulo 2

Transformaciones lineales y matrices

2.1. Representación matricial de las transformaciones lineales

Ya que es posible obtener distintas representaciones matriciales de una transformación lineal dada mediante la elección de bases distintas, parece natural intentar elegir bases de modo que la matriz resultante tenga una forma lo más sencilla posible. El teorema que sigue prueba que podemos hacer todos los elementos 0 excepto los de la diagonal que va desde el vértice superior izquierdo al inferior derecho. A lo largo de esa diagonal habrá una hilera de unos seguidos de ceros, siendo el número de unos igual al rango de la transformación. Una matriz (t_{ik}) con todos los elementos $t_{ik} = 0$ cuando $i \neq k$ se llama **matriz diagonal**.

Teorema 2.1. *Sean V y W espacios lineales de dimensión finita, con $\dim V = n$ y $\dim W = m$. Supongamos que $T \in \mathscr{L}(V, W)$ y que $r = \dim T(V)$ representa el rango de T. Existen entonces una base $(e_1, \ldots, e_n)$ para V y otra $(w_1, \ldots, w_m)$ para W tales que*

$$T(e_i) = w_i \quad \text{para } i = 1, 2, \ldots, r \tag{2.1}$$

y

$$T(e_i) = 0 \quad \text{para } i = r + 1, \ldots, n \tag{2.2}$$

Por consiguiente, la matriz (t_{ik}) de T relativa a esas bases tiene todos los elementos cero excepto los r elementos de la diagonal que valen

$$t_{11} = t_{22} = \cdots = t_{rr} = 1$$

Parte IV

Probabilidad y estadística

Capítulo 1

Probabilidad

1.1. Espacio muestral

En el estudio de la estadística, nos ocupamos fundamentalmente de la **presentación e interpretación de resultados aleatorios** que surgen en investigaciones planificadas o estudios científicos. Por ejemplo, en Estados Unidos, para justificar la instalación de un semáforo, se podría registrar el número de accidentes mensuales en la intersección de Driftwood Lane y Royal Oak Drive. En una fábrica, los artículos producidos en la línea de ensamblaje podrían clasificarse como "defectuosos" o "no defectuosos". En una reacción química, se podría medir el volumen de gas liberado al variar la concentración de un ácido. Así, quienes trabajan con estadística manejan tanto **datos numéricos**, que reflejan conteos o mediciones, como **datos categóricos**, que se clasifican según un criterio específico.

En este contexto, cualquier registro de información, ya sea numérico o categórico, se denominará **observación**. Por ejemplo, los números $2, 0, 1$ y 2, que representan el número de accidentes ocurridos de enero a abril del año pasado en la intersección mencionada, constituyen un conjunto de observaciones. Del mismo modo, las categorías N, D, N, N y D donde N indica "no defectuos" y D "defectuoso" forman un conjunto de observaciones al inspeccionar cinco artículos.

El término **experimento** en estadística describe cualquier proceso que genere un conjunto de datos. Un ejemplo sencillo es el lanzamiento de una moneda, que produce dos posibles resultados: cara o cruz. Otro ejemplo es el lanzamiento de un misil, donde se observa su velocidad en tiempos específicos. Incluso las opiniones de votantes sobre un nuevo impuesto pueden considerarse observaciones de un experimento. En estadística, nos interesan particularmente las observaciones obtenidas al repetir un experimento varias veces. En la mayoría de los casos, los resultados están influenciados por el azar y, por tanto, no se pueden predecir con certeza. Por ejemplo, si un químico repite un análisis bajo las mismas condiciones, las medidas obtenidas variarán, evidenciando un elemento de probabilidad en el procedimiento. Del mismo modo, aunque lancemos una moneda al aire repetidamente, no podemos garantizar que obtendremos cara en un lanzamiento específico, aunque sí conocemos todas las posibilidades para cada lanzamiento.

