

Facultad de Ciencias Astronómicas y Geofísicas
Universidad Nacional de La Plata



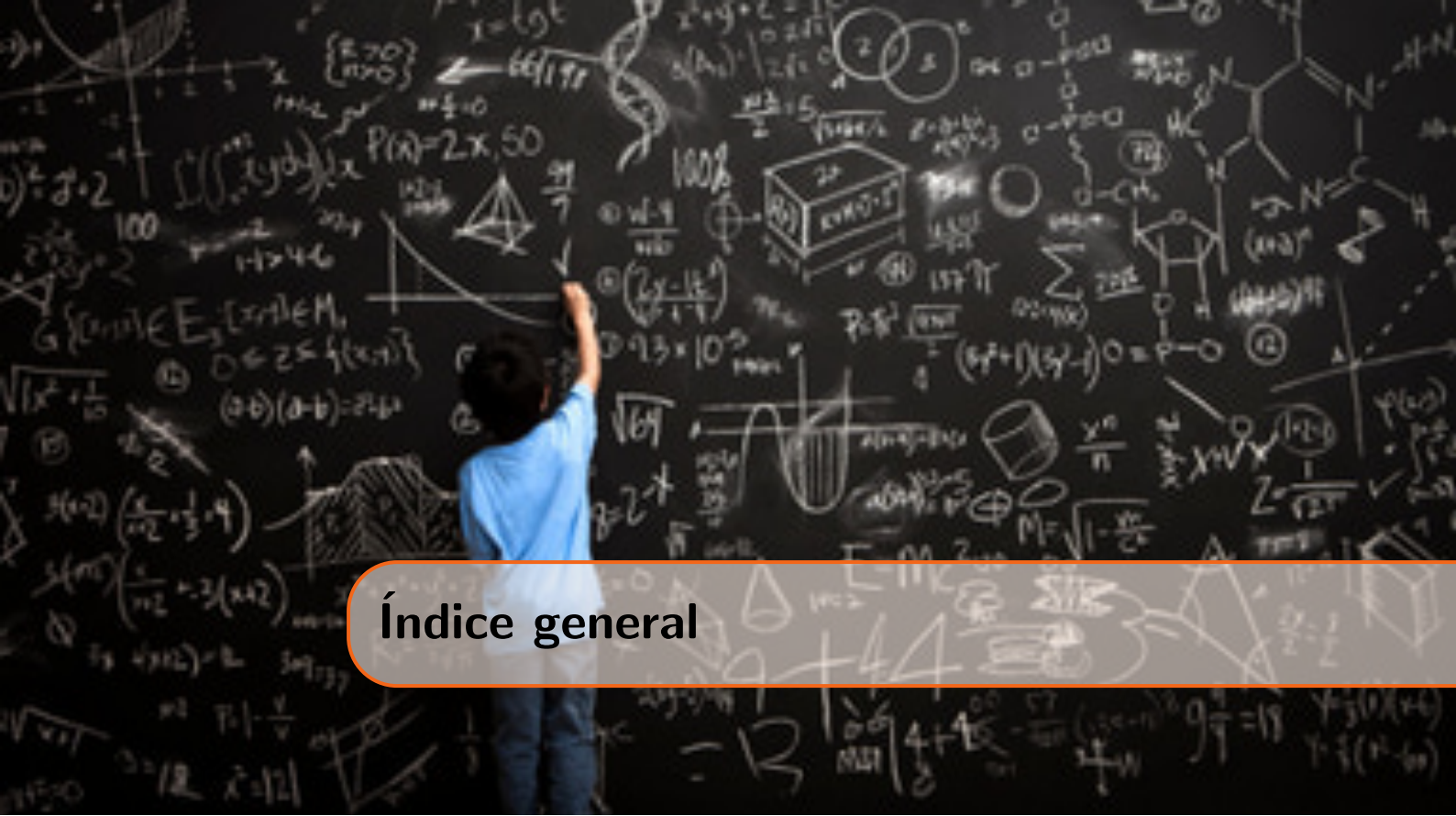
Matemáticas Avanzadas

Apuntes de Álgebra Lineal

Profesor
O.I. Miloni

Jefe de Trabajos Prácticos
N.P. Maffione

Versión preliminar

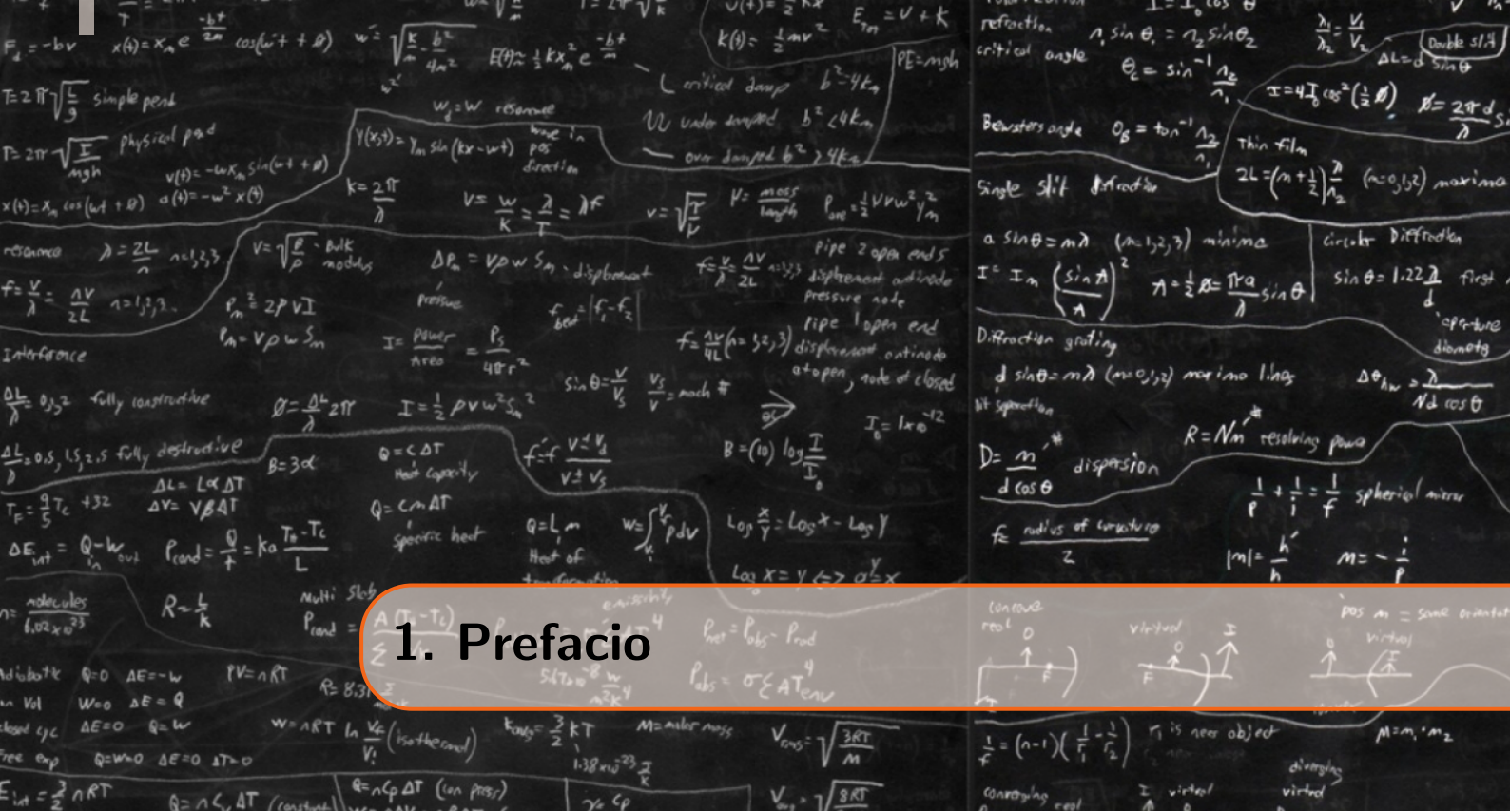


Índice general

1	Prefacio	5
2	Espacios vectoriales	7
2.1	Introducción a los espacios vectoriales	7
2.1.1	Subespacios vectoriales	9
2.1.2	Base y coordenadas. Dimensión	9
2.1.3	Invarianza y representación	11
2.1.4	Convención de Einstein para la sumatoria	12
3	Transformaciones lineales	15
3.1	Introducción a las transformaciones lineales	15
3.1.1	Matriz asociada a una transformación lineal	16
3.1.2	Cambio de base y cambio de coordenadas	18
3.1.3	El espacio de las transformaciones lineales	20
4	Formas Lineales y Espacio Dual	23
4.1	Funcionales lineales: 1-formas	23
4.1.1	El espacio doble dual	26
4.1.2	Coordenadas de vectores y de 1-formas	27
4.1.3	Aspectos metodológicos. Ejemplo: \mathbb{R}^n	28
4.1.4	Cambio de coordenadas para 1-formas	29
4.1.5	Resumen de cambio de base y coordenadas	29
5	Formas bilineales y multilineales. Tensores	31
5.1	Funcionales bilineales sobre V	31
5.1.1	Álgebra de las formas bilineales	32

5.2	Producto tensorial de 1–formas	32
5.2.1	Coordenadas de un tensor	33
5.3	Tensores cartesianos en general	35
5.3.1	Covarianza y contravarianza de un tensor	36
5.3.2	Cambio de coordenadas en tensores cartesianos	36
5.4	Componentes de 1–formas y tensores. Caso \mathbb{R}^3	38
5.4.1	Relación entre componentes y coordenadas de tensores	39
5.4.2	Aspectos metodológicos	41
5.5	Formas cuadráticas	41
6	Espacios euclídeos y espacios métricos	43
6.1	Producto interno	43
6.1.1	Axiomas de producto interno	43
6.1.2	Norma o módulo de un vector	44
6.2	Ortogonalidad	45
6.2.1	Construcción de una base ortogonal	46
6.2.2	Base ortogonal. Coeficientes de Fourier	47
6.3	El producto interno como un tensor. El tensor métrico	48
6.3.1	Aplicación: longitud de arco	48
6.4	Coordenadas covariantes de un vector contravariante.	49
7	Análisis tensorial. Operadores diferenciales	51
7.1	Introducción	51
7.2	Cálculo operacional	51
7.2.1	Multiplicación de tensores contravariantes	51
7.2.2	Derivación de tensores	52
7.3	Derivación covariante	55
7.3.1	Símbolos de Christoffel	55
7.3.2	Derivada de un vector contravariante	56
7.3.3	Derivada de un vector covariante	57
7.3.4	Derivada de un tensor dos veces contravariante	58
7.3.5	Derivada de un tensor dos veces covariante	59
7.4	Operadores diferenciales	59
7.4.1	Gradiente	60
7.4.2	Rotor	61
7.4.3	Divergencia	61
7.4.4	Laplaciano	63
7.5	Aplicaciones a la teoría de curvas	64
7.5.1	Campo de tensores sobre curvas. Derivada	65
7.5.2	Transporte paralelo	66
7.5.3	Geodésicas	66

7.6	Elementos de Geometría Riemanniana	67
7.6.1	Tensor de curvatura de Riemann	67
7.6.2	Tensor de Ricci	67
7.6.3	Escalar de Ricci	67
8	Tensores en Física	69
8.1	Energía cinética	69
8.1.1	Energía cinética de una partícula	69
8.1.2	Energía cinética de un cuerpo rígido. Tensor de inercia	70
8.2	Elasticidad. Tensor de deformación	72
8.3	Desarrollo del potencial electrostático. Tensor momento cuadrupolar	73
9	Autovalores y autovectores	77
9.1	Introducción. Definiciones	77
9.2	El polinomio característico	83
9.2.1	Multiplicidad de las raíces	83
9.3	Operador Adjunto	84
9.3.1	Matrices hermíticas. Matrices simétricas	84
9.3.2	Aspecto metodológico para la diagonalización de un operador simétrico.	85
9.4	Operadores unitarios	86
10	Formas canónicas	89
10.1	Forma canónica de Jordan	89
10.1.1	Matriz de bloque de Jordan	89
10.1.2	Matriz de Jordan	90
10.1.3	Relación entre la multiplicidad geométrica y los bloques	90
10.1.4	Esquema general de construcción de bloques de Jordan	92
	Bibliografía	95
	Libros	95
	Índice	97



1. Prefacio

Este material está dedicado a la primer unidad temática de la disciplina *Matemáticas Avanzadas* 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 Ciencias Astronómicas y Geofísicas de la Universidad Nacional de La Plata.

Matemáticas Avanzadas es una materia compuesta por las siguientes unidades temáticas:

- Álgebra Lineal:** Espacios Vectoriales. Subespacios. Base. Transformaciones lineales. Cambio de base. Álgebra de transformaciones lineales. Formas canónicas de transformaciones lineales. Álgebra 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 la 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.

Dado el enfoque tensorial que pretendemos darle a la unidad temática Álgebra Lineal, hemos decidido elaborar el siguiente material de guía para el estudio de esta primera parte.

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.

Octavio Miloni
Profesor de Matemáticas Avanzadas
4 de Abril de 2015

2. Espacios vectoriales

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

2.1 Introducción a los espacios vectoriales

Definición 2.1.1 — Espacio vectorial. Dada una estructura compuesta por un conjunto V de elementos de naturaleza arbitraria (a los que llamaremos *vectores*, $\{\vec{v}, \vec{w}, \vec{u} \dots\}$), y un conjunto numérico que sea cuerpo K (a los que en este contexto llamaremos *escalares*, $\{\lambda, \mu \dots\}$). Si podemos definir dos operaciones, de adición entre elementos de V , y de producto entre elementos de V y elementos de K , que satisfagan los siguientes axiomas:

- sobre la adición entre elementos de V :
 1. **conmutativa**: $\vec{v} \oplus \vec{w} = \vec{w} \oplus \vec{v}$,
 2. **asociativa**: $(\vec{v} \oplus \vec{w}) \oplus \vec{u} = \vec{v} \oplus (\vec{w} \oplus \vec{u})$,
 3. que exista el **elemento identidad**, denotado por $\vec{0} \in V$ (y llamado nulo o neutro), tal que: $\vec{v} \oplus \vec{0} = \vec{v}, \forall \vec{v} \in V$,
 4. que exista el **elemento inverso aditivo**, denotado por $\vec{v}' \in V$ (y llamado opuesto), tal que: $\vec{v} \oplus \vec{v}' = \vec{0}, \forall \vec{v} \in V$;
- y sobre el producto entre elementos de V y elementos de K :
 1. **compatible** con el cuerpo K de escalares: $\lambda \odot (\mu \odot \vec{v}) = (\lambda \mu) \odot \vec{v}$,
 2. **distributivo respecto a la suma de escalares**: $(\lambda + \mu) \odot \vec{v} = \lambda \odot \vec{v} \oplus \mu \odot \vec{v}$,
 3. **distributivo respecto a la suma de vectores**: $\lambda \odot (\vec{v} \oplus \vec{w}) = \lambda \odot \vec{v} \oplus \lambda \odot \vec{w}$,
 4. que exista el **elemento identidad**, denotado por $1 \in K$ (y llamado unidad) no nulo y único, tal que: $1 \odot \vec{v} = \vec{v}$,

la estructura algebraica constituida por la 4-upla (V, K, \oplus, \odot) se denomina *Espacio Vectorial* sobre el cuerpo K . Dado que las operaciones de adición de vectores y de producto por escalares son operaciones binarias, gozan de la propiedad de clausura^a, luego, además de los axiomas antes establecidos, el espacio vectorial responde a las leyes de cierre sobre:

- la adición de vectores: $\vec{v} \in V$ y $\vec{w} \in V$ entonces $\vec{v} \oplus \vec{w} \in V$,
- el producto por escalares: $\lambda \in K$ y $\vec{v} \in V$ entonces $\lambda \odot \vec{v} \in V$.

^aEsto se debe a que la adición es una operación binaria, i.e. trabaja sobre dominios y codominios que son subconjuntos del mismo conjunto, mientras que el producto por escalares es una operación binaria externa, donde si bien los dominios pertenecen a conjuntos distintos, el codominio es un subconjunto de uno de los conjuntos a los que pertenecen los dominios.

■ **Ejemplo 2.1 — Espacios vectoriales.** Sea V el conjunto de las ternas (x, y, z) de números reales¹ para las cuales definimos una suma entre vectores como:

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

y un producto por números reales como:

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

entonces la 4-upla (V, K, \oplus, \odot) 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:

$$\begin{aligned} \vec{v}_1 \oplus \vec{v}_2 &= (x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= (x_2 + x_1, y_2 + y_1, z_2 + z_1) \\ &= \vec{v}_2 \oplus \vec{v}_1 \text{ y la suma es conmutativa.} \end{aligned}$$

$$\begin{aligned} \vec{v}_1 \oplus (\vec{v}_2 \oplus \vec{v}_3) &= \\ &= (x_1, y_1, z_1) \oplus (x_2 + x_3, y_2 + y_3, z_2 + z_3) \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3), z_1 + (z_2 + z_3)) \\ &= ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3, (z_1 + z_2) + z_3) \\ &= (\vec{v}_1 \oplus \vec{v}_2) \oplus \vec{v}_3 \text{ y la suma es asociativa.} \end{aligned}$$

Si consideramos el elemento $(0, 0, 0)$ podemos notar que $(x, y, z) \oplus (0, 0, 0) = (x + 0, y + 0, z + 0) = (x, y, z)$ y la suma tiene un **elemento nulo**. Para la operación de suma entre vectores nos faltaría demostrar que para todo elemento existe un **opuesto**. En efecto, $(x, y, z) \oplus (x', y', z') = (0, 0, 0)$ entonces $x' = -x$, $y' = -y$ y $z' = -z$, esto es, existe siempre un elemento opuesto.

Sean λ y μ números reales, tenemos que:

$$\lambda \odot (\mu \odot (x, y, z)) = \lambda \odot (\mu x, \mu y, \mu z) = (\lambda \mu x, \lambda \mu y, \lambda \mu z) = (\lambda \mu) \odot (x, y, z),$$

lo que hace cumplir el primer axioma (compatibilidad con el cuerpo K) para el producto entre vectores y escalares.

