

Apuntes de Matemática Avanzada

Primera Parte

Algebra Lineal

Octavio Miloni

I
**Espacios Vectoriales y Transformaciones
Lineales**

1 Prefacio

Este material está dedicado a la disciplina *Matemática Avanzada* que se encuentra en el primer cuatrimestre del cuarto año de la carrera de Licenciatura en Meteorología y Ciencias de la Atmósfera, de la Facultad de Cs. Astronómicas y Geofísicas de la Universidad Nacional de La Plata.

Matemática Avanzada es una materia compuesta por las siguientes unidades temáticas

- **Álgebra Lineal:** Espacios Vectoriales. Subespacios. Base. Transformaciones lineales. Cambio de base. Algebra de transformaciones lineales. Formas canónicas de transformaciones lineales. Algebra tensorial. Autovalores y autovectores. Polinomio característico. Polinomio mínimo. Forma de Jordan. Sistemas de ecuaciones lineales.
- **Variable Compleja:** Números complejos. Funciones complejas elementales. Funciones analíticas. Funciones armónicas. Transformaciones conformes. Integración en el campo complejo. Propiedades. Teorema de Cauchy-Goursat. Corolarios. Series de funciones complejas. Ceros y singularidades. Teorema de Laurent. Integración en el campo real mediante el teorema de los residuos.
- **Ecuaciones Diferenciales:** Concepto de ecuación diferencial. Ecuaciones lineales de primer orden a coeficientes analíticos. Caso homogéneo. Puntos ordinarios y singulares regulares. Teorema de existencia y unicidad de las soluciones. Solución mediante series de potencias. Concepto de serie, integral y transformada de Fourier.

Esta materia tiene Análisis Matemático II como correlativa, de lo que se desprende que los estudiantes que cursen esta materia ya tendrán, como mínimo, aprobados los trabajos prácticos de las disciplinas Álgebra, Análisis Matemático I y Análisis Matemático II.

En virtud de la profundidad alcanzada en estas disciplinas, se puede afirmar sin lugar a dudas que el nivel en Matemática de los estudiantes es elevado, conjuntamente con el ritmo de estudio alcanzado a esta altura de la Carrera.

Es por esto que hemos decidido elaborar un material de guía para el estudio de la materia.

Queremos aclarar que este material de manera alguna suprime la necesidad de los libros recomendados para el estudio de la materia, sino que está pensado para hacer coherente el recorte de temas que la constituye.

2 Espacios Vectoriales

Definición: Todo conjunto V de elementos x, y, z, \dots de naturaleza arbitraria para los cuales están definidas dos operaciones, de adición y multiplicación por números (escalares los cuales se encuentran en un cuerpo) que cumplen los axiomas siguientes:

- *Suma*

- La suma es cerrada, es decir, $x + y \in V$
- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- Existe un elemento nulo, 0 tal que $x + 0 = x \forall x \in V$
- Para todo $x \in V$ existe un $x' \in V$ tal que $x + x' = 0$

- *Producto por Números*

- $\lambda \cdot x \in V$
- $1 \cdot x = x$
- $\lambda \cdot (\mu \cdot x) = (\lambda \mu) \cdot x$
- $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
- $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$

Se denomina *Espacio Vectorial*.

Ejemplo 1: Sea V el conjunto de las ternas (x, y, z) de números reales para las cuales definimos una suma:

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

y un producto por números reales como

$$\lambda \cdot (x, y, z) = (\lambda x, \lambda y, \lambda z)$$

tiene estructura de espacio vectorial.

En efecto, consideremos tres ternas $\vec{v}_1 = (x_1, y_1, z_1)$, $\vec{v}_2 = (x_2, y_2, z_2)$ y $\vec{v}_3 = (x_3, y_3, z_3)$ en V . Tenemos que

$$\begin{aligned} \vec{v}_1 + \vec{v}_2 &= (x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_2 + x_1, y_2 + y_1, z_2 + z_1) \\ &= \vec{v}_2 + \vec{v}_1 \end{aligned}$$

$$\begin{aligned} \vec{v}_1 + (\vec{v}_2 + \vec{v}_3) &= \\ &= (x_1, y_1, z_1) + (x_2 + x_3, y_2 + y_3, z_2 + z_3) \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), z_1 + (z_2 + z_3)) \\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, (z_1 + z_2) + z_3) \\ &= (\vec{v}_1 + \vec{v}_2) + \vec{v}_3 \end{aligned}$$

Si consideramos el elemento $(0, 0, 0)$ podemos notar que $(x, y, z) + (0, 0, 0) = (x + 0, y + 0, z + 0) = (x, y, z)$. Para la suma nos faltaría demostrar que para todo elemento existe un inverso aditivo (u opuesto). En efecto, $(x, y, z) + (x', y', z') = (0, 0, 0)$ entonces $x' = -x$, $y' = -y$ y $z' = -z$, esto es, existe siempre un elemento opuesto.

Para el caso del producto por un número tenemos que, a partir de la definición, $1 \cdot (x, y, z) = (1x, 1y, 1z)$, lo que hace cumplir el primer axioma. Además, sean λ y μ números reales, tenemos que $\lambda \cdot (\mu(x, y, z)) = \lambda \cdot (\mu x, \mu y, \mu z) = (\lambda \mu x, \lambda \mu y, \lambda \mu z) = (\lambda \mu) \cdot (x, y, z)$. Consideremos ahora $(\lambda + \mu)(x, y, z) = ((\lambda + \mu)x, (\lambda + \mu)y, (\lambda + \mu)z) = (\lambda x + \mu x, \lambda y + \mu y, \lambda z + \mu z) = \lambda \cdot (x, y, z) + \mu \cdot (x, y, z)$. Por último, consideremos $\lambda \cdot \{(x_1, y_1, z_1) + (x_2, y_2, z_2)\} = \lambda \cdot (x + x_1, y + y_1, z + z_1) = (\lambda(x + x_1), \lambda(y + y_1), \lambda(z + z_1)) = \lambda \cdot (x_1, y_1, z_1) + \lambda \cdot (x_2, y_2, z_2)$.

Como vemos, la comprobación de que cierto conjunto es un espacio vectorial consiste en ir comprobando cada uno de los axiomas, usando las definiciones para la suma y el producto.

Ejemplo 2: Sea V el conjunto de funciones continuas en el intervalo $[a, b]$ donde definimos una suma $f + g$ definida como $(f + g)(x) = f(x) + g(x)$ y un producto con números $\lambda \cdot f$ definido como $(\lambda \cdot f)(x) = \lambda f(x)$, es un espacio vectorial.

La demostración es elemental y se deja como ejercicio.

2.1 Base y coordenadas. Dimensión

En el ejemplo 1 efectuamos una comprobación de que las definiciones de suma y producto por números dotan al conjunto de las ternas de números reales de una estructura de espacio vectorial. Ahora realicemos un camino inverso: Comencemos con una terna de números reales (x, y, z) . Con este elemento descompongamos de la siguiente manera: A partir de las definiciones de suma y producto podemos escribir

$$\begin{aligned}(x, y, z) &= (x, 0, 0) + (0, y, 0) + (0, 0, z) \\ &= x \cdot (1, 0, 0) + y \cdot (0, 1, 0) + z \cdot (0, 0, 1)\end{aligned}$$

Esta cuenta elemental pone en evidencia una característica de este espacio: Si bien el espacio vectorial posee infinitos elementos, existen tres elementos particulares a partir de los cuales (en este caso los vectores $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$) un vector *cualquiera* puede ser escrito como una combinación de ellos. Veamos que la representación no es única. Consideremos los vectores $\mathbf{e}_1 = (1, -1, 0)$, $\mathbf{e}_2 = (0, 1, 1)$ y $\mathbf{e}_3 = (0, 0, 1)$. Notemos que el mismo vector (x, y, z) puede ser escrito como

$$(x, y, z) = \frac{(x + y)}{2} \cdot \mathbf{e}_1 + \frac{(x - y)}{2} \cdot \mathbf{e}_2 + \frac{[2z - (x + y)]}{2} \cdot \mathbf{e}_3$$

Esto significa que la forma de representación no es única, pero que la cantidad de elementos para representar un vector si. Estas son las ideas de lo que denominaremos base y dimensión.

Definición: Sea V un espacio vectorial sobre los números reales. Sea $\{\mathbf{e}_1, \mathbf{e}_2 \dots \mathbf{e}_n\}$ un conjunto de vectores de V . Se dice que el conjunto forma un sistema de generadores de V si para todo vector del espacio existen números reales $\lambda_1, \lambda_2, \dots \lambda_n$ tales que

$$\vec{v} = \lambda_1 \cdot \mathbf{e}_1 + \lambda_2 \cdot \mathbf{e}_2 + \dots + \lambda_n \cdot \mathbf{e}_n$$

En los ejemplos presentados mostramos que dado un espacio vectorial, no es único el sistema de generadores para dicho espacio. Sin embargo en los ejemplos vimos que el número de generadores pareciera que sí. En realidad esto no es así: si a un sistema de generadores le incorporamos un número de vectores determinado, como éstos al ser multiplicados por cero son el vector nulo la definición de sistema de generadores sigue siendo satisfecha. A la idea de sistema de generadores vamos a incorporar una otra idea en la que determine un criterio del número mínimo de generadores necesarios para escribir cualquier vector de un determinado espacio vectorial.

Definición: Sea V un espacio vectorial sobre los números reales. Sea $\{\mathbf{e}_1, \mathbf{e}_2 \dots \mathbf{e}_n\}$ un conjunto de vectores de V . Se dice que los vectores son linealmente independientes si

$$\vec{0} = \lambda_1 \cdot \mathbf{e}_1 + \lambda_2 \cdot \mathbf{e}_2 + \dots + \lambda_n \cdot \mathbf{e}_n$$

implica que cada número λ_j debe ser el cero.

Esta definición provee un criterio de “filtrado” de vectores que no sean necesarios para la generación de vectores. Con estas dos definiciones anteriores estableceremos la siguiente definición.

Definición: Sea V un espacio vectorial sobre los números reales. Sea $\{\mathbf{e}_1, \mathbf{e}_2 \dots \mathbf{e}_n\}$ un conjunto de vectores de V . Se dice que los vectores forman una base si son un sistema de generadores y linealmente independientes.

Definición: Sea V un espacio vectorial sobre los números reales. Si una base de V posee n elementos diremos que la dimensión es n .

En los ejemplos anteriores vimos que un mismo vector $\vec{v} = (x, y, z)$ admite al menos dos representaciones. Esto es: En la base $\{(1, 0, 0); (0, 1, 0); (0, 0, 1)\}$ lo escribimos como

$$\vec{v} = x \cdot (1, 0, 0) + y \cdot (0, 1, 0) + z \cdot (0, 0, 1)$$

y en la base $\{(1, -1, 0); (0, 1, 1); (0, 0, 1)\}$ como

$$\vec{v} = \frac{(x+y)}{2} \cdot (1, -1, 0) + \frac{(x-y)}{2} \cdot (0, 1, 1) + \frac{[2z - (x+y)]}{2} \cdot (0, 0, 1)$$

Como ente algebraico, el vector \vec{v} es único, lo que cambia es la representación en la base.

Vamos a definir como *coordenadas* de un vector a los coeficientes de los elementos en la base escogida para el espacio vectorial. De esta manera en la primera base (denominada canónica) las coordenadas son simplemente x, y, z y en la segunda base escogida las coordenadas serán

$$\frac{(x+y)}{2}, \frac{(x-y)}{2}, \frac{[z - 2(x+y)]}{2}$$

En los textos clásicos de cálculo tensorial tales como el de Santaló el estudio se centra en el comportamiento de los tensores bajo cambio de coordenadas. Sin embargo, hablar de cambio de coordenadas no es posible sin hablar de cambio de base. En este trabajo consideraremos primero cambios de base y veremos como repercute estos cambios en los cambios de coordenadas.

2.2 Invarianza y Representación

No siempre se reflexiona sobre la naturaleza de los objetos de determinado espacio vectorial y con frecuencia se tiene a confundir los objetos con sus representaciones.

Lo que se intenta aclarar es lo siguiente: Un objeto de \mathcal{R}^3 , por ejemplo el $(1, 2, 3)$, es un elemento de un conjunto y por lo tanto no depende de ninguna base. Simplemente es un elemento cuyas componentes (en tanto números que lo componen) son los números 1, 2 y 3. Decimos que este elemento es un invariante, y tiene característica de *absoluto*. Al introducir una base en el espacio lo que tendremos de manera no única son coordenadas que lo representan en una determinada base. Una vez que elegimos una base, lo que tendremos es una *representación* del vector en una determinada base, y por lo tanto, sus coordenadas.

Si para \mathcal{R}^3 consideramos las bases $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} = \{(1, 0, 0); (0, 1, 0); (0, 0, 1)\}$ y $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{(.5, 0, 0); (0, 0, -2); (0, -1, 0)\}$, el vector $(1, 2, 3)$ se representa como $1 \cdot \mathbf{e}_1 + 2 \cdot \mathbf{e}_2 + 3 \cdot \mathbf{e}_3$ como así también $2 \cdot \mathbf{e}'_1 - \frac{3}{2} \cdot \mathbf{e}'_2 - 1 \cdot \mathbf{e}'_3$ es decir

$$1 \cdot \mathbf{e}_1 + 2 \cdot \mathbf{e}_2 + 3 \cdot \mathbf{e}_3 = 2 \cdot \mathbf{e}'_1 - \frac{3}{2} \cdot \mathbf{e}'_2 - 1 \cdot \mathbf{e}'_3 = (1, 2, 3)$$

Esto significa, que dado un espacio vectorial, la elección de la base es arbitraria. Una vez elegida la base, cada vector estará representado por un conjunto de coordenadas, pero

$$\sum_{j=1}^n \lambda^j \cdot \mathbf{e}_j$$

es un invariante, es decir que no depende de la base, de hecho, si hubiéramos escogido otra base, $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$, tendríamos que se satisface la igualdad

$$\sum_{j=1}^n \lambda^j \cdot \mathbf{e}_j = \sum_{i=1}^n \beta^i \cdot \mathbf{e}'_i$$

2.3 Convenio de la suma de Einstein

Al representar un vector de un espacio vectorial de dimensión n en una determinada base, debemos utilizar el símbolo de sumatoria,

$$\vec{v} = \sum_{i=1}^n \beta^i \cdot \mathbf{e}_i$$

además esta suma es un invariante.

Einstein propuso un criterio para no escribir los símbolos de sumatoria Σ siempre y cuando se satisfagan las siguientes condiciones

- Aparezcan dos índices repetidos.
- Los índices repetidos sean uno superior y otro inferior.

Entonces,

$$a_i^i = \sum_{i=i_{min}}^{i_{max}} a_i^i = a_{i_{min}}^{i_{min}} + a_{i_{min}+1}^{i_{min}+1} + \dots + a_{i_{max}}^{i_{max}}$$

además si los elementos son los de una matriz $\mathcal{R}^{n \times n}$ y $i_{min} = 1$ y $i_{max} = n$ tendremos la traza de la matriz

Contrariamente, el término a_{ii} es simplemente el único término de la diagonal a_{ii}

Esta operación, sumar en índices repetidos (uno arriba y uno abajo) se denomina también *contracción*.

Las contracciones producen invariantes. En efecto, la expresión $\lambda^j \cdot \mathbf{e}_j$ es un invariante, ya que

$$\lambda^j \cdot \mathbf{e}_j = \sum_{j=1}^n \lambda^j \cdot \mathbf{e}_j$$

y es un elemento absoluto del espacio vectorial.

La notación de Einstein es de mucha utilidad en el estudio de la Teoría de la Relatividad, tanto Especial como General, ya que las operaciones que se realizan contienen muchas sumatorias de estas características. Más adelante se volverá en ese punto.

Ejemplo: Multiplicación de matrices

Sean $A \in \mathcal{R}^{n \times m}$ y $B \in \mathcal{R}^{m \times k}$. El elemento c_j^i (donde el supraíndice indica fila y el subíndice, columna) del producto $A \times B$ se obtendrá como

$$c_j^i = a_k^i b_j^k \quad \left(= \sum_{k=1}^m a_k^i b_j^k \right)$$

En este caso el resultado no es un invariante, ya que la producción de invariantes se obtiene cuando todos los índices están afectados a sumatorias.

El uso de la convención de Einstein es muy práctico y simplificador, pero, como en todo lo que respecta a la matemática, debemos familiarizarnos bien con él.

3 Transformaciones Lineales

En este parágrafo estudiaremos los conceptos de transformaciones lineales y sus matrices asociadas.

Definición: Sean V y W espacios vectoriales con dimensión finita. Sea T una función $T: V \rightarrow W$. Diremos que la función T es una *Transformación Lineal* si cumple

a) $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$

b) $T(\lambda \cdot \vec{v}_1) = \lambda \cdot T(\vec{v}_1)$

Ejemplo: Sea $V = \mathcal{R}^3$ y $W = \mathcal{R}^2$. Sea T una función definida como

$$T(x, y, z) = (x + y, y - z)$$

Veamos que

$$\begin{aligned} T(x + x', y + y', z + z') &= ([x + x'] + [y + y'], [y + y'] - [z + z']) \\ &= (x + x' + y + y', y + y' - z - z') \\ &= ([x + y] + [x' + y'], [y - z] + [y' - z']) \\ &= (x + y, y - z) + (x' + y', y' - z') \\ &= T(x, y, z) + T(x', y', z') \end{aligned}$$

y

$$\begin{aligned}
 T(\lambda(x, y, z)) = T(\lambda x, \lambda y, \lambda z) &= (\lambda x + \lambda y, \lambda y - \lambda z) \\
 &= (\lambda(x + y), \lambda(y - z)) \\
 &= \lambda(x + y, y - z) \\
 &= \lambda T(x, y, z)
 \end{aligned}$$

Como satisface las condiciones a) y b) la transformación es lineal.

Cuando la transformación lineal va de un espacio en sí mismo decimos que la transformación lineal es un *endomorfismo* u *operador lineal*.

Consideremos dos espacios vectoriales V y W de dimensión n y m respectivamente. Sean $\{v_1, v_2, \dots, v_n\}$ y $\{w_1, w_2, \dots, w_m\}$ bases de V y W respectivamente.

3.1 Matriz Asociada a una Transformación Lineal

Apliquemos la transformación lineal T a un vector arbitrario $\vec{v} \in V$. Como tenemos una base de V , podemos escribir

$$\vec{v} = \lambda^i v_i$$

donde por definición los números $\lambda^1, \lambda^2, \dots, \lambda^n$ son las coordenadas del vector en la base dada. Entonces, aplicando la transformación T tenemos

$$T(\vec{v}) = T(\lambda^i v_i)$$

como la transformación es lineal tendremos

$$T(\vec{v}) = \lambda^i T(v_i)$$

Ahora bien, como la transformación T aplica un elemento de V en un elemento de W , cada transformado debe poder escribirse como una combinación lineal de los elementos de la base de W . Esto significa que

$$T(\vec{v}) = \beta^j w_j$$

Por otro lado, cada transformado de los elementos de la base de V poseerán a su vez un desarrollo en la base de w

$$T(v_i) = a_i^j w_j$$

los números a_i^j indican las coordenadas del vector de la base transformado en la base de W . El motivo de incorporar dos índices es el siguiente: un índice (inferior) asociado al elemento de la base del espacio de partida transformado y un índice (el superior) asociado a la coordenada del transformado en la base del espacio de llegada.

Con esto, la transformación del elemento \vec{v} la podemos expresar

$$\begin{aligned}
 T(\vec{v}) &= \lambda^i T(v_i) \\
 &= \lambda^i a_i^j w_j \\
 &= a_i^j \lambda^i w_j
 \end{aligned}$$

Entonces, tenemos

$$\beta^j w_j = a_i^j \lambda^i w_j$$

o, lo que es equivalente

$$\left[\beta^j - a_i^j \lambda^i \right] w_j = 0$$

y como los vectores w_i son base, son linealmente independientes, con lo que

$$\left[\beta^j - a_i^j \lambda^i \right] = 0$$

o, lo que es lo mismo que

$$\beta^j = a_i^j \lambda^i$$

Esta última expresión relaciona las coordenadas del vector a transformar, con las coordenadas del vector en el espacio W . Más aún, la relación viene dada a partir de un producto de matrices, de la forma:

$$\begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^m \end{bmatrix} = \begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix}$$

Esta relación es entre coordenadas. La matriz asociada a la transformación no es intrínseca, ya que sus elementos dependen de las bases utilizadas para describir los espacios.

La matriz asociada de la transformación lineal se construye de la siguiente manera:

1. Se aplica la transformación lineal al primer elemento de la base, es decir $T(\mathbf{e}_1)$
2. Se escribe el vector transformado en la base de W
3. Las coordenadas del transformado de \mathbf{e}_1 forma la primera columna de la matriz asociada.
4. Se repite el procedimiento para los n elementos de la base de V encolumnándolos

De esta manera, se obtendrá una matriz de n columnas y m filas.

Recordemos que según la convención que hemos hecho, los elementos de una matriz, a_j^i , están posicionados en la fila i -ésima y la columna j -ésima.

3.2 Cambio de Base y Cambio de Coordenadas

Dado un espacio vectorial V de dimensión finita n . Consideremos a $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ y $\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2', \dots, \mathbf{e}_n'\}$ dos bases distintas del espacio. Para determinar el procedimiento de cambio de base y el consecuente cambio de coordenadas podemos pensar el asunto de la siguiente manera: Consideremos la transformación identidad, $I_d : V \rightarrow V$ tal que para cualquier elemento del espacio V le asigna el mismo elemento, es decir

$$I_d(\vec{v}) = \vec{v}$$

Dado que la transformación identidad es una transformación lineal tendrá una matriz asociada respecto a las bases de representación del espacio. Es claro que si utilizamos la misma base para el dominio y el codominio de la transformación identidad la matriz asociada será la matriz identidad $n \times n$. Ahora, si para el dominio usamos la base \mathcal{B} y para el codominio, la base \mathcal{B}' lo que obtendremos es que la matriz asociada ya no será la matriz identidad, sino que será la matriz conocida como *matriz de cambio de base*.

Consideremos entonces, la siguiente relación entre las bases:

$$\begin{aligned} \mathbf{e}_1 &= \Lambda_1^\ell \mathbf{e}'_\ell \\ \mathbf{e}_2 &= \Lambda_2^\ell \mathbf{e}'_\ell \\ &\vdots = \vdots \\ \mathbf{e}_n &= \Lambda_n^\ell \mathbf{e}'_\ell \end{aligned}$$

Entonces, aplicando la transformación identidad, tendremos

$$\begin{aligned} I_d(\mathbf{e}_1) &= \mathbf{e}_1 = \Lambda_1^\ell \mathbf{e}'_\ell \\ I_d(\mathbf{e}_2) &= \mathbf{e}_2 = \Lambda_2^\ell \mathbf{e}'_\ell \\ &\vdots = \vdots = \vdots \\ I_d(\mathbf{e}_n) &= \mathbf{e}_n = \Lambda_n^\ell \mathbf{e}'_\ell \end{aligned}$$

Entonces, la matriz asociada a esta transformación es

$$\begin{bmatrix} \Lambda_1^1 & \Lambda_2^1 & \Lambda_3^1 & \cdots & \Lambda_n^1 \\ \Lambda_1^2 & \Lambda_2^2 & \Lambda_3^2 & \cdots & \Lambda_n^2 \\ \Lambda_1^3 & \Lambda_2^3 & \Lambda_3^3 & \cdots & \Lambda_n^3 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \Lambda_1^n & \Lambda_2^n & \Lambda_3^n & \cdots & \Lambda_n^n \end{bmatrix}$$

Entonces, dado un vector determinado en la base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ tiene por representación $\vec{v} = \lambda^\ell \mathbf{e}_\ell$. La relación establecida a través de la transformación lineal, tendremos que el vector transformado se escribirá, en la base $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ como $\vec{v} = \beta^\ell \mathbf{e}'_\ell$. Entonces, la relación entre las coordenadas será

$$\begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^n \end{bmatrix} = \begin{bmatrix} \Lambda_1^1 & \Lambda_2^1 & \Lambda_3^1 & \cdots & \Lambda_n^1 \\ \Lambda_1^2 & \Lambda_2^2 & \Lambda_3^2 & \cdots & \Lambda_n^2 \\ \Lambda_1^3 & \Lambda_2^3 & \Lambda_3^3 & \cdots & \Lambda_n^3 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \Lambda_1^n & \Lambda_2^n & \Lambda_3^n & \cdots & \Lambda_n^n \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix}$$

o, en coordenadas,

$$\beta^\mu = \Lambda_\nu^\mu \lambda^\nu$$

En ocasiones, cuando el espacio es \mathcal{R}^n se denotan los sistemas de coordenadas como $\{x^1, x^2, \dots, x^n\}$ (donde se asume una determinada base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$) entonces, el cambio de coordenadas a un nuevo sistema $\{x'^1, x'^2, \dots, x'^n\}$ el cambio viene dado a través de

$$x'^\mu = \Lambda_\nu^\mu x^\nu$$

Es más, si tenemos un cambio de coordenadas, que relaciona las viejas con las nuevas, el cambio de base viene dado a través de

$$\mathbf{e}_\mu = \Lambda_\mu^\nu \mathbf{e}'_\nu$$

es decir, que se utilizan las mismas cantidades Λ_ν^μ pero no es estrictamente un producto de matrices.