En los tres tipos principales de estudios estadísticos —**diseños experimentales**, **estudios observacionales** y **estudios retrospectivos**— el resultado final es siempre un conjunto de datos, inevitablemente sujeto a **incertidumbre**. Aunque solo uno de ellos incluye la palabra "experimento" en su descripción, tanto el proceso de generar los datos como el de observarlos forman parte de un experimento.

Definición 1.1. Al conjunto de todos los resultados posibles de un experimento estadístico se le llama **espacio muestral** y se representa por el símbolo S.

Parte V

Análisis de datos multivariantes

Capítulo 1

Álgebra matricial

1.1. Definiciones básicas

Un conjunto de n números reales $\mathbf{x}$ puede representarse como un punto en el espacio de n dimensiones $\mathbb{R}^n$. Definiremos el vector $\mathbf{x}$ como el segmento orientado que une el origen de coordenadas con el punto $\mathbf{x}$. La orientación es importante, porque no es lo mismo el vector $\mathbf{x}$ que el $-\mathbf{x}$. Con esta correspondencia, a cada punto del espacio en $\mathbb{R}^n$ le asociamos un vector. En adelante, representaremos un vector mediante $\mathbf{x}$, para diferenciarlo del escalar x, y llamaremos $\mathbb{R}^n$ al espacio de todos los vectores de n coordenadas o componentes. En particular, un conjunto de números con todos los valores iguales se representará por un vector **constante**, que es aquél con todas sus coordenadas iguales. Un vector constante es de la forma $c\mathbf{1}$, donde c es cualquier constante y $\mathbf{1}$ es el vector con todas sus coordenadas iguales a la unidad.

En Estadística podemos asociar a los valores de una variable en n elementos un vector en $\mathbb{R}^n$, cuyo componente i-ésimo es el valor de la variable en el elemento i. Por ejemplo, si medimos las edades de tres personas en una clase y obtenemos los valores 20, 19 y 21 años, esta muestra se representa por el vector tridimensional

$$\mathbf{x} = \begin{pmatrix} 20 \\ 19 \\ 21 \end{pmatrix}$$

La **suma** (o **diferencia**) de dos vectores $\mathbf{x}$ e $\mathbf{y}$, ambos en $\mathbb{R}^n$, se define como un nuevo vector con componentes iguales a la suma (diferencia) de los componentes de los sumandos:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Es inmediato comprobar que la suma de vectores es asociativa, $\mathbf{x}+(\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}$

y conmutativa, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

La suma de dos vectores corresponde a la idea intuitva de trasladar un vector al extremo del otro y construir la línea que va desde el origen del primero al extremo del segundo. La operación suma (resta) de dos vectores da lugar a otro vector y estadísticamente corresponde a generar una nueva variable como suma (resta) de otras dos anteriores. Por ejemplo, si $\mathbf{x}$ representa el número de trabajadores varones en un conjunto de empresas e $\mathbf{y}$ el número de trabajadoras, la variable $\mathbf{x} + \mathbf{y}$ representa el número total de trabajadores y la variable $\mathbf{x} - \mathbf{y}$ la diferencia entre hombres y mujeres de cada empresa.

El producto de una constante por un vector, es un nuevo vector con componentes los del vector inicial multiplicados por la constante.

$$\mathbf{z} = k\mathbf{x} = \begin{pmatrix} kx_1 \\ \vdots \\ kx_n \end{pmatrix}$$

Multiplicar por una constante equivale a un cambio en las unidades de medición. Por ejemplo, si en lugar de medir el número de trabajadores en unidades (variable $\mathbf{x}$) lo hacemos en centenas (variable $\mathbf{z}$), entonces la variable $\mathbf{z}$ es igual a $\mathbf{x}/100$.