Consideremos ahora:

¹Al ser los elementos del conjunto V , ternas de números reales (debido a que el cuerpo sobre el cual se define el espacio vectorial es el de los números reales), en las demostraciones uno apela, a fin de cuentas, a esta estructura más básica, i.e. el cuerpo de los números reales.

$$(\lambda + \mu) \odot (x, y, z) = ((\lambda + \mu)x, (\lambda + \mu)y, (\lambda + \mu)z) = (\lambda x + \mu x, \lambda y + \mu y, \lambda z + \mu z) = \lambda \odot (x, y, z) + \mu \odot (x, y, z),$$

lo que hace cumplir el segundo axioma.

Además:

$$\begin{aligned} \lambda \odot ((x, y, z) \oplus (x_1, y_1, z_1)) &= \lambda \odot (x + x_1, y + y_1, z + z_1) \rightarrow \\ (\lambda(x + x_1), \lambda(y + y_1), \lambda(z + z_1)) &= \lambda \odot (x_1, y_1, z_1) \oplus \lambda \odot (x, y, z), \end{aligned}$$

lo que hace cumplir el tercer axioma.

Finalmente, para el caso del producto por un número tenemos que, a partir de la definición, $1 \odot (x, y, z) = (1x, 1y, 1z)$, lo que hace cumplir el cuarto y último axioma para la operación de producto entre vectores y escalares. ■

Como vemos, la comprobación de que cierta estructura es un espacio vectorial consiste en ir comprobando cada uno de sus axiomas, usando tanto las definiciones para la suma entre vectores como para el producto entre vectores y escalares.

Ejercicio 2.1 Sea V el conjunto de funciones continuas en el intervalo $[a, b]$ donde definimos una suma $f \oplus g$ como $(f \oplus g)(x) = f(x) + g(x)$ y un producto con números: $\lambda \odot f$, definido como $(\lambda \odot f)(x) = \lambda f(x)$. Demostrar que la 4-upla (V, K, \oplus, \odot) tiene estructura de espacio vectorial. ■

2.1.1 Subespacios vectoriales

Dados los espacios vectoriales (V, K, \oplus, \odot) y (W, K, \oplus, \odot) :

Teorema 2.1.1 — Subespacio. Sea $W \subset V$ tal que $\vec{v}, \vec{w} \in W$, si:

- $\lambda \odot \vec{v} \in W$,
- $\vec{v} \oplus \vec{w} \in W$,

entonces la 4-upla (W, K, \oplus, \odot) es un subespacio de (V, K, \oplus, \odot) .

2.1.2 Base y coordenadas. Dimensión

En el Ejemplo 2.1 efectuamos una comprobación de que las definiciones de suma entre vectores y producto por números dotan a la 4-upla (V, K, \oplus, \odot) (donde V es el 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) (como elemento \vec{v} del espacio vectorial (V, K, \oplus, \odot)), a la que descomponemos de la siguiente manera de acuerdo a las definiciones de suma entre vectores y producto por números que ya hemos introducido:

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

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 $\mathbf{e}_1 = (1, 1, 1)$, $\mathbf{e}_2 = (1, -1, 0)$ y $\mathbf{e}_3 = (0, 0, 1)$. Notemos que el mismo vector (x, y, z) puede ser escrito como:

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

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

Definición 2.1.2 — Sistema de generadores. Sea la 4-upla (V, K, \oplus, \odot) 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 $\vec{v} \in V$ existen números reales $\lambda^1, \lambda^2, \dots, \lambda^n$ tales que:

$$\vec{v} = \lambda^1 \odot \mathbf{e}_1 \oplus \lambda^2 \odot \mathbf{e}_2 \oplus \dots \oplus \lambda^n \odot \mathbf{e}_n.$$

Mostramos que, dado un espacio vectorial, el sistema de generadores para el conjunto de vectores que integra dicho espacio no es único. Sin embargo, vimos que el número de generadores pareciera que sí lo es. 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 “0” son el vector nulo, la definición de sistema de generadores sigue siendo satisfecha. Luego, a la idea de sistema de generadores vamos a incorporarle otra idea, en la que se determine un criterio para el número mínimo de generadores necesarios para escribir cualquier vector del conjunto V que integre el espacio vectorial.

Definición 2.1.3 — Independencia lineal. Un conjunto de vectores $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ son linealmente independientes si $\sum_{\mu=1}^n \lambda^\mu \vec{v}_\mu = \vec{0} \rightarrow \lambda^\mu = 0, \forall \mu = 1, 2, \dots, n$.

Definición 2.1.4 — Sistema de generadores linealmente independientes. Sea la 4-upla (V, K, \oplus, \odot) un espacio vectorial sobre los números reales. Sea $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ un conjunto de generadores de V . Se dice que los vectores que conforman el sistema de generadores son linealmente independientes si:

$$\vec{0} = \lambda^1 \odot \mathbf{e}_1 \oplus \lambda^2 \odot \mathbf{e}_2 \oplus \dots \oplus \lambda^n \odot \mathbf{e}_n$$

implica que cada número λ^j (con $j = 1, 2, \dots, n$) debe ser cero.

Esta definición provee un criterio de *filtrado* de los vectores que no sean necesarios para la generación de vectores. Con estas dos definiciones previas, establecemos las siguientes dos:

Definición 2.1.5 — Base. Sea la 4–upla (V, K, \oplus, \odot) 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 de vectores forman una base de V si son un sistema de generadores linealmente independientes.

Definición 2.1.6 — Dimensión. Sea la 4–upla (V, K, \oplus, \odot) un espacio vectorial sobre los números reales. Si una base de V posee n elementos, diremos que la dimensión de V es n .

Hemos visto en lo precedente que un mismo vector $\vec{v} = (x, y, z)$ admitía al menos dos representaciones. Esto es, en la base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{(1, 0, 0); (0, 1, 0); (0, 0, 1)\}$ lo escribimos como:

$$\vec{v} = x \odot (1, 0, 0) \oplus y \odot (0, 1, 0) \oplus z \odot (0, 0, 1),$$

y en la base $\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\} = \{(1, 1, 1); (1, -1, 0); (0, 0, 1)\}$, como:

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

Como ente algebraico, el vector \vec{v} es único, absoluto: un **invariante**, lo que cambia es la representación según la base (lo que llamaremos *coordenadas*).

Vamos a definir como *coordenadas* de un vector a los coeficientes de los generadores en la base escogida para el conjunto V del espacio vectorial. De esta manera, en la primera base (denominada **canónica** y en la que los coeficientes coinciden numéricamente con los valores de los elementos reales que constituyen la terna que define al vector \vec{v}) las coordenadas son simplemente los números x, y, z , en cambio en la segunda base, las coordenadas serán:

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

En los textos clásicos de cálculo tensorial tales como [L73], 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.

2.1.3 Invarianza y representación

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

Lo que se intenta aclarar es lo siguiente: un objeto de \mathbb{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 \mathbb{R}^3 consideramos las bases $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{(1, 0, 0); (0, 1, 0); (0, 0, 1)\}$ y $\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\} = \{(\frac{1}{2}, 0, 0); (0, 0, -2); (0, -1, 0)\}$, el vector $(1, 2, 3)$ se representará como $1 \odot \mathbf{e}_1 \oplus 2 \odot \mathbf{e}_2 \oplus 3 \odot \mathbf{e}_3$ como así también por $2 \odot \mathbf{e}_1' \oplus -\frac{3}{2} \odot \mathbf{e}_2' \oplus -2 \odot \mathbf{e}_3'$, es decir:

$$1 \odot \mathbf{e}_1 \oplus 2 \odot \mathbf{e}_2 \oplus 3 \odot \mathbf{e}_3 = 2 \odot \mathbf{e}_1' \oplus -\frac{3}{2} \odot \mathbf{e}_2' \oplus -2 \odot \mathbf{e}_3' = (1, 2, 3).$$

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

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

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:

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

con λ^j y β^i escalares (del cuerpo K) a determinar.

2.1.4 Convención de Einstein para la sumatoria

Al representar un vector de un conjunto V de un espacio vectorial de dimensión n en una determinada base, debemos necesariamente utilizar el símbolo de sumatoria, i.e.:

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

donde como ya hemos mencionado en la sección anterior, esta suma es un invariante.

Proposición 2.1.2 — Condiciones para la notación de Einstein. Einstein propuso un criterio para suprimir los símbolos de sumatoria (Σ), siempre y cuando se satisfagan las siguientes condiciones:

1. Aparezcan dos índices repetidos;
2. los índices repetidos sean uno superior y otro inferior.

Entonces, por ejemplo:

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

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

En cambio, el término a_{ii} es simplemente un único término de la diagonal.

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

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

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

es un elemento absoluto del conjunto V que integra la estructura de 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 dentro de este marco contienen muchas sumatorias de estas características. Más adelante se volverá en este punto.

■ **Ejemplo 2.2 — Multiplicación de matrices.** Sean $\mathbf{A} \in \mathbb{R}^{n \times m}$ y $\mathbf{B} \in \mathbb{R}^{m \times l}$. El elemento c_j^i (donde el supraíndice indica fila con $i \in [1, n]$ y el subíndice indica la columna, con $j \in [1, l]$) del producto $\mathbf{A} \times \mathbf{B}$ 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

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

3.1 Introducción a las transformaciones lineales

Ahora estudiaremos los conceptos correspondientes a transformaciones lineales y las matrices asociadas a sus representaciones.

Definición 3.1.1 — Transformación lineal. 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:

- $T(\vec{v}_1 \oplus \vec{v}_2) = T(\vec{v}_1) \oplus T(\vec{v}_2)$,
- $T(\lambda \odot \vec{v}_1) = \lambda \odot T(\vec{v}_1)$.

En las expresiones de la izquierda de la igualdad, las operaciones binarias (\oplus, \odot) son sobre la 4-upla (V, K, \oplus, \odot) (con V el conjunto dominio de T), cuando en las de la derecha, las mismas operaciones son sobre la 4-upla (W, K, \oplus, \odot) (con W el conjunto imagen o codominio).

Definición 3.1.2 — Núcleo. El núcleo de T está definido como:

$$\text{Nu}(T) = \{\vec{v} \in V : T(\vec{v}) = \vec{0}\}.$$

Definición 3.1.3 — Imagen. El imagen de T está definida como:

$$\text{Im}(T) = \{\vec{w} \in W : \exists \vec{v} \in V, T(\vec{v}) = \vec{w}\}.$$

Observaciones respecto de las transformaciones lineales:

1. el nulo del dominio está en el núcleo de T : $T(\vec{0}_V) = \vec{0}_W$,
2. el núcleo de T es un subespacio de (V, K, \oplus, \odot) ,
3. la imagen de T es un subespacio de (W, K, \oplus, \odot) .

■ **Ejemplo 3.1 — Transformaciones lineales.** Sea $V = \mathbb{R}^3$ y $W = \mathbb{R}^2$. Sea T una transformación definida como $T(x, y, z) = (x + y, y - z)$ y $\vec{v}_1 = (x, y, z)$, $\vec{v}_2 = (x', y', z') \in V$. Veamos que:

$$\begin{aligned} T(\vec{v}_1 \oplus \vec{v}_2) &= T(x + x', y + y', z + z') &= ([x + x'] + [y + y'], [y + y'] - [z + z']) \\ &= (x + x' + y + y', y + y' - z - z') \\ &= ([x + y] + [x' + y'], [y - z] + [y' - z']) \\ &= (x + y, y - z) \oplus (x' + y', y' - z') \\ &= T(x, y, z) \oplus T(x', y', z') = T(\vec{v}_1) \oplus T(\vec{v}_2), \end{aligned}$$

y

$$\begin{aligned} T(\lambda \odot \vec{v}_1) &= T(\lambda x, \lambda y, \lambda z) &= (\lambda x + \lambda y, \lambda y - \lambda z) \\ &= (\lambda(x + y), \lambda(y - z)) \\ &= \lambda(x + y, y - z) \\ &= \lambda \odot T(x, y, z) = \lambda \odot T(\vec{v}_1). \end{aligned}$$

Luego, como T satisface las condiciones antes expuestas, la transformación T es lineal. ■

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

■ **Ejemplo 3.2 — Núcleo e imagen.** Sea $V = \mathbb{R}^2$ y $W = \mathbb{R}^3$ y T una transformación $T: V \rightarrow W$ definida como $T(x, y) = (2x, x + y, x - 2y)$. Hallar el núcleo y la imagen de T .

Para encontrar el núcleo se debe cumplir que $(2x, x + y, x - 2y) = \vec{0}_W$, luego $x, y = 0$, entonces $\text{Nu}(T) = \{\vec{0}_V\}$. Por otro lado, notemos que si un vector de \mathbb{R}^3 está en la imagen, entonces es de la forma: $(2x, x + y, x - 2y)$ que puede descomponerse como $x \odot (2, 1, 1) \oplus y \odot (0, 1, -2)$. Luego como $(2, 1, 1)$ y $(0, 1, -2)$ son linealmente independientes, forman una base de la imagen y esta tiene dimensión 2. ■

3.1.1 Matriz asociada a una transformación lineal

Consideremos dos espacios vectoriales V y W de dimensión n y m , respectivamente. Sean $\mathcal{B}_V = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ y $\mathcal{B}_W = \{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_m\}$ bases de V y W , respectivamente.

Apliquemos la transformación lineal T a un vector arbitrario $\vec{v} \in V$. Como tenemos una base de V , podemos escribir $\vec{v} = \lambda^i \odot \mathbf{e}_i$, donde, por definición, los números $\lambda^1, \lambda^2, \dots, \lambda^n$ son las coordenadas del vector en la base \mathcal{B}_V dada. Entonces, aplicando la transformación T tenemos $T(\vec{v}) = T(\lambda^i \odot \mathbf{e}_i)$, y como la transformación es lineal tendremos: $T(\vec{v}) = \lambda^i \odot T(\mathbf{e}_i)$. Ahora bien, como la transformación T aplica un elemento de V en un elemento de W , cada transformado debe poder reescribirse como una combinación lineal de los elementos de la base de W . Esto significa que $T(\vec{v}) = \beta^j \odot \mathbf{e}'_j$. Por otro lado, cada transformado de los elementos de la base de V (generadores) poseerán a su vez un desarrollo en la base de W : $T(\mathbf{e}_i) = [T]_i^j \odot \mathbf{e}'_j$, donde los números $[T]_i^j$ 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 (el inferior) asociado al elemento de la base del espacio de partida que estamos transformando 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 como:

$$\begin{aligned} T(\vec{v}) &= \lambda^i \odot T(\mathbf{e}_i) \\ &= \lambda^i [T]_i^j \odot \mathbf{e}_j' \\ &= [T]_i^j \lambda^i \odot \mathbf{e}_j'. \end{aligned}$$

Entonces, tenemos que $\beta^j \odot \mathbf{e}_j' = [T]_i^j \lambda^i \odot \mathbf{e}_j'$ o, lo que es equivalente $[\beta^j - [T]_i^j \lambda^i] \odot \mathbf{e}_j' = 0$ y como los vectores \mathbf{e}_j' son base, son linealmente independientes, con lo que $[\beta^j - [T]_i^j \lambda^i] = 0$ o, lo que es lo mismo:

$$[T]_i^j \lambda^i = \beta^j. \quad (3.1)$$

Esta última expresión relaciona las coordenadas del vector a transformar en la base \mathcal{B}_V de V , i.e. λ^i , con las coordenadas del vector transformado en la base \mathcal{B}_W de W , i.e. β^j . Más aún, la relación viene dada a partir de un producto de matrices, de la forma:

$$\begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^m \end{bmatrix} = \begin{bmatrix} [T]_1^1 & [T]_2^1 & \cdots & [T]_n^1 \\ [T]_1^2 & [T]_2^2 & \cdots & [T]_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ [T]_1^m & [T]_2^m & \cdots & [T]_n^m \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix}$$

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.

Entonces, la matriz asociada a la representación de la transformación lineal se construye de la siguiente manera:

1. Se aplica la transformación lineal T al primer elemento de la base de partida: $T(\mathbf{e}_1)$;
2. se escribe el vector transformado en la base de llegada W : $T(\mathbf{e}_1) = [T]_1^j \odot \mathbf{e}_j'$;
3. las coordenadas del transformado de \mathbf{e}_1 , $[T]_1^j$ forman 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 $m \times n$ (m filas y n columnas).

R Recordemos que según la convención que hemos mencionado, el elemento $[T]_i^j$ de una matriz, estará posicionados en la fila i -ésima y la columna j -ésima.

■ **Ejemplo 3.3 — Matriz asociada a una transformación lineal.** Sea $T: \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^3$ definida explícitamente por:

$$T \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = (2a - b, b - c, 2c - d).$$

Hallar la matriz asociada a la transformación en la base canónica para cada espacio.

Aplicamos entonces la transformación lineal T sobre cada elemento de la base canónica de $\mathbb{R}^{2 \times 2}$:

$$\begin{aligned} T \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] &= (2, 0, 0) = 2 \odot (1, 0, 0), \\ T \left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right] &= (-1, 1, 0) = -1 \odot (1, 0, 0) \oplus 1 \odot (0, 1, 0), \\ T \left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right] &= (0, -1, 2) = -1 \odot (0, 1, 0) \oplus 2 \odot (0, 0, 1), \\ T \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] &= (0, 0, -1) = -1 \odot (0, 0, 1), \end{aligned}$$

con lo cual, agrupando las coordenadas en la base canónica del espacio de llegada (\mathbb{R}^3) en formato columna, tendremos que la matriz asociada a la transformación lineal T se puede escribir como:

$$[T] = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 & -1 \end{bmatrix}$$

■

3.1.2 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 mismo espacio. Para determinar el procedimiento de cambio de base y el consecuente cambio de coordenadas podemos pensar el asunto de la siguiente manera: sea 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 V . 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](#).