Notemos que el cambio de coordenadas de las viejas a las nuevas es similar al cambio de base pero de las nuevas a las viejas.

Si tuviéramos un cambio de base de las viejas a las nuevas:

$$\mathbf{e}'_\nu = \Phi_\nu^\mu \mathbf{e}_\mu$$

podemos aplicar

$$\mathbf{e}'_\nu = \Phi_\nu^\mu \Lambda_\mu^\lambda \mathbf{e}'_\lambda$$

Entonces, se debe cumplir

$$\Phi_\nu^\mu \Lambda_\mu^\lambda = \delta_\nu^\lambda$$

lo que implica que la matriz Φ es la matriz inversa de la matriz Λ .

Resumiendo:

Esquema para transformación de coordenadas

1. Supongamos que damos un cambio de coordenadas: $x'^\mu = \Lambda_\nu^\mu x^\nu$
2. El cambio de base asociado será

$$\mathbf{e}'_\nu = [\Lambda^{-1}]_\nu^\mu \mathbf{e}_\mu$$

notemos además que este cálculo no es un producto de matrices

3. La relación entre la base nueva con la vieja es

$$\mathbf{e}_\nu = \Lambda_\nu^\mu \mathbf{e}'_\mu$$

Notemos que esta relación es muy parecida al cambio de coordenadas, sólo que esta relaciona base nueva con vieja y el cambio de coordenadas, al revés. Este cálculo no es un producto de matrices.

4. La relación entre las coordenadas nuevas con las viejas, es

$$x^\nu = [\Lambda^{-1}]_\mu^\nu x'^\mu$$

Este cálculo es un producto de matrices.

Otra forma de establecer el cambio de coordenadas o bases es a partir de la relación invariante

$$\vec{v} = x^\mu \mathbf{e}_\mu = x'^\mu \mathbf{e}'_\mu$$

Si contamos con el cambio $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} \rightarrow \mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ dado a través de $\mathbf{e}_\nu = \Lambda_\nu^\mu \mathbf{e}'_\mu$ entonces, reemplazando en la relación invariante

$$x^\nu \Lambda_\nu^\mu \mathbf{e}'_\mu = x'^\mu \mathbf{e}'_\mu \rightarrow [x^\nu \Lambda_\nu^\mu - x'^\mu] \mathbf{e}'_\mu = 0$$

y como los \mathbf{e}_μ son base, tendremos

$$x'^\mu = \Lambda_\nu^\mu x^\nu$$

De esta misma manera se obtienen todas las relaciones ya escritas.

3.3 El espacio de las Transformaciones Lineales

Vamos a estudiar el conjunto de todas las transformaciones lineales entre un espacio V y otro espacio W definidos en un cuerpo K . Llamaremos a este conjunto $L(V, W)$. Definimos la suma entre transformaciones lineales de la siguiente manera: Sean f y g en $L(V, W)$ $f + g$ y $\lambda \cdot f$, con $\lambda \in K$

$$(f + g)(\vec{v}) = f(\vec{v}) + g(\vec{v}), \quad (\lambda \cdot f)(\vec{v}) = \lambda \cdot f(\vec{v})$$

Se deja como ejercicio probar que con esta definición $L(V, W)$ es un espacio vectorial sobre K .

Un aspecto interesante a estudiar es qué dimensión tiene este espacio $L(V, W)$. Supongamos que V tiene dimensión n y que W tiene dimensión m .

Sean $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ una base de V y $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_m\}$ una base para W .

Consideremos en este espacio $n \times m$ transformaciones E_{ν}^{μ} con $\mu = 1, 2, \dots, n$ y $\nu = 1, 2, \dots, m$ definidas como

$$E_{\nu}^{\mu}(\mathbf{e}_{\lambda}) = \delta_{\lambda}^{\mu} \mathbf{e}'_{\nu} = \begin{cases} 0 & \lambda \neq \mu \\ \mathbf{e}'_{\nu} & \lambda = \mu \end{cases}$$

Consideremos ahora una transformación lineal cualquiera en $L(V, W)$, T . Sean los a_{ν}^{μ} los elementos de la matriz asociada de T en las bases \mathcal{B} y \mathcal{B}' . Tenemos entonces,

$$T(\mathbf{e}_{\lambda}) = a_{\nu}^{\mu} \mathbf{e}'_{\nu}$$

Aprovechando la definición de las funciones E_{ν}^{μ} podríamos escribir

$$T(\mathbf{e}_{\lambda}) = a_{\nu}^{\mu} \mathbf{e}'_{\nu} = a_{\nu}^{\mu} E_{\nu}^{\mu}(\mathbf{e}_{\lambda})$$

Además, como $L(V, W)$ es un espacio vectorial, podemos reescribir la última ecuación como

$$[T - a_{\nu}^{\mu} E_{\nu}^{\mu}](\mathbf{e}_{\lambda}) = 0$$

lo que implica que la función $[T - a_{\nu}^{\mu} E_{\nu}^{\mu}]$ es la función nula, ya que es nula para todos los elementos de la base. Entonces tenemos que

$$T = a_{\nu}^{\mu} E_{\nu}^{\mu}$$

lo que implica que las funciones E_{ν}^{μ} generan $L(V, W)$. Analicemos la independencia lineal. Consideremos la combinación lineal

$$b_{\nu}^{\mu} E_{\nu}^{\mu} = 0 \quad (\text{función nula})$$

Aplicando a todos los elementos de la base de V se obtiene que los coeficientes deben ser los nulos, por lo que se demuestra de esa manera la independencia lineal. Esto implica que las funciones E_{ν}^{μ} forman una base para $L(V, W)$ resultando entonces un espacio de dimensión $n \times m$.

II

Formas Lineales y Espacio Dual

4 Funcionales Lineales. 1-formas

Vamos a considerar ahora un tipo particular de transformaciones lineales: las funcionales lineales, llamadas también formas lineales o 1-formas.

Dado V un espacio vectorial. Por funcional lineal vamos a entender a toda transformación lineal $f : V \rightarrow \mathcal{R}$, esto es, aquellas transformaciones que a cada vector le asocia un número real.

Ejemplo: Sea $V = \mathcal{R}^3$. Consideremos la función

$$f(x, y, z) = 2x + y - z.$$

Es trivial demostrar que es una funcional lineal. Calculemos el valor asociado a cada elemento de la base canónica.

$$\begin{aligned} f(1, 0, 0) &= 2 \\ f(0, 1, 0) &= 1 \\ f(0, 0, 1) &= -1 \end{aligned}$$

Lo que nos va a interesar de las funcionales lineales es que ellas mismas poseen na estructura de espacio vectorial. Para ello debemos definir una suma y un producto por números. Consideremos el conjunto de todas las funcionales lineales sobre el espacio V . Dadas f_1 y f_2 funcionales lineales sobre V . y $\lambda \in \mathcal{R}$ definimos

Suma:

$$(f_1 + f_2)(\vec{v}) = f_1(\vec{v}) + f_2(\vec{v})$$

Producto por números

$$(\lambda f)(\vec{v}) = \lambda \cdot f(\vec{v})$$

Definidos de esta manera, se puede demostrar que el conjunto de todas las funcionales lineales sobre un espacio vectorial dado es un espacio vectorial. Este espacio es denominado *espacio dual* asociado a V y se lo denota V^* .

Es decir

$$V^* = \{f/f : V \rightarrow \mathcal{R}, \text{lineal}\}$$

Como todo espacio vectorial, al espacio dual se le puede encontrar varias bases. La base que nos interesa construir es una particular denominada *base dual*.

Ya hemos demostrado que el espacio de transformaciones lineales de V en W tiene dimensión $n \times m$, donde n es la dimensión de V y m , de W . En particular, para funcionales lineales el espacio de llegada es el espacio vectorial \mathcal{R} (si, es trivial demostrar que \mathcal{R} es un espacio vectorial de dimensión 1 y que la base canónica es $\mathcal{B}_{\mathcal{R}} = \{1\}$).

Por lo tanto la dimensión del espacio dual a un espacio V de dimensión n será $n \times 1 = n$.

Vamos ahora a construir una base para el espacio dual.

Como el espacio dual tiene dimensión n propondremos como base al conjunto:

$$\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$$

de manera tal de que cada funcional lineal $f \in V^*$ puede escribirse como

$$f = a_\mu \mathbf{dx}^\mu$$

a los elementos de la base los definiremos a partir de la propia base del espacio V .

Sea $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ una base para el espacio V , definiremos los elementos de la base $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$ a partir de las relaciones

$$\mathbf{dx}^\mu(\mathbf{e}_\nu) = \delta_\nu^\mu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}$$

Veamos que efectivamente es una base. Para demostrar esto, tengamos en cuenta que si el espacio tiene dimensión n y en ese espacio tenemos n vectores linealmente independientes, entonces pueden ser una base para el espacio.

Entonces, primero comprobemos que los elementos de \mathcal{B}^* son linealmente independientes. En efecto, consideremos una combinación lineal de los elementos de \mathcal{B}^* cuyo resultado sea la funcional nula.

$$a_\mu \mathbf{dx}^\mu = 0$$

si aplicamos a cada elemento de la base de V , por ejemplo \mathbf{e}_1 obtenemos

$$\begin{aligned} [a_\mu \mathbf{dx}^\mu](\mathbf{e}_1) &= a_\mu \mathbf{dx}^\mu(\mathbf{e}_1) \\ &= a_\mu \delta_1^\mu = a_1 = 0 \end{aligned}$$

Esto significa que el a_1 tiene que ser cero. Si aplicamos a cada elemento de la base de V obtenemos que cada coeficiente debe ser cero, lo que implica que todos tienen que ser ceros. Entonces, los elementos de \mathcal{B}^* son linealmente independientes.

Ahora, para completar la demostración, restaría probar que generan V^* . Sea f una 1-forma cualquiera de V^* . Consideremos la combinación lineal

$$\alpha_\mu \mathbf{dx}^\mu + \beta f = 0$$

Si todos los coeficientes fueran cero, entonces tendríamos un conjunto de $n + 1$ vectores linealmente independiente. Entonces, no puede ser LI. Si no son LI, entonces son dependientes, con lo que existen algunos de los números que no sean cero (además β no puede ser cero, porque implicaría que todos lo sean). Entonces, podemos escribir

$$f = -\frac{1}{\beta} \alpha_\mu \mathbf{dx}^\mu = \beta_\mu \mathbf{dx}^\mu$$

Lo que indica que cualquier vector del dual es generado por los $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$.

4.0.1 El Espacio Doble Dual

Si consideramos al espacio V^* dual de V como un espacio vectorial de dimensión n podemos preguntarnos cuál sería su propio espacio dual.

Si aprovechamos la definición de espacio dual, necesitamos encontrar funcionales lineales que a cada funcional del espacio V^* le asocie un número real.

Consideremos una funcional $f \in V^*$ y sea $\vec{v} \in V$; sabemos que $f(\vec{v}) \in \mathcal{R}$. Si consideramos a \vec{v} como una funcional en V^{**} tal que a cada funcional f le asocie un número real de la forma:

$$\vec{v}(f) \equiv f(\vec{v})$$

Tenemos entonces las funcionales sobre V^* . Tenemos entonces una identificación entre el espacio V y el dual del dual, V^{**} .

Por como identificamos los elementos de V^{**} la base del espacio doble dual deberá ser:

$$\mathcal{B}^{**} = \{\mathbf{e}_1; \mathbf{e}_2; \dots; \mathbf{e}_n\}$$

ya que

$$\mathbf{e}_\mu(\mathbf{dx}^\nu) \equiv \mathbf{dx}^\nu(\mathbf{e}_\mu) = \delta_\nu^\mu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}$$

4.1 Ejemplos de 1-formas

Ejemplo 1. Traza de una matriz cuadrada.

Dada una matriz $n \times n$, de elementos a_ν^μ , $\mu, \nu = 1, 2, \dots, n$ la traza

$$Tr(\mathbf{A}) = a_\mu^\mu$$

es una 1-forma.

Ejemplo 2. Valor numérico de un polinomio.

Sea V el espacio de los polinomios de grado menor o igual que n , $\mathcal{R}_n[x]$. Sea $t \in \mathcal{R}$. Definamos la transformación $L_t : \mathcal{R}_n[x] \rightarrow \mathcal{R}$ a través de

$$L_t(p(x)) = p(t)$$

L_t es una 1-forma.

Ejemplo 3. Funcional Integral.

Sea $[a, b]$ un intervalo cerrado de los reales y sea $C_{[a,b]}$ el espacio de las funciones reales continuas en $[a, b]$. Sea $f \in C_{[a,b]}$, la expresión

$$L(f) = \int_a^b f(t) dt$$

es una 1-forma.

Ejemplo 4. La diferencial de una función diferenciable.

Sea $f : \mathcal{R}^n \rightarrow \mathcal{R}$ una función diferenciable en un punto \mathbf{a} . Decimos que f es diferenciable en el punto \mathbf{a} si para todo \vec{h} podemos escribir la variación de f como

$$f(\mathbf{a} + \vec{h}) - f(\mathbf{a}) = T(\vec{h}) + \varepsilon(\vec{h}) \|\vec{h}\|$$

donde T es una transformación lineal y la función ε satisface $\lim_{\vec{h} \rightarrow \vec{0}} \varepsilon(\vec{h}) = 0$

La función $T(\vec{h})$ es una 1-forma y se expresa como

$$T(\vec{h}) = \left. \frac{\partial f}{\partial x^\mu} \right|_{\mathbf{a}} h^\mu$$

y se denomina *diferencial de la función*. En general, se denota a la 1-forma como

$$T = \left. \frac{\partial f}{\partial x^\mu} \right|_{\mathbf{a}} dx^\mu$$

Es a partir de esta 1-forma la elección de la notación para los elementos de la base dual.

De esta manera, la diferencial de la función aplica el vector desplazamiento \vec{h} a un número real a partir de

$$T(\vec{h}) = \left[\left. \frac{\partial f}{\partial x^\mu} \right|_{\mathbf{a}} dx^\mu \right] (\vec{h}) = \left. \frac{\partial f}{\partial x^\mu} \right|_{\mathbf{a}} dx^\mu(\vec{h}) = \left. \frac{\partial f}{\partial x^\mu} \right|_{\mathbf{a}} h^\mu$$

5 Coordenadas de vectores y de 1-formas

La definición de Espacio Dual y sus bases permite obtener las coordenadas de un vector (o 1-forma) por directa aplicación de 1-formas o vectores (vectores como elementos del doble dual).

A partir de las definiciones de las base del espacio dual y del espacio doble dual (coincidente con el propio espacio original) podemos obtener un resultado que permite identificar las coordenadas de los vectores y de las 1-formas.

En efecto, sea V un espacio vectorial de dimensión finita, n y sea V^* su espacio dual. Un vector cualquiera del espacio \vec{v} se escribe de manera invariante como $v^\mu \mathbf{e}_\mu$. Entonces, aplicando una elemento de la base dual \mathbf{dx}^ν obtenemos

$$\mathbf{dx}^\nu(\vec{v}) = \mathbf{dx}^\nu(v^\mu \mathbf{e}_\mu) = v^\mu \mathbf{dx}^\nu(\mathbf{e}_\mu) = v^\mu \delta_\nu^\mu = v^\nu$$

Lo que significa que podemos expresar a un vector de la forma

$$\vec{v} = \mathbf{dx}^\mu(\vec{v}) \mathbf{e}_\mu$$

De manera análoga tenemos que dado un elemento del espacio dual, f , éste es expandible en la base dual

$$f = f_\mu \mathbf{dx}^\mu$$

Aplicamos el elemento \mathbf{e}_ν de la base del doble dual a esta 1-forma,

$$\mathbf{e}_\nu(f) = \mathbf{e}_\nu(f_\mu \mathbf{dx}^\mu) = f_\mu \mathbf{e}_\nu(\mathbf{dx}^\mu) = f_\mu \mathbf{dx}^\mu(\mathbf{e}_\nu) = f_\mu \delta_\nu^\mu = f_\nu$$

Esto significa que podemos escribir cualquier 1-forma, f

$$f = \mathbf{e}_\mu(f) \mathbf{dx}^\mu = f(\mathbf{e}_\mu) \mathbf{dx}^\mu$$

A partir de las definiciones y las propiedades obtenidas, podemos notar que la aplicación de un elemento de la base dual a un determinado vector (en el sentido más amplio del término) nos devuelve directamente la coordenada asociada. El justamente por esta razón que las 1-formas también son llamadas como funciones coordenadas.

5.1 Aspectos Metodológicos. Caso \mathcal{R}^n

Dada una base para \mathcal{R}^n , \mathbf{e} para la obtención de la base dual debemos darnos cuenta que toda funcional lineal en este espacio tiene la forma

$$f = a_1 x^1 + a_2 x^2 + \cdots + a_n x^n$$

donde las cantidades x_j son las componentes de cada uno de los elementos de la base del espacio. Con lo cual, para determinar cada elemento de la base dual debemos aplicar a cada elemento de la base y hacerle valer 1 o 0 según corresponda. De esta manera, lo que obtendremos serán los elementos de la base dual.

5.2 Ejemplos

Ejemplo 1. Para \mathcal{R}^2 , consideremos la base $\mathcal{B} = \{(-1, 1); (1, 0)\}$. Determinar

- La base dual
- Las coordenadas del vector $\vec{v} = (3, 5)$
- Las coordenadas de la 1-forma $f(x, y) = 2x + 3y$

Solución. a) Para determinar la base dual, simplemente debemos aplicar la definición. Primero, como las 1-formas en \mathcal{R}^2 tienen la forma $f(x, y) = ax + by$ lo que debemos plantear es si: $\mathbf{dx}^1 = a_1 x + b_1 y$ (expresión en componentes) y $\mathbf{dx}^2 = a_2 x + b_2 y$ Para que satisfagan las propiedades de base dual, tenemos que deben satisfacer:

$$\mathbf{dx}^1(-1, 1) = 1, \quad \mathbf{dx}^1(1, 0) = 0, \quad \mathbf{dx}^2(-1, 1) = 0, \quad \mathbf{dx}^2(1, 0) = 1$$

entonces, las relaciones serán

$$\begin{aligned} \mathbf{dx}^1(-1, 1) &= a_1(-1) + b_1(1) = 1 \\ \mathbf{dx}^1(1, 0) &= a_1(1) + b_1(0) = 0 \\ \mathbf{dx}^2(-1, 1) &= a_2(-1) + b_2(1) = 0 \\ \mathbf{dx}^2(1, 0) &= a_2(1) + b_2(0) = 1 \end{aligned}$$

Resultando,

$$a_1 = 0, \quad b_1 = 1, \quad a_2 = 1, \quad b_2 = 1$$

Entonces, los elementos de la base dual son:

$$\mathbf{dx}^1(x, y) = y, \quad \mathbf{dx}^2(x, y) = x + y$$

b) Para conocer las coordenadas del vector, simplemente aplicamos las 1-formas al vector y cada resultado será la coordenada. Es decir, si $\vec{v} = v^1(-1, 1) + v^2(1, 0)$ tendremos que $v^1 = \mathbf{dx}^1(\vec{v})$ y $v^2 = \mathbf{dx}^2(\vec{v})$ Entonces,

$$v^1 = \mathbf{dx}^1(3, 5) = 5, \quad v^2 = \mathbf{dx}^2(3, 5) = 3 + 5 = 8$$

c) La 1-forma $f(x, y) = 2x + 3y$ será expresada como $f(x, y) = (a_1 \mathbf{dx}^1 + a_2 \mathbf{dx}^2)(x, y)$ para conocer las coordenadas del funcional, aplicamos $f(-1, 1)$ y $f(1, 0)$. Entonces,

$$a_1 = f(-1, 1) = 2 \cdot (-1) + 3 \cdot 1 = 1, \quad a_2 = f(1, 0) = 2 \cdot (1) + 3 \cdot 0 = 2$$

Entonces, la funcional f se escribe como

$$f = \mathbf{dx}^1 + 2 \mathbf{dx}^2$$

6 Cambio de Coordenadas para 1-formas

Consideremos un espacio vectorial de dimensión n con base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Esta base tendrá asociada una base dual $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$. Consideremos ahora un cambio de base

$$\mathbf{e}'_\mu = \Lambda_\mu^\nu \mathbf{e}_\nu$$

Queremos ver qué expresión tendrá el cambio de coordenadas

$$\mathbf{dx}'^\alpha = \Phi_\beta^\alpha \mathbf{dx}^\beta$$

Para ver cómo se relaciona, apliquemos a ambos miembros la 1-forma al elemento de la base \mathbf{e}'_λ , esto es

$$\mathbf{dx}'^\alpha(\mathbf{e}'_\lambda) = \Phi_\beta^\alpha \mathbf{dx}^\beta(\mathbf{e}'_\lambda)$$

aplicando el cambio de base

$$\mathbf{dx}'^\alpha(\mathbf{e}'_\lambda) = \Phi_\beta^\alpha \mathbf{dx}^\beta(\Lambda_\lambda^\nu \mathbf{e}_\nu) = \Phi_\beta^\alpha \Lambda_\lambda^\nu \mathbf{dx}^\beta(\mathbf{e}_\nu) = \Phi_\beta^\alpha \Lambda_\lambda^\nu \delta_\nu^\beta$$

Entonces,

$$\delta_\lambda^\alpha = \Phi_\nu^\alpha \Lambda_\lambda^\nu$$

Lo que implica que la matriz de cambio $\Phi = \Lambda^{-1}$. Con esto, tenemos que para el cambio de coordenadas de funcionales, tendremos, para una forma $f = f_\mu \mathbf{dx}^\mu$ una representación en una nueva base $\mathcal{B}'^* = \{\mathbf{dx}'^1, \mathbf{dx}'^2, \dots, \mathbf{dx}'^n\}$ como $f = f'_\mu \mathbf{dx}'^\mu$ donde

$$f'_\mu = \Lambda_\mu^\nu f_\nu$$

7 Resumen de Cambio de Base y Coordenadas

Vamos a esquematizar en un cuadro las relaciones que se obtuvieron a partir de un cambio de base, y sus repercusiones en el cambio de coordenadas, en el cambio de base del espacio dual y de coordenadas en el espacio dual.

Consideremos un espacio V de dimensión n , con base original $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Si efectuamos un cambio de base, de la base original a una nueva base $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ podemos resumir todos los cambios resultantes

Espacio	Base	Coordenadas
Espacio V	$\mathbf{e}'_\mu = \Lambda_\mu^\nu \mathbf{e}'_\nu$	$x'^\mu = [\Lambda^{-1}]^\mu_\nu x^\nu$
Espacio V^*	$\mathbf{dx}'^\mu = [\Lambda^{-1}]^\mu_\nu \mathbf{dx}^\nu$	$f'_\mu = \Lambda_\mu^\nu f_\nu$

III

Formas Bilineales y Multilineales. Tensores

8 Formas Bilineales sobre V

Dado un espacio vectorial V , de dimensión n , vamos a estudiar las funciones del tipo $f : V \times V \rightarrow \mathcal{R}$, es decir asocian a un par de vectores \vec{u} y \vec{v} un número real. Es decir, que la función aplica $f(\vec{u}; \vec{v})$ a un número real. Si además, con respecto a cada argumento la función es lineal, es decir

- $f(\vec{u}; \lambda \cdot \vec{v}_1 + \vec{v}_2) = \lambda f(\vec{u}; \vec{v}_1) + f(\vec{u}; \vec{v}_2)$
- $f(\lambda \cdot \vec{u}_1 + \vec{u}_2; \vec{v}) = \lambda f(\vec{u}_1; \vec{v}) + f(\vec{u}_2; \vec{v})$

Se dice que la función es *bilineal*.