Llamaremos **vector transpuesto $\mathbf{x}'$**, de otro $\mathbf{x}$, a un vector con las mismas componentes, pero escritas ahora en fila:

$$\mathbf{x}' = (x_1, \ldots, x_n)$$

Al transponer un vector columna se obtiene un vector fila. Generalmente los vectores fila se utilizan para describir los valores de p variables distintas en un mismo elemento de una población.

El **producto escalar o interno** de dos vectores $\mathbf{x}$, $\mathbf{y}$, ambos en $\mathbb{R}^n$, que escribiremos $\mathbf{x}'\mathbf{y}$ o $\mathbf{y}'\mathbf{x}$, es el escalar obtenido al sumar los productos de sus componentes.

$$\mathbf{x}'\mathbf{y} = \mathbf{y}'\mathbf{x} = \sum_{i=1}^{n} x_i y_i$$

Se llamará **norma (cuadrática)** o longitud de un vector $\mathbf{x}$, a la raíz cuadrada positiva del producto escalar $\mathbf{x}'\mathbf{x}$. Se escribe $\|\mathbf{x}\|$:

$$\|x\| = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$$

La norma es la longitud del segmento que une el origen con el punto $\mathbf{x}$.

El producto escalar puede calcularse también como el producto de las normas de los vectores por el coseno del ángulo que forman. Para ilustrar este concepto consideremos los vectores

$$\mathbf{x} = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} a \\ c \end{pmatrix}$$

Observemos que el producto escalar es $\mathbf{x}'\mathbf{y} = a^2$ y que este mismo resultado se obtiene multiplicando la norma de ambos vectores, $\|\mathbf{x}\| = a$, $\|\mathbf{y}\| = \sqrt{a^2 + c^2}$ por el coseno del ángulo θ que forman, dado por $\cos\theta = a/\sqrt{a^2 + c^2}$. Observemos que el producto escalar puede también expresarse como el producto de la norma de un vector por la proyección del otro sobre él. Si uno de los vectores tiene norma uno, el producto escalar es directamente la proyección del otro vector sobre él.

Se demuestra en general que:

$$|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$$

que se conoce como la **desigualdad de Cauchy-Schwarz**. Esta desigualdad permite definir el **ángulo** entre dos vecotres $\mathbf{x}$ e $\mathbf{y}$ de $\mathbb{R}^n$ (no nulos) por la relación:

$$\cos\theta = \frac{\mathbf{x}'\mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$$

Si dos variables tienen media cero, el coseno del ángulo que forman es su **coeficiente de correlación**.

Dos vectores son **ortogonales**, o perpendiculares, si y sólo si, su producto escalar es cero. Por la definición de ángulo

$$\mathbf{x}'\mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$$

siendo θ el ángulo que forman los vectores. Si $\theta = 90°$ el coseno es cero y también lo será su producto escalar.

El producto escalar tiene una clara interpretación estadística. Para describir una variable tomamos su media. Para describir un vector podemos tomar su proyección sobre el vector constante. El vector constante de norma unidad en dimensión n es $\frac{1}{\sqrt{n}}\mathbf{1}$, y la proyección de $\mathbf{x}$ sobre este vector es $\frac{1}{\sqrt{n}}\mathbf{1}'\mathbf{x} = \sum x_i/\sqrt{n} = \bar{x}\sqrt{n}$. El vector constante resultante de esta proyección es

$$\bar{x}\sqrt{n}\left(\frac{1}{\sqrt{n}}\right)\mathbf{1} = \bar{x}\mathbf{1}$$

Por tanto, la media es el escalar que define el vector obtenido al proyectar el vector de datos sobre la dirección constante. También puede interpretarse como la norma estandarizada del vector obtenido al proyectar los datos en la dirección del vector constante, donde para estandarizar la norma de un vector dividiremos siempre por $\sqrt{n}$, siendo n la dimensión del espacio.