Entonces, aplicando la transformación identidad, tendremos:

$$\begin{aligned} I_d(\mathbf{e}_1) &= \mathbf{e}_1 = [I_d]_1^\ell \odot \mathbf{e}'_\ell \\ I_d(\mathbf{e}_2) &= \mathbf{e}_2 = [I_d]_2^\ell \odot \mathbf{e}'_\ell \\ &\vdots = \vdots = \vdots \\ I_d(\mathbf{e}_n) &= \mathbf{e}_n = [I_d]_n^\ell \odot \mathbf{e}'_\ell. \end{aligned}$$

Y la matriz asociada a esta transformación es:

$$\begin{bmatrix} [I_d]_1^1 & [I_d]_2^1 & [I_d]_3^1 & \cdots & [I_d]_n^1 \\ [I_d]_1^2 & [I_d]_2^2 & [I_d]_3^2 & \cdots & [I_d]_n^2 \\ [I_d]_1^3 & [I_d]_2^3 & [I_d]_3^3 & \cdots & [I_d]_n^3 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ [I_d]_1^n & [I_d]_2^n & [I_d]_3^n & \cdots & [I_d]_n^n \end{bmatrix}$$

Luego, dado un vector determinado en la base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ que tiene por representación $\vec{v} = \lambda^\ell \mathbf{e}_\ell$, la relación a través de la transformación lineal identidad establece

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:

$$\begin{bmatrix} \beta^1 \\ \beta^2 \\ \vdots \\ \beta^n \end{bmatrix} = \begin{bmatrix} [I_d]_1^1 & [I_d]_2^1 & [I_d]_3^1 & \cdots & [I_d]_n^1 \\ [I_d]_1^2 & [I_d]_2^2 & [I_d]_3^2 & \cdots & [I_d]_n^2 \\ [I_d]_1^3 & [I_d]_2^3 & [I_d]_3^3 & \cdots & [I_d]_n^3 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ [I_d]_1^n & [I_d]_2^n & [I_d]_3^n & \cdots & [I_d]_n^n \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^n \end{bmatrix}$$

o, en coordenadas:

$$\beta^\mu = [I_d]_\nu^\mu \lambda^\nu,$$

que no es más que la expresión dada por la Eq. (3.1) cuando $T \equiv I_d$.

En ocasiones, cuando el espacio es \mathbb{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 $\{x'^1, x'^2, \dots, x'^n\}$ (donde se asume una determinada base $\mathcal{B}' = \{\mathbf{e}_1', \mathbf{e}_2', \dots, \mathbf{e}_n'\}$) viene dado a través de $x'^\mu = [I_d]_\nu^\mu x^\nu$.

Es más, si tenemos un cambio de coordenadas, que relaciona las viejas con las nuevas coordenadas, el cambio de base viene dado a través de $\mathbf{e}_\nu = [I_d]_\mu^\nu \odot \mathbf{e}'_\mu$, es decir, que se utilizan las mismas cantidades $[I_d]_\nu^\mu$ pero no es estrictamente un producto de matrices (en este último caso estamos multiplicando escalares, $[I_d]_\nu^\mu$, por vectores de la base nueva, \mathbf{e}'_μ , para obtener un vector de la base vieja, \mathbf{e}_ν).

Notemos, además, que el cambio de las viejas a las nuevas coordenadas es similar al cambio de base, pero de los nuevos a los viejos vectores de la base.

R Si tuviéramos un cambio de los viejos a los nuevos vectores de la base: $\mathbf{e}'_\nu = \Phi_\nu^\mu \odot \mathbf{e}_\mu$, podríamos aplicar que $\mathbf{e}'_\nu = \Phi_\nu^\mu [I_d]_\mu^\lambda \odot \mathbf{e}'_\lambda$. Entonces, se debe cumplir que $\Phi_\nu^\mu [I_d]_\mu^\lambda = \delta_\nu^\lambda$, lo que implica que la matriz Φ es la matriz inversa de la matriz \mathbf{I}_d debido a que la expresión $\Phi_\nu^\mu [I_d]_\mu^\lambda = \delta_\nu^\lambda$ indica un producto de matrices.

Resumen. Esquema de transformación de coordenadas.

Descripción. Paso a paso:

1. El cambio de coordenadas: $x'^\mu = [I_d]_\nu^\mu x^\nu$, es un producto de matrices;
2. el cambio de base asociado será $\mathbf{e}'_\nu = [I_d^{-1}]_\nu^\mu \odot \mathbf{e}_\mu$. Notemos que este cálculo no es un producto de matrices;
3. la relación entre la base nueva y la base vieja es $\mathbf{e}_\nu = [I_d]_\mu^\nu \odot \mathbf{e}'_\mu$. Notemos que esta relación es muy parecida al cambio de coordenadas, sólo que esta obtiene los vectores de la base vieja en función de aquéllos de la base nueva, cuando en el cambio de coordenadas es al revés, se obtienen las coordenadas en la base nueva, en función de las coordenadas en la base vieja.

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

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

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 $\mathbf{e}_\nu = [I_d]_\nu^\mu \odot \mathbf{e}'_\mu$, entonces, reemplazando en la relación invariante:

$$x^\nu [I_d]_\nu^\mu \odot \mathbf{e}'_\mu = x'^\mu \odot \mathbf{e}'_\mu \rightarrow (x^\nu [I_d]_\nu^\mu - x'^\mu) \odot \mathbf{e}'_\mu = \vec{0},$$

y como los \mathbf{e}'_μ son base, tendremos que $x'^\mu = [I_d]_\nu^\mu x^\nu$. De esta misma manera se obtienen todas las relaciones ya escritas.

Generalizando. Sea T una transformación lineal donde $\mathbf{T}_{\mathcal{B}_W \mathcal{B}_V}$ señala la matriz asociada a las bases de V y W . Sea \mathbf{I}_V la matriz cambio de base en V que relaciona las bases \mathcal{B}_V y \mathcal{B}'_V y \mathbf{I}_W la matriz cambio de base en W que relaciona las bases \mathcal{B}_W y \mathcal{B}'_W . Entonces, la matriz asociada a la transformación lineal en las nuevas bases, $\mathbf{T}_{\mathcal{B}'_W \mathcal{B}'_V}$ se obtiene de:

$$\mathbf{T}_{\mathcal{B}'_W \mathcal{B}'_V} = \mathbf{I}_W^{-1} \mathbf{T}_{\mathcal{B}_W \mathcal{B}_V} \mathbf{I}_V.$$

Claramente, en el caso que $\mathcal{B}_V = \mathcal{B}_W = \mathcal{B}$ y $\mathcal{B}'_V = \mathcal{B}'_W = \mathcal{B}'$, $\mathbf{I}_V = \mathbf{I}_W = \mathbf{I}_d$ y $\mathbf{T}_{\mathcal{B}'} = \mathbf{I}_d^{-1} \mathbf{T}_{\mathcal{B}} \mathbf{I}_d$.

Teorema 3.1.1 — Dimensión de las transformaciones lineales. Sea $T : V \rightarrow W$, T lineal, donde $\dim(V) = n$ y $\dim(W) = m$. Entonces: $\dim[\text{Nu}(T)] + \dim[\text{Im}(T)] = \dim(V)$.

3.1.3 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 $f, g, f \oplus g$ y $\lambda \odot f$, con $\lambda \in K$, en $L(V, W)$, entonces:

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

Ejercicio 3.1 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:

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

Consideremos ahora una transformación lineal cualquiera en $L(V, W)$, e.g. T . Sean $[T]_\nu^\mu$ los elementos de la matriz asociada de T en las bases \mathcal{B} y \mathcal{B}' . Tenemos entonces,

$$T(\mathbf{e}_\lambda) = [T]_\lambda^\mu \cdot \mathbf{e}'_\mu.$$

Aprovechando la definición de las transformaciones E_ν^μ podríamos escribir:

$$T(\mathbf{e}_\lambda) = [T]_\lambda^\mu \cdot \mathbf{e}'_\mu = [T]_\lambda^\mu \odot E_\mu^\lambda(\mathbf{e}_\lambda).$$

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

$$\left[T - [T]_{\lambda}^{\mu} \odot E_{\mu}^{\lambda} \right] (\mathbf{e}_{\lambda}) = 0,$$

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

$$T(\cdot) = [T]_{\lambda}^{\mu} \odot E_{\mu}^{\lambda}(\cdot),$$

y las transformaciones E_{μ}^{λ} generan $L(V, W)$. Analicemos la independencia lineal. Consideremos la combinación lineal:

$$b_{\nu}^{\mu} \odot E_{\mu}^{\nu}(\cdot) = 0(\cdot) \quad (\text{i.e. pensémosla como la función nula}).$$

Aplicando a todos los elementos de la base de V se obtiene que los coeficientes b_{ν}^{μ} deben ser todos nulos, por lo que se demuestra de esa manera la independencia lineal de E_{μ}^{ν} .

Esto implica que las funciones E_{μ}^{ν} forman una base para $L(V, W)$ resultando entonces un espacio de dimensión $n \times m$.

Funcionales lineales: 1-formas

El espacio doble dual
Coordenadas de vectores y de 1-formas
Aspectos metodológicos. Ejemplo: \mathbb{R}^n
Cambio de coordenadas para 1-formas
Resumen de cambio de base y coordenadas

4. Formas Lineales y Espacio Dual

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

4.1 Funcionales lineales: 1-formas

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

Definición 4.1.1 — Funcional Lineal. Sea la 4-upla (V, K, \oplus, \odot) un espacio vectorial sobre el cuerpo K , por funcional lineal vamos a entender a toda transformación lineal $f: V \rightarrow K$, esto es, aquellas transformaciones que a cada vector le asocian un escalar.

R Las funcionales lineales siempre se escriben como una combinación lineal de las **componentes** de los vectores del espacio.

■ **Ejemplo 4.1 — Funcionales Lineales.** A continuación enumeramos algunos ejemplos de funcionales lineales:

1. Sea $V = \mathbb{R}^4$, la forma general de una funcional lineal es la siguiente: $f(x, y, z, w) = ax + by + cz + dw$, donde a, b, c, d pueden ser nulos.
2. Sea $V = \mathbb{R}^{2 \times 2}$, la forma general de una funcional lineal es la siguiente: $f \left[\begin{pmatrix} x & y \\ z & w \end{pmatrix} \right] = ax + by + cz + dw$, donde a, b, c, d también pueden ser nulos.
3. Sea $V = \mathbb{R}^3$. Consideremos la función $f(x, y, z) = 2x + y - z$. Es trivial demostrar que es una funcional lineal sobre la 4-upla $(V, \mathbb{R}, \oplus, \odot)$ (donde el cuerpo numérico es el de los reales), para ello calculemos el valor asociado a cada elemento de la base canónica:

¹Los términos: funcionales lineales, formas lineales o 1-formas, los usaremos de forma indistinta a lo largo del texto.

$$\begin{cases} f(1,0,0) = 2, \\ f(0,1,0) = 1, \\ f(0,0,1) = -1. \end{cases}$$

Luego, como cualquier $\vec{v} \in V$ puede escribirse como combinación lineal de los vectores de la base canónica (que son los que acabamos de ver cómo se transforman bajo f) y f es lineal (lo que resulta fácilmente demostrable), f cumple con asociar a todo vector de V un número real, *ergo* f es un funcional lineal sobre la 4-upla $(V, \mathbb{R}, \oplus, \odot)$. ■

Lo que nos va a interesar de las funcionales lineales es que ellas mismas poseen una estructura de espacio vectorial. Pero para ello, recordemos de la Sección 2.1, deberemos definir una suma y un producto por escalares.

Consideremos el conjunto de todas las funcionales lineales sobre el espacio vectorial definido por la 4-upla (V, K, \oplus, \odot) . Dadas f_1 y f_2 funcionales lineales sobre V y $\lambda \in K$, definimos como suma: $(f_1 \oplus f_2)(\vec{v}) = f_1(\vec{v}) + f_2(\vec{v})$, y como producto por escalares: $(\lambda \odot f)(\vec{v}) = \lambda \cdot f(\vec{v})$.

Definidos de esta manera, se puede demostrar que el conjunto de todas las funcionales lineales sobre un espacio vectorial dado es un espacio vectorial. A este espacio se lo denomina **espacio dual** asociado a V , y se lo denota como V^* . En otras palabras: $V^* = \{f/f : V \rightarrow K, \text{ transformación lineal}\}$.

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 K (e.g. \mathbb{R}^2).

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 proponemos 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:

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

Hicimos uso explícito en la notación de \odot dado que el resultado $f \in V^*$ y no a K .

²En este caso, es trivial demostrar que \mathbb{R} es un espacio vectorial de dimensión 1 y que la base canónica es $\mathcal{B}_{\mathbb{R}} = \{1\}$.

R Nótese la ubicación de los índices en la notación de Einstein. Recuerde que estamos definiendo una base para cierto tipo de transformaciones lineales, por lo que sus elementos deben ser también transformaciones lineales.

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:

$$\mathbf{dx}^\mu(\mathbf{e}_\nu) = \delta_\nu^\mu = \begin{cases} 1, & \mu = \nu \\ 0, & \mu \neq \nu. \end{cases} \quad (4.1)$$

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:

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

Si se lo aplicamos a cada elemento de la base de V , e.g. \mathbf{e}_1 , obtenemos:

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

Y 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 $\alpha_\mu \odot \mathbf{dx}^\mu + \beta \odot f = \mathbf{0}$. Si todos los coeficientes fueran cero, entonces tendríamos un conjunto de V^* de $n+1$ vectores linealmente independientes. Entonces, todos no pueden ser linealmente independientes. Si no son linealmente independientes, entonces son dependientes, con lo que existen algunos de los números que no sean cero³. Ergo, podemos escribir $f = -\frac{1}{\beta} \alpha_\mu \odot \mathbf{dx}^\mu = \beta_\mu \odot \mathbf{dx}^\mu$, lo que indica que cualquier vector del dual es generado por los $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$.

Ejercicio 4.1 Considere la base dual para \mathbb{R}^3 : $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \dots, \mathbf{dx}^n\}$, a cuyos vectores de la base se los define como

$$\mathbf{dx}^1(x, y, z) = 2x - y + z; \quad \mathbf{dx}^2(x, y, z) = x - 2y + 3z; \quad \mathbf{dx}^3(x, y, z) = -x + y - z.$$

Hallar:

- La base de \mathbb{R}^3 cuya dual es la dada,
- las coordenadas del vector $(6, 7, 8)$ en la base encontrada en el ítem anterior,

³Nótese que, además, β no puede ser cero, porque implicaría que todos lo sean, y ya hemos dicho que de ser así tendríamos un conjunto de V^* de $n+1$ vectores linealmente independientes.

- Las coordenadas de la funcional lineal $f(x, y, z) = x + y + z$ en \mathcal{B}^* .

4.1.1 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 escalar del cuerpo K .

Consideremos una funcional $f \in V^*$ y sea $\vec{v} \in V$; sabemos que $f(\vec{v}) \in K$. Si consideramos a \vec{v} como una funcional en V^{**} tal que a cada funcional f le asocie un escalar del cuerpo K de la forma $\vec{v}(f) \equiv f(\vec{v})$, tendremos una identificación entre el espacio V y el dual del dual (el doble dual), V^{**} .

Por cómo 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:

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

- **Ejemplo 4.2 — Traza de una matriz cuadrada.** Dada una matriz $n \times n$ (e.g. de la representación matricial de un operador lineal), de elementos a^μ_ν , $\mu, \nu = 1, 2, \dots, n$, la traza:

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

es una 1-forma. ■

- **Ejemplo 4.3 — Valor numérico de un polinomio.** Sea V el espacio de los polinomios de grado menor o igual que n , $\mathbb{R}_n[x]$. Sea $t \in \mathbb{R}$ el cuerpo numérico de los reales. Definamos la transformación $L_t: \mathbb{R}_n[x] \rightarrow \mathbb{R}$ a través de:

$$L_t(p) = p(t),$$

ergo L_t es una 1-forma. ■

- **Ejemplo 4.4 — 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 (y una extremadamente importante). ■

- **Ejemplo 4.5 — La diferencial de una función diferenciable.** Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable en un punto a (que denotaremos por \vec{a}). Decimos que f es diferenciable en el punto a si para todo \vec{h} podemos escribir la variación de f como:

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

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

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

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

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

$$T = \left. \frac{\partial f}{\partial x^\mu} \right|_a \odot \mathbf{dx}^\mu,$$

donde, claramente, las componentes son el gradiente de la función, evaluado en el punto a . ■

R 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:

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

4.1.2 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, i.e. vectores como elementos del doble dual).

A partir de las definiciones de las bases del espacio dual (V^*) y del espacio doble dual (coincidente con el propio espacio original, $V^{**} \equiv V$), podremos 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 V , e.g. \vec{v} , se escribe de manera invariante como $v^\mu \odot \mathbf{e}_\mu$. Entonces, aplicándole un elemento de la base dual (**nótese la diferencia: debe ser un elemento de la base dual, y no sólo un elemento de una base del espacio dual**), e.g. \mathbf{dx}^ν obtenemos:

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

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

$$\vec{v} = \mathbf{dx}^\mu(\vec{v}) \odot \mathbf{e}_\mu. \quad (4.2)$$

De manera análoga tenemos que dado un elemento del espacio dual, f , éste es expandible en la base dual como: $f = f_\mu \odot \mathbf{dx}^\mu$. Apliquemos el elemento \mathbf{e}_ν de la base del doble dual a esta 1-forma:

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

Esto significa que podemos escribir cualquier 1-forma, f , como:

$$f = \mathbf{e}_\mu(f) \odot \mathbf{dx}^\mu = f(\mathbf{e}_\mu) \odot \mathbf{dx}^\mu. \quad (4.3)$$

R 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. Es justamente por esta razón que las 1-formas también son llamadas **funciones coordenadas**.

4.1.3 Aspectos metodológicos. Ejemplo: \mathbb{R}^n

Dada una base para \mathbb{R}^n , para la obtención de la base dual debemos darnos cuenta que toda funcional lineal en este espacio tiene la forma $f(x^1, x^2, \dots, x^n) = a_1 x^1 + a_2 x^2 + \dots + a_n x^n$, donde las cantidades x^j son las coordenadas del vector al que le estoy aplicando f , mientras que los a_j tienen que ver con la transformación. Con lo cual, para determinar cada elemento de la base dual, deberemos aplicar f a cada elemento de la base y hacerle valer 1 o 0 según corresponda (tal y como dice la definición, Ec. 4.1, e.g. en el caso que $f(\mathbf{e}_\mu) = 1$ cuando $\mu = 1$ y 0 cuando $\mu \neq 1$, entonces $f \equiv \mathbf{dx}^1$). De esta manera, lo que obtendremos, constructivamente, serán los elementos de la base dual.

■ **Ejemplo 4.6 — Metodología: caso $n = 2$.** Para \mathbb{R}^2 , consideremos la base $\mathcal{B} = \{(-1, 1); (1, 0)\}$. Determinar:

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

1) Para determinar la base dual, simplemente debemos aplicar la definición, Ec. (4.1). Primero, como las 1-formas en \mathbb{R}^2 tienen la forma $f(x, y) = ax + by$ (por el punto 3.) lo que debemos plantear es si: $\mathbf{dx}^1 = a_1 x + b_1 y$ y $\mathbf{dx}^2 = a_2 x + b_2 y$, para que cumplan con las propiedades de base dual, 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 quedarán como:

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

lo que constituye un sistema de 4×4 , que arroja por resultado los siguientes valores para los coeficientes: $a_1 = 0$, $b_1 = 1$, $a_2 = 1$, $b_2 = 1$.