Podemos notar que, dado un espacio vectorial V y su dual V^* una forma bilineal puede ser obtenida de la siguiente manera: Sean f y g dos elementos del espacio dual V^* . Entonces,

$$F(\vec{u}, \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$$

(donde notemos que el producto se realiza en el cuerpo) es una forma bilineal. Más aún, toda forma bilineal puede ser escrita como producto de dos 1-formas.

Consideremos, como ejemplo, la forma bilineal sobre \mathcal{R}^2

$$F(\vec{u}; \vec{v}) = 2u_x v_x - 4u_x v_y + u_y v_x - 2u_y v_y$$

puede escribirse como

$$F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$$

donde

$$f(\vec{u}) = 2u_x + u_y, \quad g(\vec{v}) = v_x - 2v_y$$

Veamos que esto no es una particularidad, sino que cualquier forma bilineal puede ser escrita como "producto" de dos 1-formas. El encomillado viene de que lo que se ve como producto es la aplicación en los vectores, pero no sabemos que es producto de 1-formas.

8.1 Álgebra de las Formas Bilineales

Consideremos el conjunto de todas las formas bilineales $F : V \times V \rightarrow \mathcal{R}$. Vamos a dotar a este conjunto de una suma y un producto por un escalar, de la siguiente manera

- **Suma.** Dadas dos formas bilineales F y G , se define la suma a través de

$$[F + G](\vec{u}; \vec{v}) = F(\vec{u}; \vec{v}) + G(\vec{u}; \vec{v})$$

- **Producto por escalar.** El producto por un escalar se define

$$[\lambda \cdot F](\vec{u}; \vec{v}) = \lambda \cdot F(\vec{u}; \vec{v})$$

Estas definiciones permiten comprobar que el conjunto de las formas bilineales poseen una estructura de espacio vectorial. De deja como ejercicio comprobar que la dimensión de este espacio es n^2

Aplicando estas definiciones de suma y producto para formas bilineales, y el hecho de poder expresar la forma bilineal como el producto

$$F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$$

Hagamos

$$F(\vec{u}; \vec{v}) = f_\mu \mathbf{dx}^\mu (u^\nu \mathbf{e}_\nu) \cdot g_\alpha \mathbf{dx}^\alpha (v^\beta \mathbf{e}_\beta) = f_\mu g_\alpha u^\nu v^\beta \mathbf{dx}^\mu (\mathbf{e}_\nu) \cdot \mathbf{dx}^\alpha (\mathbf{e}_\beta)$$

Entonces,

$$F(\vec{u}; \vec{v}) = f_\mu g_\alpha u^\nu v^\beta \delta_\nu^\mu \cdot \delta_\beta^\alpha = f_\mu g_\alpha u^\mu v^\alpha$$

Esta es, entonces, una de las maneras de calcular la aplicación bilineal. Vamos a construir una base para el espacio de formas bilineales.

9 Producto Tensorial de 1-formas

A partir de poder escribir la aplicación bilineal como producto de aplicar a cada argumento,

$$F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$$

se define el *producto tensorial de dos 1-formas* f y g y se denota $f \otimes g$ a través de

$$(f \otimes g)(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$$

Algunas propiedades del producto tensorial

- $(f_1 + f_2) \otimes g = f_1 \otimes g + f_2 \otimes g$
- $f \otimes (g_1 + g_2) = f \otimes g_1 + f \otimes g_2$
- $(\lambda \cdot f) \otimes g = \lambda \cdot (f \otimes g)$
- $f \otimes (\lambda \cdot g) = \lambda \cdot (f \otimes g)$

Lo que no siempre se cumple es la conmutatividad, esto es, en general $f \otimes g \neq g \otimes f$.

Con estas definiciones se puede probar que el conjunto *producto tensorial de 1-formas* es un espacio vectorial. Además, posee dimensión n^2

Vamos a proponer como base de este espacio al conjunto

$$\mathcal{B} = \{\mathbf{dx}^\mu \otimes \mathbf{dx}^\nu\} \quad \mu, \nu = 1, 2, \dots, n$$

Más aún, el espacio definido de esta manera se lo denomina espacio producto tensorial $V^* \otimes V^*$.

Los elementos de este espacio se denominan tensores de tipo $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$

9.1 Coordenadas de un tensor de $V^* \otimes V^*$

Dada una forma bilineal $F : V \times V \rightarrow \mathcal{R}$ esta forma pertenece al espacio generado por la base $\mathcal{B} = \{\mathbf{dx}^\mu \otimes \mathbf{dx}^\nu\} \quad \mu, \nu = 1, 2, \dots, n$ entonces, F puede escribirse como

$$F = f_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$$

Veamos cuales deben ser las coordenadas $f_{\mu\nu}$. Aplicando a ambos miembros del desarrollo elementos de la base de V , tenemos:

$$F(\mathbf{e}_\alpha; \mathbf{e}_\beta) = f_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu (\mathbf{e}_\alpha; \mathbf{e}_\beta)$$

Por definición de producto tensorial de 1-formas, tenemos

$$F(\mathbf{e}_\alpha; \mathbf{e}_\beta) = f_{\mu\nu} \mathbf{dx}^\mu(\mathbf{e}_\alpha) \cdot \mathbf{dx}^\nu(\mathbf{e}_\beta) = f_{\mu\nu} \delta_\alpha^\mu \delta_\beta^\nu$$

Entonces,

$$f_{\alpha\beta} = F(\mathbf{e}_\alpha; \mathbf{e}_\beta)$$

Entonces,

$$F = F(\mathbf{e}_\mu; \mathbf{e}_\nu) \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$$

Si aplicamos esta forma bilineal a un par de vectores cualesquiera, \vec{u} y \vec{v} tenemos

$$F(\vec{u}; \vec{v}) = f_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu(u^\alpha \mathbf{e}_\alpha; v^\beta \mathbf{e}_\beta)$$

resultando

$$F(\vec{u}; \vec{v}) = f_{\mu\nu} u^\alpha v^\beta \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu(\mathbf{e}_\alpha; \mathbf{e}_\beta) = f_{\mu\nu} u^\alpha v^\beta \mathbf{dx}^\mu(\mathbf{e}_\alpha) \cdot \mathbf{dx}^\nu(\mathbf{e}_\beta) = f_{\mu\nu} u^\alpha v^\beta \delta_\mu^\alpha \delta_\nu^\beta$$

Entonces

$$F(\vec{u}; \vec{v}) = f_{\mu\nu} u^\mu v^\nu = F(\mathbf{e}_\mu; \mathbf{e}_\nu) u^\mu v^\nu$$

9.2 Coordenadas de un tensor de $V \otimes V$

Así como definimos e identificamos el espacio doble dual con el propio espacio, podemos entonces, definir una forma bilineal que asocie a dos 1-formas a un número real: $F : V^* \times V^* \rightarrow \mathcal{R}$, definida como $F(f, g)$ donde f y g son 1-formas. De manera análoga a lo que vimos en la sección anterior, podemos escribir esta forma bilineal como producto tensorial de dos 1-formas de V^{**} es decir

$$F(f, g) = (\vec{u} \otimes \vec{v})(f, g) = \vec{u}(f) \cdot \vec{v}(g) = f(\vec{u}) \cdot g(\vec{v})$$

A partir del álgebra del producto tensorial, podemos notar que

$$\{\mathbf{e}_\mu \otimes \mathbf{e}_\nu\}, \quad \mu, \nu = 1, 2, \dots, n$$

es una base para el espacio $V \otimes V$. Tomando esta base para el espacio, podemos notar que todo tensor en $V \otimes V$ puede escribirse como

$$T = t^{\mu\nu} (\mathbf{e}_\mu \otimes \mathbf{e}_\nu)$$

de manera tal que para conocer las coordenadas en esta base sólo es necesario aplicar esta forma bilineal *-tensor-* a los elementos de la base de V^*

$$T(\mathbf{dx}^\alpha, \mathbf{dx}^\beta) = t^{\alpha\beta}$$

10 Tensores Cartesianos en General

Sea V un espacio vectorial de dimensión n . Sea V^* el espacio dual. Sean $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ y $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$ las bases del espacio y del dual.

Definimos como *tensor del tipo* $\binom{p}{q}$ a toda forma multilineal

$$T : \underbrace{V \times V \times \dots \times V}_{p\text{-veces}} \times \underbrace{V^* \times V^* \times \dots \times V^*}_{q\text{-veces}} \rightarrow \mathcal{R}$$

la cual es definida a partir del producto tensorial. Es decir

$$T \in \underbrace{V^* \otimes V^* \otimes \dots \otimes V^*}_{p\text{-veces}} \otimes \underbrace{V \otimes V \otimes \dots \otimes V}_{q\text{-veces}}$$

De esta manera, las formas bilineales sobre V ,

$$f_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$$

es un tensor del tipo $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ y las formas bilineales sobre V^* ,

$$f^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

son tensores del tipo $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$

La definición de tensor en general, permite definir un tensor de tipo $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ que aplica a un elemento de V y un elemento de V^* un número real, con lo que podremos escribirlo

$$T = T_\mu^\nu \mathbf{dx}^\mu \otimes \mathbf{e}_\nu$$

donde las coordenadas se obtienen aplicando la forma a los elementos de la base

$$T_\mu^\nu = T(\mathbf{e}_\mu, \mathbf{dx}^\nu)$$

En general, un tensor del tipo $\begin{pmatrix} p \\ q \end{pmatrix}$ se escribe como

$$T = T_{\mu_1 \mu_2 \dots \mu_p}^{\nu_1 \nu_2 \dots \nu_q} \mathbf{dx}^{\mu_1} \otimes \mathbf{dx}^{\mu_2} \otimes \dots \otimes \mathbf{dx}^{\mu_p} \otimes \mathbf{e}_{\nu_1} \otimes \mathbf{e}_{\nu_2} \otimes \dots \otimes \mathbf{e}_{\nu_q}$$

donde las coordenadas se obtienen a partir de aplicarlos a los elementos de la base

$$T_{\mu_1 \mu_2 \dots \mu_p}^{\nu_1 \nu_2 \dots \nu_q} = T(\mathbf{e}_{\mu_1}, \mathbf{e}_{\mu_2}, \dots, \mathbf{e}_{\mu_p}; \mathbf{dx}^{\nu_1}, \mathbf{dx}^{\nu_2}, \dots, \mathbf{dx}^{\nu_q})$$

Con esta definición, tanto los vectores como las 1-formas serán tensores. Los vectores -visto como 1-formas sobre V^* - son tensores del tipo $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ y las 1-formas sobre V (es decir los elementos de V^*) serán tensores del tipo $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

10.1 Covarianza y Contravarianza de un tensor

A partir de la definición de tensor de tipo $\begin{pmatrix} p \\ q \end{pmatrix}$ se dice que es p veces covariante y q veces contravariante. Entonces, un tensor del tipo $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ es de la forma

$$T = T_{\alpha\beta}^\gamma \mathbf{dx}^\alpha \otimes \mathbf{dx}^\beta \otimes \mathbf{e}_\gamma$$

y la cantidad de subíndices indican la cantidad de veces covariante y los supraíndices, la cantidad de veces contravariante.

11 Cambio de coordenadas en tensores cartesianos

Consideremos un cambio de base

$$\mathbf{e}'_\nu = \Lambda_\nu^\mu \mathbf{e}_\mu, \quad \mathbf{e}_\nu = [\Lambda^{-1}]_\nu^\mu \mathbf{e}'_\mu$$

Como hemos visto, este cambio de base, produce un cambio de coordenadas

$$x'^\mu = [\Lambda^{-1}]_\nu^\mu x^\nu, \quad x^\mu = \Lambda_\nu^\mu x'^\nu$$

A su vez, en el espacio dual induce a los cambios de base y coordenadas

$$\mathbf{dx}'^\nu = [\Lambda^{-1}]_\nu^\mu \mathbf{dx}^\mu, \quad f'_\nu = \Lambda_\nu^\mu f_\mu$$

junto con los cambios inversos

$$\mathbf{dx}^\nu = \Lambda_\mu^\nu \mathbf{dx}'^\mu, \quad f_\nu = [\Lambda^{-1}]_\nu^\mu f'_\mu$$

Tomemos un tensor dos veces covariante. Este objeto admite una representación invariante $T = T_{\alpha\beta} \mathbf{dx}^\alpha \otimes \mathbf{dx}^\beta$ lo que significa que esta representación no depende del sistema de coordenadas (o base elegida, que es equivalente), por lo que podemos escribir

$$T'_{\mu\nu} \mathbf{dx}'^\mu \otimes \mathbf{dx}'^\nu = T_{\alpha\beta} \mathbf{dx}^\alpha \otimes \mathbf{dx}^\beta$$

Reemplazando el cambio de los elementos \mathbf{dx}^α en función de la base nueva tenemos

$$T'_{\mu\nu} \mathbf{dx}'^\mu \otimes \mathbf{dx}'^\nu = T_{\alpha\beta} \Lambda_\mu^\alpha \mathbf{dx}'^\mu \otimes \Lambda_\nu^\beta \mathbf{dx}'^\nu = T_{\alpha\beta} \Lambda_\mu^\alpha \Lambda_\nu^\beta \mathbf{dx}'^\mu \otimes \mathbf{dx}'^\nu$$

entonces,

$$\left[T'_{\mu\nu} - \Lambda_\mu^\alpha \Lambda_\nu^\beta T_{\alpha\beta} \right] \mathbf{dx}'^\mu \otimes \mathbf{dx}'^\nu = 0$$

y como los $\mathbf{dx}'^\mu \otimes \mathbf{dx}'^\nu$ son LI, tenemos que el cambio de coordenadas para un tensor dos veces covariante resulta

$$T'_{\mu\nu} = \Lambda_\mu^\alpha \Lambda_\nu^\beta T_{\alpha\beta}$$

Tomemos un tensor dos veces contravariante. De manera invariante, podemos escribir

$$T'^{\mu\nu} \mathbf{e}'_\mu \otimes \mathbf{e}'_\nu = T^{\alpha\beta} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta$$

Reemplazando el cambio de base (la base sin primar con respecto a las primadas), tenemos

$$T'^{\mu\nu} \mathbf{e}'_\mu \otimes \mathbf{e}'_\nu = T^{\alpha\beta} [\Lambda^{-1}]_\alpha^\mu [\Lambda^{-1}]_\beta^\nu \mathbf{e}'_\mu \otimes \mathbf{e}'_\nu$$

Agrupando como antes, y aprovechando la independencia lineal de los $\mathbf{e}'_\mu \otimes \mathbf{e}'_\nu$ tenemos

$$T'^{\mu\nu} = [\Lambda^{-1}]_\alpha^\mu [\Lambda^{-1}]_\beta^\nu T^{\alpha\beta}$$

Tensor una vez covariante y una vez contravariante. Este objeto, se escribe como

$$T'^{\mu\nu} \mathbf{dx}'^\mu \otimes \mathbf{e}'_\nu = T^\beta_\alpha \mathbf{dx}^\alpha \otimes \mathbf{e}_\beta$$

De la misma manera que se hizo para los casos anteriores, reemplazamos el cambio de bases del espacio y del dual (de las sin primar con respecto a las primadas)

$$T'^{\nu} \mathbf{dx}'^{\mu} \otimes \mathbf{e}'_{\nu} = T^{\beta}_{\alpha} [\Lambda^{-1}]^{\nu}_{\beta} \Lambda^{\alpha}_{\mu} \mathbf{dx}'^{\mu} \otimes \mathbf{e}'_{\nu}$$

Con lo cual, se obtiene

$$T'^{\nu} = [\Lambda^{-1}]^{\nu}_{\beta} \Lambda^{\alpha}_{\mu} T^{\beta}_{\alpha}$$

Procedimiento esquemático para cambio de coordenadas de tensores.

A partir de estos ejemplos se infiere lo siguiente: Dado un tensor p – veces covariante y q – veces contravariante, los cambios de coordenadas del tensor se realizan según el esquema

- Por cada índice covariante $[]_{\mu}$ el cambio contiene una matriz Λ

$$[]'_{\mu} = \Lambda^{\alpha}_{\mu} []_{\alpha}$$

- Por cada índice contravariante $[]^{\mu}$ el cambio contiene una matriz Λ^{-1}

$$[]'^{\mu} = [\Lambda^{-1}]^{\mu}_{\alpha} []^{\alpha}$$

Siguiendo este criterio, dado un tensor p – veces covariante y q – veces contravariante cuyas coordenadas en un sistema son

$$T^{\mu_1 \mu_2 \mu_3 \dots \mu_q}_{\nu_1 \nu_2 \nu_3 \dots \nu_p}$$

el cambio de coordenadas queda

$$T'^{\mu_1 \mu_2 \mu_3 \dots \mu_q}_{\nu_1 \nu_2 \nu_3 \dots \nu_p} = \Lambda^{\beta_1}_{\nu_1} \Lambda^{\beta_2}_{\nu_2} \dots \Lambda^{\beta_p}_{\nu_p} [\Lambda^{-1}]^{\mu_1}_{\alpha_1} [\Lambda^{-1}]^{\mu_2}_{\alpha_2} \dots [\Lambda^{-1}]^{\mu_q}_{\alpha_q} T^{\alpha_1 \alpha_2 \alpha_3 \dots \alpha_q}_{\beta_1 \beta_2 \beta_3 \dots \beta_p}$$

En algunos textos clásicos, tales como el libro *Vectores y Tensores* de Luis Santaló (Eudeba, 1973), las nociones de vectores y tensores se orientan al comportamiento de las coordenadas de los mismos a partir de cambios de coordenadas. Así, por ejemplo, en el texto citado define: "...” *Ciertas cantidades* $T^{\mu_1 \mu_2 \mu_3 \dots \mu_q}_{\nu_1 \nu_2 \nu_3 \dots \nu_p}$ *son las coordenadas de un tensor* p *veces covariante y* q *veces contravariante si frente a un cambio de coordenadas* $x'^{\alpha} = B^{\alpha}_{\beta} x^{\beta}$ *las coordenadas se transforman como*

$$T'^{\mu_1 \mu_2 \mu_3 \dots \mu_q}_{\nu_1 \nu_2 \nu_3 \dots \nu_p} = \Lambda^{\beta_1}_{\nu_1} \Lambda^{\beta_2}_{\nu_2} \dots \Lambda^{\beta_p}_{\nu_p} B^{\mu_1}_{\alpha_1} B^{\mu_2}_{\alpha_2} \dots B^{\mu_q}_{\alpha_q} T^{\alpha_1 \alpha_2 \alpha_3 \dots \alpha_q}_{\beta_1 \beta_2 \beta_3 \dots \beta_p}$$

...”

Aquí, las matrices Λ y B son inversas una de otra.

A este criterio Santaló lo define como un *criterio de tensorialidad* de las magnitudes. Entonces, se define un objeto por lo que le ocurre cuando se cambian las coordenadas.

En nuestro abordaje, definimos las magnitudes de manera intrínseca y luego los cambios de base producen una regla de cambio de coordenadas en los tensores.

El libro de Luis Santaló trabaja casi en su totalidad en coordenadas.

Vectores y Tensores ha sido -y en gran medida lo sigue siendo- una fuente inagotable de consulta para el estudio de estos temas, aunque el lenguaje se haya modificado.

12 Formas Cuadráticas

Dada una forma bilineal sobre un espacio V , es decir, que la forma asigna a dos vectores de V un número real. En términos de la aplicación, tenemos que una forma bilineal F , tiene por expresión $F(\vec{u}; \vec{v}) \in \mathcal{R}$. Una *forma cuadrática* se obtiene a partir de calcular la forma bilineal para un mismo vector $F(\vec{v}; \vec{v})$.

Para el ejemplo ya visto, en el cual,

$$F(\vec{u}; \vec{v}) = 2u_x v_x - 4u_x v_y + u_y v_x - 2u_y v_y$$

si calculamos

$$F(\vec{v}; \vec{v}) = 2v_x^2 - 4v_x v_y + v_y v_x - 2v_y^2 = 2v_x^2 - 3v_x v_y - 2v_y^2$$

Entonces, esta forma bilineal define una forma cuadrática (llamando $(v_x, v_y) = (x, y)$)

$$F(\vec{v}; \vec{v}) = Q(x, y) = 2x^2 - 3xy - 2y^2$$

Las formas p -lineales sobre un espacio vectorial V si se aplica a un mismo vector se obtiene una expresión polinómica homogénea de orden p , lo que significa que cada término tiene una potencia total igual a p .

En el caso de formas cuadráticas, cada término es homogéneo de orden 2, como se ve en el ejemplo $Q(x, y) = 2x^2 - 3xy - 2y^2$ donde cada término es cuadrático (los términos son x^2, xy, y^2 los cuales sumando los índices da 2).

En la búsqueda de extremos relativos para una función de n variables reales al efectuar el desarrollo de Taylor para estudiar su comportamiento alrededor de un punto crítico, donde sus derivadas parciales se anulan tenemos que podemos escribir, si \vec{x}_0 es el vector donde la función tiene un punto crítico, el desarrollo de Taylor es

$$f(\vec{x}) = f(\vec{x}_0) + \left. \frac{\partial f}{\partial x^\mu} \right|_{\vec{x}_0} (x^\mu - x_0^\mu) + \frac{1}{2} h_{\mu\nu}(\vec{x}_0) (x^\mu - x_0^\mu) (x^\nu - x_0^\nu) + \dots$$

donde $h_{\mu\nu}(\vec{x}_0)$ son los elementos de la matriz *Hessiana*.

Podemos notar que el tercer término del desarrollo es una forma cuadrática. Conocer propiedades de esta forma cuadrática permitirá caracterizar los puntos críticos.

Si en el punto (vector) \vec{x}_0 las derivadas parciales son nulas, podemos escribir

$$f(\vec{x}) - f(\vec{x}_0) = \frac{1}{2} h_{\mu\nu}(\vec{x}_0) (x^\mu - x_0^\mu) (x^\nu - x_0^\nu) + \dots$$

Lo que significa que la naturaleza del punto crítico dependerá del signo que tenga la forma cuadrática.

Más adelante estudiaremos un método que permitirá obtener el signo de una forma cuadrática a partir de un método que procura, entre otras cosas, llevar expresiones de formas cuadráticas a formas canónicas.

IV

Espacios Euclídeos. Espacios Métricos

13 Producto Interno

En los estudios básicos sobre vectores, hemos visto las operaciones *producto escalar* y *producto vectorial* entre vectores del espacio tridimensional.

Dado un par de vectores de vectores de \mathcal{R}^3 , $\vec{u} = (u_x, u_y, u_z)$ y $\vec{v} = (v_x, v_y, v_z)$ conocemos la definición de producto escalar como

$$\vec{u} \cdot \vec{v} = u_x v_x + u_y v_y + u_z v_z$$

Con esta definición se pueden identificar ángulos de inclinaciones relativas, criterios de perpendicularidad, etc. Sin embargo, al trabajar con vectores de manera más general que ternas o pares ordenados, es necesario definir un producto interno entre elementos de un espacio vectorial de forma tal que sea aplicable a la variedad de elementos que ahora tienen la cualidad de *vectores*, como por ejemplo, \mathcal{R}^n , $\mathcal{R}^{n \times m}$, las funciones continuas en determinado intervalo, etc.

Como hicimos para espacios vectoriales, definiremos las operaciones no sobre un espacio en particular, sino que nos centraremos en las propiedades que deben satisfacer para ser llamadas como tales.

En ese sentido, un producto interno sobre un espacio determinado será una operación que deberá satisfacer determinadas propiedades. Esto da más libertad para definir propiamente la operación y, para el caso de \mathcal{R}^3 el producto escalar ya conocido será un caso particular de producto interno, que llamaremos *producto interno canónico*.