La variabilidad de los datos se mide por la **desviación típica**, que es la distancia estandarizada entre el vector de datos y el vector constante. La proyección del vector de datos sobre la dirección del vector constante produce el vector $\bar{x}\mathbf{1}$, y la norma del vector diferencia, $\mathbf{x} - \bar{x}\mathbf{1}$, mide la distancia entre el vector de datos y el vector constante. La norma estandarizada, dividiendo por la raíz de la dimensión del espacio es:

$$\frac{1}{\sqrt{n}}\left\|x - \bar{x}\mathbf{1}\right\| = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} = \sigma_{\mathbf{x}}$$

La medida de dependencia lineal entre dos variables, $\mathbf{x}$, $\mathbf{y}$ es la **covarianza**. La covarianza es el producto escalar promedio de los dos vectores medidos en desviaciones a la media, o tomando sus diferencias respecto a la proyección sobre el vector constante. Si promediamos el producto escalar de estos vectores

$$\frac{1}{n}(\mathbf{x} - \bar{x}\mathbf{1})'(\mathbf{y} - \bar{y}) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n} = \mathrm{Cov}(\mathbf{x}, \mathbf{y})$$

se obtiene la covarianza. Obsérvese que si las variables tienen media cero, el producto escalar de los dos vectores, dividido por n, es directamente la covarianza. En estos casos, el **coeficiente de correlación** es el coseno del ángulo entre los vectores que las representan:

$$\rho = \frac{\mathrm{Cov}(\mathbf{x}, \mathbf{y})}{\sigma_{\mathbf{x}}\sigma_{y}} = \frac{\frac{\mathbf{x}'\mathbf{y}}{n}}{\frac{\|x\|}{\sqrt{n}}\frac{\|y\|}{\sqrt{n}}} = \frac{\mathbf{x}'\mathbf{y}}{\|x\|\,\|y\|} = \cos\theta$$

que es la interpretación geométrica del coeficiente de correlación. La implicación estadística de la ortogonalidad es la incorrelación. Si dos variables tienen media cero y son ortogonales, es decir, los vectores que las caracterizan forman un ángulo de 90 grados, $\rho = 0$, las variables están incorreladas.

Parte VI

Diseño y análisis de modelos de probabilidad

Capítulo 1

De la medida a la probabilidad

1.1. Generación de estructuras

Como en la axiomática de la probabilidad el punto de partida para poder definir la función de probabilidad es considerar una clase de subconjuntos de Ω que posea una estructura flexible y operativa (será el álgebra o σ-álgebra, según las necesidades de descripción), es importante preguntarse si a partir de una clase arbitraria $\mathscr{C}$ de subconjuntos de Ω se puede engendrar (o generar) una clase con estructura determinada (la que se necesite) que contenga a la clase $\mathscr{C}$.

Por otro lado como $\mathscr{P}(\Omega)$ es σ-álgebra, dada cualquier clase $\mathscr{C}$ de Ω siempre existirá alguna clase, con estructura dada, que la contiene. Una cuestión de interés en el desarrollo es la siguiente, entre todas las clases $\mathscr{C}_i$ $(i \in I)$ con estructura dada que contienen a $\mathscr{C}$ $(\mathscr{C} \subset \mathscr{C}_i)$, ¿existe una $\mathscr{C}_m$ $(m \in I)$ que sea la mínima en el sentido de estar contenida en cualquier otra con la misma estructura que contenga a $\mathscr{C}$?, es decir, $\mathscr{C} \subset \mathscr{C}_m \subset \mathscr{C}_i$ $(i \in I)$. En caso afirmativo, esta clase $\mathscr{C}_m$ se llama la *mínima clase engendrada o generada por la clase original* $\mathscr{C}$. Intuitivamente, el problema consiste en "añadir" el mínimo número de subconjuntos a la clase $\mathscr{C}$ hasta conseguir la estructura deseada (en nuestro caso sólo consideraremos el álgebra, σ-álgebra y clase monótona). La estructura de clase monótona es importante por actuar como puente entre el álgebra y la σ-álgebra. Este problema tiene dos tipos de resultados: los que sólo garantizan la existencia de $\mathscr{C}_m$ y los que, además muestran un método de construcción. Salvo casos particulares no existen métodos constructivos para "construir" la mínima clase engendrada con estructura dada.