Entonces, los elementos de la base dual son:

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

2) Para conocer las coordenadas del vector, simplemente aplicamos las 1-formas (funciones coordenadas) al vector (ver Ec. 4.2). Es decir, si $\vec{v} = v^1 \odot (-1, 1) \oplus v^2 \odot (1, 0)$ tendremos que $v^1 = \mathbf{dx}^1(\vec{v})$ y $v^2 = \mathbf{dx}^2(\vec{v})$. Luego,

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

Corroboremos: $v^1 \odot (-1, 1) \oplus v^2 \odot (1, 0) = 5 \odot (-1, 1) \oplus 8 \odot (1, 0) = (-5 + 8, 5) = (3, 5)$.

3) La 1-forma $f(x, y) = 2x + 3y$ será expresada como $f(x, y) = [a_1 \odot \mathbf{dx}^1 \oplus a_2 \odot \mathbf{dx}^2](x, y)$. Para conocer las coordenadas del funcional, aplicamos $f(-1, 1) = a_1$ y $f(1, 0) = a_2$ (ver Ec. 4.3). Entonces,

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

Y la funcional f en la base dual $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2\}$, se escribe como

$$f = 1 \odot \mathbf{dx}^1 \oplus 2 \odot \mathbf{dx}^2.$$

■

4.1.4 Cambio de coordenadas para 1-formas

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

$$\mathbf{e}'_\lambda = \Lambda_\lambda^v \odot \mathbf{e}_v.$$

Nótese que, en este caso, $[I_d]^{-1v}_\lambda = \Lambda_\lambda^v$, con $[I_d]_\lambda^v$ los elementos de la matriz asociada a la transformación identidad señalada en la Sección 3.1.2. Queremos ver, entonces, qué expresión tendrá el cambio de coordenadas para 1-formas. Sea:

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

para ver cómo se relacionan las coordenadas, apliquemos la 1-forma al elemento de la base \mathbf{e}'_λ a ambos miembros, esto es:

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

aplicando el cambio de base,

$$\mathbf{dx}'^\alpha(\mathbf{e}'_\lambda) = \Phi_\beta^\alpha \mathbf{dx}^\beta(\Lambda_\lambda^v \odot \mathbf{e}_v) = \Phi_\beta^\alpha \Lambda_\lambda^v \mathbf{dx}^\beta(\mathbf{e}_v) = \Phi_\beta^\alpha \Lambda_\lambda^v \delta_v^\beta.$$

Entonces,

$$\delta_\lambda^\alpha = \Phi_\beta^\alpha \Lambda_\lambda^v.$$

Lo que implica que la matriz de cambio, cumple con $\Phi = \Lambda^{-1}$. Con esto, como las coordenadas transforman con la inversa de la matriz con la que transforman las bases, para el cambio de coordenadas de funcionales, tendremos para una 1-forma $f = f_v \odot \mathbf{dx}^v$ una representación en una nueva base $\mathcal{B}'^* = \{\mathbf{dx}^{1'}, \mathbf{dx}^{2'}, \dots, \mathbf{dx}^{n'}\}$ como $f = f'_\mu \odot \mathbf{dx}^{\mu'}$ donde:

$$f'_\mu = \Lambda_\mu^v f_v.$$

4.1.5 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 como sigue:

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

- Funcionales bilineales sobre V
- Algebra de las formas bilineales
- Producto tensorial de 1-formas
- Coordenadas de un tensor
- Tensores cartesianos en general
 - Covarianza y contravarianza de un tensor
 - Cambio de coordenadas en tensores cartesianos
- Componentes de 1-formas y tensores. Caso \mathbb{R}^3
 - Relación entre componentes y coordenadas de tensores
 - Aspectos metodológicos
- Formas cuadráticas

5. Formas bilineales y multilineales. Tensores

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

5.1 Funcionales bilineales sobre V

Definición 5.1.1 — Funcional bilineal. Dado un espacio vectorial V , de dimensión n , vamos a estudiar las funciones del tipo $f : V \times V \rightarrow K$, i.e. aquéllas que asocian a un par de vectores \vec{u} y \vec{v} un escalar. En otras palabras, la función aplica $f(\vec{u}; \vec{v})$ a un escalar. Si además, con respecto a cada argumento la función es lineal:

$$f[\vec{u}; \lambda \odot \vec{v}_1 \oplus \vec{v}_2] = \lambda f(\vec{u}; \vec{v}_1) + f(\vec{u}; \vec{v}_2)$$

$$f[\lambda \odot \vec{u}_1 \oplus \vec{u}_2; \vec{v}] = \lambda f(\vec{u}_1; \vec{v}) + f(\vec{u}_2; \vec{v})$$

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 porque las 1-formas están aplicadas sobre vectores, i.e. son escalares) es una forma bilineal.

■ **Ejemplo 5.1 — Funcionales bilineales.** Es trivial demostrar que la siguiente es una forma bilineal sobre \mathbb{R}^2 : $F(\vec{u}; \vec{v}) = 2u_x v_x - 4u_x v_y + u_y v_x - 2u_y v_y$, y que además puede reescribirse como $F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$ donde $f(\vec{u}) = 2u_x + u_y$ y $g(\vec{v}) = v_x - 2v_y$. ■

R Más aún, notemos que esto no es una particularidad, sino que cualquier forma bilineal puede ser escrita como *producto* de dos 1-formas. La *itálica* viene de que lo que se ve como producto (\cdot) es en realidad el resultado de la aplicación de cada 1-forma sobre vectores, lo que da por resultado sendos escalares. Pero aún no sabemos qué es un producto de 1-formas que aún no se aplican sobre vectores.

5.1.1 Álgebra de las formas bilineales

Consideremos el conjunto de todas las formas bilineales $F : V \times V \rightarrow K$. 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:

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

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

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

Estas definiciones permiten comprobar que el conjunto de las formas bilineales poseen una estructura de espacio vectorial.

Ejercicio 5.1 Se deja como ejercicio comprobar que la dimensión de este espacio es n^2 , siendo n la dimensión de V . ■

Aplicando estas definiciones de suma y producto para formas bilineales, y el hecho de poder expresar la forma bilineal como el producto de dos 1-formas, $F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$, reescribamos:

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

Entonces,

$$F(\vec{u}; \vec{v}) = f_\mu g_\alpha u^\nu v^\beta \delta_\nu^\mu \cdot \delta_\beta^\alpha = f_\mu g_\alpha u^\mu v^\alpha = f_\nu g_\beta u^\nu v^\beta, \quad (5.1)$$

es un invariante. Esta es, entonces, una de las maneras de calcular la aplicación bilineal.

Vamos a construir una base para el espacio de formas bilineales.

5.2 Producto tensorial de 1-formas

Definición 5.2.1 — Producto tensorial de 1-formas. A partir de poder escribir la aplicación bilineal como producto en K (e.g. en \mathbb{R}) de la aplicación lineal sobre cada uno de sus argumentos, $F(\vec{u}; \vec{v}) = f(\vec{u}) \cdot g(\vec{v})$, se define el **producto tensorial de dos 1-formas** f y g , denotándolo como $f \otimes g$, a través de la expresión:

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

Algunas propiedades del producto tensorial:

- $(f_1 \oplus f_2) \otimes g = f_1 \otimes g \oplus f_2 \otimes g$;
- $f \otimes (g_1 \oplus g_2) = f \otimes g_1 \oplus f \otimes g_2$;
- $(\lambda \odot f) \otimes g = \lambda \odot (f \otimes g)$;
- $f \otimes (\lambda \odot g) = \lambda \odot (f \otimes g)$.

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 y que, además, posee dimensión n^2 . Más aún, el espacio definido

de esta manera se lo denomina **espacio producto tensorial** $V^* \otimes V^*$ y los elementos de este espacio se denominan tensores de tipo $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ (concepto que se desarrollará más adelante).

Luego, vamos a proponer como base de este espacio al conjunto:

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

5.2.1 Coordenadas de un tensor

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

Dada una forma bilineal $F : V \times V \rightarrow \mathbb{R}$, como mencionamos, esta forma pertenece al espacio generado por la base $\mathcal{B}^* = \{\mathbf{dx}^\mu \otimes \mathbf{dx}^\nu\}$ con $\mu, \nu = 1, 2, \dots, n$, luego F puede escribirse como $F = f_{\mu\nu} \odot \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$. Veamos, entonces, cuáles deben ser las coordenadas $f_{\mu\nu}$.

Aplicando a ambos miembros del desarrollo elementos de la base de V^1 , tenemos:

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

Y por definición de producto tensorial de 1-formas:

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

entonces:

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

Lo que nos permite, finalmente, reescribir $F = f_{\mu\nu} \odot \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu$ como,

$$F = F(\mathbf{e}_\mu; \mathbf{e}_\nu) \odot \mathbf{dx}^\mu \otimes \mathbf{dx}^\nu. \quad (5.2)$$

Si aplicamos esta forma bilineal (**tensor**) a un par de vectores cualesquiera, \vec{u} y \vec{v} , tendremos:

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

resultando,

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

Finalmente, en el caso general, tendremos:

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

donde recuperamos la expresión dada por la Ec. (5.1).

¹Notemos que para determinar coordenadas, el procedimiento es siempre el mismo: aplicar la forma sobre los elementos de la base.

■ **Ejemplo 5.2** Hallar las coordenadas del tensor dos veces covariante $F \in V^* \times V^* : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ definido como: $F = g \otimes h$, con $g(x, y) = 3x + 2y$ y $h(x, y) = x - 4y$.

Para obtener las coordenadas, sólo debemos aplicar el tensor a los elementos de la base de V . Suponiendo que esta base es la canónica ($\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$), tendremos que las coordenadas del tensor son:

$$\begin{aligned} f_{11} &= F(\mathbf{e}_1, \mathbf{e}_1) = 3 \cdot 1 = 3; \\ f_{12} &= F(\mathbf{e}_1, \mathbf{e}_2) = -12; \\ f_{21} &= F(\mathbf{e}_2, \mathbf{e}_1) = 2; \\ f_{22} &= F(\mathbf{e}_2, \mathbf{e}_2) = -8. \end{aligned}$$

Luego, el tensor F puede ser escrito en la base $\mathcal{B}^* \times \mathcal{B}^* = \{\mathbf{dx}^1 \otimes \mathbf{dx}^1, \mathbf{dx}^1 \otimes \mathbf{dx}^2, \mathbf{dx}^2 \otimes \mathbf{dx}^1, \mathbf{dx}^2 \otimes \mathbf{dx}^2\}$, siendo $\{\mathbf{dx}^1, \mathbf{dx}^2\}$ la base dual de la base canónica, como: $F = 3 \odot \mathbf{dx}^1 \otimes \mathbf{dx}^1 - 12 \odot \mathbf{dx}^1 \otimes \mathbf{dx}^2 + 2 \odot \mathbf{dx}^2 \otimes \mathbf{dx}^1 - 8 \odot \mathbf{dx}^2 \otimes \mathbf{dx}^2$. ■

Ejercicio 5.2 Escribir el tensor F para la base de \mathbb{R}^2 , $\mathcal{B} = \{(1, -1), (0, 2)\}$. ■

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 escalar, i.e.: $F : V^* \times V^* \rightarrow K$, 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

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

es una base para el espacio $V \otimes V$. Tomando esta base para el espacio, podemos notar que toda forma bilineal en $V \otimes V$ puede escribirse como:

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

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^* :

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

■ **Ejemplo 5.3** Sea $T \in V \times V : \mathbb{R}^2 \times \mathbb{R}^2$, definido explícitamente como $T = (-1, 4) \otimes (2, 1)$. Hallar las coordenadas del mismo si la base del dual es $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2\}$ con $\mathbf{dx}^1(x, y) = 2x - y$ y $\mathbf{dx}^2(x, y) = -x + y$.

Aplicando el tensor T a cada elemento de la base dual (i.e. de V^*), tendremos inmediatamente las coordenadas:

$$\begin{aligned} t^{11} &= T(\mathbf{dx}^1, \mathbf{dx}^1) = (-1, 4) \otimes (2, 1)(\mathbf{dx}^1, \mathbf{dx}^1) = \mathbf{dx}^1(-1, 4) \cdot \mathbf{dx}^1(2, 1) = (-6) \cdot 3 = -18; \\ t^{12} &= T(\mathbf{dx}^1, \mathbf{dx}^2) = (-1, 4) \otimes (2, 1)(\mathbf{dx}^1, \mathbf{dx}^2) = \mathbf{dx}^1(-1, 4) \cdot \mathbf{dx}^2(2, 1) = 6; \\ t^{21} &= T(\mathbf{dx}^2, \mathbf{dx}^1) = (-1, 4) \otimes (2, 1)(\mathbf{dx}^2, \mathbf{dx}^1) = \mathbf{dx}^2(-1, 4) \cdot \mathbf{dx}^1(2, 1) = 15; \\ t^{22} &= T(\mathbf{dx}^2, \mathbf{dx}^2) = (-1, 4) \otimes (2, 1)(\mathbf{dx}^2, \mathbf{dx}^2) = \mathbf{dx}^2(-1, 4) \cdot \mathbf{dx}^2(2, 1) = -5. \end{aligned}$$

Y el tensor podrá ser escrito como: $T = -18 \odot \mathbf{e}_1 \otimes \mathbf{e}_1 + 6 \odot \mathbf{e}_1 \otimes \mathbf{e}_2 + 15 \odot \mathbf{e}_2 \otimes \mathbf{e}_1 - 5 \odot \mathbf{e}_2 \otimes \mathbf{e}_2$. ■

5.3 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.

Definición 5.3.1 — Tipo de tensores. Decimos tensor del tipo $\begin{pmatrix} p \\ q \end{pmatrix}$ a toda forma multilineal:

$$T : \underbrace{V \times V \times \dots \times V}_{p\text{-veces}} \times \underbrace{V^* \times V^* \times \dots \times V^*}_{q\text{-veces}} \rightarrow K, \quad (5.3)$$

definida a partir del producto tensorial. Es decir:

$$T \in \underbrace{V^* \otimes V^* \otimes \dots \otimes V^*}_{p\text{-veces}} \otimes \underbrace{V \otimes V \otimes \dots \otimes V}_{q\text{-veces}}. \quad (5.4)$$

Nótese el *intercambio* que existe entre los espacios de las dos expresiones. Esto se debe a que, en la Ec. (5.3), los espacios corresponden a aquéllos a los cuales pertenecen los elementos sobre los cuales actúan otros elementos que pertenecen a los espacios denotados en la Ec. (5.4).

De esta manera, las formas bilineales sobre V :

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

podemos generalizarlas bajo el concepto de tensores del tipo $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ y las formas bilineales sobre V^* :

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

serán entonces, tensores del tipo $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

La definición de tensor, en general, permite definir un tensor del tipo $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ que aplica a un elemento de V y un elemento de V^* un elemento del cuerpo K (e.g. un número real, elemento de \mathbb{R}), con lo que, e.g., podremos escribirlo como:

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

donde las coordenadas se obtienen, como es usual, aplicando el tensor a los elementos de la base:

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

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

$$T = T_{\mu_1 \mu_2 \dots \mu_p}^{v_1 v_2 \dots v_q} \odot \mathbf{dx}^{\mu_1} \otimes \mathbf{dx}^{\mu_2} \otimes \dots \otimes \mathbf{dx}^{\mu_p} \otimes \mathbf{e}_{v_1} \otimes \mathbf{e}_{v_2} \otimes \dots \otimes \mathbf{e}_{v_q},$$

donde sus coordenadas se obtienen a partir de aplicarlo a los elementos de la base:

$$T_{\mu_1 \mu_2 \dots \mu_p}^{v_1 v_2 \dots v_q} = T(\mathbf{e}_{\mu_1}, \mathbf{e}_{\mu_2}, \dots, \mathbf{e}_{\mu_p}; \mathbf{dx}^{v_1}, \mathbf{dx}^{v_2}, \dots, \mathbf{dx}^{v_q}).$$

Con esta definición, tanto los vectores como las 1-formas serán tensores. Los vectores, vistos 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 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

5.3.1 Covarianza y contravarianza de un tensor

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

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

la cantidad de subíndices que aparece en las coordenadas indican la cantidad de veces que es covariante, mientras que los supraíndices, la cantidad de veces que es contravariante.

5.3.2 Cambio de coordenadas en tensores cartesianos

Consideremos un cambio de base:

$$\mathbf{e}'_v = \Lambda_v^{\mu} \mathbf{e}_{\mu}, \quad \mathbf{e}_{\mu} = [\Lambda^{-1}]_{\mu}^v \mathbf{e}'_v.$$

Como hemos visto, este cambio de base produce, respectivamente, el siguiente cambio de coordenadas:

$$x'^v = [\Lambda^{-1}]_{\mu}^v x^{\mu}, \quad x^{\mu} = \Lambda_v^{\mu} x'^v.$$

A su vez, en el espacio dual induce los cambios de base:

$$\mathbf{dx}'^v = [\Lambda^{-1}]_{\mu}^v \mathbf{dx}^{\mu}, \quad \mathbf{dx}^{\mu} = \Lambda_v^{\mu} \mathbf{dx}'^v,$$

junto con los siguientes cambios de coordenadas:

$$f'_v = \Lambda_v^{\mu} f_{\mu}, \quad f_{\mu} = [\Lambda^{-1}]_{\mu}^v f'_v.$$

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

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

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

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

entonces,

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

y como los $\mathbf{dx}'^{\mu} \otimes \mathbf{dx}'^{\nu}$ son linealmente independientes, 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, ahora, un tensor dos veces contravariante. De manera invariante, podemos volver a escribir:

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

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

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

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

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

Finalmente, veamos un tensor una vez covariante y una vez contravariante. Este objeto, se escribe como:

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

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):

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

con lo cual, se obtiene:

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

Procedimiento esquemático para cambio de coordenadas de tensores.