13.1 Axiomas de producto interno

Definición de producto interno. Dado un espacio vectorial V sobre el cuerpo de los números reales. Una operación $\langle \vec{u} | \vec{v} \rangle$ es un producto interno siempre y cuando se satisfagan las siguientes propiedades

- $\langle \vec{u} | \vec{v} \rangle$ es un número real
- $\langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle$ (conmutatividad sólo para espacios sobre los reales)
- $\langle \lambda \vec{u}_1 + \vec{u}_2 | \vec{v} \rangle = \lambda \langle \vec{u}_1 | \vec{v} \rangle + \langle \vec{u}_2 | \vec{v} \rangle$
- $\langle \vec{v} | \vec{v} \rangle > 0 \quad \forall \vec{v} \neq \vec{0}, \quad \langle \vec{v} | \vec{v} \rangle = 0 \leftrightarrow \vec{v} = \vec{0}$

Ejercicio. Comprobar que el producto interno canónico (producto escalar ya conocido) en \mathcal{R}^2 definido como

$$\vec{u} \cdot \vec{v} = u_x v_x + u_y v_y$$

es efectivamente un producto interno.

Ejercicio. Comprobar que la operación binaria en \mathcal{R}^2 definida como

$$\langle \vec{u} | \vec{v} \rangle = u_x v_x - u_y v_x - u_x v_y + 4 u_y v_y$$

es un producto interno.

Ejemplos de producto interno.

Producto interno en el espacio de funciones. Consideremos el espacio de funciones continuas en el intervalo $[-\pi, \pi]$. Un producto interno en este espacio está definido a partir de

$$\langle f|g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$$

Ejercicio. Comprobar que así definido es efectivamente un producto interno en el espacio de funciones continuas.

Producto interno en el espacio de matrices $n \times n$. Para el espacio de matrices $\mathcal{R}^{n \times n}$ se define el producto

$$\langle \mathbf{A}|\mathbf{B} \rangle = a_{\nu}^{\mu} b_{\mu}^{\nu}$$

Ejercicio. Comprobar que efectivamente es un producto interno.

Los espacios vectoriales dotados de un producto interno se los denominan *espacios producto interno* o *espacios euclídeos*.

13.2 Norma o módulo de un vector

Una vez definido un producto interno, podemos definir la norma de un vector de un espacio vectorial \vec{v} como

$$\|\vec{v}\| = \sqrt{\langle \vec{v}|\vec{v} \rangle} \quad \text{o} \quad \|\vec{v}\|^2 = \langle \vec{v}|\vec{v} \rangle$$

Teorema. Si V es un espacio producto interno se cumple

- $\|\lambda \vec{v}\| = |\lambda| \|\vec{v}\|$
- $\|\vec{v}\| > 0$, para $\vec{v} \neq \vec{0}$
- $|\langle \vec{u}|\vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$ *Desigualdad de Cauchy-Schwarz*
- $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ *Desigualdad triangular*

Las demostraciones de los dos primeros puntos son inmediatas a partir de la definición de producto interno, y se deja como ejercicio. Para demostrar la desigualdad de Cauchy-Schwarz consideremos el vector

$$\vec{w} = \vec{u} - \frac{\langle \vec{u}|\vec{v} \rangle}{\|\vec{v}\|^2} \vec{v}$$

Si calculamos el cuadrado de la norma del vector \vec{w} es un número positivo. Entonces,

$$0 \leq \|\vec{w}\|^2 = \langle \vec{u} - \frac{\langle \vec{u}|\vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} | \vec{u} - \frac{\langle \vec{u}|\vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \rangle$$

Haciendo las cuentas,

$$0 \leq \|\vec{u}\|^2 - \frac{|\langle \vec{u} | \vec{v} \rangle|^2}{\|\vec{v}\|^2}$$

de donde se obtiene la desigualdad de Cauchy - Schwarz.

$$|\langle \vec{u} | \vec{v} \rangle|^2 \leq \|\vec{u}\|^2 \|\vec{v}\|^2 \quad \rightarrow \quad |\langle \vec{u} | \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$$

Calculemos ahora $\|\vec{u} + \vec{v}\|^2$. Tenemos que

$$\|\vec{u} + \vec{v}\|^2 = \langle \vec{u} + \vec{v} | \vec{u} + \vec{v} \rangle = \langle \vec{u} | \vec{u} \rangle + \langle \vec{v} | \vec{v} \rangle + 2 \langle \vec{u} | \vec{v} \rangle$$

Pero además, tenemos $\langle \vec{u} | \vec{v} \rangle \leq |\langle \vec{u} | \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$, con lo que

$$\|\vec{u} + \vec{v}\|^2 \leq \langle \vec{u} | \vec{u} \rangle + \langle \vec{v} | \vec{v} \rangle + 2 \|\vec{u}\| \|\vec{v}\| = (\|\vec{u}\| + \|\vec{v}\|)^2$$

De donde se obtiene la desigualdad triangular.

13.3 Ortogonalidad

El producto escalar ya conocido (el canónico) induce la noción de ángulo entre vectores. De hecho, un criterio de perpendicularidad entre vectores se obtiene a partir de la nulidad del producto escalar.

La ortogonalidad, ahora, será un concepto estrictamente algebraico, ya que para espacios vectoriales generales no existen ángulos entre vectores.

No obstante -y como pasa casi siempre en matemática- tomamos estas ideas para generalizarlas. Ahora vamos a definir la *ortogonalidad* de vectores cualesquiera (de cualquier naturaleza) a partir del producto interno.

Definición. Sea V un espacio producto interno. Sean \vec{u} y \vec{v} dos vectores de V . Se dice que \vec{u} y \vec{v} son ortogonales si y sólo si

$$\langle \vec{u} | \vec{v} \rangle = 0$$

Definición. Sea V un espacio producto interno. Sea $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ de vectores no nulos. Se dice que un conjunto ortogonal, si los vectores son ortogonales de a pares, esto es

$$\langle \vec{v}_\mu | \vec{v}_\nu \rangle = 0 \quad \text{si} \quad \mu \neq \nu$$

Ortogonalidad implica independencia lineal. Un conjunto ortogonal es linealmente independiente.

En efecto, consideremos el conjunto ortogonal $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$. Calculemos

$$a^\mu \vec{v}_\mu = 0$$

calculemos para algún \vec{v}_ν el producto interno

$$\langle a^\mu \vec{v}_\mu | \vec{v}_\nu \rangle = 0$$

$$a^\mu \langle \vec{v}_\mu | \vec{v}_\nu \rangle = 0$$

Como el conjunto es ortogonal, tenemos que son ortogonales de a pares, sólo el término en el que se multiplica $\langle \vec{v}_\nu | \vec{v}_\nu \rangle$ es no nulo, obteniendo

$$a^\nu \|\vec{v}_\nu\|^2 = 0$$

y por definición de producto interno, $\|\vec{v}_\nu\|^2 > 0$, se debe cumplir $a^\nu = 0$. Si repetimos el procedimiento para todos los elementos del conjunto, obtenemos que todos los coeficientes de la combinación lineal deben ser cero, por lo que son linealmente independientes.

Definición. Sea V un espacio producto interno. Sea $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ un conjunto de vectores de V . El conjunto se llama ortonormal si y sólo si

$$\langle \vec{v}_\mu | \vec{v}_\nu \rangle = \delta_{\mu\nu} = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}$$

Esto significa que un conjunto ortonormal es un conjunto ortogonal cuyos elementos poseen norma unidad.

13.4 Construcción de una base ortogonal

Unos de los resultados más importantes para espacios producto interno es la posibilidad de contar con una base ortogonal. Ya sabiendo que un conjunto ortogonal es linealmente independiente, poder construir n (donde n es la dimensión del espacio) vectores ortogonales nos garantiza una base.

El procedimiento es el denominado proceso de ortogonalización de Gram-Schmidt.

Comencemos con un conjunto $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$. Vamos a construir otro conjunto $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ mediante el siguiente algoritmo:

$$\begin{aligned} \vec{v}_1 &= \vec{u}_1 \\ \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2 | \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 \end{aligned}$$

Notemos que ahora $\langle \vec{v}_1 | \vec{v}_2 \rangle = 0$ con lo que hemos construido dos vectores ortogonales. Ahora calculemos

$$\vec{v}_3 = \vec{u}_3 - \frac{\langle \vec{u}_3 | \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{u}_3 | \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2$$

Ahora notemos que \vec{v}_3 es ortogonal a \vec{v}_1 y a \vec{v}_2 . Este procedimiento se repite hasta construir una base ortogonal.

Resumiendo el algoritmo de Gram-Schmidt

$$\begin{aligned}\vec{v}_1 &= \vec{u}_1 \\ \vec{v}_2 &= \vec{u}_2 - \frac{\langle \vec{u}_2 | \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 \\ \vec{v}_3 &= \vec{u}_3 - \frac{\langle \vec{u}_3 | \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{u}_3 | \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2 \\ &\vdots \\ \vec{v}_m &= \vec{u}_m - \sum_{\ell=1}^{m-1} \frac{\langle \vec{u}_m | \vec{v}_\ell \rangle}{\|\vec{v}_\ell\|^2} \vec{v}_\ell\end{aligned}$$

Con este procedimiento, si hubiéramos partido de una base para el espacio, el proceso de ortogonalización de Gram-Schmidt nos conduce a la construcción de una base ortogonal.

13.5 Base Ortogonal. Coeficientes de Fourier

Consideremos un espacio vectorial real de dimensión n . Sea $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ una base ortogonal. Sea $\vec{v} \in V$, entonces

$$\vec{v} = v^\mu \mathbf{e}_\mu$$

Calculemos $\langle \vec{v} | \mathbf{e}_\nu \rangle$. Calculando el producto, tenemos

$$\langle \vec{v} | \mathbf{e}_\nu \rangle = \langle v^\mu \mathbf{e}_\mu | \mathbf{e}_\nu \rangle = v^\mu \langle \mathbf{e}_\mu | \mathbf{e}_\nu \rangle = v^\nu \|\mathbf{e}_\nu\|^2$$

Entonces, las coordenadas las obtenemos calculando

$$v^\nu = \frac{\langle \vec{v} | \mathbf{e}_\nu \rangle}{\|\mathbf{e}_\nu\|^2}$$

Entonces, la expresión del vector es

$$\vec{v} = \frac{\langle \vec{v} | \mathbf{e}_\mu \rangle}{\|\mathbf{e}_\mu\|^2} \mathbf{e}_\mu$$

Aquí no parece ser consistente la convención de Einstein. En realidad $\frac{\langle \vec{v} | \mathbf{e}_\mu \rangle}{\|\mathbf{e}_\mu\|^2}$ son las coordenadas contravariantes del vector \vec{v} , sólo que el cálculo no posee un supraíndice. Hay que tener cuidado entonces con esta notación.

Si además la base es ortogonal, tenemos

$$\vec{v} = \langle \vec{v} | \mathbf{e}_\mu \rangle \mathbf{e}_\mu$$

Las coordenadas de un vector obtenidas a partir de estas relaciones se denominan *coeficientes de Fourier*. Cuando se estudien las series de Fourier se retomarán estas ideas de gran utilidad.

Cuando la dimensión no sea finita todas estas ideas seguirán siendo válidas y serán retomadas para la construcción de las series de Fourier que no es otra cosa que la representación del espacio de funciones integrables en un determinado intervalo del eje real.

Ya lo veremos más en detalle, pero una función integrable en el intervalo $[-\pi, \pi]$ puede ser desarrollada como

$$f(x) = \sum_{\ell=0}^{\infty} \frac{\langle f | \cos(\ell t) \rangle}{\|\cos(\ell t)\|^2} \cos(\ell x) + \frac{\langle f | \sin(\ell t) \rangle}{\|\sin(\ell t)\|^2} \sin(\ell x)$$

donde sólo es necesario conocer el producto interno en este espacio.

14 El producto interno como un tensor. El tensor métrico

A partir de la definición de producto interno podemos notar que la primera propiedad nos dice que es un número real (en espacios reales). Además, con la tercera propiedad y la segunda se puede demostrar que es una aplicación que es lineal en cada argumento. Entonces, es una forma bilineal. Esto significa que el producto interno es un tensor dos veces covariante.

Si $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ es la base del espacio y $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$ es la base dual, tendremos que el producto interno visto como un tensor dos veces covariante puede expresarse como

$$\mathbf{g} = g_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$$

Como vimos, las coordenadas $g_{\mu\nu}$ del tensor en la base $\{\mathbf{dx}^\mu \otimes \mathbf{dx}^\nu\}$ se obtienen aplicando el tensor a los elementos de la base del espacio. Esto es

$$g_{\mu\nu} = \mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \langle \mathbf{e}_\mu | \mathbf{e}_\nu \rangle$$

Las n^2 cantidades $g_{\mu\nu}$ son denominadas coordenadas del tensor métrico o directamente *métrica*. De la definición de producto interno, tenemos que el tensor es simétrico (es decir que en cualquier sistema de coordenadas, $g_{\mu\nu} = g_{\nu\mu}$)

Calculemos nuevamente el producto interno entre dos vectores \vec{u} y \vec{v} . Tenemos

$$\langle \vec{u} | \vec{v} \rangle = \mathbf{g}(\vec{u}, \vec{v}) = g_{\mu\nu} \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu(\vec{u}, \vec{v})$$

aplicando la definición de producto tensorial, tenemos

$$\langle \vec{u} | \vec{v} \rangle = \mathbf{g}(\vec{u}, \vec{v}) = g_{\mu\nu} \mathbf{dx}^\mu(\vec{u}) \mathbf{dx}^\nu(\vec{v}) = g_{\mu\nu} u^\mu v^\nu$$

14.1 Aplicación: Longitud de arco

En diversas oportunidades nos encontramos ante la necesidad de hallar la longitud de una determinada curva parametrizada con una función vectorial

$$\vec{r}(t) = x^\mu(t) \mathbf{e}_\mu, \quad a \leq t \leq b$$

En situaciones particulares, tales como trabajar en la base canónica de \mathcal{R}^3 y para el producto interno canónico tenemos que la longitud de arco se calcula

$$\ell = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt$$

Esta formulación particular puede representarse de una manera más general

$$\ell = \int_a^b \sqrt{\left\langle \frac{d\vec{r}}{dt} \middle| \frac{d\vec{r}}{dt} \right\rangle} dt$$

que en una métrica que no sea la canónica, se escribirá

$$\ell = \int_a^b \sqrt{g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} dt$$

Esta expresión es más general que la que normalmente se ve en los textos elementales de cálculo y geometría.

15 1-Forma asociada. Coordenadas Covariantes de un vector contravariante.

Sea \vec{u} un vector de un espacio vectorial V . Como el producto interno es bilineal, entonces si fijamos un vector, por ejemplo el vector \vec{u} , podemos definir una 1-forma $g_{\vec{u}} : V \rightarrow \mathcal{R}$ de la siguiente manera:

$$g_{\vec{u}}(v) = \langle \vec{u} | \vec{v} \rangle$$

Como funcional, pertenece al espacio dual de V , el V^* y tendrá una representación

$$g_{\vec{u}} = a_{\mu} \mathbf{d}x^{\mu}$$

donde para obtener las coordenadas a_{μ} debemos calcular la 1-forma a los elementos de la base. Tendremos entonces

$$g_{\vec{u}}(\mathbf{e}_{\nu}) = a_{\mu} \mathbf{d}x^{\mu}(\mathbf{e}_{\nu})$$

Ahora bien, por un lado tenemos $\mathbf{d}x^{\mu}(\mathbf{e}_{\nu}) = \delta_{\nu}^{\mu}$. Y por otro lado tenemos que

$$g_{\vec{u}}(\mathbf{e}_{\nu}) = \langle \vec{u} | \mathbf{e}_{\nu} \rangle = u^{\mu} g_{\mu\nu}$$

Reemplazando en la forma original, tenemos

$$g_{\vec{u}} = u^{\mu} g_{\mu\nu} \mathbf{d}x^{\nu}$$

Esto significa que dado un vector \vec{u} del espacio V le podemos asignar a través del producto interno una 1-forma $g_{\vec{u}}$ cuyas coordenadas en la base dual son $u^{\nu} g_{\mu\nu}$. Como las 1-formas poseen coordenadas covariantes, lo que obtuvimos es a partir de un vector contravariante un vector covariante

$$u_{\mu} = g_{\mu\nu} u^{\nu}$$

Esto se conoce como *coordenadas covariantes de un vector contravariante*. Esta transformación de vector a 1-forma también se lo conoce como *bajada de índice* y para llevarla a cabo es necesario multiplicar por el tensor métrico.

Bajada de Índices. El procedimiento descubierto en el punto anterior permite generalizarlo a tensores más generales. Por ejemplo, dado un tensor del tipo $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ cuyas coordenadas serán en una base $T_{\alpha\beta\gamma}^{\mu\nu}$ puede transformarse en un tensor del tipo $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ con coordenadas obtenidas a partir de multiplicar por las coordenadas del tensor métrico y contraer en un índice

$$T_{\alpha\beta\gamma\sigma}^{\mu} = T_{\alpha\beta\gamma}^{\mu\nu} g_{\nu\sigma}$$

Subida de Índices. Para asociar un vector (contravariante) a una 1-forma, será necesario calcular la inversa de la matriz asociada a las coordenadas del tensor métrico.

Definimos $g^{\mu\nu}$ a la inversa de $g_{\alpha\beta}$, con lo cual la ecuación que relaciona las coordenadas será

$$g^{\mu\nu} g_{\nu\alpha} = \delta_{\alpha}^{\mu}$$

Con estas cantidades podríamos asociar a una 1-forma un vector contravariante

$$f^\mu = f_\nu g^{\mu\nu}$$

Tomando el tensor del tipo $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ podemos transformarlo en uno del tipo $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ cuyas coordenadas serán

$$T_{\mu\nu}^{\alpha\beta\gamma} = T_{\mu\nu\xi}^{\alpha\beta} g^{\xi\gamma}$$

El procedimiento de subida y bajada de índices se realiza multiplicando por $g_{\mu\nu}$ (si se quiere bajar índice) o por $g^{\mu\nu}$ (si se quiere subir índice) y se efectúa un producto por cada índice que se desea modificar.

Si queremos pasar del tensor original $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ a uno del tipo $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ debemos hacer

$$T_{\mu}^{\alpha\beta\gamma\delta} = T_{\mu\eta\xi}^{\alpha\beta} g^{\xi\gamma} g^{\eta\delta}$$

V

Aplicaciones: Algunos Tensores en Física

16 Energía Cinética. Tensor de Inercia

16.1 Energía Cinética de una partícula

En cursos iniciales de mecánica clásica definimos la *Energía Cinética* de una partícula de masa m como

$$T = \frac{1}{2} m v^2$$

donde v es la velocidad de la partícula, asumiendo que el movimiento es rectilíneo.

Si la partícula puede moverse en el espacio \mathcal{R}^3 , la expresión de la energía cinética deber expresarse como

$$T = \frac{1}{2} m \|\vec{v}\|^2$$

donde ahora debemos considerar la norma al cuadrado del vector velocidad, y no el cuadrado del escalar velocidad para el movimiento unidimensional.

Al tener definida la energía cinética a través de una norma, debemos tener en cuenta que la norma de un vector es el producto del vector por sí mismo, por lo que

$$T = \frac{1}{2} m \langle \vec{v} | \vec{v} \rangle$$

Lo que implica que la energía cinética es una forma bilineal que se aplica a un mismo vector, la velocidad de a partícula. Esto es, la energía cinética es una forma cuadrática.

Ahora bien, por ser una forma cuadrática, es una forma bilineal, y por lo tanto un tensor dos veces covariante, ya que toma dos vectores (aunque sea el mismo, lo toma dos veces) y le asigna un número real.

Si no asumimos que el producto interno es el canónico, podríamos tener un tensor métrico $g_{\mu\nu}$ cuya representación matricial no sea diagonal, por lo que podríamos escribir

$$T = \frac{1}{2} m g_{\mu\nu} v^\mu v^\nu$$

Más aún, al escribir el vector velocidad \vec{v} de manera invariante $v^\mu \mathbf{e}_\mu$ la expresión de la energía cinética es válida para cualquier sistema de coordenadas arbitrario, no necesariamente en la base canónica que produce velocidades "lineales".

Este abordaje es fundamental para trabajar con las denominadas *coordenadas generalizadas* que representan los sistemas dinámicos a partir de q^1, q^2, q^3 . En un sistema de coordenadas generalizadas, la energía cinética en general se escribe entonces

$$T = \frac{1}{2} m g_{\mu\nu} \dot{q}^\mu \dot{q}^\nu$$

El momento generalizado (que en coordenadas cartesianas es la cantidad de movimiento) se define a través de la relación

$$p_\mu = \frac{\partial T}{\partial \dot{q}^\mu} = m g_{\mu\nu} \dot{q}^\nu$$

Notemos que como la sumatoria es doble, \dot{q}^μ aparece dos veces, por ese motivo no está más el factor $\frac{1}{2}$

A partir de lo que hemos visto respecto a subida y bajada de índices podemos notar que el momento generalizado es un vector covariante, es decir una 1-forma, ya que se obtiene contrayendo el tensor métrico con el vector velocidad.

Es particularmente curioso si pensamos en coordenadas cartesianas: El momento lineal (cantidad de movimiento) se obtiene simplemente multiplicando la masa por la velocidad. Ahora, desde una perspectiva tensorial, esta simple operación transforma a un vector contravariante (la velocidad) a un vector covariante (el momento).

Este tipo de curiosidades son producidas al trabajar en espacios cartesianos, con las bases canónicas y con el producto interno canónico. En espacios definidos de manera tan particular producen estas características que no permiten ver la naturaleza tensorial de las cantidades.

16.2 Energía Cinética de un Cuerpo Rígido. Tensor de Inercia

Calculemos la energía cinética de rotación de un cuerpo rígido compuesto por N partículas. Consideremos un cuerpo rígido constituido por N partículas. Sea $\vec{\omega}$ la velocidad angular del cuerpo. Como estamos en presencia de un cuerpo rígido, todas las partículas que lo constituyen están afectados a esa velocidad angular. La energía cinética de rotación de cada partícula de masa m_ℓ se calcula como

$$T_\ell = \frac{1}{2} m_\ell \langle \vec{r}_\ell \times \vec{\omega} | \vec{r}_\ell \times \vec{\omega} \rangle$$

donde \times indica el producto vectorial.

A partir de la definición de producto vectorial, podemos escribir la siguiente identidad

$$\langle \vec{r}_\ell \times \vec{\omega} | \vec{r}_\ell \times \vec{\omega} \rangle = \langle (\vec{\omega} \times \vec{r}_\ell) \times \vec{\omega} | \vec{r}_\ell \rangle$$

que a su vez puede escribirse

$$\langle \vec{r}_\ell \times \vec{\omega} | \vec{r}_\ell \times \vec{\omega} \rangle = [|\vec{\omega}|^2 \vec{r}_\ell - \langle \vec{\omega} | \vec{r}_\ell \rangle \vec{\omega}] \cdot \vec{r}_\ell$$

Entonces, la energía cinética para la partícula ℓ -ésima se puede escribir

$$T_\ell = \frac{1}{2} m_\ell [|\vec{\omega}|^2 \vec{r}_\ell - \langle \vec{\omega} | \vec{r}_\ell \rangle \vec{\omega}] \cdot \vec{r}_\ell$$

$$T_\ell = \frac{1}{2} m_\ell [|\vec{\omega}|^2 |\vec{r}_\ell|^2 - \langle \vec{\omega} | \vec{r}_\ell \rangle^2]$$

Reemplazando, obtenemos

$$T_\ell = \frac{1}{2} m_\ell \left[g_{\mu\nu} \omega^\mu \omega^\nu g_{\alpha\beta} x_\ell^\alpha x_\ell^\beta - \omega^\mu \omega^\nu x_\ell^\alpha x_\ell^\beta g_{\mu\alpha} g_{\nu\beta} \right]$$

Entonces, para la partícula ℓ -ésima se puede escribir

$$\begin{aligned} T_\ell &= \frac{1}{2} m_\ell \left[g_{\mu\nu} g_{\alpha\beta} x_\ell^\alpha x_\ell^\beta - x_\ell^\alpha x_\ell^\beta g_{\mu\alpha} g_{\nu\beta} \right] \omega^\mu \omega^\nu \\ &= \frac{1}{2} m_\ell \left[(g_{\mu\nu} g_{\alpha\beta} - g_{\mu\alpha} g_{\nu\beta}) x_\ell^\alpha x_\ell^\beta \right] \omega^\mu \omega^\nu \end{aligned}$$

Sumando sobre todas las partículas que constituyen el cuerpo, tenemos que la energía cinética se puede escribir como

$$T = \frac{1}{2} I_{\mu\nu} \omega^\mu \omega^\nu$$

donde

$$I_{\mu\nu} = \sum_{\ell=1}^N m_{\ell} (g_{\mu\nu} g_{\alpha\beta} - g_{\mu\alpha} g_{\nu\beta}) x_{\ell}^{\alpha} x_{\ell}^{\beta}$$

Si el cuerpo rígido está compuesto por una distribución continua de materia, la sumatoria puede extenderse a una integral de la siguiente manera: $\sum_{\ell=1}^N m_{\ell} \rightarrow \iiint_{\mathcal{V}} \rho dx dy dz$ Con lo cual, las coordenadas del tensor de energía cinética se pueden escribir como

$$I_{\mu\nu} = \iiint_{\mathcal{V}} \rho (g_{\mu\nu} g_{\alpha\beta} - g_{\mu\alpha} g_{\nu\beta}) x^{\alpha} x^{\beta} dx dy dz$$

Para el caso del producto interno canónico, en donde las coordenadas del tensor métrico son 1 en la diagonal, podemos escribir las coordenadas del tensor energía cinética (al que comúnmente se lo denomina *tensor de inercia*)

$$I_{\mu\nu} = \iiint_{\mathcal{V}} \rho [\delta_{\mu\nu}(x^2 + y^2 + z^2) - x_{\mu} x_{\nu}] dx dy dz$$

donde, dado la simetría producida por lo diagonal de la métrica, las coordenadas covariantes del vector posición coinciden con las coordenadas del punto; esto es, $x_1 = x$, $x_2 = y$, $x_3 = z$

Repasando, energía cinética de rotación de un cuerpo rígido se puede escribir

$$T = \frac{1}{2} I_{\mu\nu} \omega^{\mu} \omega^{\nu}$$

donde ya sabemos calcular las cantidades $I_{\mu\nu}$. A partir de esta expresión, podemos obtener la expresión para el momento asociado a la velocidad angular, llamado *momento angular*, a través de

$$L_{\nu} = \frac{\partial T}{\partial \omega^{\nu}} = I_{\mu\nu} \omega^{\mu}$$

17 Elasticidad. Tensor de Deformación

Sea $\vec{u} : \mathcal{R}^3 \rightarrow \mathcal{R}^3$ el campo vectorial que asigna a cada punto de un medio continuo el vector desplazamiento.