En lo que sigue, sea Ω un conjunto base y $\mathscr{C} \subset \mathscr{P}(\Omega)$ una clase no vacía.

- Se define el álgebra engendrada (generada) por la clase $\mathscr{C}$, y la representaremos $\mathscr{A}_0(\mathscr{C})$, a la mínima clase, con estructura de álgebra, que contiene a $\mathscr{C}$.

- Se define la σ-álgebra engendrada (generada) por la clase $\mathscr{C}$, y la representaremos $\sigma(\mathscr{C})$, a la mínima clase, con estructura de σ-álgebra, que contiene a $\mathscr{C}$.

- Se define la clase monótona engendrada (generada) por la clase $\mathscr{C}$, y la representaremos $\mathscr{M}(\mathscr{C})$, a la mínima clase, con estructura de clase monótona, que contiene a $\mathscr{C}$.

Lema 1.1 (**Intersección de estructuras**). *La intersección arbitraria de clases con la misma estructura es otra clase con la misma estructura. Es decir, si $\{\mathscr{C}_i\}_{i\in I}$ es una familia de clases (de Ω) con $\mathscr{C}_i$ $(\forall i \in I)$ anillo, σ-anillo, álgebra, σ-álgebra o clase monótona, entonces $\mathscr{C} = \bigcap_{i\in I} C_i$ es anillo, σ-anillo, álgebra, σ-álgebra o clase monótona.*

Demostración. La propia definición de las estructuras garantizan la invariancia de las mismas bajo la operación de intersección. En efecto, supongamos que $\left\{\mathscr{A}_0^{(i)}\right\}_{i\in I}$ es una familia de álgebras definidas sobre el mismo espacio Ω. Definamos la clase $\mathscr{A}_0 = \bigcap_{i\in I} \mathscr{A}_0^{(i)}$ y veamos que es álgebra.

1. Si $A_j \in \mathscr{A}_0$ $(j = 1, 2, \ldots, n)$, entonces $A_j \in \mathscr{A}_0^{(i)}$ $(j = 1, 2, \ldots, n; \forall i \in I)$. Como cada $\mathscr{A}_0^{(i)}$ es álgebra, $\bigcup_{j=1}^n A_j \in \mathscr{A}_0^{(i)}$ $(\forall i \in I)$, luego $\bigcup_{j=1}^n A_j \in \mathscr{A}_0$.

2. Si $A \in \mathscr{A}_0$ entoncces $A \in \mathscr{A}_0^{(i)}$ $(\forall i \in I)$, luego $\overline{A} \in \mathscr{A}_0^{(i)}$ $(\forall i \in I)$. Por tanto, $\overline{A} \in \mathscr{A}_0$.

El mismo método de demostración se sigue para el resto de las estructuras. □

El problema de la existencia de las clases engendradas se aborda de este modo por medio de la siguiente

Proposición 1.1 (**Existencia de mínimas estructuras**). *Dada una clase arbitraria $\mathscr{C}$ de subconjuntos de Ω, siempre existen $\mathscr{A}_0(\mathscr{C})$, $\sigma(\mathscr{C})$ y $\mathscr{M}(\mathscr{C})$.*

Bibliografía

[1] Tom M. Apostol. *Análisis matemático, segunda edición.* Editorial Reverté S.A., 1972.

[2] Tom M. Apostol. *Calculus. Volumen 2, segunda edición.* Editorial Reverté S.A., 1985.

[3] Manuel Castellet, Irene Llerena. *Álgebra lineal y geometría.* Editorial Reverté, 1994.

[4] Daniel Peña. *Análisis de datos multivariantes.* Editorial McGrawHill, 2002.