A partir de los casos detallados en la Sección 5.3.2 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

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

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

$$[]'^{\nu} = [\Lambda^{-1}]_{\beta}^{\nu} []^{\beta}.$$

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

$$T_{\alpha_1 \alpha_2 \dots \alpha_p}^{\beta_1 \beta_2 \dots \beta_q},$$

el cambio de coordenadas queda:

$$T_{\mu_1 \mu_2 \dots \mu_p}^{\nu_1 \nu_2 \dots \nu_q} = \Lambda_{\mu_1}^{\alpha_1} \Lambda_{\mu_2}^{\alpha_2} \dots \Lambda_{\mu_p}^{\alpha_p} [\Lambda^{-1}]_{\beta_1}^{\nu_1} [\Lambda^{-1}]_{\beta_2}^{\nu_2} \dots [\Lambda^{-1}]_{\beta_q}^{\nu_q} T_{\alpha_1 \alpha_2 \dots \alpha_p}^{\beta_1 \beta_2 \dots \beta_q}.$$

R En algunos textos clásicos, tales como [L73], 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 $T_{\mu_1 \mu_2 \dots \mu_p}^{\nu_1 \nu_2 \dots \nu_q}$ son las coordenadas de un tensor p veces covariante y q veces contravariante si frente a un cambio de coordenadas $x'^\alpha = B_\beta^\alpha x^\beta$ las coordenadas se transforman como $T_{\mu_1 \mu_2 \dots \mu_p}^{\nu_1 \nu_2 \dots \nu_q} = A_{\mu_1}^{\alpha_1} A_{\mu_2}^{\alpha_2} \dots A_{\mu_p}^{\alpha_p} B_{\beta_1}^{\nu_1} B_{\beta_2}^{\nu_2} \dots B_{\beta_q}^{\nu_q} T_{\alpha_1 \alpha_2 \dots \alpha_p}^{\beta_1 \beta_2 \dots \beta_q}$.

Aquí, las matrices **A** y **B** son inversas una de otra. A este criterio, el autor 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.

En otras palabras, el libro [L73] trabaja casi en su totalidad en coordenadas. El libro 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.

5.4 Componentes de 1-formas y tensores. Caso \mathbb{R}^3

Una 1-forma perteneciente al espacio dual de \mathbb{R}^3 , al actuar sobre un vector $\vec{v} = (x, y, z)$ (de componentes x, y, z), tiene la siguiente expresión general:

$$f(x, y, z) = a \cdot x + b \cdot y + c \cdot z, \text{ con } a, b, c \in \mathbb{R}.$$

Notemos que para cada valor constante de f , tenemos la **ecuación de un plano** con vector normal $\vec{n} = (a, b, c)$.

Notemos, además, que podemos asociar los números a, b, c con las *componentes* de la 1-forma.

Si queremos ver la 1-forma sin necesidad de ver como actúa sobre un vector, podemos, entonces, definir:

$$f = \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} \text{ con } a, b, c \text{ las componentes de la 1-forma,}$$

asociación de componentes que es perfectamente lógica. En efecto, si consideramos a V como el dual de V^* , tenemos que:

$$\vec{v}(f) \equiv f(\vec{v}) = a \cdot x + b \cdot y + c \cdot z = x \cdot a + y \cdot b + z \cdot c$$

donde si a, b, c son las componentes de la 1-forma (o en este caso, del *vector*), podemos asociar a x, y, z como las componentes del vector (o en este caso, de la 1-forma) \vec{v} .

5.4.1 Relación entre componentes y coordenadas de tensores

Del mismo modo que relacionamos las componentes y coordenadas de vectores, podemos relacionar coordenadas y componentes de 1-formas.

Veamos. Dada la base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ del espacio \mathbb{R}^3 , tenemos:

$$\vec{v} = (x, y, z) = v^1 \cdot \mathbf{e}_1 + v^2 \cdot \mathbf{e}_2 + v^3 \cdot \mathbf{e}_3,$$

donde podemos encontrar, en función de la definición de los vectores base,

$$v^1 = v^1(x, y, z),$$

$$v^2 = v^2(x, y, z),$$

$$v^3 = v^3(x, y, z).$$

Por otro lado, debemos poder escribir, para 1-formas:

$$f = \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = f_1 \mathbf{dx}^1 + f_2 \mathbf{dx}^2 + f_3 \mathbf{dx}^3$$

donde podremos encontrar las siguientes relaciones entre coordenadas y componentes:

$$f_1 = f_1(a, b, c),$$

$$f_2 = f_2(a, b, c),$$

$$f_3 = f_3(a, b, c).$$

Notemos que para encontrar estas relaciones será necesario aplicar a vectores, ya que las 1-formas se manifiestan mediante la aplicación a vectores.

■ **Ejemplo 5.4 — Componentes de 1-formas.** Consideremos para \mathbb{R}^3 la siguiente base: $\mathcal{B} = \{(1, 0, 0); (1, -1, 0); (-1, 1, 1)\}$. Esta base induce la base dual $\mathcal{B}^* = \{\mathbf{dx}^1, \mathbf{dx}^2, \mathbf{dx}^3\}$, construida al aplicar cada elemento de ella sobre los vectores de la base \mathcal{B} :

$$dx^1(x, y, z) = x + y,$$

$$dx^2(x, y, z) = -y + z,$$

$$dx^3(x, y, z) = z.$$

Con lo que, inaugurando el uso de las componentes, podemos reescribir la base \mathcal{B}^* como:

$$\mathcal{B}^* = \left\{ \begin{Bmatrix} 1 \\ 1 \\ 0 \end{Bmatrix}, \begin{Bmatrix} 0 \\ -1 \\ 1 \end{Bmatrix}, \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} \right\}.$$

Notemos que un vector $\vec{v} = (x, y, z)$ se puede escribir como $\vec{v} = v^1 \odot (1, 0, 0) \oplus v^2 \odot (1, -1, 0) \oplus v^3 \odot (-1, 1, 1)$, donde al relacionar las coordenadas con las componentes, podemos escribir:

$$\vec{v} = (x + y) \odot (1, 0, 0) \oplus (z - y) \odot (1, -1, 0) \oplus z \odot (-1, 1, 1).$$

Escribamos ahora una 1-forma como combinación de los elementos de la base dual:

$$f = \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = f_1 \odot \mathbf{dx}^1 \oplus f_2 \odot \mathbf{dx}^2 \oplus f_3 \odot \mathbf{dx}^3 = f_1 \odot \begin{Bmatrix} 1 \\ 1 \\ 0 \end{Bmatrix} \oplus f_2 \odot \begin{Bmatrix} 0 \\ -1 \\ 1 \end{Bmatrix} \oplus f_3 \odot \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} = \begin{Bmatrix} f_1 \\ f_1 - f_2 \\ f_2 + f_3 \end{Bmatrix}.$$

- R** Podemos definir una suma usual y un producto usual entre representaciones en componentes de 1-formas. Para sumas y productos no usuales habría que probarlos. En otras palabras, habría que ver el comportamiento entre representaciones de elementos del dual para espacios que no sean n -uplas.

Entonces, $f_1 = a$, $f_2 = a - b$, $f_3 = b + c - a$. Luego $f = a \odot \mathbf{dx}^1 \oplus (a - b) \odot \mathbf{dx}^2 \oplus (b + c - a) \odot \mathbf{dx}^3$. En efecto: $f(x, y, z) = a \cdot \mathbf{dx}^1(x, y, z) + (a - b) \cdot \mathbf{dx}^2(x, y, z) + (b + c - a) \cdot \mathbf{dx}^3(x, y, z) = a \cdot (x + y) + (a - b) \cdot (-y + z) + (b + c - a) \cdot z = ax + by + cz$. ■

Con la definición de componentes de una 1-forma, podemos escribir la representación en coordenadas como función de las componentes, análogamente con lo que hacemos con vectores.

Finalmente, esta representación en componentes, nos permitirá escribir expresiones explícitas de tensores de cualquier tipo, actuando en componentes.

■ **Ejemplo 5.5 — Componentes de tensores.** Consideremos \mathbb{R}^3 y su dual. Consideremos las bases utilizadas para el ejemplo anterior. Esto es:

$$\mathcal{B} = \{(1, 0, 0); (1, -1, 0); (-1, 1, 1)\} \text{ y } \mathcal{B}^* = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Los vectores y las 1-formas, a su vez, se escriben, respectivamente, como:

$$\vec{v} = (x, y, z) = (x + y) \odot \mathbf{e}_1 \oplus (z - y) \odot \mathbf{e}_2 \oplus z \odot \mathbf{e}_3 \text{ y } f = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \odot \mathbf{dx}^1 \oplus (a - b) \odot \mathbf{dx}^2 \oplus (b + c - a) \odot \mathbf{dx}^3$$

Consideremos ahora, un tensor del tipo $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, el cual se puede escribir de la siguiente manera: $\mathbf{T} = T_{\nu}^{\mu} \odot \mathbf{dx}^{\nu} \otimes \mathbf{e}_{\mu}$. Supongamos, para simplificar, que $\mathbf{T} = 3 \odot \mathbf{dx}^1 \otimes \mathbf{e}_1 \oplus 2 \odot \mathbf{dx}^2 \otimes \mathbf{e}_3$, i.e. tiene muchas coordenadas nulas sólo para hacer el ejemplo no tan largo.

Para obtener **la expresión explícita, i.e., en componentes**, apliquemos el tensor a un vector \vec{v} y a una 1-forma f : $\mathbf{T}(\vec{v}; f) = 3 \cdot \mathbf{dx}^1(\vec{v}) \cdot \mathbf{e}_1(f) + 2 \cdot \mathbf{dx}^2(\vec{v}) \cdot \mathbf{e}_3(f)$. Resolvemos, haciendo uso de:

$$\begin{aligned} \mathbf{dx}^1(\vec{v}) &= \mathbf{dx}^1[(x + y) \odot \mathbf{e}_1 \oplus (z - y) \odot \mathbf{e}_2 \oplus z \odot \mathbf{e}_3] &= (x + y), \\ \mathbf{dx}^2(\vec{v}) &= \mathbf{dx}^2[(x + y) \odot \mathbf{e}_1 \oplus (z - y) \odot \mathbf{e}_2 \oplus z \odot \mathbf{e}_3] &= (z - y), \\ \mathbf{e}_1(f) &= \mathbf{e}_1[a \odot \mathbf{dx}^1 \oplus (a - b) \odot \mathbf{dx}^2 \oplus (b + c - a) \odot \mathbf{dx}^3] &= a, \\ \mathbf{e}_3(f) &= \mathbf{e}_3[a \odot \mathbf{dx}^1 \oplus (a - b) \odot \mathbf{dx}^2 \oplus (b + c - a) \odot \mathbf{dx}^3] &= (b + c - a), \end{aligned}$$

con lo que la expresión explícita del tensor será:

$$\mathbf{T}(\vec{x}; f) = 3(x + y) \cdot a + 2(z - y) \cdot (b + c - a).$$

■

5.4.2 Aspectos metodológicos

La relación entre las coordenadas y las componentes se puede obtener de la siguiente manera:

1. Escribimos la 1-forma en cuestión, en componentes: $f = \begin{Bmatrix} a \\ b \\ c \end{Bmatrix}$;

2. si, además, cada elemento de la base dual está dado por sus componentes:

$$\mathbf{dx}^1 = \begin{Bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{Bmatrix}, \quad \mathbf{dx}^2 = \begin{Bmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{Bmatrix}, \quad \mathbf{dx}^3 = \begin{Bmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{Bmatrix},$$

3. entonces, para la 1-forma $f = f_\mu \odot \mathbf{dx}^\mu$, se pueden relacionar coordenadas y componentes de la siguiente manera:

$$\begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = f_1 \odot \begin{Bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{Bmatrix} \oplus f_2 \odot \begin{Bmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{Bmatrix} \oplus f_3 \odot \begin{Bmatrix} \alpha_3 \\ \beta_3 \\ \gamma_3 \end{Bmatrix}.$$

4. Ahora bien, teniendo en cuenta que, independientemente de cómo esté definida la suma y el producto por un número en el conjunto de vectores, **la suma y el**

producto por un número en el conjunto de los elementos $\begin{Bmatrix} a \\ b \\ c \end{Bmatrix}$ **son las**

usuales. En efecto, esto se obtiene en virtud de que al actuar una 1-forma sobre las componentes de un vector, el resultado es un número real con la estructura usual.

5. Luego, con estas consideraciones, tenemos la siguiente igualdad, componente a componente:

$$\begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = \begin{Bmatrix} f_1 \alpha_1 + f_2 \alpha_2 + f_3 \alpha_3 \\ f_1 \beta_1 + f_2 \beta_2 + f_3 \beta_3 \\ f_1 \gamma_1 + f_2 \gamma_2 + f_3 \gamma_3 \end{Bmatrix},$$

6. la que brinda la relación entre las f_μ (coordenadas) y las componentes de la 1-forma f , a partir de resolver el siguiente sistema lineal:

$$\begin{cases} f_1 \cdot \alpha_1 + f_2 \cdot \alpha_2 + f_3 \cdot \alpha_3 & = & a, \\ f_1 \cdot \beta_1 + f_2 \cdot \beta_2 + f_3 \cdot \beta_3 & = & b, \\ f_1 \cdot \gamma_1 + f_2 \cdot \gamma_2 + f_3 \cdot \gamma_3 & = & c. \end{cases}$$

7. En virtud de que \mathbf{dx}^1 , \mathbf{dx}^2 y \mathbf{dx}^3 son linealmente independientes, el sistema admite solución única:

$$\begin{aligned} f_1 &= f_1(a, b, c), \\ f_2 &= f_2(a, b, c), \\ f_3 &= f_3(a, b, c). \end{aligned}$$

R Este abordaje permite encontrar las relaciones sin la necesidad de aplicar $f(x, y, z)$.

5.5 Formas cuadráticas

Definición 5.5.1 — Funcional cuadrático. Dada una forma bilineal sobre un espacio V , es decir, que la forma asigna a dos vectores de V un elemento del cuerpo K . En términos de la aplicación, tenemos que una forma bilineal F , tiene por expresión $F(\vec{u}; \vec{v}) \in K$. Una **forma cuadrática** se obtiene a partir de calcular la forma bilineal para un mismo vector $F(\vec{v}; \vec{v})$.

■ **Ejemplo 5.6 — Funcionales cuadráticos.** Para el Ejemplo 5.1, 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: $F(\vec{v}; \vec{v}) \rightarrow Q(x, y) = 2x^2 - 3xy - 2y^2$ (i.e. $Q(\lambda x, \lambda y) = \lambda^2 Q(x, y)$). ■

En el caso de 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 5.6: $Q(x, y) = 2x^2 - 3xy - 2y^2$, donde cada término es cuadrático (los términos son x^2, xy, y^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 como:

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

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 (por definición de punto crítico), 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.

Producto interno

Axiomas de producto interno

Norma o módulo de un vector

Ortogonalidad

Construcción de una base ortogonal

Base ortogonal. Coeficientes de Fourier

El producto interno como un tensor. El tensor métrico

Aplicación: longitud de arco

Coordenadas covariantes de un vector contra-variante.

6. Espacios euclídeos y espacios métricos

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

6.1 Producto interno

En el caso particular del espacio tridimensional, dado un par de vectores de \mathbb{R}^3 , $\vec{u} = (u_x, u_y, u_z)$ y $\vec{v} = (v_x, v_y, v_z)$ ¹, reconocemos 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, \mathbb{R}^n , $\mathbb{R}^{n \times m}$, las funciones continuas en determinado intervalo, etc.

- R** 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, lo que da más libertad para definir propiamente la operación y, para el caso de \mathbb{R}^3 , el producto escalar ya conocido será un caso particular de producto interno, que llamaremos *producto interno canónico*.

6.1.1 Axiomas de producto interno

Definición 6.1.1 — 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;

¹Nótese que no se utilizó la notación de Einstein porque estas no son coordenadas, sino componentes.

- $\langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle$ (conmutatividad sólo para espacios sobre los reales^a);
- $\langle \lambda \odot \vec{u}_1 \oplus \vec{u}_2 | \vec{v} \rangle = \lambda \langle \vec{u}_1 | \vec{v} \rangle + \langle \vec{u}_2 | \vec{v} \rangle$;
- $\langle \vec{v} | \vec{v} \rangle > 0 \quad \forall \vec{v} \neq \vec{0}, \quad \langle \vec{v} | \vec{v} \rangle = 0 \leftrightarrow \vec{v} = \vec{0}$.

^aSi el cuerpo fuera el de los complejos, esta propiedad se traduciría en: $\langle \vec{u} | \vec{v} \rangle = \overline{\langle \vec{v} | \vec{u} \rangle}$, donde la barra indica conjugación compleja. Además, $\langle \vec{u} | \lambda \odot \vec{v} \rangle = \bar{\lambda} \langle \vec{u} | \vec{v} \rangle$, luego el producto interno en complejos no es estrictamente una forma bilineal.

Ejercicio 6.1 Comprobar que el producto interno canónico (producto escalar ya conocido) en \mathbb{R}^2 definido como:

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

es efectivamente un producto interno. ■

Ejercicio 6.2 Comprobar que la operación binaria en \mathbb{R}^2 definida como:

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

es un producto interno. ■

■ **Ejemplo 6.1 — 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 6.3 Comprobar que, así definido, es efectivamente un producto interno en el espacio de las funciones continuas. ■

■ **Ejemplo 6.2 — Producto interno en el espacio de matrices $n \times n$.** Para el espacio de matrices $\mathbb{R}^{n \times n}$ se define el producto:

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

■

Ejercicio 6.4 Comprobar que, efectivamente, es un producto interno. ■

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

6.1.2 Norma o módulo de un vector

Definición 6.1.2 — Norma de un vector. Una vez definido un producto interno, podemos definir la **norma** de un vector de un espacio vectorial V , \vec{v} , como:

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

Teorema 6.1.1 Si V es un espacio producto interno, se cumple:

- $\|\lambda \odot \vec{v}\| = |\lambda| \|\vec{v}\|$;
- $\|\vec{v}\| > 0$, para $\vec{v} \neq \vec{0}$;

- $|\langle \vec{u} | \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$, Desigualdad de Cauchy–Schwarz;
- $\|\vec{u} \oplus \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$, Desigualdad triangular.