Sean $P(\mathbf{x})$ y $Q(\mathbf{y})$ dos puntos del medio en cuestión separados una distancia h en la dirección $\vec{\ell}$

Entonces, tenemos

$$\mathbf{y} = \mathbf{x} + h \vec{\ell}$$

Además tenemos que $\|\vec{PQ}\| = h$. Supongamos que P se desplaza a $P'(\mathbf{x}') = P'(\mathbf{x} + \vec{u}_1)$ y que Q se desplaza a $Q'(\mathbf{y}') = Q'(\mathbf{y} + \vec{u}_2)$

Entonces tenemos que

$$\begin{aligned} P'Q' &= \mathbf{y} + \vec{u}_2 - \mathbf{x} + \vec{u}_1 &= \mathbf{x} + h \vec{\ell} + \vec{u}_2 - (\mathbf{x} + \vec{u}_1) \\ & &= h \vec{\ell} + \vec{u}_2 - \vec{u}_1 \end{aligned}$$

Notemos que $\vec{u}_2 = \vec{u}(\mathbf{y})$ y $\vec{u}_1 = \vec{u}(\mathbf{x})$.
O lo que es lo mismo $\vec{u}_2 = \vec{u}(\mathbf{x} + h \vec{\ell})$

Entonces

$$\vec{u}_2 - \vec{u}_1 = \vec{u}(\mathbf{x} + h\vec{\ell}) - \vec{u}(\mathbf{x})$$

Desarrollando a primer orden en Taylor tenemos

$$\begin{aligned}\vec{u}_2 - \vec{u}_1 &= \vec{u}(\mathbf{x} + h\vec{\ell}) - \vec{u}(\mathbf{x}) \\ &= \vec{u}(\mathbf{x}) + \frac{\partial u^\mu}{\partial x^\nu} h \ell^\nu \mathbf{e}_\mu - \vec{u}(\mathbf{x}) \\ &= \frac{\partial u^\mu}{\partial x^\nu} h \ell^\nu \mathbf{e}_\mu\end{aligned}$$

Calculemos ahora la norma de $P'\vec{Q}'$.

$$\|P'\vec{Q}'\|^2 = \langle P'\vec{Q}' | P'\vec{Q}' \rangle$$

Reemplazando, tenemos

$$\begin{aligned}\|P'\vec{Q}'\|^2 &= \langle P'\vec{Q}' | P'\vec{Q}' \rangle \\ &= \langle h\vec{\ell} + u_2 - u_1 | h\vec{\ell} + u_2 - u_1 \rangle \\ &= h^2 \langle \vec{\ell} | \vec{\ell} \rangle + 2h \langle \vec{\ell} | (u_2 - u_1) \rangle + \\ &+ \langle (u_2 - u_1) | (u_2 - u_1) \rangle\end{aligned}$$

A primer orden en

$$\|(\vec{u}_2 - \vec{u}_1)\|$$

tenemos

$$\begin{aligned}\|P'\vec{Q}'\|^2 &= \langle P'\vec{Q}' | P'\vec{Q}' \rangle \\ &= \langle h\vec{\ell} + \vec{u}_2 - \vec{u}_1 | h\vec{\ell} + \vec{u}_2 - \vec{u}_1 \rangle \\ &= h^2 \langle \vec{\ell} | \vec{\ell} \rangle + 2h \langle \vec{\ell} | (\vec{u}_2 - \vec{u}_1) \rangle\end{aligned}$$

Reemplazando el desarrollo de Taylor para $\vec{u}_2 - \vec{u}_1$ obtenemos

$$\begin{aligned}\|P'\vec{Q}'\|^2 &= h^2 \underbrace{\langle \vec{\ell} | \vec{\ell} \rangle}_{\text{vale 1}} + 2h \langle \vec{\ell} | \frac{\partial u^\mu}{\partial x^\nu} h \ell^\nu \mathbf{e}_\mu \rangle \\ &= h^2 + 2h \langle \ell^\alpha \mathbf{e}_\alpha | \frac{\partial u^\mu}{\partial x^\nu} h \ell^\nu \mathbf{e}_\mu \rangle \\ &= h^2 \left(1 + 2\ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} \langle \mathbf{e}_\alpha | \mathbf{e}_\mu \rangle \right) \\ &= h^2 \left(1 + 2\ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} g_{\alpha\mu} \right)\end{aligned}$$

Aplicando raíz cuadrada a ambos miembros, obtenemos

$$\|P'\vec{Q}'\| = h \sqrt{1 + 2\ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} g_{\alpha\mu}}$$

Aplicando el desarrollo de Taylor para la raíz cuadrada

$$\sqrt{1+x} \approx 1 + \frac{x}{2}$$

obtenemos

$$\|\vec{P}'Q'\| = h \left(1 + \ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} g_{\alpha\mu} \right)$$

Definimos la **deformación** del sólido como

$$D = \frac{\|\vec{P}'Q'\| - \|\vec{P}Q\|}{\|\vec{P}Q\|}$$

Entonces, reemplazando, obtenemos

$$\begin{aligned} \frac{\|\vec{P}'Q'\| - \|\vec{P}Q\|}{\|\vec{P}Q\|} &= \frac{h \left(1 + \ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} g_{\alpha\mu} \right) - h}{h} \\ &= \ell^\alpha \ell^\nu \frac{\partial u^\mu}{\partial x^\nu} g_{\alpha\mu} \end{aligned}$$

Notemos que el último factor, es la coordenada covariante del vector u con lo cual, podemos reescribir la ecuación

$$\frac{\|\vec{P}'Q'\| - \|\vec{P}Q\|}{\|\vec{P}Q\|} = \ell^\alpha \ell^\nu \frac{\partial u_\alpha}{\partial x^\nu}$$

La doble sumatoria en α y en ν tiene un grado de simetría de manera tal que podemos definir

$$e_{\alpha\nu} = \frac{1}{2} \left(\frac{\partial u_\alpha}{\partial x^\nu} + \frac{\partial u_\nu}{\partial x^\alpha} \right)$$

Y define las coordenadas del *tensor de deformación*

La deformación se calcula

$$D = e_{\alpha\nu} \ell^\alpha \ell^\nu$$

18 Desarrollo del Potencial Electroestático. Tensor Momento Cuadrupolar

Si se conoce la distribución de carga eléctrica en el espacio \mathcal{R}^3 el potencial eléctrico de determinará con la integral

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \iiint_{\mathcal{R}^3} \frac{\rho(\vec{r}')}{\|\vec{r} - \vec{r}'\|} dx' dy' dz'$$

Tenemos, además,

$$\|\vec{r} - \vec{r}'\| = \sqrt{\|\vec{r} - \vec{r}'\|^2}$$

donde

$$\|\vec{r} - \vec{r}'\|^2 = \|\vec{r}\|^2 - 2 \langle \vec{r} | \vec{r}' \rangle + \|\vec{r}'\|^2$$

Para el cálculo del potencial denominado exterior (donde $\|\vec{r}\| \gg \|\vec{r}'\|$) podemos hacer un desarrollo bajo esta asunción.

Antes de hacer el desarrollo, consideremos

$$\|\vec{r} - \vec{r}'\|^2 = \|\vec{r}\|^2 - 2 \langle \vec{r} | \vec{r}' \rangle + \|\vec{r}'\|^2 = \|\vec{r}\|^2 \left[1 - 2 \frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\|^2} + \frac{\|\vec{r}'\|^2}{\|\vec{r}\|^2} \right]$$

Entonces,

$$\frac{1}{\|\vec{r} - \vec{r}'\|} = \frac{1}{\|\vec{r}\|} \frac{1}{\sqrt{1 - 2 \frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\|^2} + \frac{\|\vec{r}'\|^2}{\|\vec{r}\|^2}}}$$

A partir de la desigualdad de Cauchy-Schwarz, tenemos que $|\langle \vec{r} | \vec{r}' \rangle| \leq \|\vec{r}\| \|\vec{r}'\|$ entonces, podemos definir una cantidad menor que 1 (que puede ser el coseno de un ángulo),

$$\frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\| \|\vec{r}'\|} = \cos(\zeta)$$

y si llamamos $t = \frac{\|\vec{r}'\|}{\|\vec{r}\|}$ podemos desarrollar la función

$$\frac{1}{\|\vec{r} - \vec{r}'\|} = \frac{1}{\|\vec{r}\|} \frac{1}{\sqrt{1 - 2t \cos(\zeta) + t^2}} \approx \frac{1}{\|\vec{r}\|} \left\{ 1 + \cos(\zeta) t + \frac{1}{2} [3 \cos^2(\zeta) - 1] t^2 \right\}$$

reemplazando, obtenemos,

$$\frac{1}{\|\vec{r} - \vec{r}'\|} \approx \frac{1}{\|\vec{r}\|} \left\{ 1 + \frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\| \|\vec{r}'\|} \frac{\|\vec{r}'\|}{\|\vec{r}\|} + \frac{1}{2} \left[3 \left[\frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\| \|\vec{r}'\|} \right]^2 - 1 \right] \frac{\|\vec{r}'\|^2}{\|\vec{r}\|^2} \right\}$$

Simplificando,

$$\begin{aligned} \frac{1}{\|\vec{r} - \vec{r}'\|} &\approx \frac{1}{\|\vec{r}\|} \left\{ 1 + \frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\|^2} + \frac{1}{2} \left[3 \frac{\langle \vec{r} | \vec{r}' \rangle^2}{\|\vec{r}\|^4} - \frac{\|\vec{r}'\|^2}{\|\vec{r}\|^2} \right] \right\} \\ \frac{1}{\|\vec{r} - \vec{r}'\|} &\approx \frac{1}{\|\vec{r}\|} + \frac{\langle \vec{r} | \vec{r}' \rangle}{\|\vec{r}\|^3} + \frac{1}{2\|\vec{r}\|^5} [3\langle \vec{r} | \vec{r}' \rangle^2 - \|\vec{r}\|^2 \|\vec{r}'\|^2] \end{aligned}$$

Escribamos en función de las coordenadas del tensor métrico $g_{\mu\nu}$ los términos

$$\langle \vec{r} | \vec{r}' \rangle^2 = g_{\mu\alpha} g_{\nu\beta} x^\mu x^\nu x'^\alpha x'^\beta$$

$$\|\vec{r}\|^2 \|\vec{r}'\|^2 = g_{\mu\nu} g_{\alpha\beta} x^\mu x^\nu x'^\alpha x'^\beta$$

Con estas expresiones, podemos escribir la aproximación $\frac{1}{\|\vec{r} - \vec{r}'\|}$ como

$$\frac{1}{\|\vec{r} - \vec{r}'\|} \approx \frac{1}{\|\vec{r}\|} + \frac{1}{\|\vec{r}\|^3} g_{\mu\nu} x^\mu x'^\nu + \frac{1}{2\|\vec{r}\|^5} [3g_{\mu\alpha} g_{\nu\beta} - g_{\mu\nu} g_{\alpha\beta}] x^\mu x^\nu x'^\alpha x'^\beta$$

Finalmente, si multiplicamos por $\frac{1}{4\pi\epsilon_0} \rho(\vec{r}')$ e integramos en todo \mathcal{R}^3 Obtenemos la aproximación del potencial electrostático

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{Q}{\|\vec{r}\|} + \frac{1}{\|\vec{r}\|^3} P_\mu x^\mu + \frac{1}{2\|\vec{r}\|^5} Q_{\mu\nu} x^\mu x^\nu + \dots \right]$$

donde

$$Q = \iiint_{\mathcal{R}^3} \rho(\vec{r}') dx' dy' dz' \quad \text{que representa la carga total}$$

$$P_\mu = \iiint_{\mathcal{R}^3} \rho(\vec{r}') g_{\mu\nu} x'^\nu dx' dy' dz' \quad \text{que representa el momento dipolar}$$

y

$$Q_{\mu\nu} = \iiint_{\mathcal{R}^3} \rho(\vec{r}') [3g_{\mu\alpha}g_{\nu\beta} - g_{\mu\nu}g_{\alpha\beta}] x'^\alpha x'^\beta dx' dy' dz'$$

las cantidades $Q_{\mu\nu}$ son las coordenadas del denominado tensor momento cuadrupolar o simplemente cuadrupolo.

VI

Análisis Tensorial. Operadores Diferenciales.

19 Introducción

Para completar la exposición del capítulo de tensores, es necesario introducir las *coordenadas curvilíneas* ya que en muchos problemas aparecen cambios no lineales de coordenadas tales como polares, esféricas, etc.

En esta extensión y en un abordaje tensorial, vamos a estudiar de manera invariante los operadores diferenciales que ya fueron vistos en cursos de cálculo vectorial (Análisis Matemático II) los cuales fueron definidos exclusivamente para coordenadas cartesianas y que a partir de esta restricción se pierde la naturaleza geométrica de los objetos. El gradiente, por ejemplo, es presentado en los cursos elementales como un simple vector, pero ya hemos visto que su naturaleza geométrica no es la de un vector (contravariante) sino que sus coordenadas son las de una 1-forma. De la misma manera el rotor, la divergencia y laplaciano fueron definidos para coordenadas cartesianas y en la métrica euclídea (Pitágoras) lo que no permite enmarcar de manera covariante las expresiones.

En este capítulo introduciremos las definiciones del cálculo vectorial de manera covariante, es decir, independiente del sistema de coordenadas, lo que nos permitirá trabajar a partir de definiciones que no están forzadas a ser válidas únicamente en \mathcal{R}^3 .

20 Cálculo operacional

Vamos a introducir operaciones entre tensores, en principio particulares, para luego extender a tensores en general. Comenzaremos con los diferentes productos para luego estudiar

operaciones diferenciales.

Multiplicación de tensores contravariantes

Sean \mathbf{A} y \mathbf{B} dos tensores dos veces contravariantes,

$$\mathbf{A} = A^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu, \quad \mathbf{B} = B^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

- I. **Producto punto.** Definimos el producto punto (extensión del producto interno entre vectores contravariantes)

$$\mathbf{A} \cdot \mathbf{B} = [A^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu] \cdot [B^{\alpha\beta} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta] = A^{\mu\nu} B^{\alpha\beta} \mathbf{e}_\mu \otimes \underbrace{\mathbf{e}_\nu \cdot \mathbf{e}_\alpha}_{\langle \mathbf{e}_\nu | \mathbf{e}_\alpha \rangle} \otimes \mathbf{e}_\beta = A^{\mu\nu} B^{\alpha\beta} g_{\nu\alpha} \mathbf{e}_\mu \otimes \mathbf{e}_\beta$$

Además, $B^{\alpha\beta} g_{\nu\alpha} = B_\nu^\beta$

$$\mathbf{A} \cdot \mathbf{B} = A^{\mu\nu} B_\nu^\beta \mathbf{e}_\mu \otimes \mathbf{e}_\beta$$

Notemos que si el espacio admite una base ortonormal, $B_\nu^\beta = B^{\beta\nu}$. Pero si no escribimos los índices adecuadamente, no podemos aplicar la convención de Einstein.

- II. **Producto tensorial.** Para este mismo caso de tensores dos veces contravariantes, el producto tensorial se define

$$\mathbf{A} \otimes \mathbf{B} = [A^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu] \otimes [B^{\alpha\beta} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta] = A^{\mu\nu} B^{\alpha\beta} \mathbf{e}_\mu \otimes \mathbf{e}_\nu \otimes \mathbf{e}_\alpha \otimes \mathbf{e}_\beta$$

Otra de las operaciones necesarias para el planteo de las ecuaciones dinámicas son las operaciones vinculadas a la derivación, que dan lugar a los *operadores diferenciales*, principalmente el carácter tensorial del *gradiente*, del *rotor*, de la *divergencia* y del *laplaciano*.

21 Derivación de tensores. Operadores diferenciales

El concepto de derivación proviene de cambio, de tasa de variación. Esto significa que es necesario definir *campos tensoriales*, esto es tensores dependientes de la posición.

La incorporación de la variación espacial, introduce otro problema: el de que las bases tampoco sean constantes.

Coordenadas curvilíneas. Estamos acostumbrados a que la base de un espacio vectorial es una entidad fija, la cual forma de ambiente para el espacio en consideración. Sin embargo al considerar sistemas de coordenadas obtenidos a partir de cambios no lineales (como el caso de las coordenadas polares para \mathcal{R}^2) las bases inducidas por el cambio de coordenadas cambian en cada punto.

Consideremos un espacio V de dimensión n , con base original $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Si efectuamos un cambio de base, de la base original a una nueva base $\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2', \dots, \mathbf{e}_n'\}$ podemos resumir todos los cambios resultantes

Espacio	Base	Coordenadas
Espacio V	$\mathbf{e}'_{\mu} = \Lambda^{\nu}_{\mu} \mathbf{e}_{\nu}$	$x'^{\mu} = [\Lambda^{-1}]^{\mu}_{\nu} x^{\nu}$
Espacio V^*	$\mathbf{dx}'^{\mu} = [\Lambda^{-1}]^{\mu}_{\nu} \mathbf{dx}^{\nu}$	$f'_{\mu} = \Lambda^{\nu}_{\mu} f_{\nu}$

Si ahora pensamos más en cambio de coordenadas que en cambio de base, tenemos que llamando $\Phi^{\mu}_{\nu} = [\Lambda^{-1}]^{\mu}_{\nu}$

Consideremos un cambio de coordenadas

$$x'^{\mu} = \Phi^{\mu}_{\nu} x^{\nu}$$

Este cambio induce un cambio de base a través de

$$\mathbf{e}'_{\mu} = [\Phi^{-1}]^{\nu}_{\mu} \mathbf{e}_{\nu}$$

Con esta reescritura, podemos escribir la correspondiente tabla, pero usando la matriz Φ

Espacio	Coordenadas	Bases
Espacio V	$x'^{\mu} = \Phi^{\mu}_{\nu} x^{\nu}$	$\mathbf{e}'_{\mu} = [\Phi^{-1}]^{\nu}_{\mu} \mathbf{e}_{\nu}$
Espacio V^*	$f'_{\mu} = [\Phi^{-1}]^{\nu}_{\mu} f_{\nu}$	$\mathbf{dx}'^{\mu} = \Phi^{\mu}_{\nu} \mathbf{dx}^{\nu}$

En el contexto de cambios lineales, estas matrices son constantes, las bases no cambian con la posición. Esto es, la matriz Φ , es constante.

Ahora consideremos *cambios de coordenadas no lineales*. En muchos ejemplos de mecánica o geometría, las simetrías de los problemas inducen a cambios de coordenadas que simplifican las ecuaciones y reducen la cantidad de variables. Por ejemplo, para el caso de \mathcal{R}^2 tenemos *las coordenadas polares*

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \theta &= \arctan \left[\frac{y}{x} \right] \end{aligned}$$

En este caso no es posible considerar una matriz Φ que caracterice el cambio de coordenadas. Para este tipo de cambio de coordenadas, aplicamos una linealización, de manera tal de considerar cambios lineales *locales*.

Calculemos los diferenciales

$$\begin{aligned} dr &= \frac{x}{\sqrt{x^2 + y^2}} dx + \frac{y}{\sqrt{x^2 + y^2}} dy = \cos(\theta) dx + \sin(\theta) dy \\ d\theta &= -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = -\frac{\sin(\theta)}{r} dx + \frac{\cos(\theta)}{r} dy \end{aligned}$$

El cambio de coordenadas local y lineal será representado ahora por la matriz Φ cuyas componentes son

$$\Phi_{\nu}^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}}, \quad (\text{las nuevas respecto de las viejas})$$

y por lo tanto, el cambio de base inducido será a partir de la matriz inversa

$$[\Phi^{-1}]_{\mu}^{\nu} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \quad (\text{las viejas respecto de las nuevas})$$

Para este cambio particular

$$\Phi = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\frac{\sin(\theta)}{r} & \frac{\cos(\theta)}{r} \end{bmatrix}, \quad \Phi^{-1} = \begin{bmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{bmatrix}$$

Con las matrices, podemos obtener el cambio de base inducido

$$\begin{aligned} \mathbf{e}'_1 &= \mathbf{e}_r = \cos(\theta) \mathbf{e}_1 + \sin(\theta) \mathbf{e}_2 \\ \mathbf{e}'_2 &= \mathbf{e}_{\theta} = -r \sin(\theta) \mathbf{e}_1 + r \cos(\theta) \mathbf{e}_2 \end{aligned}$$

(Es necesario recordar como se recorren los índices para el cambio de base, ya que no es un producto de matrices)

Podemos notar que \mathbf{e}_r es un vector unitario en dirección radial y que \mathbf{e}_{θ} es un vector, cuya norma depende del punto y cuya dirección es perpendicular a \mathbf{e}_r , y orientado positivamente. Notemos que además, como ya se ha adelantado, los vectores base son variables, es decir, cambian punto a punto

$$\begin{aligned} \frac{\partial \mathbf{e}_r}{\partial r} &= 0 \\ \frac{\partial \mathbf{e}_r}{\partial \theta} &= \frac{1}{r} \mathbf{e}_{\theta} \\ \frac{\partial \mathbf{e}_{\theta}}{\partial r} &= -\frac{1}{r} \mathbf{e}_{\theta} \\ \frac{\partial \mathbf{e}_{\theta}}{\partial \theta} &= -r \mathbf{e}_r \end{aligned}$$

Al aparecer esta posibilidad de que las bases pueden ser funciones de la posición, generará un nuevo concepto de derivada, ya que para bases "constantes" la derivación de vectores sólo se efectúa sobre las coordenadas dejando inalterada las bases.

Antes de pasar al estudio de derivación, consideremos en el ejemplo de coordenadas polares el tensor métrico.