Ejercicio 6.5 Las demostraciones de los dos primeros puntos son inmediatas a partir de la definición de producto interno, y se deja como ejercicio. ■

En cambio, para demostrar la desigualdad de Cauchy–Schwarz, consideremos el vector:

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

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

$$0 \leq \|\vec{w}\|^2 = \left\langle \vec{u} \oplus -\frac{\langle \vec{u} | \vec{v} \rangle}{\|\vec{v}\|^2} \odot \vec{v} \mid \vec{u} \oplus -\frac{\langle \vec{u} | \vec{v} \rangle}{\|\vec{v}\|^2} \odot \vec{v} \right\rangle.$$

Haciendo las cuentas (nótese que hay un valor absoluto involucrado: $|\langle \vec{u} | \vec{v} \rangle|$),

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

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:

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

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

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

de donde se obtiene la desigualdad triangular.

R En algunos casos, es posible definir una norma de vectores sin necesidad de tener definido un producto interno. En ese caso, decimos que el espacio es un **espacio normado**.

6.2 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, tomaremos estas ideas para generalizarlas. Ahora definiremos la *ortogonalidad* de dos vectores cualesquiera (i.e. de cualquier naturaleza) a partir del producto interno.

Definición 6.2.1 — Ortogonalidad. 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 6.2.2 — Conjunto ortogonal. Sea V un espacio producto interno. Sea $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ un conjunto de vectores no nulos. Se dice que S es un conjunto ortogonal, si los vectores son ortogonales de a pares, esto es $\langle \vec{v}_\mu | \vec{v}_\nu \rangle = 0$ si $\mu \neq \nu$.

Proposición 6.2.1 — Ortogonalidad implica independencia lineal. Un conjunto ortogonal es linealmente independiente.

En efecto, consideremos el conjunto ortogonal $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$. Tomemos $a^\mu \odot \vec{v}_\mu = 0$ y calculemos para algún \vec{v}_ν el producto interno: $\langle a^\mu \odot \vec{v}_\mu | \vec{v}_\nu \rangle = 0 \rightarrow a^\mu \langle \vec{v}_\mu | \vec{v}_\nu \rangle = 0$. Como el conjunto es ortogonal, tenemos que son ortogonales de a pares, i.e. 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 como por definición de producto interno, $\|\vec{v}_\nu\|^2 > 0$, se debe cumplir que $a^\nu = 0$. 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 los vectores del conjunto son linealmente independientes.

Definición 6.2.3 — Conjunto ortonormal. 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.

6.2.1 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**, y se describe a continuación.

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

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

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

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

y 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:

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

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 para el mismo espacio.

6.2.2 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 $\vec{v} = v^\mu \odot \mathbf{e}_\mu$. Calculemos $\langle \vec{v} | \mathbf{e}_\nu \rangle$, tenemos $\langle \vec{v} | \mathbf{e}_\nu \rangle = \langle v^\mu \odot \mathbf{e}_\mu | \mathbf{e}_\nu \rangle = v^\mu \langle \mathbf{e}_\mu | \mathbf{e}_\nu \rangle = v^\nu \|\mathbf{e}_\nu\|^2$. Entonces, las coordenadas las obtenemos calculando:

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

Y la expresión del vector es, entonces:

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

R 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 ortonormal, tenemos que $\vec{v} = \langle \vec{v} | \mathbf{e}_\mu \rangle \odot \mathbf{e}_\mu$. 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 que resultan ser 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.

6.3 El producto interno como un tensor. El tensor métrico

A partir de la definición de producto interno (Sección 6.1.1) podemos notar que la primera propiedad nos dice que es un número real (en espacios reales). Además, con la segunda (conmutatividad) y tercera (linealidad) propiedad 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 su base dual, tendremos que el producto interno visto como un tensor dos veces covariante puede expresarse como:

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

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, i.e. 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:

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

y aplicando la definición de producto tensorial, tendremos finalmente que:

$$g_{\mu\nu} \mathbf{dx}^\mu (\vec{u}) \mathbf{dx}^\nu (\vec{v}) = g_{\mu\nu} u^\mu v^\nu.$$

6.3.1 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 $\vec{r}(t) = x^\mu(t) \odot \mathbf{e}_\mu$, $a \leq t \leq b$. En situaciones particulares, tales como trabajar en la base canónica de \mathbb{R}^3 y para el producto interno canónico tenemos que la longitud de arco se calcula como:

$$\ell = \int_a^b \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2 + \left(\frac{dx^3}{dt}\right)^2} dt.$$

Esta formulación particular puede representarse de una manera más general como:

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

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

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

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

6.4 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 $g_{\vec{u}} : V \rightarrow \mathbb{R}$ de la siguiente manera: $g_{\vec{u}}(\vec{v}) = \langle \vec{u} | \vec{v} \rangle$. Como funcional, pertenece al espacio dual de V , el V^* y consecuentemente, tendrá una representación: $g_{\vec{u}} = a_{\mu} \odot \mathbf{dx}^{\mu}$ donde para obtener las coordenadas a_{μ} deberemos calcular la 1-forma sobre los elementos de la base. Tendremos entonces:

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

Ahora bien, por un lado tenemos que $\mathbf{dx}^{\mu}(\mathbf{e}_\nu) = \delta_{\nu}^{\mu}$, ergo $g_{\vec{u}}(\mathbf{e}_\nu) = a_{\nu}$. Y por otro lado tenemos que:

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

con lo cual $a_{\nu} = u^{\mu} g_{\mu\nu}$. Y reemplazando en la forma original,

$$g_{\vec{u}} = a_{\mu} \odot \mathbf{dx}^{\mu} = u^{\nu} g_{\nu\mu} \odot \mathbf{dx}^{\mu}.$$

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 $u^{\nu} g_{\nu\mu}$. Como las 1-formas poseen coordenadas covariantes, lo que obtuvimos es: a partir de un vector contravariante de coordenadas u^{ν} , un vector covariante de coordenadas $a_{\mu} = g_{\nu\mu} u^{\nu}$.

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.

Definición 6.4.1 — Bajada de índices. El procedimiento descubierto en el punto anterior permite generalizarlo a tensores arbitrarios: dado un tensor del tipo $\binom{3}{2}$ cuyas coordenadas serán en una base $T_{\alpha\beta\gamma}^{\mu\nu}$, puede transformarse en un tensor del tipo $\binom{4}{1}$ con coordenadas obtenidas a partir de multiplicar por las coordenadas del tensor métrico y contraer en un índice:

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

De manera análoga se pueden obtener las coordenadas contravariantes de un vector covariante (i.e. de una 1-forma).

Definición 6.4.2 — 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:

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

■ **Ejemplo 6.3 — Subida de índices.** 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:

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

Ahora bien, si queremos pasar del tensor original $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ a uno del tipo $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ debemos hacer:

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

■

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.

Introducción

Cálculo operacional

- Multiplicación de tensores contravariantes
- Derivación de tensores

Derivación covariante

- Símbolos de Christoffel
- Derivada de un vector contravariante
- Derivada de un vector covariante
- Derivada de un tensor dos veces contravariante
- Derivada de un tensor dos veces covariante

Operadores diferenciales

- Gradiente
- Rotor
- Divergencia
- Laplaciano

Aplicaciones a la teoría de curvas

- Campo de tensores sobre curvas. Derivada
- Transporte paralelo
- Geodésicas

Elementos de Geometría Riemanniana

- Tensor de curvatura de Riemann
- Tensor de Ricci
- Escalar de Ricci

7. Análisis tensorial. Operadores diferenciales

Bibliografía recomendada para el capítulo: [B85; D76; L73; S68].

7.1 Introducción

Para completar la exposición sobre tensores, es necesario introducir las [coordenadas curvilíneas](#), ya que en muchos problemas aparecen cambios no lineales de coordenadas, tales como pueden ser los que involucran a las coordenadas 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 (e.g. Análisis Matemático II), los que fueron definidos exclusivamente para coordenadas cartesianas. No obstante, 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 el 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, en cambio, introduciremos las definiciones brindadas dentro de los cursos de cálculo vectorial de manera covariante, i.e., independiente del sistema de coordenadas, lo que nos permitirá trabajar a partir de definiciones que no están forzadas a ser válidas únicamente en \mathbb{R}^3 . Más aún, este marco nos permitirá conservar la naturaleza geométrica de estos objetos, perdida en enfoques menos generales.

7.2 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.

7.2.1 Multiplicación de tensores contravariantes

Producto punto. Sean **A** y **B** dos tensores dos veces contravariantes,

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

definimos el producto punto (extensión del producto interno entre vectores contravariantes) como sigue:

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

Notemos que el producto punto contrae dos índices, por lo que si se lo aplica sobre dos tensores una vez contravariantes (vectores), obtendremos un escalar (coincidiendo con el producto escalar usual). Por otro lado, recordemos que $B^{\alpha\beta} g_{\nu\alpha} = B_\nu^\beta$, luego:

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

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 (en este caso la sumatoria sobre ν).

Producto tensorial. Para este mismo caso de tensores dos veces contravariantes, el producto tensorial se define como:

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

- R** El producto punto sólo puede efectuarse entre tensores del mismo tipo, es decir, es un tipo de producto interno. El resultado de la operación resulta un tensor del mismo tipo de los factores. En cambio, el producto tensorial puede efectuarse entre tensores de diferentes tipos y el resultado es un tensor cuya contravarianza es la suma de las contravarianzas, y lo mismo con la covarianza.

7.2.2 Derivación de tensores

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](#), cabe destacar, principalmente, el carácter tensorial del [gradiente](#), del [rotor](#), de la [divergencia](#) y del [laplaciano](#).

El concepto de derivación proviene de cambio, de tasa de variación. Esto significa que es necesario definir [campos tensoriales](#), i.e., tensores dependientes de la posición.

Sin embargo, la incorporación de la variación espacial, introduce otro problema: **las bases tampoco son constantes**.

Estamos acostumbrados a que la base de un espacio vectorial es una entidad fija, la que establece el 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 \mathbb{R}^2), las bases inducidas por estos cambios 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 como sigue (ver Sección 4.1.5):

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

Si ahora nos enfocamos más desde una perspectiva de cambio de coordenadas, en vez de cambio de base, llamaremos $[\Lambda^{-1}]_\nu^\mu = \Phi_\nu^\mu$, luego, un cambio de coordenadas estará dado por $x'^\mu = \Phi_\nu^\mu x^\nu$ donde este cambio induce lo propio entre bases, a través de $\mathbf{e}'_\mu = [\Phi^{-1}]_\mu^\nu \mathbf{e}_\nu$.

Con esta reescritura, podemos escribir la tabla anterior, pero usando la matriz Φ y no Λ :

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

En el contexto de cambios lineales, las matrices Φ y Λ son constantes, i.e. las bases no cambian con la posición.

Coordenadas curvilíneas

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 \mathbb{R}^2 tenemos *las coordenadas polares* (r, θ) inducidas por la transformación no lineal:

$$\begin{cases} r &= \sqrt{x^2 + y^2}, \\ \theta &= \arctan\left(\frac{y}{x}\right). \end{cases}$$

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

Calculemos los diferenciales:

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

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

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

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

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

Para el caso particular de coordenadas polares, tendremos:

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

Con estas matrices, podemos obtener el cambio inducido de las bases:

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

R Es necesario recordar cómo 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, además, que como ya se ha adelantado los vectores base son variables, i.e., cambian punto a punto:

$$\begin{cases} \frac{\partial \mathbf{e}_r}{\partial r} = 0, \\ \frac{\partial \mathbf{e}_r}{\partial \theta} = \frac{1}{r} \odot \mathbf{e}_{\theta}, \\ \frac{\partial \mathbf{e}_{\theta}}{\partial r} = \frac{1}{r} \odot \mathbf{e}_{\theta}, \\ \frac{\partial \mathbf{e}_{\theta}}{\partial \theta} = -r \odot \mathbf{e}_r. \end{cases} \quad (7.1)$$

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 efectuaba sobre las coordenadas, dejando inalterada las bases.

■ **Ejemplo 7.1 — Tensor métrico en coordenadas polares.** Antes de pasar al estudio de la derivación, consideremos el tensor métrico en el ejemplo de coordenadas polares.

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

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

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 cartesiadas¹

¹Nótese que entendemos para las bases sin primar (las viejas), la siguiente denotación: $g_{11} = g_{xx}$ y $g_{22} = g_{yy}$.

es la identidad:

$$g'_{11} = g'_{rr} = [\Phi^{-1}]_1^1 [\Phi^{-1}]_1^1 g_{11} + [\Phi^{-1}]_1^2 [\Phi^{-1}]_1^2 g_{22} = \cos^2(\theta) + \sin^2(\theta) = 1,$$

$$g'_{22} = g'_{\theta\theta} = [\Phi^{-1}]_2^1 [\Phi^{-1}]_2^1 g_{11} + [\Phi^{-1}]_2^2 [\Phi^{-1}]_2^2 g_{22} = r^2 \sin^2(\theta) + r^2 \cos^2(\theta) = r^2.$$

Entonces, el tensor métrico se puede representar de manera invariante como:

$$\mathbf{g} = \mathbf{dx}^1 \otimes \mathbf{dx}^1 \oplus \mathbf{dx}^2 \otimes \mathbf{dx}^2 = \mathbf{dx}'^1 \otimes \mathbf{dx}'^1 \oplus r^2 \odot \mathbf{dx}'^2 \otimes \mathbf{dx}'^2,$$

o matricialmente, las coordenadas se pueden acomodar de la siguiente forma:

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

■

7.3 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 en [L73].

7.3.1 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 será una combinación lineal de los elementos de la propia base.

Esta representación la efectuaremos de manera operacional, i.e., **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:

- el vector base que se está derivando,
- la coordenada respecto de la cual se está derivando,
- 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 como:

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

donde las cantidades $\Gamma_{\mu\nu}^\alpha$ son los [símbolos de Christoffel](#).

R 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, de la siguiente manera:

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

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

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

entonces, con esta notación, los Christoffel los podemos reescribir como:

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

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

■ **Ejemplo 7.2 — Christoffel para polares.** A modo de ejemplo, en el caso de coordenadas polares tendremos:

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

los primeros dos símbolos responden a que $\frac{\partial \mathbf{e}_r}{\partial r} = 0$, el segundo par de ellos a que $\frac{\partial \mathbf{e}_r}{\partial \theta} = \frac{1}{r} \odot \mathbf{e}_\theta$ como que $\frac{\partial \mathbf{e}_\theta}{\partial r} = \frac{1}{r} \odot \mathbf{e}_\theta$ (i.e. recordemos la simetría para los Christoffel ya mencionada), y el tercer y último par de símbolos a que $\frac{\partial \mathbf{e}_\theta}{\partial \theta} = -r \odot \mathbf{e}_r$, tal y como se dejó explícito en el sistema manifiesto en la Ec. 7.1. ■

7.3.2 Derivada de un vector contravariante

Consideremos un vector contravariante $\vec{v} = v^\mu \odot \mathbf{e}_\mu$. Teniendo en cuenta que ahora los elementos de la base pueden tener derivadas parciales no nulas, podemos calcular:

$$\frac{\partial \vec{v}}{\partial x^\nu} = \frac{\partial v^\mu}{\partial x^\nu} \odot \mathbf{e}_\mu \oplus v^\mu \odot \frac{\partial \mathbf{e}_\mu}{\partial x^\nu},$$

y reemplazando la expresión para la derivada de los vectores base:

$$\frac{\partial \vec{v}}{\partial x^\nu} = \frac{\partial v^\mu}{\partial x^\nu} \odot \mathbf{e}_\mu \oplus v^\mu \Gamma_{\mu\nu}^\alpha \odot \mathbf{e}_\alpha.$$

Luego, si cambiamos los índices μ por α en el segundo término (que están sumandose, lo que los transforma en índices mudos), podemos agrupar como:

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

y, finalmente, utilizando la notación para la derivada parcial, con una coma en el subíndice para indicar la coordenada espacial respecto de la cual se deriva el vector, escribimos:

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

R 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 de coordenadas, i.e. **la derivada covariante de una coordenada contravariante** es:

$$v^\mu{}_{;v} = v^\mu{}_{,v} + \Gamma_{v\alpha}^\mu v^\alpha, \quad (7.2)$$

Con esta notación, **la derivada parcial de un vector contravariante** la denotamos como:

$$\frac{\partial \vec{v}}{\partial x^v} = v^\mu{}_{;v} \odot \mathbf{e}_\mu.$$

R 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 Tullio Levi-Civita, y lo denominó **Cálculo Diferencial Absoluto**.

7.3.3 Derivada de un vector covariante

Un vector covariante, como sabemos, es una 1-forma, i.e.:

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

Ahora bien, plantear la derivación parcial de la misma manera que lo hemos hecho para vectores contravariantes es más complicado, ya que deberíamos tener calculada la cantidad:

$$\frac{\partial \mathbf{dx}^\mu}{\partial x^v}.$$

Sin embargo, 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. Luego: $[v^\mu f_\mu]_{;v} = [v^\mu f_\mu]_{,v}$. Además, bajo la imposición de que la derivada covariante respete la regla del producto, tenemos que: $v^\mu_{;v} f_\mu + v^\mu f_{\mu;v} = v^\mu_{,v} f_\mu + v^\mu f_{\mu,v}$. Reemplazando la derivada covariante de las coordenadas contravariantes de \vec{v} (en rojo) según lo visto en la Ec. 7.2, tenemos:

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

simplificando

$$\Gamma_{\alpha v}^\mu v^\alpha f_\mu + v^\mu f_{\mu;v} = v^\mu f_{\mu,v}.$$

En el primer término del miembro de la izquierda podemos cambiar el índice α con μ , obteniendo:

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

De donde encontramos la derivada covariante de una coordenada covariante:

$$f_{\mu;v} = f_{\mu,v} - \Gamma_{\mu v}^\alpha f_\alpha.$$

Luego, la derivada parcial de un vector covariante se escribe como:

$$\frac{\partial f}{\partial x^v} = f_{\mu;v} \odot \mathbf{d}x^\mu.$$

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

7.3.4 Derivada de un tensor dos veces contravariante

Consideremos un tensor dos veces contravariante: $\mathbf{T} = t^{\mu\nu} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu$. Calculemos la derivada parcial respecto a x^λ :

$$\frac{\partial \mathbf{T}}{\partial x^\lambda} = t^{\mu\nu}_{;\lambda} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu \oplus t^{\mu\nu} \odot \frac{\partial}{\partial x^\lambda} [\mathbf{e}_\mu \otimes \mathbf{e}_\nu] = t^{\mu\nu}_{;\lambda} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu \oplus t^{\mu\nu} \odot \frac{\partial \mathbf{e}_\mu}{\partial x^\lambda} \otimes \mathbf{e}_\nu \oplus t^{\mu\nu} \odot \mathbf{e}_\mu \otimes \frac{\partial \mathbf{e}_\nu}{\partial x^\lambda}.$$

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

$$\frac{\partial \mathbf{T}}{\partial x^\lambda} = [t^{\mu\nu}_{;\lambda} + t^{\alpha\nu} \Gamma_{\alpha\lambda}^\mu + t^{\mu\alpha} \Gamma_{\alpha\lambda}^\nu] \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu.$$

Entonces, definiendo la derivada covariante de una coordenada dos veces contravariante como:

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

podemos reescribir la derivada parcial de un tensor dos veces contravariante como:

$$\frac{\partial \mathbf{T}}{\partial x^\lambda} = t^{\mu\nu}_{;\lambda} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu.$$

7.3.5 Derivada de un tensor dos veces covariante

Análogamente a lo que hicimos para obtener la derivada covariante de las coordenadas de 1-formas (Sección 7.3.3), encontraremos la derivada covariante de las coordenadas de tensores dos veces covariantes a partir de construir un escalar por contracción: $t_{\mu\nu} u^{\mu\nu}$ es un escalar, por lo que $[t_{\mu\nu} u^{\mu\nu}]_{;\lambda} = [t_{\mu\nu} u^{\mu\nu}]_{,\lambda}$.