Como el tensor métrico es un tensor dos veces covariante, en la nueva representación el cambio de coordenadas de \mathbf{g} viene dado por

$$g'_{\mu\nu} = [\Phi^{-1}]_{\nu}^{\alpha} [\Phi^{-1}]_{\mu}^{\beta} g_{\alpha\beta} \quad (\text{recordemos que los supraíndices indican filas})$$

Entonces, en la nueva base (y por consiguiente en las nuevas coordenadas) el tensor métrico tiene por coordenadas, teniendo en cuenta que en las coordenadas cartesianas es la identidad

$$\begin{aligned} g'_{11} = g'_{rr} &= [\Phi^{-1}]_1^1 [\Phi^{-1}]_1^1 g_{11} + [\Phi^{-1}]_1^2 [\Phi^{-1}]_1^2 g_{22} = \cos^2(\theta) + \sin^2(\theta) = 1 \\ g'_{22} = g'_{\theta\theta} &= [\Phi^{-1}]_2^1 [\Phi^{-1}]_2^1 g_{11} + [\Phi^{-1}]_2^2 [\Phi^{-1}]_2^2 g_{22} = r^2 \sin^2(\theta) + r^2 \cos^2(\theta) = r^2 \end{aligned}$$

Entonces, el tensor métrico se puede representar

$$\mathbf{g} = \mathbf{dx}^1 \otimes \mathbf{dx}^1 + \mathbf{dx}^2 \otimes \mathbf{dx}^2 = \mathbf{dx}'^1 \otimes \mathbf{dx}'^1 + r^2 \mathbf{dx}'^2 \otimes \mathbf{dx}'^2$$

o de manera matricial, las coordenadas se pueden representar

$$g'_{\mu\nu} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}$$

22 Derivación covariante

La derivación de vectores, en el marco general de coordenadas curvilíneas, debe ser reformulada habida cuenta que los vectores base ya no pueden considerarse como fijos o constantes. El ejemplo de coordenadas polares ya puso de manifiesto esta propiedad variable de la base.

Existen varias formulaciones para la introducción de la derivación covariante, desde abordajes formales en el marco de la geometría diferencial, así como también desde un punto más operacional, sin necesidad de entrar en los detalles técnicos relacionados a variedades diferenciales, etc.

Nuestro abordaje del tema será más bien operacional, trabajando en coordenadas, al estilo del presentado por Santaló en sus *Vectores y Tensores con sus aplicaciones*.

Derivación de un vector base. Símbolos de Christoffel

A partir de la posibilidad de que los vectores de la base pueden cambiar por la posición, tendrá sentido la expresión

$$\frac{\partial \mathbf{e}_\mu}{\partial x_\nu}$$

Más aún, este elemento deberá ser un vector del espacio, por lo que es una combinación lineal de los elementos de la propia base.

Esta representación la efectuaremos de manera operacional, es decir, vamos a definir cantidades que sean las coordenadas de los vectores base derivados. Para esto, debemos tener un conjunto de índices que nos puedan indicar por un lado, el vector base que se está derivando, por otro lado la coordenada respecto de la cual se está derivando y por último la coordenada en la base.

En función de lo establecido, vamos a definir unos elementos $\Gamma_{\mu\nu}^\alpha$ tales que la derivada de un vector base se puede escribir

$$\frac{\partial \mathbf{e}_\mu}{\partial x_\nu} = \Gamma_{\mu\nu}^\alpha \mathbf{e}_\alpha$$

las cantidades $\Gamma_{\mu\nu}^\alpha$ son los *símbolos de Christoffel*. Como hemos definido los *Christoffel* estos no son las coordenadas de un tensor.

A partir de consideraciones geométricas en las que no entraremos en detalles, los símbolos de Christoffel se obtienen a partir de la métrica

$$\Gamma_{\mu\nu}^\alpha = \frac{g^{\alpha\beta}}{2} \left[\frac{\partial g_{\mu\beta}}{\partial x^\nu} + \frac{\partial g_{\nu\beta}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\beta} \right]$$

Es frecuente para simplificar notación, identificar las derivadas parciales con un subíndice de la forma

$$\frac{\partial A}{\partial x^\mu} \equiv A_{,\mu}$$

con esta notación, los Christoffel los podemos escribir como

$$\Gamma_{\mu\nu}^\alpha = \frac{g^{\alpha\beta}}{2} [g_{\mu\beta,\nu} + g_{\nu\beta,\mu} - g_{\mu\nu,\beta}] \quad \text{notemos que} \quad \Gamma_{\mu\nu}^\alpha = \Gamma_{\nu\mu}^\alpha$$

Las n^2 (en realidad la simetría reduce a $\frac{n(n+1)}{2}$) coordenadas del tensor métrico, junto con los n^3 (también, por la simetría de los subíndices se reduce a $\frac{n^2(n+1)}{2}$) Christoffel son necesarios calcularlos inicialmente, ya que a partir de estas cantidades se puede realizar todo estudio geométrico .

A modo de ejemplo, en el caso de coordenadas polares tenemos

$$\Gamma_{rr}^r = 0, \quad \Gamma_{rr}^\theta = 0; \quad \Gamma_{r\theta}^r = 0, \quad \Gamma_{r\theta}^\theta = \frac{1}{r}; \quad \Gamma_{\theta r}^r = 0; \quad \Gamma_{\theta r}^\theta = \frac{1}{r}.$$

22.1 Derivada de un vector contravariante

Consideremos un vector contravariante $\mathbf{v} = v^\mu \mathbf{e}_\mu$. Considerando ahora que los elementos de la base pueden tener derivadas parciales no nulas podemos calcular

$$\frac{\partial \mathbf{v}}{\partial x^\nu} = \frac{\partial v^\mu}{\partial x^\nu} \mathbf{e}_\mu + v^\mu \frac{\partial \mathbf{e}_\mu}{\partial x^\nu}$$

reemplazando la expresión para la derivada de los vectores base

$$\frac{\partial \mathbf{v}}{\partial x^\nu} = \frac{\partial v^\mu}{\partial x^\nu} \mathbf{e}_\mu + v^\mu \Gamma_{\mu\nu}^\alpha \mathbf{e}_\alpha$$

cambiando los índices μ por α (que están sumandose, lo que los transforma en índices mudos) podemos agrupar

$$\frac{\partial \mathbf{v}}{\partial x^\nu} = \left[\frac{\partial v^\mu}{\partial x^\nu} + \Gamma_{\nu\alpha}^\mu v^\alpha \right] \mathbf{e}_\mu$$

Utilizando la notación para la derivada parcial, con una coma en el subíndice para indicar la coordenada espacial respecto de la cual se deriva la derivada del vector la escribimos

$$\frac{\partial \mathbf{v}}{\partial x^\nu} = [v^\mu_{,\nu} + \Gamma_{\nu\alpha}^\mu v^\alpha] \mathbf{e}_\mu$$

Esta expresión compatibiliza con el cálculo vectorial elemental, en el cual las derivaciones de vectores sólo afectaban a las coordenadas (o componentes, para ser más precisos).

Notemos que si la base es fija, los Christoffel son nulos y por lo tanto recuperamos los resultados del cálculo vectorial elemental, relacionado a la derivación de vectores.

Las derivadas de las coordenadas del vector tienen ahora un carácter *absoluto*, de la misma manera que en el cálculo vectorial elemental, sólo que ahora para preservar las direcciones no

hay que considerar la derivada parcial, sino este nuevo tipo de derivada, denominada *derivada covariante*.

La derivada covariante, entonces, es una derivada *absoluta*, que mide la variación de un vector, en cada una de sus direcciones. Vamos a denotar la derivación covariante con un punto y coma, para diferenciarla de la derivación parcial

$$v^\mu_{;\nu} = v^\mu_{,\nu} + \Gamma^\mu_{\nu\alpha} v^\alpha$$

Y con esta notación, la derivada parcial la denotamos

$$\frac{\partial \mathbf{v}}{\partial x^\nu} = v^\mu_{;\nu} \mathbf{e}_\mu$$

Con la introducción de la derivación covariante, la derivada de un vector es equivalente a la obtenida en el cálculo vectorial elemental.

Un cálculo diferencial basado en estas ideas de invarianza fue desarrollado por Tulio Levi-Civita y lo denominó *Cálculo Diferencial Absoluto*.

22.2 Derivada de un vector covariante

Un vector covariante, como sabemos es una 1-forma,

$$f = f_\mu \mathbf{dx}^\mu$$

Plantear la derivación parcial de la misma manera que lo hemos hecho para vectores contravariantes es más complicado, ya que deberíamos tener calculada

$$\frac{\partial \mathbf{dx}^\mu}{\partial x^\nu}$$

Esta complejidad la podemos obviar teniendo en cuenta que

$$v^\mu f_\mu$$

es un escalar, por lo que derivar covariantemente esta expresión coincide con la derivada parcial.

$$[v^\mu f_\mu]_{;\nu} = [v^\mu f_\mu]_{,\nu}$$

Además, bajo la imposición de la derivada covariante respete la regla del producto, tenemos

$$v^\mu_{;\nu} f_\mu + v^\mu f_{\mu;\nu} = v^\mu_{,\nu} f_\mu + v^\mu f_{\mu,\nu}$$

Reemplazando la derivada covariante de las coordenadas contravariantes de \mathbf{v} tenemos

$$v^\mu_{,\nu} f_\mu + \Gamma^\mu_{\alpha\nu} v^\alpha f_\mu + v^\mu f_{\mu;\nu} = v^\mu_{,\nu} f_\mu + v^\mu f_{\mu,\nu}$$

simplificando

$$\Gamma^\mu_{\alpha\nu} v^\alpha f_\mu + v^\mu f_{\mu;\nu} = v^\mu f_{\mu,\nu}$$

En el primer término del miembro de la izquierda podemos cambiar el índice α con μ y obtenemos

$$v^\mu [\Gamma^\alpha_{\mu\nu} f_\alpha + f_{\mu;\nu} - f_{\mu,\nu}] = 0$$

De donde obtenemos la derivada covariante de una coordenada covariante

$$f_{\mu;\nu} = f_{\mu,\nu} - \Gamma^\alpha_{\mu\nu} f_\alpha$$

Este método nos permite extender la derivación covariante de tensores de diferentes tipos de contravarianza y covarianza.

22.3 Derivada de un tensor dos veces contravariante

Consideremos un tensor dos veces contravariante

$$\mathbf{T} = t^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

Calculemos la derivada parcial respecto a x^λ

$$\begin{aligned} \frac{\partial \mathbf{T}}{\partial x^\lambda} &= t^{\mu\nu}{}_{,\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\nu + t^{\mu\nu} \frac{\partial}{\partial x^\lambda} [\mathbf{e}_\mu \otimes \mathbf{e}_\nu] \\ \frac{\partial \mathbf{T}}{\partial x^\lambda} &= t^{\mu\nu}{}_{,\lambda} \mathbf{e}_\mu \otimes \mathbf{e}_\nu + t^{\mu\nu} \frac{\partial \mathbf{e}_\mu}{\partial x^\lambda} \otimes \mathbf{e}_\nu + t^{\mu\nu} \mathbf{e}_\mu \otimes \frac{\partial \mathbf{e}_\nu}{\partial x^\lambda} \end{aligned}$$

Reemplazando las derivadas de los vectores base y cambiando convenientemente los subíndices, obtenemos

$$\frac{\partial \mathbf{T}}{\partial x^\lambda} = \left[t^{\mu\nu}{}_{,\lambda} + t^{\alpha\nu} \Gamma_{\alpha\lambda}^\mu + t^{\mu\alpha} \Gamma_{\alpha\lambda}^\nu \right] \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

Entonces

$$t^{\mu\nu}{}_{;\lambda} = t^{\mu\nu}{}_{,\lambda} + t^{\alpha\nu} \Gamma_{\alpha\lambda}^\mu + t^{\mu\alpha} \Gamma_{\alpha\lambda}^\nu$$

22.4 Derivada de un tensor dos veces covariante

Análogamente a lo que hicimos para obtener la derivada covariante de las coordenadas de 1-formas, vamos a construir un escalar por contracción

$$t_{\mu\nu} w^{\mu\nu}$$

es un escalar, por lo que

$$[t_{\mu\nu} w^{\mu\nu}]_{;\lambda} = [t_{\mu\nu} w^{\mu\nu}]_{,\lambda}$$

Reemplazando las derivadas, agrupando convenientemente y simplificando las expresiones se obtiene

$$t_{\mu\nu;\lambda} = t_{\mu\nu,\lambda} - \Gamma_{\mu\lambda}^\alpha t_{\alpha\nu} - \Gamma_{\nu\lambda}^\alpha t_{\mu\alpha}$$

En general, un tensor del tipo $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ se obtiene combinando las propiedades anteriores

$$t^\mu{}_{\nu;\lambda} = t^\mu{}_{\nu,\lambda} + \Gamma_{\alpha\lambda}^\mu t^\alpha{}_\nu - \Gamma_{\nu\lambda}^\alpha t^\mu{}_\alpha$$

Es necesario cuidado para compensar los índices, pero en general por cada índice de contravarianza se suman términos con símbolos de Christoffel y por cada índice de covarianza se restan.

Relación entre los símbolos de Christoffel y el determinante de la métrica. Sea g el determinante de las coordenadas del tensor métrico. A partir de la relación de las derivadas de un determinante y la relación que resulta los símbolos de Christoffel, se puede obtener (después de un engorroso cálculo)

$$\Gamma_{\beta\gamma}^\alpha = \frac{1}{2} \frac{\partial [\log(g)]}{\partial x^\gamma} = \frac{1}{\sqrt{g}} \frac{\partial [\sqrt{g}]}{\partial x^\gamma}$$

23 Aplicación al cálculo vectorial. Gradiente, Rotor, Divergencia y Laplaciano

En este abordaje tensorial calculemos los operadores gradiente, rotor, divergencia y laplaciano de campos vectoriales, invariante ante cambio de coordenadas.

En los cursos de cálculo vectorial introductorios, las definiciones de los operadores están definidos en circunstancias muy particulares: El espacio es \mathcal{R}^3 y el sistema de coordenadas es el cartesiano (con base canónica) y la métrica viene a partir del producto interno canónico.

El hecho de tener definida una derivación independiente del sistema de coordenadas (y por lo tanto de las bases) posibilita la generalización de estos operadores, tanto en el espacio (que ya puede extenderse a espacios más generales denominados *variedades*), el sistema de coordenadas y la métrica.

Gradiente. El gradiente de una función es una 1-forma diferencial, definida a partir de la derivada direccional, cuyas coordenadas son las derivadas parciales. Además, como las derivadas covariantes de un escalar coinciden con las derivadas parciales, tenemos que el gradiente se escribe

$$\nabla\phi = \phi_{,\mu} \mathbf{dx}^\mu$$

Es muy común considerar al gradiente como un vector, es decir contravariante. Si fuera de esta manera, tendríamos que si consideráramos como un vector

$$\nabla\phi = \phi_{,\mu} \mathbf{e}_\mu$$

daría lo mismo en cualquier sistema de coordenadas,

$$\nabla\phi = \frac{\partial\phi}{\partial x} \mathbf{e}_x + \frac{\partial\phi}{\partial y} \mathbf{e}_y = \frac{\partial\phi}{\partial r} \mathbf{e}_r + \frac{\partial\phi}{\partial\theta} \mathbf{e}_\theta$$

lo cual es falso.

Primer parámetro diferencial de Beltrami. El *primer parámetro diferencial de Beltrami* es el cuadrado de la norma del gradiente, por lo que se obtiene a partir de la expresión

$$\Delta_1\phi = g^{\mu\nu} \phi_{,\mu} \phi_{,\nu}$$

Si quisiéramos considerar un vector contravariante a partir del gradiente, deberíamos utilizar el inverso de la métrica para subir los índices, de esta manera, el *gradiente contravariante* será

$$\vec{\nabla}\phi = g^{\mu\nu} \phi_{,\mu} \mathbf{e}_\nu$$

Indicamos con una barra al operador para indicar que es el gradiente contravariante.

Para el caso de coordenadas polares en el plano, notemos que

$$\vec{\nabla}\phi = \frac{\partial\phi}{\partial x} \mathbf{e}_x + \frac{\partial\phi}{\partial y} \mathbf{e}_y = \frac{\partial\phi}{\partial r} \mathbf{e}_r + \frac{1}{r^2} \frac{\partial\phi}{\partial\theta} \mathbf{e}_\theta$$

Rotor. El rotor de un vector covariante es un vector dos veces covariantes obtenido por derivación covariante

$$\mathbf{Rot}(\tilde{v}) = [v_{\mu;\nu} - v_{\nu;\mu}] \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$$

El uso del tilde es para poner énfasis que el vector es covariante. Si reemplazamos por la definición de derivada covariante, tenemos

$$\mathbf{Rot}(\tilde{v}) = [v_{\mu,\nu} - \Gamma_{\mu\nu}^{\alpha} v_{\alpha} - v_{\nu,\mu} + \Gamma_{\nu\mu}^{\alpha} v_{\alpha}] \mathbf{dx}^{\mu} \otimes \mathbf{dx}^{\nu}$$

En virtud de la simetría de los Christoffel tenemos

$$\mathbf{Rot}(\tilde{v}) = [v_{\mu,\nu} - v_{\nu,\mu}] \mathbf{dx}^{\mu} \otimes \mathbf{dx}^{\nu}$$

Aquí también surge una diferencia con relación a lo visto en cursos de cálculo vectorial elemental, ya que al definir el rotor en coordenadas cartesianas, no aparece su naturaleza covariante, sino que, debido a que la métrica es la euclídea componentes covariantes y contravariantes se confunden, puesto que, como números, son los mismos.

Divergencia. El operador divergencia es un escalar que se obtiene a partir de un vector contravariante y su definición absoluta es

$$\mathbf{Div}(\vec{v}) = v^{\mu}_{;\mu}$$

Calculemos primero la derivada covariante de la coordenada v^{μ} del vector

$$v^{\mu}_{;\nu} = v^{\mu}_{,\nu} + \Gamma_{\nu\alpha}^{\mu} v^{\alpha} = v^{\mu}_{,\nu} + \frac{1}{\sqrt{g}} \frac{\partial[\sqrt{g}]}{\partial x^{\alpha}} v^{\alpha}$$

Ahora igualemos $\nu = \mu$ y en virtud de que en el segundo término α es mudo lo podemos escribir como

$$\mathbf{Div}(\vec{v}) = v^{\mu}_{;\mu} = v^{\mu}_{,\mu} + \frac{1}{\sqrt{g}} \frac{\partial[\sqrt{g}]}{\partial x^{\mu}} v^{\mu}$$

Notemos que es posible simplificar la expresión en la forma

$$\mathbf{Div}(\vec{v}) = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{\mu}} [\sqrt{g} v^{\mu}]$$

recordemos que g es el determinante de las coordenadas de la métrica. Notemos que en coordenadas cartesianas es la suma

$$\mathbf{Div}(\vec{v}) = \frac{\partial v^x}{\partial x} + \frac{\partial v^y}{\partial y} + \frac{\partial v^z}{\partial z}$$

Ejemplo. Coordenadas Esféricas. Para el cambio de coordenadas

$$\begin{aligned} x &= r \cos(\theta) \sin(\varphi) \\ y &= r \sin(\theta) \sin(\varphi) \\ z &= r \cos(\varphi) \end{aligned}$$

Mediante cálculo directo, el elemento de distancia

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2(\theta) d\varphi^2$$

lo que las coordenadas del tensor métrico en esféricas se puede escribir como

$$(g)_{\mu\nu} = \begin{bmatrix} g_{rr} & 0 & 0 \\ 0 & g_{\theta\theta} & 0 \\ 0 & 0 & g_{\varphi\varphi} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2(\varphi) \end{bmatrix}$$

Entonces, el inverso se obtiene de manera trivial

$$(g)^{\mu\nu} = \begin{bmatrix} g^{rr} & 0 & 0 \\ 0 & g^{\theta\theta} & 0 \\ 0 & 0 & g^{\varphi\varphi} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2(\varphi)} \end{bmatrix}$$

El determinante de la métrica es, estas coordenadas, $g = r^4 \sin^2(\varphi)$. Entonces la divergencia se puede calcular como

$$\mathbf{Div}(\vec{v}) = \frac{1}{r^2 \sin(\varphi)} \left\{ \frac{\partial [r^2 \sin(\varphi) v^r]}{\partial r} + \frac{\partial [r^2 \sin(\varphi) v^\theta]}{\partial \theta} + \frac{\partial [r^2 \sin(\varphi) v^\varphi]}{\partial \varphi} \right\}$$

En muchas ocasiones relacionadas a la física teórica, la divergencia es aplicada a un vector gradiente. Sin embargo, ya hemos hecho la aclaración de que el gradiente es una 1-forma y por lo tanto sus coordenadas son covariantes. La divergencia está definida para vectores contravariantes, por lo que si queremos calcular la divergencia de un gradiente (y obtener el laplaciano) será necesario convertirlo en contravariante para luego si poder calcular su divergencia.

Laplaciano. El *Laplaciano*, también llamado *segundo parámetro diferencial de Beltrami* será obtenido a partir de la divergencia de un vector gradiente transformado en contravariante, entonces, dado un campo escalar ϕ

$$\nabla^2(\phi) = \mathbf{Div}(g^{\mu\nu} \phi_{,\nu}) = (g^{\mu\nu} \phi_{,\nu})_{;\mu}$$

En su forma compacta, utilizando el determinante de la métrica, podemos escribir

$$\nabla^2(\phi) = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \left[\sqrt{g} g^{\mu\nu} \frac{\partial \phi}{\partial x^\nu} \right]$$

Calculemos el Laplaciano en esféricas. Tenemos el determinante y la matriz inversa, así que estamos avanzados. Entonces,

$$\nabla^2(\phi) = \frac{1}{r^2 \sin(\varphi)} \frac{\partial}{\partial x^\mu} \left[r^2 \sin(\varphi) g^{\mu\nu} \frac{\partial \phi}{\partial x^\nu} \right]$$

Entonces, reemplazando, obtenemos

$$\nabla^2(\phi) = \frac{1}{r^2 \sin(\varphi)} \left\{ \frac{\partial}{\partial r} \left[r^2 \sin(\varphi) g^{rr} \frac{\partial \phi}{\partial r} \right] + \frac{\partial}{\partial \theta} \left[r^2 \sin(\varphi) g^{\theta\theta} \frac{\partial \phi}{\partial \theta} \right] + \frac{\partial}{\partial \varphi} \left[r^2 \sin(\varphi) g^{\varphi\varphi} \frac{\partial \phi}{\partial \varphi} \right] \right\}$$

Reemplazando los elementos de la inversa de la métrica, tenemos la expresión final

$$\nabla^2(\phi) = \frac{1}{r^2 \sin(\varphi)} \left\{ \frac{\partial}{\partial r} \left[r^2 \sin(\varphi) \frac{\partial \phi}{\partial r} \right] + \frac{\partial}{\partial \theta} \left[\sin(\varphi) \frac{\partial \phi}{\partial \theta} \right] + \frac{\partial}{\partial \varphi} \left[\frac{1}{\sin(\varphi)} \frac{\partial \phi}{\partial \varphi} \right] \right\}$$

Esta expresión es muy engorrosa para hacerla por cambio de variables, pero de manera tensorial es relativamente sencilla de obtener.

23.1 Los operadores ∇ y $\vec{\nabla}$

En los cursos elementales de cálculo vectorial, se presenta al operador ∇ de la siguiente manera

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

Esta definición encierra un conjunto de particularidades que a la hora de plantear una formulación invariante debe ser modificada.

La primera confusión radica en suponer que ∇ es un vector, es decir contravariante. Como la derivación adiciona una coordenada covariante, el operador ∇ es un operador 1-forma

$$\nabla = [\]_{;\mu} \mathbf{dx}^\mu$$

$[\]_{;\mu}$ indica derivación covariante

Tensor gradiente. Como en muchas expresiones vectoriales es necesario definir el "vector gradiente" vamos a definir el vector $\vec{\nabla}$ aprovechando la subida de índices dada por la métrica

$$\vec{\nabla} = g^{\nu\mu} [\]_{;\mu} \mathbf{e}_\nu$$

Con estas definiciones, podemos definir el gradiente pero ahora de un tensor \mathbf{T} , de cualquier tipo a partir de la expresión

$$\mathbf{Grad}(\) = \vec{\nabla} \otimes (\)$$

Esta expresión incorpora una coordenada contravariante al tensor al que se lo aplica.

En particular, si se aplica a una campo escalar ϕ tenemos

$$\mathbf{Grad}(\phi) = \vec{\nabla} \otimes (\phi) = g^{\mu\nu} [\phi]_{;\nu} \mathbf{e}_\mu = g^{\mu\nu} \frac{\partial \phi}{\partial x^\nu} \mathbf{e}_\mu$$

ya que la derivada covariante y parcial coinciden para campos escalares.