Reemplazando las derivadas, agrupando convenientemente y simplificando las expresiones, se obtiene la derivada covariante de una coordenada dos veces covariante:

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

Entonces la derivada parcial de un tensor dos veces covariante será:

$$\frac{\partial \mathbf{T}}{\partial x^{\lambda}} = t_{\mu\nu;\lambda} \odot \mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu}.$$

Por último, y generalizando en alguna medida lo encontrado en los apartados anteriores, un tensor del tipo $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ se obtiene combinando las propiedades vistas en las secciones 7.3.2, 7.3.3, 7.3.4 y 7.3.5:

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

i.e. al margen del necesario cuidado para compensar los índices, en general por cada índice de contravarianza se suman términos con símbolos de Christoffel y por cada índice de covarianza se restan.

Terminemos entonces con un ejemplo general de aplicación de derivada covariante.

■ **Ejemplo 7.3 — Derivada covariante de un tensor $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.** Calculemos la derivada covariante de un tensor tres veces covariante y dos veces contravariante:

$$\begin{aligned} T_{\nu_1\nu_2\nu_3;\alpha}^{\mu_1\mu_2} &= T_{\nu_1\nu_2\nu_3,\alpha}^{\mu_1\mu_2} + \\ &+ \Gamma_{\beta\alpha}^{\mu_1} T_{\nu_1\nu_2\nu_3}^{\beta\mu_2} + \Gamma_{\beta\alpha}^{\mu_2} T_{\nu_1\nu_2\nu_3}^{\mu_1\beta} \\ &- \Gamma_{\nu_1\alpha}^{\beta} T_{\beta\nu_2\nu_3}^{\mu_1\mu_2} - \Gamma_{\nu_2\alpha}^{\beta} T_{\nu_1\beta\nu_3}^{\mu_1\mu_2} - \Gamma_{\nu_3\alpha}^{\beta} T_{\nu_1\nu_2\beta}^{\mu_1\mu_2}. \end{aligned}$$

■

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 de la relación que resulta con los símbolos de Christoffel, se puede obtener (después de un engorroso cálculo) la siguiente igualdad que resultará de utilidad:

$$\Gamma_{\gamma\alpha}^{\alpha} = \frac{\partial[\ln(\sqrt{g})]}{\partial x^{\gamma}} = \frac{1}{\sqrt{g}} \frac{\partial[\sqrt{g}]}{\partial x^{\gamma}} = \frac{1}{2g} \frac{\partial g}{\partial x^{\gamma}}. \quad (7.3)$$

7.4 Operadores diferenciales

Introduzcamos de manera general los operadores diferenciales, teniendo en cuenta su naturaleza geométrica, i.e. a partir de los productos (\bullet, \otimes) definidos entre tensores

(Sección 7.2.1) y la definición de derivada covariante (Sección 7.3), podremos extender los operadores diferenciales que se vieron en los cursos de Análisis Vectorial.

7.4.1 Gradiente

Al operador $\tilde{\nabla}$ se lo define como:

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

donde, para el caso de una función escalar $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, el operador nabla aplicado a la función resulta en una 1-forma diferencial que llamaremos **gradiente**, definida a partir de la derivada direccional y cuyas coordenadas son las derivadas parciales de la función (las derivadas covariantes de un escalar coinciden con las derivadas parciales):

$$\tilde{\nabla}[\phi] = [\phi]_{;\mu} \odot \mathbf{dx}^\mu.$$

Es muy común considerar al gradiente como un vector, i.e. un vector contravariante. Si fuera de esta manera:

$$\vec{\nabla}[\phi] = [\phi]_{;\mu} \odot \mathbf{e}_\mu,$$

y daría lo mismo en cualquier sistema de coordenadas, i.e.:

$$\vec{\nabla}[\phi] = \frac{\partial \phi}{\partial x} \odot \mathbf{e}_x \oplus \frac{\partial \phi}{\partial y} \odot \mathbf{e}_y = \frac{\partial \phi}{\partial r} \odot \mathbf{e}_r \oplus \frac{\partial \phi}{\partial \theta} \odot \mathbf{e}_\theta, \quad (7.4)$$

lo que es claramente falso. 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** (o vector gradiente de uso común en Análisis Vectorial) será:

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

donde indicamos con una barra al operador para indicar que es el gradiente contravariante, definido a partir del operador $\tilde{\nabla}$:

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

■ **Ejemplo 7.4 — Gradiente contravariante en coordenadas polares.** Para el caso de coordenadas polares en el plano, notemos que:

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

que es la expresión correcta para el (vector) gradiente (recuerde que $g_{\theta\theta} = r^2$), a diferencia de lo que teníamos en la Ec. 7.4. ■

Con las definiciones de los operadores $\tilde{\nabla}$ y $\vec{\nabla}$ podemos definir el gradiente de tensores en general:

$$\mathbf{Grad}[\] = \vec{\nabla} \otimes [].$$

R El operador **Grad** puede ser aplicado a tensores de cualquier tipo.

■ **Ejemplo 7.5 — Gradiente de un tensor del tipo $\binom{1}{2}$.** Aplicando el gradiente al tensor \mathbf{T} :

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

directamente obtenemos:

$$\mathbf{Grad}[\mathbf{T}] = \tilde{\nabla} \otimes [T_{\alpha}^{\mu\nu} \odot \mathbf{dx}^{\alpha} \otimes \mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu}] = T_{\alpha;\beta}^{\mu\nu} \odot \mathbf{dx}^{\beta} \otimes \mathbf{dx}^{\alpha} \otimes \mathbf{e}_{\mu} \otimes \mathbf{e}_{\nu},$$

que es un tensor del tipo $\binom{2}{2}$. ■

Definición 7.4.1 — 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 siguiente expresión:

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

7.4.2 Rotor

El rotor de un vector covariante es un vector dos veces covariante, obtenido por derivación covariante:

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

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

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

y en virtud de la simetría de los Christoffel, podemos escribir:

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

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 euclídeana, componentes covariantes y contravariantes se confunden, puesto que, como números, son los mismos.

7.4.3 Divergencia

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

$$\mathbf{Div}(\tilde{v}) = v^{\mu}_{;\mu}. \quad (7.5)$$

Recordemos primero la derivada covariante de la coordenada v^{μ} del vector \tilde{v} :

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

para la divergencia sólo interesa el caso donde $\nu = \mu$, i.e. $v^{\mu}_{;\mu}$:

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

donde hicimos uso de la Ec. 7.3, recordemos que g es el determinante de la métrica y, en virtud de que en el segundo término, α es mudo, lo podemos escribir como:

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

o de manera simplificada:

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

■ **Ejemplo 7.6 — Divergencia en coordenadas cartesianas.** Notemos que en coordenadas cartesianas, la cantidad responde a la suma:

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

que es la expresión conocida del Análisis Vectorial. ■

■ **Ejemplo 7.7 — Divergencia en coordenadas esféricas.** A partir del cambio de coordenadas usual dado por:

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

mediante cálculo directo, el elemento de distancia se obtiene fácilmente:

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

lo que evidencia que las coordenadas del tensor métrico en esféricas se pueden escribir como:

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

y entonces, el inverso se obtiene de manera trivial:

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

Luego, el determinante de la métrica es, en estas coordenadas, $g = r^4 \sin^2(\varphi)$ y la divergencia se puede calcular como:

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

que es la expresión familiar de la divergencia para esféricas. ■

Con las definiciones de los operadores $\tilde{\nabla}$ y $\vec{\nabla}$, también podemos definir la divergencia de tensores en general:

$$\mathbf{Div}[\] = \vec{\nabla} \bullet [\].$$

R A diferencia del operador **Grad**, el operador **Div** puede ser aplicado sólo a tensores contravariantes.

■ **Ejemplo 7.8 — Divergencia de un tensor del tipo $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$.** Aplicando la divergencia al tensor $\mathbf{T} = T^{\mu\nu} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu$, tendremos:

$$\mathbf{Div}[\mathbf{T}] = \vec{\nabla} \bullet [T^{\mu\nu} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu] = g^{\alpha\beta} T^{\mu\nu}{}_{;\alpha} \langle \mathbf{e}_\beta | \mathbf{e}_\mu \rangle \odot \mathbf{e}_\nu = T^{\mu\nu}{}_{;\mu} \odot \mathbf{e}_\nu$$

que es un tensor del tipo $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Notemos que para un vector contravariante, la divergencia se escribe como expresa la Ec. 7.5. ■

R 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.**

7.4.4 Laplaciano

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

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

Más aún, en su forma compacta, utilizando el determinante de la métrica, podemos escribir:

$$\nabla^2(\phi) = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \left[\sqrt{g} g^{\mu\nu} \frac{\partial \phi}{\partial x^\nu} \right]. \quad (7.6)$$

■ **Ejemplo 7.9 — Laplaciano en coordenadas esféricas.** Calculemos el Laplaciano en coordenadas esféricas, para ello tenemos el determinante y la matriz inversa del Ejemplo 7.7, 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, expandiendo, 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. ■

Como hemos hecho en las secciones anteriores con el gradiente **Grad** y la divergencia **Div**, también vamos a definir el Laplaciano haciéndonos eco de las definiciones de producto entre tensores.

A partir del vector gradiente y la divergencia, introducimos como Laplaciano al operador determinado por la siguiente relación:

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

R En virtud de las definiciones y las restricciones para los productos, el Laplaciano sólo puede calcularse a tensores estrictamente contravariantes.

■ **Ejemplo 7.10 — Laplaciano de un tensor del tipo $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$.** Sea **T** dos veces contravariante, entonces, el Laplaciano será:

$$\begin{aligned} \nabla^2[\mathbf{T}] &= \vec{\nabla} \bullet \vec{\nabla} \otimes [T^{\mu\nu} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu] = \vec{\nabla} \bullet [g^{\alpha\beta} T^{\mu\nu}{}_{;\beta} \odot \mathbf{e}_\alpha \otimes \mathbf{e}_\mu \otimes \mathbf{e}_\nu] \\ &= g^{\lambda\sigma} [g^{\alpha\beta} T^{\mu\nu}{}_{;\beta}]_{;\sigma} \langle \mathbf{e}_\lambda | \mathbf{e}_\alpha \rangle \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu \\ &= g^{\lambda\sigma} g_{\lambda\alpha} [g^{\alpha\beta} T^{\mu\nu}{}_{;\beta}]_{;\sigma} \odot \mathbf{e}_\mu \otimes \mathbf{e}_\nu. \end{aligned}$$

■

■ **Ejemplo 7.11 — Laplaciano de un campo escalar.** Consideremos ahora un campo escalar ϕ . Entonces, al Laplaciano podremos escribirlo como:

$$\begin{aligned} \nabla^2[\phi] &= \vec{\nabla} \bullet \vec{\nabla} \otimes [\phi] = \vec{\nabla} \bullet \left[g^{\alpha\beta} \frac{\partial \phi}{\partial x^\beta} \mathbf{e}_\alpha \right] = g^{\mu\nu} \left[g^{\alpha\beta} \frac{\partial \phi}{\partial x^\beta} \right]_{;\nu} \langle \mathbf{e}_\mu | \mathbf{e}_\alpha \rangle \\ &= \underbrace{g^{\mu\nu} g_{\mu\alpha}}_{\delta_\alpha^\nu} \left[g^{\alpha\beta} \frac{\partial \phi}{\partial x^\beta} \right]_{;\nu} = \left[g^{\nu\beta} \frac{\partial \phi}{\partial x^\beta} \right]_{;\nu}, \end{aligned}$$

y, utilizando la relación de los símbolos de Christoffel con la raíz del determinante, Ec. 7.3, arribamos finalmente a la expresión indicada en la Ec. 7.6. ■

7.5 Aplicaciones a la teoría de curvas

Definición 7.5.1 — Curvas. En un espacio S de Riemann de dimensión n (en general se denomina *variedad Riemanniana*), i.e., un espacio de dimensión n donde tenemos una métrica $g_{\mu\nu}$, vamos a denotar por curva a una función $x^\alpha : \mathbb{R} \rightarrow S$ definida a partir de:

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

Con estas funciones, se puede definir luego una tangente a la curva:

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

7.5.1 Campo de tensores sobre curvas. Derivada

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

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

lo que es equivalente a decir: con métrica $g_{\mu\nu}$.

Ahora bien, consideremos 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:

$$\frac{d\mathbf{T}}{dt} = \frac{\partial \mathbf{T}}{\partial x^\alpha} \odot \frac{dx^\alpha}{dt}. \quad (7.7)$$

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

■ **Ejemplo 7.12 — Tensor dos veces contravariante.** Para ilustrar la definición de derivada a lo largo de una curva, consideremos primero un tensor dos veces contravariante:

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

Entonces, dada una curva como la ya definida, la derivada sigue la expresión dada por la Ec. 7.7, que se puede escribir en coordenadas como:

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

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

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

■

El ejemplo anterior alcanza para ilustrar cómo se calcula la derivada para tensores de diferentes tipos: **siempre se utiliza la derivación covariante en las coordenadas.**

■ **Ejemplo 7.13 — Derivada de un campo vectorial.** Antes que nada, queremos hacer notar que este ejemplo está incluido naturalmente en el Ejemplo 7.12, pero dado que estas herramientas pedagógicas son excelentes para la asimilación de ideas, igualmente calculemos la derivada a lo largo de la curva. Al vector \vec{v} lo escribimos de manera tensorial, i.e. como un tensor una vez contravariante: $\vec{v} = v^\mu \odot \mathbf{e}_\mu$. Entonces, la derivada en cada punto de la curva es:

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

Finalmente, las coordenadas del vector cambian a lo largo de la curva según la siguiente ley:

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

■

7.5.2 Transporte paralelo

Una aplicación de lo visto en la sección anterior (Sección 7.5.1) es lo que se denomina como **transporte paralelo** de un vector a lo largo de una curva sobre la variedad Riemanniana. Consideremos una curva sobre una variedad Riemanniana. Consideremos, además, 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 entre 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 [B85].

Definición 7.5.2 — Transporte paralelo. Dado el vector \vec{v} , este será **transportado paralelamente** a lo largo de la curva si y sólo si:

$$\frac{d\vec{v}}{dt} = 0 \rightarrow v^\mu{}_{;\alpha} \frac{dx^\alpha}{dt} = [v^\mu{}_{,\alpha} + \Gamma_{\nu\alpha}^\mu v^\nu] \frac{dx^\alpha}{dt} = 0, \quad \forall \mu = 1, 2, \dots, n, \quad (7.8)$$

es decir, el vector no cambia en su transporte a lo largo de la curva.

7.5.3 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 que se tienen y por ende no profundizaremos en el tema.

Definición 7.5.3 — Geodésicas. Otra formulación para la definición de geodésica es aquella curva cuya tangente se transporta paralelamente sobre ella.

La anterior definición es una interpretación que pretende generalizar el concepto de recta, la cual en un espacio plano satisface esta definición. Luego, con esta definición, y recordando la condición que satisface un vector que se transporta paralelamente (Ec. 7.8), 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:

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

que en formato simplificado se escribe como:

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

7.6 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 (dándole estructura a dichas variedades).

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.

7.6.1 Tensor de curvatura de Riemann

Sin entrar en los detalles a partir de los cuales se obtiene, vamos a tomar como definición del [tensor de curvatura de Riemann](#), al tensor cuyas coordenadas $R_{\beta\mu\nu}^\alpha$ se obtienen a partir de derivadas segundas de la métrica, de la siguiente manera:

$$R_{\beta\mu\nu}^\alpha = \frac{1}{2} g^{\alpha\sigma} [g_{\sigma\nu, \beta\mu} - g_{\sigma\mu, \beta\nu} + g_{\beta\mu, \sigma\nu} - g_{\beta\nu, \sigma\mu}].$$

Una variedad se denomina *plana* si la curvatura de Riemann es nula, i.e.:

$$R_{\beta\mu\nu}^\alpha = 0.$$

A partir de la bajada de índices podemos definir el tensor:

$$R_{\alpha\beta\mu\nu} = g_{\alpha\lambda} R_{\beta\mu\nu}^\lambda,$$

el cual es un tensor de curvatura de Riemann completamente covariante, y como tal satisface las siguientes identidades:

1. $R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}$,
2. $R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0$ (que corresponde a la propiedad cíclica).

7.6.2 Tensor de Ricci

A partir del tensor de Riemann, se define el [tensor de Ricci](#) mediante la contracción:

$$R_{\alpha\beta} = R_{\alpha\mu\beta}^\mu = R_{\beta\alpha}.$$

7.6.3 Escalar de Ricci

Con el tensor de Ricci, se define el [escalar de Ricci](#) a través de:

$$R = g^{\alpha\beta} R_{\alpha\beta}.$$

Los parámetros de curvatura mencionados en las Secciones 7.6.1, 7.6.2 y 7.6.3 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.

Energía cinética

Energía cinética de una partícula

Energía cinética de un cuerpo rígido. Tensor de inercia

Elasticidad. Tensor de deformación

Desarrollo del potencial electrostático. Tensor momento cuadrupolar

8. Tensores en Física

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

8.1 Energía cinética

8.1.1 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 el escalar velocidad de la partícula, asumiendo movimiento rectilíneo.