Divergencia de un vector contravariante. Como ya fuera definida, la divergencia de un campo vectorial contravariante (única definición para la divergencia) está definida como una contracción

$$\mathbf{Div}(\vec{v}) = v^\mu_{;\mu}$$

A partir del operador $\vec{\nabla}$ podemos, al igual que se define en cálculo vectorial como producto interior, en la forma

$$\begin{aligned} \mathbf{Div}(\vec{v}) &= \vec{\nabla} \cdot \vec{v} = g^{\mu\nu} [\]_{;\mu} \mathbf{e}_\nu \cdot v^\alpha \mathbf{e}_\alpha \\ &= g^{\mu\nu} v^\alpha_{;\mu} \langle \mathbf{e}_\nu | \mathbf{e}_\alpha \rangle \\ &= g^{\mu\nu} g_{\mu\alpha} v^\alpha_{;\mu} \\ &= \delta^\nu_\alpha v^\alpha_{;\mu} \\ &= v^\nu_{;\nu} \end{aligned}$$

como en la definición original.

Laplaciano. Con la definición del vector $\vec{\nabla}$ podemos definir el Laplaciano como operador

$$\nabla^2 [\] = \vec{\nabla} \cdot \vec{\nabla} \otimes [\]$$

En particular, el laplaciano de un campo escalar, ϕ es

$$\begin{aligned}
 \nabla^2[\phi] &= \vec{\nabla} \cdot \vec{\nabla} \otimes [\phi] \\
 &= \vec{\nabla} \cdot g^{\mu\nu}[\phi]_{;\nu} \mathbf{e}_\mu = \vec{\nabla} \cdot g^{\mu\nu}[\phi]_{,\nu} \mathbf{e}_\mu \\
 &= g^{\alpha\beta}[\]_{;\beta} \mathbf{e}_\alpha \cdot g^{\mu\nu}[\phi]_{,\nu} \mathbf{e}_\mu \\
 &= g^{\alpha\beta} [g^{\mu\nu} \phi_{,\nu}]_{;\beta} \langle \mathbf{e}_\alpha | \mathbf{e}_\mu \rangle \\
 &= g^{\alpha\beta} g_{\alpha\mu} [g^{\mu\nu} \phi_{,\nu}]_{;\beta} = \delta_\mu^\beta [g^{\mu\nu} \phi_{,\nu}]_{;\beta} \\
 &= [g^{\mu\nu} \phi_{,\nu}]_{;\mu}
 \end{aligned}$$

que es la expresión que se había dado para el laplaciano.

24 Aplicaciones a la teoría de curvas

Curvas. En un espacio S de Riemann de dimensión n (en general se denomina *variedad Riemanniana*), es decir, donde tenemos definida una métrica $g_{\mu\nu}$, vamos definir una curva a una función $\mathcal{R} \rightarrow S$ definida a través de

$$x^\alpha(t), \quad a \leq t \leq b, \quad \alpha = 1, 2, \dots, n$$

Con estas funciones se puede definir un tangente a la curva

$$\frac{d\mathbf{x}}{dt} = \frac{dx^\alpha}{dt} \mathbf{e}_\alpha$$

24.1 Campo de Tensores sobre curvas. Derivada.

Consideremos un campo tensorial \mathbf{T} sobre una variedad Riemanniana con métrica definida por el elemento de arco

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

Esto es equivalente a decir con métrica $g_{\mu\nu}$.

Vamos a considerar el campo tensorial definido en los puntos de la curva, a través de la composición de funciones

$$\mathbf{T}(t) = \mathbf{T}[\mathbf{x}(t)]$$

La derivada del tensor en cada punto de la curva la calculamos a través de la regla de la cadena

$$\frac{d\mathbf{T}}{dt} = \frac{\partial \mathbf{T}}{\partial x^\alpha} \frac{dx^\alpha}{dt}$$

Lo que debemos tener en cuenta es que la derivada parcial de un tensor se aplica como derivación covariante en sus coordenadas.

Ejemplos. Para ilustrar la definición de derivada a lo largo de una curva, consideremos primero un tensor dos veces contravariantes

$$\mathbf{T} = T^{\mu\nu} \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

Entonces, dada una curva como la ya definida, la derivada

$$\frac{d\mathbf{T}}{dt} = \frac{\partial \mathbf{T}}{\partial x^\alpha} \frac{dx^\alpha}{dt}$$

se puede escribir en coordenadas

$$\frac{d\mathbf{T}}{dt} = T^{\mu\nu}{}_{;\alpha} \frac{dx^\alpha}{dt} \mathbf{e}_\mu \otimes \mathbf{e}_\nu$$

Es decir, que en coordenadas, la derivada se escribe como

$$T^{\mu\nu}{}_{;\alpha} \frac{dx^\alpha}{dt}$$

Este ejemplo alcanza para ilustrar cómo se calcula la derivada para tensores de diferentes tipos. Siempre se utiliza la derivación covariante en las coordenadas.

Derivada de un campo vectorial. Este ejemplo está incluido naturalmente en el anterior, pero igualmente calculemos la derivada a lo largo de la curva. El vector \mathbf{v} lo escribimos de manera tensorial, como tensor una vez contravariante

$$\mathbf{v} = v^\mu \mathbf{e}_\mu$$

Entonces, la derivada en cada punto de la curva es

$$\frac{d\mathbf{v}}{dt} = v^\mu{}_{;\alpha} \frac{dx^\alpha}{dt} \mathbf{e}_\mu$$

Entonces, las coordenadas del vector cambian a lo largo de la curva según la ley

$$v^\mu{}_{;\alpha} \frac{dx^\alpha}{dt} = [v^\mu{}_{,\alpha} + \Gamma^\mu_{\nu\alpha} v^\nu] \frac{dx^\alpha}{dt}$$

24.2 Transporte paralelo

Una aplicación de lo anterior es el *transporte paralelo* de un vector a lo largo de una curva sobre la variedad Riemanniana. Consideremos una curva sobre una variedad Riemanniana. Consideremos un campo vectorial definido en la variedad y definamos la manera de transportar paralelamente al vector sobre la curva.

De manera intuitiva podemos asumir que la manera de transportar paralelamente a un vector determinado desde dos puntos de la curva es aquella en la que en cada punto la derivada del vector en cuestión en la dirección a la curva es nula. Existen varios abordajes para este tema, los cuales pueden verse en la bibliografía, pero este en particular está inspirado en lo planteado por Schutz [4]. Entonces, dado el vector \mathbf{v} , este será *transportado paralelamente* a lo largo de la curva si y sólo si

$$\frac{d\mathbf{v}}{dt} = 0, \quad \rightarrow \quad v^\mu{}_{;\alpha} \frac{dx^\alpha}{dt} = [v^\mu{}_{,\alpha} + \Gamma^\mu_{\nu\alpha} v^\nu] \frac{dx^\alpha}{dt} = 0 \quad \forall \mu = 1, 2, \dots, n$$

24.3 Geodésicas

Existen varias maneras de definir las curvas geodésicas. La idea primigenia es, dada una variedad Riemanniana, de todas las curvas que unen dos puntos de la variedad, llamaremos *geodésica* a aquella de menor longitud.

Esta definición implica minimizar una función que es la longitud de la curva entre los puntos en cuestión.

$$\int_a^b \sqrt{g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} dt$$

donde la incógnita es justamente la curva $x^\alpha(t)$.

Ahora, esta formulación exige conocimientos de *cálculo variacional*, que no supondremos.

Otra formulación para la definición de geodésica es aquella curva cuyo tangente se transporta paralelamente sobre ella. Esta definición es una interpretación de la generalización de la recta, la cual en un espacio plano satisface esta definición.

Con esta definición, recordando la condición de que un vector se transporte paralelamente es

$$[v^\mu{}_{,\alpha} + \Gamma^\mu_{\nu\alpha} v^\nu] \frac{dx^\alpha}{dt} = 0 \quad \forall \mu = 1, 2, \dots, n$$

Si ahora el vector a ser transportado paralelamente es el propio tangente, tenemos que

$$v^\mu = \frac{dx^\mu}{dt}$$

entonces, la ecuación diferencial para una curva geodésica será

$$\underbrace{\left[\frac{dx^\mu}{dt} \right]_{,\alpha}}_{\frac{d^2 x^\mu}{dt^2}} + \Gamma^\mu_{\nu\alpha} \frac{dx^\nu}{dt} \frac{dx^\alpha}{dt} = 0 \quad \forall \mu = 1, 2, \dots, n$$

Entonces, la ecuación diferencial para la geodésica será

$$\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\alpha} \frac{dx^\nu}{dt} \frac{dx^\alpha}{dt} = 0 \quad \forall \mu = 1, 2, \dots, n$$

25 Elementos de Geometría Riemanniana

Con el objeto de concluir esta introducción al análisis tensorial vamos a estudiar aspectos geométricos de las variedades Riemannianas. Tales aspectos serán los que caractericen a las variedades y donde la curvatura tendrá un rol fundamental.

Nuestra hipótesis central es una variedad en la cual está definida una métrica caracterizada por los elementos $g_{\mu\nu}$. A partir de estos elementos, vamos a definir diferentes cantidades tensoriales que caracterizan la variedad.

Tensor de Curvatura de Riemann. Sin entrar en los detalles a partir de los cuales se obtiene, vamos a tomar como definición al *tensor de curvatura de Riemann*, al tensor cuyas coordenadas $R^\alpha_{\beta\mu\nu}$ se obtienen a partir de la métrica y sus derivadas segundas

$$R^\alpha_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\sigma} [g_{\sigma\nu, \beta\mu} - g_{\sigma\mu, \beta\nu} + g_{\beta\mu, \sigma\nu} - g_{\beta\nu, \sigma\mu}]$$

Una variedad se denomina *plana* si la curvatura de Riemann es nula, es decir

$$R^\alpha_{\beta\mu\nu} = 0$$

A partir de la bajada de índices podemos definir el tensor

$$R_{\alpha\beta\mu\nu} = g_{\alpha\lambda} R^\lambda_{\beta\mu\nu}$$

Este tensor de curvatura de Riemann completamente covariante satisface las siguientes propiedades

I.

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}$$

II.

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0$$

Tensor de Ricci. A partir del tensor de Riemann, se define el *tensor de Ricci* mediante la contracción

$$R_{\alpha\beta} = R^{\mu}_{\alpha\mu\beta} = R_{\beta\alpha}$$

Escalar de Ricci. Con el tensor de Ricci, se define el *escalar de Ricci* a través de

$$R = g^{\alpha\beta} R_{\alpha\beta}$$

Estos parámetros de curvatura son de aplicación en la *Teoría General de la Relatividad* de Einstein, a partir de la cual es que se establece la curvatura del espacio-tiempo.

VII

Autovalores y Autovectores. Formas Canónicas.

26 Introducción. Definiciones

En diversos problemas nos encontramos con representaciones matriciales. Sistema de ecuaciones lineales, sistema de ecuaciones diferenciales ordinarias y problemas de optimización son algunos ejemplos. Podemos notar que para los casos en los cuales las matrices asociadas son cuadradas, una representación diagonal de las mismas, esto es, que las matrices involucradas sean de la forma

$$\mathbf{A} = \begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}$$

En particular, un sistema de ecuaciones diferenciales lineales ordinarias

$$\begin{aligned} \frac{dx}{dt} &= a_1 x + b_1 y \\ \frac{dy}{dt} &= a_2 x + b_2 y \end{aligned}$$

Resulta simple de resolver si está en una forma diagonal, ya que

$$\begin{aligned} \frac{dx}{dt} &= d_1 x \quad \rightarrow \quad x(t) = x_0 e^{d_1 t} \\ \frac{dy}{dt} &= d_2 y \quad \rightarrow \quad y(t) = y_0 e^{d_2 t} \end{aligned}$$

son dos ecuaciones completamente desacopladas cuya resolución es elemental.

En este capítulo nos dedicaremos a estudiar técnicas a partir de las cuales podamos determinar la posibilidad de expresar las matrices de manera diagonal. O, en su defecto, escribir las matrices *lo más diagonal posible*.

Definición. Sea V un espacio vectorial de dimensión finita sobre un cuerpo K . Consideremos un operador lineal, es decir, una transformación lineal $T : V \rightarrow V$. Un autovalor de T es un número $\lambda \in K$, tal que existe un vector no nulo de V , \vec{v} que satisface

$$T(\vec{v}) = \lambda \vec{v}$$

Al vector \vec{v} que satisface la ecuación se lo denomina autovector.

A partir de la definición, podemos notar que no necesariamente un autovalor posee un único autovector asociado.

Definición. La colección de todos los autovectores asociados a un mismo autovalor λ forman un subespacio de V que se llama espacio propio asociado al autovalor λ .

En efecto, dado un autovalor λ podemos notar que si \vec{u} y \vec{v} son autovectores asociados a λ entonces, el vector $\vec{w} = c\vec{u} + \vec{v}$ es también un autovector asociado a λ

$$T(\vec{w}) = T(c\vec{u} + \vec{v}) = cT(\vec{u}) + T(\vec{v}) = c\lambda\vec{u} + \lambda\vec{v} = \lambda(c\vec{u} + \vec{v}) = \lambda\vec{w}$$

A partir de estas definiciones, podemos probar el siguiente resultado:

Teorema. Sea T un operador lineal sobre un espacio V de dimensión finita y sea λ un escalar. Las siguientes afirmaciones son equivalentes:

- a) λ es un autovalor de T
- b) El operador $T - \lambda I_d$ es singular
- c) Si $[T]$ es la matriz asociada de T respecto a una determinada base, entonces

$$\det([T] - \lambda I_{n \times n}) = 0$$

Este teorema es fundamental porque establece la metodología de trabajo con relación al estudio de autovalores y autovectores.

Un aspecto a resaltar es el siguiente: Las afirmaciones a) y b) se refieren a T como operador, mientras que c) trabaja con la matriz asociada.

Para demostrar el Teorema, basta con demostrar que a) \rightarrow b), b) \rightarrow c) y que c) \rightarrow a)

- a) \rightarrow b)

Si λ es un autovalor, existe un autovector no nulo, \vec{v} tal que $T(\vec{v}) = \lambda \vec{v}$. Dado que tenemos en el espacio de operadores lineales operaciones suma y producto, podemos escribir, como operador

$$(T - \lambda I_d)(\vec{v}) = 0 \quad (\text{Operador nulo})$$

Esto implica que el operador $(T - \lambda I_d)$ admite un vector no nulo en el núcleo. Entonces, no puede ser biyectivo. Por lo tanto, por definición, es singular.

- b) \rightarrow c)

Dada una base para el espacio V , $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, al escribir el operador $(T - \lambda I_d)$ tendrá como matriz asociada, a $[(T - \lambda I_d)]_{\mathcal{B}}$. Dado que el operador es singular, tendremos un vector no nulo \vec{v} en el núcleo de T . Si $\vec{v} = v^\mu \mathbf{e}_\mu$ podemos escribir la representación matricial

$$[(T - \lambda I_d)]_{\mathcal{B}} \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Entonces, por la teoría de sistema de ecuaciones lineales, si un sistema homogéneo admite otra solución que la trivial no puede tener solución única, por lo tanto, $\det[T - \lambda I_d] = 0$

- c) \rightarrow a)

Si $\det[T - \lambda I_d] = 0$, entonces, existe un vector columna no nulo $\vec{v} = v^\mu \mathbf{e}_\mu$ que es solución del sistema lineal

$$[(T - \lambda I_d)]_{\mathcal{B}} \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Considerando la matriz $[(T - \lambda I_d)]$ como la matriz de una transformación lineal, tenemos que para este vector \vec{v} se cumple

$$(T - \lambda I_d) \vec{v} = 0$$

o lo que es lo mismo,

$$T(\vec{v}) = \lambda \vec{v}$$

lo que implica que λ es un autovalor.

Esto concluye la demostración del Teorema.

El teorema establece la siguiente metodología de análisis, para el estudio de autovalores y autovectores:

1. Dada la matriz de un operador lineal, con el inciso *c*) del teorema calculamos los autovectores resolviendo la ecuación algebraica

$$\det [T] - \lambda I_d = 0$$

Notemos que al calcular el determinante, lo que se obtiene es una ecuación en λ ,

$$p(\lambda) = \det [T] - \lambda I_d = 0$$

Esta ecuación se denomina *ecuación característica*, y al polinomio $p(\lambda)$, *polinomio característico*. Para una matriz $n \times n$, la ecuación característica será una ecuación de grado n .

2. Una vez obtenidos los autovalores, los autovectores, serán los elementos de

$$\text{Nu}(T - \lambda I_d) \quad \text{es decir,} \quad (T - \lambda I_d)(\vec{v}) = 0$$

Ejemplo. (Grossman, *Álgebra Lineal*, Ed. Mc Graw-Hill, 2008) Consideremos la matriz $\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$. La ecuación característica es

$$\det \left(\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = 0$$

$$\det \left(\begin{bmatrix} 4 - \lambda & 2 \\ 3 & 3 - \lambda \end{bmatrix} \right) = (4 - \lambda)(3 - \lambda) - 6 = \lambda^2 - 7\lambda + 6 = 0$$

La solución de la ecuación característica es $\lambda_1 = 1$ y $\lambda_2 = 6$. Para obtener las coordenadas de los autovectores hacemos, para el primer autovalor $\lambda = 1$

$$\begin{bmatrix} 4 - 1 & 2 \\ 3 & 3 - 1 \end{bmatrix} \begin{bmatrix} v_1^1 \\ v_1^2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} v_1^1 \\ v_1^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

La solución para el primer caso es

$$3v_1^1 + 2v_1^2 = 0 \quad \rightarrow \quad v_1^2 = -\frac{3}{2}v_1^1$$

Entonces, si $\{\mathbf{e}_1, \mathbf{e}_2\}$ es la base del espacio con la que se construyó la matriz del operador, el subespacio propio asociado al autovalor λ_1 es

$$S_{\lambda_1} = \overline{\left\{ \mathbf{e}_1 - \frac{3}{2}\mathbf{e}_2 \right\}} \quad \text{o, equivalentemente,} \quad S_{\lambda_1} = \overline{\{2\mathbf{e}_1 - 3\mathbf{e}_2\}}$$

Análogamente, para el autovalor $\lambda_2 = 6$ tenemos

$$\begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} v_2^1 \\ v_2^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuya solución (que obviamente no es única) es $v_2^1 = v_2^2$ lo que implica que el espacio propio asociado al autovalor $\lambda_2 = 6$ es el generado por este autovector

$$S_{\lambda_2} = \overline{\{\mathbf{e}_1 + \mathbf{e}_2\}}$$

Observación. Podemos afirmar que el espacio V es de dimensión 2, pero no sabemos absolutamente nada de la naturaleza de los objetos que constituyen el espacio. Nada indica que el espacio es \mathcal{R}^2 . Sólo podemos escribir los autovectores en función de la base hipotética que dió lugar a la matriz.

Podemos notar (y se deja como ejercicio) que los vectores $2\mathbf{e}_1 - 3\mathbf{e}_2$ y $\mathbf{e}_1 + \mathbf{e}_2$ son linealmente independientes. Por lo que en este caso, los autovectores constituyen una base para el espacio V . Si deseamos obtener la matriz del operador en la base $\{\mathbf{e}'_1, \mathbf{e}'_2\}$ donde $\mathbf{e}'_1 = 2\mathbf{e}_1 - 3\mathbf{e}_2$ y $\mathbf{e}'_2 = \mathbf{e}_1 + \mathbf{e}_2$ como hemos visto, tendremos que tener en cuenta o que ya hemos visto,

$$\mathbf{e}'_\mu = \Lambda^\nu_\mu \mathbf{e}_\nu, \quad \text{donde} \quad \mathbf{\Lambda} = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}$$

La matriz inversa de $\mathbf{\Lambda}$ sera

$$\mathbf{\Lambda}^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

Con estas matrices, la matriz asociada a la transformación lineal en esta nueva base se obtiene como

$$[T]_{\mathcal{B}'} = \mathbf{\Lambda}^{-1} [T]_{\mathcal{B}} \mathbf{\Lambda}$$

reemplazando

$$[T]_{\mathcal{B}'} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix}$$

cuyo resultado es

$$[T]_{\mathcal{B}'} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Podemos notar que si expresamos la matriz de la transformación en una base de autovectores el resultado es una matriz diagonal, cuyos elementos de la diagonal son los autovalores. Esto no siempre es posible, pero cuando se puede realizar se dice que el operador es *diagonalizable*.

El ejemplo estudiado es claramente particular, pero contiene un aspecto general de los autovectores: la independencia lineal. Vamos a ver un resultado general.

Teorema. *Sea T un operador lineal sobre un espacio V de dimensión finita. Autovectores asociados a distintos autovalores son linealmente independientes.*

Vamos a demostrar este teorema por el principio de Inducción.

Sean \vec{u} y \vec{v} son dos vectores asociados a diferentes autovalores λ_1 y λ_2 . Consideremos la combinación lineal que de el vector nulo

$$\alpha \vec{v}_1 + \beta \vec{v}_2 = \vec{0}$$

Aplicando la transformación lineal tenemos

$$\alpha T(\vec{v}_1) + \beta T(\vec{v}_2) = \alpha \lambda_1 \vec{v}_1 + \beta \lambda_2 \vec{v}_2 = \vec{0}$$

Si la combinación original es multiplicada por $-\lambda_2$ y restamos ambas ecuaciones, tenemos

$$(\lambda_2 - \lambda_1) \alpha \vec{v}_1 = 0$$

como el vector es no nulo, y los autovalores también, nos queda que α debe ser cero. Lo que implicará que β también deba serlo. Con lo cual, son linealmente independientes.

Si ahora suponemos que $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ -autovectores asociados a distintos autovalores- son linealmente independientes, veamos qué ocurre con $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}$ suponiendo que los primeros k son LI. Consideremos la combinación lineal

$$\sum_{\mu=1}^{k+1} \alpha_{\mu} v_{\mu} = 0$$

Aplicando la transformación y sabiendo que $T(\vec{v}_{\mu}) = \lambda_{\mu} \vec{v}_{\mu}$, tenemos

$$\sum_{\mu=1}^{k+1} \alpha_{\mu} \lambda_{\mu} v_{\mu} = 0$$

Si a la combinación original la multiplicamos por λ_{k+1} y restamos, tenemos

$$\sum_{\mu=1}^k \alpha^{\mu} (\lambda_{\mu} - \lambda_{k+1}) \vec{v}_{\mu} = 0 \quad (\text{notemos que el último término se cancela})$$

Además, por hipótesis inductiva, los k vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ son LI, tenemos que los $\alpha^1, \alpha^2, \dots, \alpha^k$ deben ser cero. Pero si todos estos coeficientes deben ser cero, implica que α^{k+1} también deba serlo. Por lo que son linealmente independientes.

Este Teorema, da lugar a la siguiente definición.

Definición. Sea T un operador lineal sobre un espacio V de dimensión finita. Se dice que T es diagonalizable si existe una base de V de autovectores.

Más aún, la matriz asociada al operador T en esta base de autovectores es diagonal, cuyos elementos de la diagonal son los autovalores.

Observación. Sabemos que una matriz asociada a una transformación lineal es obtenida a partir de la aplicación de la transformación a los elementos de la base del espacio. Quiere

decir que, en cierto sentido, la matriz no es un ente absoluto de la transformación lineal, sino que depende de la base.

Vamos a ver ahora, con el siguiente resultado, que los autovalores son intrínsecos al operador lineal y que no dependen de la base elegida para el espacio vectorial.

Teorema. *Las matrices semejantes poseen los mismos autovalores.*

Para demostrarlo, consideremos un operador lineal para el cual en una determinada base su matriz asociada posee un autovalor λ . Esto significa que si la matriz del operador es \mathbf{A} tenemos que $\det(\mathbf{A} - \lambda I) = 0$

Si \mathbf{B} es una matriz semejante a \mathbf{A} tenemos

$$\mathbf{B} = \mathbf{\Lambda}^{-1} \mathbf{A} \mathbf{\Lambda}$$

donde $\mathbf{\Lambda}$ es la matriz de cambio de base. Calculemos $\det(\mathbf{B} - \lambda I)$.

Tenemos

$$\det(\mathbf{B} - \lambda I) = \det(\mathbf{\Lambda}^{-1} \mathbf{A} \mathbf{\Lambda} - \lambda I) = \det(\mathbf{\Lambda}^{-1} \mathbf{A} \mathbf{\Lambda} - \lambda \mathbf{\Lambda}^{-1} \mathbf{\Lambda})$$

Tomando factor común, $\mathbf{\Lambda}^{-1}$ a izquierda, y $\mathbf{\Lambda}$ a derecha, tenemos

$$\det(\mathbf{\Lambda}^{-1} \mathbf{A} \mathbf{\Lambda} - \lambda I) = \det(\mathbf{\Lambda}^{-1} [\mathbf{A} - \lambda I] \mathbf{\Lambda})$$

Entonces, si $\det(\mathbf{A} - \lambda I) = 0$ implica que $\det(\mathbf{B} - \lambda I) = 0$. Entonces, λ es también autovalor de \mathbf{B} .

26.1 El Polinomio Característico. Multiplicidad de las raíces.

La ecuación característica es una ecuación algebraica (polinómica) del tipo $p(\lambda) = 0$. El Teorema Fundamental del Álgebra garantiza que la ecuación posee n raíces, las cuales pueden estar repetidas.

En el caso de que todas las raíces sean distintas, podemos garantizar que el operador es diagonalizable, puesto que al tener n autovalores distintos, tendrá n autovectores distintos. Además, como autovectores asociados a distintos autovalores son linealmente independientes constituirán una base del espacio y la matriz asociada en esa base será diagonal, con los autovalores en la diagonal. Este es el caso del ejemplo que fue analizado.

En el caso de que no todos los autovalores sean distintos, habrá autovalores con multiplicidad mayor a uno.

En general, el polinomio característico podrá escribirse como

$$p(\lambda) = (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} (\lambda - \lambda_3)^{d_3} \dots (\lambda - \lambda_k)^{d_k}$$

donde, claramente,

$$d_1 + d_2 + \dots + d_k = n$$

Cada d_k es denominada multiplicidad algebraica

Como vimos, dos autovectores asociados a distintos autovalores son linealmente independientes. Este teorema no prohíbe la independencia lineal entre vectores asociados a un mismo autovalor.

Definición. La dimensión del subespacio asociado a un determinado autovalor es denominada multiplicidad geométrica.

Si la multiplicidad geométrica coincide con la multiplicidad algebraica, el operador será diagonalizable. De otro modo, no lo será.

26.2 Operador Adjunto

Aplicando lo que hemos visto para operadores lineales en un espacio euclídeo, vamos a definir un operador llamado *adjunto*. Vamos a definirlo a partir del siguiente teorema que garantiza su existencia y unicidad.

Teorema. Sea T un operador lineal en un espacio producto interno de dimensión finita, V . Existe un único operador lineal T^* sobre V , que denominaremos adjunto, tal que

$$\langle T(\vec{v}) | \vec{u} \rangle = \langle \vec{v} | T^*(\vec{u}) \rangle$$

Demostración. Como vimos, el producto interno es una forma bilineal, pero si dejamos un vector fijo, define una 1-forma. Entonces, dado un \vec{v} , podemos afirmar que $f : V \rightarrow \mathcal{R}$ tal que para algún $\vec{u} \in V$ la aplicación $\vec{v} \rightarrow \langle T(\vec{v}) | \vec{u} \rangle$ define una 1-forma.

Consideremos para V una base ortonormal, es decir, $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ donde $\langle \mathbf{e}_\mu | \mathbf{e}_\nu \rangle = \delta_{\mu\nu}$ lo que implica que las coordenadas del tensor métrico son $g_{\mu\nu} = \delta_{\mu\nu}$

Calculemos, $\langle T(\vec{v}) | \vec{u} \rangle$,

$$\langle T(\vec{v}) | \vec{u} \rangle = \langle T_\nu^\mu v^\nu \mathbf{e}_\mu | u^\lambda \mathbf{e}_\lambda \rangle = T_\nu^\mu v^\nu u^\lambda g_{\mu\lambda} = T_\nu^\mu v^\nu u^\mu$$

Debido a que la base es ortonormal, las coordenadas covariantes y contravariantes coinciden, por lo que podemos escribir

$$\langle T(\vec{v}) | \vec{u} \rangle = T_\nu^\mu v_\nu u^\mu = v_\nu T_\nu^\mu u^\mu = v^\lambda T_\nu^\mu u^\mu g_{\lambda\nu}$$

Si llamamos $\mathbf{T}^* = \mathbf{T}^t$ tendremos que $T_\mu^* \nu = T_\nu^\mu$ Con esta definición, $T_\nu^\mu u^\mu$ es un producto de matrices, con lo que podemos escribir

$$\langle T(\vec{v}) | \vec{u} \rangle = v^\lambda T_\mu^* \nu u^\mu g_{\lambda\nu} = v^\lambda T_\mu^* \nu u^\mu \langle \mathbf{e}_\lambda | \mathbf{e}_\nu \rangle = \langle \vec{v} | T^*(\vec{u}) \rangle$$

Llegamos a que

$$\langle T(\vec{v}) | \vec{u} \rangle = \langle \vec{v} | T^*(\vec{u}) \rangle$$

lo que indica no sólo que existe el operador T^* sino que la matriz asociada es la matriz transpuesta a la matriz de T . En el caso más general, para espacios vectoriales sobre los reales, $\langle \vec{v} | \vec{u} \rangle = v^\lambda \bar{u}^\mu g_{\lambda\mu}$ lo que llevaría, en un análisis completamente análogo a la existencia del operador adjunto, cuya matriz asociada

$$\mathbf{T}^* = \bar{\mathbf{T}}^t$$

26.3 Matrices Hermíticas. Matrices Simétricas

Definición. Vamos a definir como operador hermítico a todo operador que satisfice

$$T^* = T$$

Podemos notar que para el caso de espacios reales, los operadores hermíticos son simétricos.

En términos matriciales,

$$T^t = T$$

Para matrices hermíticas, tenemos los siguientes resultados:

Teorema. *Los operadores hermíticos tienen autovalores reales.*

En efecto, consideremos el autovalor λ y su autovector \vec{v}

$$\begin{aligned} \langle T(\vec{v}) | \vec{v} \rangle &= \langle \vec{v} | T^*(\vec{v}) \rangle \\ \langle T(\vec{v}) | \vec{v} \rangle &= \langle \vec{v} | T(\vec{v}) \rangle \\ \langle \lambda \vec{v} | \vec{v} \rangle &= \langle \vec{v} | \lambda \vec{v} \rangle \\ \lambda \langle \vec{v} | \vec{v} \rangle &= \bar{\lambda} \langle \vec{v} | \vec{v} \rangle \end{aligned}$$

Entonces, $\lambda = \bar{\lambda}$ lo que implica que $\lambda \in \mathcal{R}$

Otro resultado importante lo aporta el siguiente teorema

Teorema. *Los autovectores asociados a distintos autovalores son ortogonales.*

En efecto, sean \vec{v} y \vec{u} autovectores asociados a los autovalores λ y ξ , respectivamente. Calculemos,

$$\begin{aligned} \langle T(\vec{v}) | \vec{u} \rangle &= \langle \vec{v} | T^*(\vec{u}) \rangle \\ \langle T(\vec{v}) | \vec{u} \rangle &= \langle \vec{v} | T(\vec{u}) \rangle \\ \langle \lambda \vec{v} | \vec{u} \rangle &= \langle \vec{v} | \xi \vec{u} \rangle \\ \lambda \langle \vec{v} | \vec{u} \rangle = \xi \langle \vec{v} | \vec{u} \rangle &\rightarrow (\lambda - \xi) \langle \vec{v} | \vec{u} \rangle = 0 \end{aligned}$$

y como los autovalores son distintos, tenemos que \vec{v} y \vec{u} son ortogonales.

Finalmente, podemos afirmar

Teorema. operador hermítico es diagonalizable. *En términos de matrices reales, tenemos que toda matriz simétrica es diagonalizable.*

La demostración de este teorema tiene una complejidad estrictamente técnica y dejamos para que sea leída en un texto de la bibliografía recomendada.

Notemos que una matriz simétrica no necesariamente posee todos los autovalores distintos, lo que significa que lo que hay que demostrar es que la multiplicidad geométrica para cada autovalor coincide con la algebraica.

Aspecto Metodológico para la Diagonalización de un operador simétrico.

Con la garantía de que a partir de una matriz simétrica puedo obtener una base de autovectores, si nos encontramos con n autovalores distintos, la propia obtención de los n autovectores nos proveerán de una base en la cual el operador es diagonal.

Si tenemos que raíces múltiples para los autovectores, un mecanismo es el siguiente:

- Obtener los autovalores
- Para cada autovalor obtener una base para el espacio propio asociado
- En el caso de que la dimensión del espacio sea mayor o igual a 2, deberemos hallar una base ortogonal a partir del proceso, por ejemplo, de Gram-Schmidt

26.4 Operadores Unitarios

Definición. Un operador se dice unitario si preserva producto interno, esto es,

$$\langle T(\vec{u})|T(\vec{v})\rangle = \langle \vec{u}|\vec{v}\rangle$$

A partir de esta definición, podemos afirmar

Teorema. Un operador unitario tiene por inversa su adjunto.

Demostración. En efecto, por ser T unitario, tenemos

$$\langle T(\vec{u})|T(\vec{v})\rangle = \langle \vec{u}|\vec{v}\rangle$$

ahora, por definición de adjunto,

$$\langle T(\vec{u})|T(\vec{v})\rangle = \langle u|T^*(T(\vec{v}))\rangle = \langle \vec{u}|\vec{v}\rangle$$

Con lo que obtenemos que su adjunto es el inverso de T .

Para el caso de operadores sobre espacios euclídeos sobre los reales, tendremos que el inverso de un operador adjunto es su transpuesto.

Ejemplo. Rotación en \mathcal{R}^2 Consideremos una rotación de ángulo θ

$$\begin{aligned} T(\mathbf{e}_x) &= \cos(\theta)\mathbf{e}_x + \sin(\theta)\mathbf{e}_y \\ T(\mathbf{e}_y) &= -\sin(\theta)\mathbf{e}_x + \cos(\theta)\mathbf{e}_y \end{aligned}$$

La matriz asociada será, entonces,

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Notemos que al calcular la inversa obtenemos

$$R_\theta^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = R_\theta^t = R_{(-\theta)}$$

que no solo coincide con la transpuesta, sino también en las rotaciones, una rotación de ángulo $-\alpha$ es la operación inversa a rotar α .

Este resultado es particularmente útil para cambio de coordenadas ortogonales.

Dado un espacio vectorial V de dimensión finita para el que contamos con una base ortogonal. Si la matriz de un operador lineal en esta base es T entonces al realizar un cambio de coordenadas ortogonales cuya matriz es \mathbf{Q} la matriz asociada en la nueva base será

$$\mathbf{T}' = \mathbf{Q}^t \mathbf{T} \mathbf{Q}$$

Matrices relacionadas a través de esta relación se las denomina *ortogonalmente semejantes*.

27 Forma Canónica de Jordan

Hemos visto que cuando la multiplicidad geométrica, esto es, la dimensión del espacio propio asociado al autovalor λ_ℓ coincide con la multiplicidad algebraica -la multiplicidad de la raíz- la matriz es diagonalizable.

Si estos números no coinciden, es porque

$$\text{multiplicidad}_{\text{geométrica}} < \text{multiplicidad}_{\text{algebraica}}$$

Veamos que se puede hacer en estos casos para expresar el operador (en su representación matricial) en una forma canónica que por supuesto ya no será diagonal.

27.1 Matriz de bloque de Jordan

Para un escalar dado, λ se define la matriz de bloques de Jordan, $B_k(\lambda)$ a la matriz

$$B_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & \cdots & \lambda & 1 \\ 0 & 0 & & \cdots & 0 & \lambda \end{bmatrix} = \lambda I_{k \times k} + N_k$$

donde

$$N_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & \cdots & 0 & 1 \\ 0 & 0 & & \cdots & 0 & 0 \end{bmatrix}$$

Podemos notar que $(N_k)^k = 0$ y se denomina nilpotente. En general, toda matriz triangular con ceros en la diagonal es nilpotente.

27.2 Matriz de Jordan

Una matriz de Jordan, está compuesta por matrices de bloque de Jordan de la forma:

$$J = \begin{bmatrix} B_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & B_2(\lambda_1) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & B_\ell(\lambda_\ell) \end{bmatrix}$$

Observación. Una matriz diagonal es una matriz de Jordan. En efecto, está compuesta por matrices de bloques de Jordan de 1×1

Ejemplos.

a) $\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 7 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} & 0 \\ 0 & 7 \end{bmatrix}$ es decir, compuesta por dos matrices de bloque de Jordan, el

primero de 2×2 y el segundo de 1×1

$$b) \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & 0 & 0 & 0 \\ 0 & [3] & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} & 0 \\ 0 & 0 & 0 & \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} \end{bmatrix}$$

Este ejemplo es interesante, porque la aparición de los 3 en la diagonal podría haber generado una confusión con respecto al bloque de Jordan.

Un caso similar para $n=3$ sería

$$c) \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} [3] & 0 \\ 0 & \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \end{bmatrix} \quad \text{esta matriz está compuesta por dos bloques de Jordan.}$$

27.3 Relación entre la multiplicidad geométrica y los bloques

Por cada autovalor, la cantidad de bloques de Jordan indicará la multiplicidad geométrica que el mismo tenga. Para el inciso c) del ejemplo anterior, la matriz posee un único autovalor, el 3, que posee una multiplicidad geométrica 2, dando como resultado la matriz presentada.

Cuando tenemos una matriz diagonalizable, lo que procuramos son los autovectores base del espacio que posibilitan que la matriz del operador en esa base sea diagonal. Eso era posible porque las multiplicidades geométricas y algebraicas para cada autovalor coincidían.

En el caso general, el concepto de diagonalización se extiende a través del siguiente resultado

Teorema. Sea \mathbf{A} una matriz cuadrada. Entonces siempre existe una matriz $\mathbf{\Lambda}$ tal que

$$\mathbf{\Lambda}^{-1} \mathbf{A} \mathbf{\Lambda} = \mathbf{J}$$

donde \mathbf{J} es una matriz de Jordan.

Lo que afirma este teorema es que siempre existe una base en la cual la representación de un operador lineal sea una matriz de Jordan.

El problema consiste entonces en determinar esa base. Consideremos el aspecto metodológico para llevar una matriz a la forma de Jordan.

Consideremos una matriz $\mathbf{A} \in \mathcal{R}^{n \times n}$ cuya ecuación característica es

$$(\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \dots (\lambda - \lambda_r)^{d_r} = 0$$

donde $d_1 + d_2 + \dots + d_r = n$

Procedimiento para la obtención del bloque de Jordan de una matriz 2×2

Sea \mathbf{A} una matriz 2×2 cuyo único autovalor sea λ cuya multiplicidad algebraica sea 2, pero que el subespacio asociado a λ tenga dimensión 1. Sea \vec{v}_1 un autovector asociado a λ . Consideremos \vec{v}_2 un vector definido a través de la relación

$$(\mathbf{A} - \lambda I)\vec{v}_2 = \vec{v}_1$$

Debemos demostrar que

- a) Que \vec{v}_2 existe, es decir, que $\mathbf{A} - \lambda I$ tiene inversa
 b) \vec{v}_2 y \vec{v}_1 son linealmente independientes y que la matriz asociada en la base $\{\vec{v}_1, \vec{v}_2\}$ es una matriz de Jordan.

a) Para demostrar la existencia del \vec{v}_2 , tomemos un vector \vec{u} que no sea autovector de \mathbf{A} . Sea \vec{w} definido a través de $\vec{w} = (\mathbf{A} - \lambda I)\vec{u}$. Notemos que \vec{w} no es nulo, ya que \vec{u} no es un autovector.

Calculemos $(\mathbf{A} - \lambda I)\vec{w}$.

$$(\mathbf{A} - \lambda I)\vec{w} = (\mathbf{A} - \lambda I)(\mathbf{A} - \lambda I)\vec{u} = (\mathbf{A} - \lambda I)^2\vec{u}$$

Ahora bien, el polinomio característico asociado a \mathbf{A} es

$$p(x) = (x - \lambda)^2$$

El Teorema de Cayley-Hamilton establece que el polinomio característico evaluado en el operador da como resultado el operador nulo,

$$p(\mathbf{A}) = 0$$

con lo que demostramos que $(\mathbf{A} - \lambda I)\vec{w} = 0$ o lo que es lo mismo que \vec{w} es un autovector.

b) Calculemos

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 = \vec{0}$$

Aplicamos el operador $(\mathbf{A} - \lambda I)$,

$$(\mathbf{A} - \lambda I)(c_1 \vec{v}_1 + c_2 \vec{v}_2) = \vec{0}$$

entonces,

$$c_1 (\mathbf{A} - \lambda I) \vec{v}_1 + c_2 (\mathbf{A} - \lambda I) \vec{v}_2 = \vec{0}$$

Como \vec{v}_1 es un autovector, $(\mathbf{A} - \lambda I) \vec{v}_1 = \vec{0}$ y además, $(\mathbf{A} - \lambda I) \vec{v}_2 = \vec{v}_1$ entonces,

$$(\mathbf{A} - \lambda I)(c_1 \vec{v}_1 + c_2 \vec{v}_2) = c_1 \vec{0} + c_2 \vec{v}_1 = \vec{0}$$

Con lo que $c_2 = 0$ y entonces $c_1 = 0$. Entonces, son linealmente independientes.

Finalmente, obtengamos la matriz del operador, pero en la base $\{\vec{v}_1, \vec{v}_2\}$. Dado que $\mathbf{A}\vec{v}_1 = \lambda \vec{v}_1$ al escribir la matriz asociada en la base elegida tendremos que la primera columna será $\begin{bmatrix} \lambda \\ 0 \end{bmatrix}$. Dado que la definición de \vec{v}_2 es $(\mathbf{A} - \lambda I)\vec{v}_2 = \vec{v}_1$ tenemos que $\mathbf{A}\vec{v}_2 = \vec{v}_1 + \lambda \vec{v}_2$ con

lo que la segunda columna será $\begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ por lo que la matriz asociada en esta base será:

$$\mathbf{A}_{\{\vec{v}_1, \vec{v}_2\}} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

esto es, la forma de Jordan.

27.4 Esquema General de Construcción de Bloques de Jordan

Para el caso general, $n \times n$ nos ocuparemos de la construcción de la forma de Jordan, pero a través de los bloques. Es decir, consideremos una matriz \mathbf{A} cuyo polinomio característico admite una factorización

$$p(\lambda) = (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} (\lambda - \lambda_3)^{d_3} \dots (\lambda - \lambda_r)^{d_r}$$

Vamos a trabajar por separado con los autovalores.

Empecemos, por ejemplo, por el λ_1 . Si la multiplicidad geométrica coincide con la algebraica, el primer bloque simplemente es una matriz diagonal, con elementos λ_1 en la misma. Para este caso, al encontrar el subespacio asociado a λ_1 obtendremos la base y en esta base la matriz será diagonal.

Supongamos que la multiplicidad geométrica de λ_1 sea $s_1 < d_1$. La multiplicidad algebraica será un número mayor o igual que 2, caso contrario, el bloque de Jordan, será una matriz 1×1 con el autovalor en la diagonal (donde habrá un solo autovector asociado). Si $d_1 = 2$ y la multiplicidad geométrica es 1, construimos el bloque de Jordan a partir del procedimiento ya visto.

En el caso en que $d_1 = 3$ los posibles valores para la multiplicidad geométrica serán 1 o 2, ya que si coinciden la matriz de bloques de Jordan estará compuesta por 3 bloques de 1×1 . Los casos de interés serán entonces, cuando la multiplicidad geométrica de λ_1 sean 1 o 2.

Caso de multiplicidad geométrica 1. Para este caso, tendremos un solo autovector asociado a λ_1 , con lo cual, debemos construir dos vectores adicionales para formar una base del subespacio asociado con dimensión 3 (que es la multiplicidad geométrica).

Sea $\vec{v}_1^{(\lambda_1)}$ el único autovector asociado a λ_1 . Con este vector, construyamos dos vectores auxiliares $\vec{v}_2^{(\lambda_1)}$ y $\vec{v}_3^{(\lambda_1)}$ de la forma

$$(\mathbf{A} - \lambda_1 I) \vec{v}_2^{(\lambda_1)} = \vec{v}_1^{(\lambda_1)} \quad \text{y} \quad (\mathbf{A} - \lambda_1 I) \vec{v}_3^{(\lambda_1)} = \vec{v}_2^{(\lambda_1)}$$

Es simple demostrar (y se deja como ejercicio) que los tres vectores son linealmente independientes. Entonces, para este subespacio, el bloque de Jordan será:

$$B_{\lambda_1} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}$$

Caso de multiplicidad geométrica 2. En este caso, tendremos dos vectores linealmente independientes y necesitamos construir un vector adicional, linealmente independiente, para construir una base que esté asociada a este autovalor y que permita construir una matriz de bloques de Jordan para este autovalor. Consideremos $\vec{v}_1^{(\lambda_1)}$ y $\vec{v}_2^{(\lambda_1)}$ los dos vectores base del subespacio asociado a λ_1 . Aquí, la extensión a un tercer vector admite dos posibilidades: La primera, dejar $\vec{v}_1^{(\lambda_1)}$ y construir el $\vec{v}_3^{(\lambda_1)}$ a partir de

$$(\mathbf{A} - \lambda_1 I) \vec{v}_3^{(\lambda_1)} = \vec{v}_2^{(\lambda_1)}$$

La independencia lineal se prueba de la misma manera que hicimos para el caso 2×2 . De esta manera, la matriz de bloque de Jordan asociada a λ_1 queda

$$B_{\lambda_1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}$$

El otro caso posible era construir el $\vec{v}_3^{\lambda_1}$ con el $\vec{v}_1^{\lambda_1}$ de la forma $(\mathbf{A} - \lambda_1 \mathbf{I})\vec{v}_3^{(\lambda_1)} = \vec{v}_1^{(\lambda_1)}$ pero esta elección nos obliga a reordenar los vectores en la forma $\{\vec{v}_1^{(\lambda_1)}, \vec{v}_3^{(\lambda_1)}, \vec{v}_2^{(\lambda_1)}\}$ y la matriz de bloques toma la forma

$$B_{\lambda_1} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$$

Este es el mecanismo a partir del cual obtenemos las matrices de bloques de Jordan. Debemos tener en cuenta las multiplicidades geométricas y algebraicas para las extensiones de las bases de manera tal de tener un número de vectores base coincidentes con la multiplicidad algebraica.

Una vez culminado el procedimiento para el autovalor λ_1 pasamos al autovalor λ_2 .

Algoritmo general. Consideremos un autovalor con multiplicidad geométrica d , con $d > 3$. La multiplicidad geométrica que nos interesará será 1, 2, 3 hasta $d - 1$, ya que coincidente con la multiplicidad geométrica carece de interés y estamos en el caso de bloque diagonal.

Multiplicidad 1. En este caso, tenemos que construir a partir de un solo vector vectores que completen una base cuya cantidad de elementos coincide con la multiplicidad algebraica. Entonces, tendremos un solo autovector con el que comenzamos el procedimiento

$$\begin{aligned} (\mathbf{A} - \lambda \mathbf{I})\vec{v}_2 &= \vec{v}_1 && \rightarrow && \text{obtenemos el } \vec{v}_2 \\ (\mathbf{A} - \lambda \mathbf{I})\vec{v}_3 &= \vec{v}_2 && \rightarrow && \text{obtenemos el } \vec{v}_3 \\ &\vdots &= &\vdots && \\ (\mathbf{A} - \lambda \mathbf{I})\vec{v}_d &= \vec{v}_{d-1} && \rightarrow && \text{obtenemos el } \vec{v}_d \end{aligned}$$

El procedimiento concluye cuando construimos $d - 1$ vectores adicionales y completamos una base de d vectores.

En este caso, el bloque queda

$$B_\lambda = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

Multiplicidad 2. Para este caso, existen múltiples posibilidades. En todas las posibilidades tendremos dos bloques de Jordan, pero pueden tener diferentes tamaños. Uno, por ejemplo, es considerar un bloque de tamaño 1 para el primer autovector, y construir con el

algoritmo presentado los vectores restantes, de esta manera, tendremos una matriz de bloques

$$B_\lambda = \begin{bmatrix} \lambda & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

Otra posibilidad es con el primer autovector construir otro vector a partir del procedimiento y con el otro autovector construimos los restantes. De esta manera, la matriz de bloques tendrá la forma

$$B_\lambda = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

De esta manera, siempre habrá dos bloques, pero en función de cómo se extienda la base tendremos diferentes tamaños para los mismos.

En resumen, la cantidad de bloques coincidirá con la multiplicidad algebraica del autovalor.