Si la partícula, en cambio, puede moverse en el espacio \mathbb{R}^3 , la expresión de la energía cinética debe escribirse 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 como lo hicimos 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 la 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, como escribimos el vector velocidad \vec{v} de manera invariante $v^\mu \odot \mathbf{e}_\mu$ (sólo al hacer esto podemos aplicar la forma bilineal del tensor métrico, pues esta se aplica sobre componentes), 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** al representar a los sistemas dinámicos a partir de, e.g. q^1, q^2, q^3 en \mathbb{R}^3 . En un sistema de coordenadas generalizadas, la energía cinética se escribe, entonces, como:

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

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, y además como el tensor métrico es simétrico no está más el factor $1/2$ ¹.

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.

R 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.

8.1.2 Energía cinética de un cuerpo rígido. Tensor de inercia

Calculemos la energía cinética de rotación de un cuerpo rígido constituido por N partículas y sea $\vec{\omega}$ la velocidad angular del cuerpo. Como estamos en presencia de un cuerpo rígido, todas las partículas que lo constituyen están afectadas a esa velocidad angular. Entonces, 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 respecto al producto mixto:

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

¹En efecto, e.g. para \mathbb{R}^2 tendríamos: $T = \frac{1}{2} m [g_{11} \dot{q}^1 \dot{q}^1 + g_{22} \dot{q}^2 \dot{q}^2 + g_{12} \dot{q}^1 \dot{q}^2 + g_{21} \dot{q}^2 \dot{q}^1]$.

que a su vez puede escribirse como:

$$\langle (|\vec{\omega}|^2 \vec{r}_\ell - \langle \vec{\omega} | \vec{r}_\ell \rangle \vec{\omega}) | \vec{r}_\ell \rangle.$$

Entonces, la energía cinética para la partícula ℓ -ésima:

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

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

Y reemplazando, obtenemos:

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

donde el subíndice l indica la partícula y no tiene nada que ver con los índices para los vectores o covectores. Entonces, para la partícula ℓ -ésima se puede escribir,

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

Y sumando sobre todas las partículas que constituyen el cuerpo:

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

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 sobre el volumen \mathcal{V} , 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 y 0 en el resto (i.e. $g_{\mu\nu} = \delta_{\mu\nu}$), podemos escribir las coordenadas² del tensor energía cinética (al que comunmente se lo denomina **tensor de inercia**), de la siguiente manera:

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

donde hemos llamado a las coordenadas contravariantes del vector: $x^1 = x$, $x^2 = y$, $x^3 = z$ (no confundir la notación con elevar al cuadrado). No obstante, dada la simetría producida por lo diagonal de la métrica (recuerde que $g_{\mu\nu} = \delta_{\mu\nu}$), las coordenadas covariantes del vector posición coinciden con las coordenadas del punto; esto es, también se cumple

²Notemos que el tensor I es una forma bilineal, por lo que se lo aplica sobre dos vectores de una base contravariante para encontrar sus coordenadas $I_{\mu\nu}$. En dicha base, ω^μ y ω^ν son las coordenadas del vector velocidad angular (asociadas a la métrica $g_{\mu\nu}$), luego la contracción de índices entre estas cantidades es lo que construye la cantidad invariante (en este caso, escalar) de la energía cinética.

que: $x_1 = x$, $x_2 = y$, $x_3 = z$.

Repasando, la energía cinética de rotación de un cuerpo rígido toma la siguiente forma:

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

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

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

8.2 Elasticidad. Tensor de deformación

Sea $\vec{u} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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, $\|\vec{PQ}\| = h$. Supongamos entonces, que P se desplaza a $P'(\mathbf{x}') = P'(\mathbf{x} + \vec{u}_1)$ y que Q se desplaza a $Q'(\mathbf{y}') = Q'(\mathbf{y} + \vec{u}_2)$.

Luego:

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

Notemos que $\vec{u}_1 = \vec{u}(\mathbf{x})$ y $\vec{u}_2 = \vec{u}(\mathbf{y})$, dado que son los vectores desplazamiento de \mathbf{x} y \mathbf{y} , respectivamente. 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

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

donde el índice ν indica la componente (no especificamos la base) de la derivada direccional, mientras que el μ , la componente de \vec{u} . Calculemos ahora la norma de $\vec{P'Q'}$:

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

y a primer orden en $\|(\vec{u}_2 - \vec{u}_1)\|$, $\|\vec{P'Q'}\|^2 = h^2 \langle \vec{\ell} | \vec{\ell} \rangle + 2h \langle \vec{\ell} | (\vec{u}_2 - \vec{u}_1) \rangle$.

Por otro lado, reemplazando el desarrollo de Taylor para $\vec{u}_2 - \vec{u}_1$, también obtenemos:

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

Aplicando raíz cuadrada a ambos miembros:

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

y finalmente, si usamos el desarrollo de Taylor para la raíz cuadrada: $\sqrt{1+x} \approx 1 + \frac{x}{2}$, podremos escribir:

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

Definimos, entonces, la **deformación** del sólido como:

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

Entonces, reemplazando,

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

Notemos que el último factor, es la coordenada covariante del vector \vec{u} , con lo cual podemos reescribir la ecuación como:

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

donde si llamamos:

$$e_{\alpha\nu} = \frac{\partial u_\alpha}{\partial x^\nu},$$

define las coordenadas del **tensor de deformación**, luego, la deformación se calcula como:
 $D = e_{\alpha\nu} \ell^\alpha \ell^\nu$.

8.3 Desarrollo del potencial electrostático. Tensor momento cuadrupolar

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

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

donde las variables primadas indican la posición de la distribución de la carga, o punto carga, las sin primar, la posición donde se quiere calcular el campo, o punto campo.

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 simplificación.

Antes de hacer el desarrollo, consideremos:

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

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 (ver Sección 6.1.2), 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,

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

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

$$\begin{aligned} \langle\vec{r}|\vec{r}'\rangle^2 &= g_{\mu\alpha}g_{\nu\beta}x^\mu x^\nu x'^\alpha x'^\beta, \\ \|\vec{r}\|^2\|\vec{r}'\|^2 &= g_{\mu\nu}g_{\alpha\beta}x^\mu x^\nu x'^\alpha x'^\beta. \end{aligned}$$

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 \mathbb{R}^3 , obtenemos la aproximación del potencial electrostático:

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

donde

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

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

y

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

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

Introducción. Definiciones
El polinomio característico
 Multiplicidad de las raíces
Operador Adyunto
 Matrices hermiticas. Matrices simétricas
 Aspecto metodológico para la diagonalización
 de un operador simétrico.
Operadores unitarios

9. Autovalores y autovectores

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

9.1 Introducción. Definiciones

En diversos problemas nos encontramos con representaciones matriciales. Sistemas de ecuaciones lineales, sistemas 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, es factible una representación diagonal de las mismas, esto es, que las matrices involucradas sean de la forma:

$$\mathbf{A} = \begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}, \tag{9.1}$$

lo que, como veremos a continuación, puede simplificar muchísimo la resolución de problemas.

En particular, un sistema de ecuaciones diferenciales lineales ordinarias, como e.g.:

$$\begin{aligned} \frac{dx}{dt} &= a_1 x + b_1 y, \\ \frac{dy}{dt} &= a_2 x + b_2 y, \end{aligned}$$

resulta simple de resolver si está en una forma diagonal (en otras palabras, si la matriz asociada al sistema es 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. Otro problema puede ser el de exponenciar una matriz, e.g. sea \mathbf{A} de la Ex. 9.1, encontrar \mathbf{A}^{100} . Al ser diagonal, no es necesario efectuar 99 productos de matrices, sino simplemente elevar a la 100 cada una de sus entradas no nulas:

$$\mathbf{A}^{100} = \begin{pmatrix} d_1^{100} & 0 & 0 & \cdots & 0 \\ 0 & d_2^{100} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & d_n^{100} \end{pmatrix}, \quad (9.2)$$

lo que facilita enormemente cualquier algoritmo computacional que involucre este tipo de operaciones.

En este capítulo, entonces, nos dedicaremos a estudiar técnicas a partir de las cuales podamos determinar la posibilidad de expresar las matrices cuadradas de manera diagonal, ó, en su defecto, escribir estas matrices lo más diagonal posible.

Definición 9.1.1 — Autovalores y autovectores. 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 , e.g. \vec{v} , que satisface:

$$T(\vec{v}) = \lambda \odot \vec{v}.$$

Además, al vector \vec{v} que satisface la ecuación anterior se lo denomina **autovector**.

R A partir de la definición, podemos hacer las siguientes observaciones: un autovalor puede ser nulo; un autovector es no nulo; no necesariamente un autovalor posee un único autovector asociado.

Definición 9.1.2 — Espacio propio. 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 $\vec{w} = c \odot \vec{u} + \vec{v}$ es también un autovector asociado a λ , veamos:

$$T(\vec{w}) = T(c \odot \vec{u} \oplus \vec{v}) = c \odot T(\vec{u}) \oplus T(\vec{v}) = c \lambda \odot \vec{u} \oplus \lambda \odot \vec{v} = \lambda \odot (c \odot \vec{u} \oplus \vec{v}) = \lambda \odot \vec{w}.$$

A partir de estas definiciones, podemos probar el siguiente resultado:

Teorema 9.1.1 Sea T un operador lineal sobre un espacio V de dimensión finita n , y sea λ un escalar del cuerpo K , las siguientes afirmaciones son equivalentes:

1. λ es un autovalor de T ;
2. el operador $T - \lambda I_d$ es singular;
3. si \mathbf{T} es la matriz asociada a T respecto a una determinada base, entonces:

$$\det(\mathbf{T} - \lambda \mathbf{I}_d) = 0$$

con \mathbf{I}_d la matriz identidad de dimensión $n \times n$.

Este teorema es fundamental porque establece la metodología de trabajo con relación al estudio de autovalores y autovectores.

- R** Un aspecto a resaltar es el siguiente: la primera y la segunda afirmación del teorema se refieren a T como operador, mientras que la tercera trabaja con su representación, en este caso la matriz asociada.

Para demostrar el Teorema 9.1.1, basta con demostrar que de 1) \rightarrow 2), de 2) \rightarrow 3) y que de 3) \rightarrow 1). Comencemos:

- 1) \rightarrow 2). Si λ es un autovalor, existe un autovector no nulo, \vec{v} , tal que $T(\vec{v}) = \lambda \odot \vec{v}$. Dado que tenemos en el espacio de operadores lineales, operaciones suma y producto, podemos escribir como operador:

$$(T - \lambda I_d)(\vec{v}) = 0 \quad \text{al operador nulo.} \quad (9.3)$$

Esto implica que el operador $(T - \lambda I_d)$ admite un vector no nulo en su núcleo, entonces, no puede ser biyectivo. Por lo tanto, por definición, es singular.

- 2) \rightarrow 3). Dada una base para el espacio V , $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, al escribir el operador $(T - \lambda I_d)$, este 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 $\vec{v} = v^\mu \odot \mathbf{e}_\mu$ podemos escribir la representación matricial como:

$$[(T - \lambda I_d)]_{\mathcal{B}} \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (9.4)$$

Entonces, por la teoría de sistema de ecuaciones lineales, si un sistema homogéneo admite otra solución que no sea la trivial, el sistema no puede tener solución única, por lo tanto, $\det(\mathbf{T} - \lambda \mathbf{I}_d) = 0$.

- 3) \rightarrow 1). Si $\det(\mathbf{T} - \lambda \mathbf{I}_d) = 0$, entonces existe un vector columna no nulo $\vec{v} = v^\mu \odot \mathbf{e}_\mu$ que es solución del sistema lineal dado por la Ec. 9.4. Ahora bien, considerando la matriz $(\mathbf{T} - \lambda \mathbf{I}_d)$, como la matriz de una transformación lineal, tenemos que para este vector \vec{v} , se cumple la Ec. 9.3, o lo que es lo mismo:

$$T(\vec{v}) = \lambda \odot \vec{v},$$

lo que implica que λ es un autovalor.

Esto concluye la demostración del Teorema 9.1.1.

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 3) del Teorema 9.1.1 calculamos los autovectores resolviendo la ecuación algebraica:

$$\det(\mathbf{T} - \lambda \mathbf{I}_d) = 0.$$

Notemos que al calcular el determinante, lo que se obtiene es una ecuación en λ ,

$$p(\lambda) = \det(\mathbf{T} - \lambda \mathbf{I}_d) = 0. \quad (9.5)$$

Esta ecuación se denomina **ecuación característica**, y al polinomio $p(\lambda)$, **polinomio característico**. Queda claro que 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:

$$\text{Nu}(T - \lambda I_d), \text{ i.e. los } \vec{v} \text{ tal que } (T - \lambda I_d)(\vec{v}) = 0,$$

o en otras palabras, los espacios propios los obtendremos a partir de los generadores (autovectores) extraídos de la ecuación anterior. Estos generadores del núcleo estarán dados en coordenadas.

■ **Ejemplo 9.1 — Ejercicio tomado del [S07].** Consideremos la matriz:

$$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}.$$

La ecuación característica es:

$$\det\left(\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0,$$

operando,

$$\det\left(\begin{bmatrix} 4-\lambda & 2 \\ 3 & 3-\lambda \end{bmatrix}\right) = (4-\lambda)(3-\lambda) - 6 = \lambda^2 - 7\lambda + 6 = 0.$$

Las soluciones a la ecuación característica son: $\lambda_1 = 1$ y $\lambda_2 = 6$. Para obtener las coordenadas de los autovectores, hacemos, para el primero de los autovalores, i.e. $\lambda = 1$:

$$\begin{bmatrix} 4-1 & 2 \\ 3 & 3-1 \end{bmatrix} \begin{bmatrix} v_1^1 \\ v_1^2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} v_1^1 \\ v_1^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

donde los subíndices para las coordenadas del autovector indican el autovalor asociado. La solución (que obviamente no es única) para este 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:

$$S_{\lambda_1} = \overline{\{\mathbf{e}_1 \oplus -\frac{3}{2} \odot \mathbf{e}_2\}} \quad \text{o, equivalentemente,} \quad S_{\lambda_1} = \overline{\{2 \odot \mathbf{e}_1 \oplus -3 \odot \mathbf{e}_2\}}.$$

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 es $v_2^1 = v_2^2$, lo que implica que el espacio propio asociado al autovalor $\lambda_2 = 6$, es:

$$S_{\lambda_2} = \overline{\{\mathbf{e}_1 \oplus \mathbf{e}_2\}}.$$

■

R 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 \mathbb{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 $2 \odot \mathbf{e}_1 \oplus -3 \odot \mathbf{e}_2$ y $\mathbf{e}_1 \oplus \mathbf{e}_2$ 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 $\mathcal{B}' = \{\mathbf{e}'_1, \mathbf{e}'_2\}$ (i.e. en la base de autovectores) donde $\mathbf{e}'_1 = 2 \odot \mathbf{e}_1 \oplus -3 \odot \mathbf{e}_2$ y $\mathbf{e}'_2 = \mathbf{e}_1 \oplus \mathbf{e}_2$, tendremos que tener en cuenta que,

$$\mathbf{e}'_\mu = \Lambda_\mu^\nu \mathbf{e}_\nu, \text{ donde } \Lambda = \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix},$$

con Λ la matriz cuyas entradas son las coordenadas de los autovectores en la base \mathcal{B} dispuestas por columnas. Luego, la matriz inversa de Λ será:

$$\Lambda^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}.$$

Con estas matrices, la matriz asociada a la transformación lineal en esta nueva base se obtiene como:

$$\mathbf{T}_{\mathcal{B}'} = \Lambda^{-1} \mathbf{T}_{\mathcal{B}} \Lambda,$$

reemplazando,

$$\mathbf{T}_{\mathcal{B}'} = \begin{bmatrix} \frac{1}{5} & -\frac{1}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix},$$

cuyo resultado es:

$$\mathbf{T}_{\mathcal{B}'} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

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. Ahora veremos un resultado general.

Teorema 9.1.2 — Independencia Lineal. 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{v}_1 y \vec{v}_2 dos vectores asociados a diferentes autovalores λ_1 y λ_2 . Consideremos la combinación lineal que de el vector nulo:

$$\alpha \odot \vec{v}_1 \oplus \beta \odot \vec{v}_2 = \vec{0}.$$

Aplicando la transformación lineal T , tenemos:

$$\alpha \odot T(\vec{v}_1) \oplus \beta \odot T(\vec{v}_2) = \alpha \lambda_1 \odot \vec{v}_1 \oplus \beta \lambda_2 \odot \vec{v}_2 = \vec{0}.$$

Si la combinación original es multiplicada por $-e.g. \lambda_2$ y restamos ambas ecuaciones, tendremos:

$$(\lambda_2 - \lambda_1) \alpha \odot \vec{v}_1 = \vec{0},$$

como el vector \vec{v}_1 es no nulo, y los autovalores también, nos queda que α debe ser cero, lo que implicará que β también deba serlo. En otras palabras, los vectores \vec{v}_1 y \vec{v}_2 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 $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k, \vec{v}_{k+1}$. Consideremos la combinación lineal:

$$\sum_{\mu=1}^{k+1} \alpha^\mu \odot \vec{v}_\mu = \vec{0}.$$

Aplicando la transformación T y sabiendo que $T(\vec{v}_\mu) = \lambda_\mu \odot \vec{v}_\mu$, tendremos que:

$$\sum_{\mu=1}^{k+1} \alpha^\mu \lambda_\mu \odot \vec{v}_\mu = \vec{0}.$$

Si a la combinación original la multiplicamos por λ_{k+1} y restamos, obtenemos:

$$\sum_{\mu=1}^k \alpha^\mu (\lambda_\mu - \lambda_{k+1}) \odot \vec{v}_\mu = \vec{0}, \text{ notemos que el último término se cancela.}$$

Además, por hipótesis inductiva, los k vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ son LI, por lo que 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 $\vec{v}_\mu, \mu = 1, \dots, k+1$ son linealmente independientes.

Este Teorema, da lugar a la siguiente definición:

Definición 9.1.3 — Diagonalización. 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 (tal y como ya fue mencionado).

- R 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.
- R 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 9.1.3 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} , tendremos que $\det(\mathbf{A} - \lambda \mathbf{I}_d) = 0$. Por otro lado, si \mathbf{B} es una matriz semejante a \mathbf{A} , tendremos:

$$\mathbf{B} = \Lambda^{-1} \mathbf{A} \Lambda,$$

donde Λ es la matriz de cambio de base (cuyas columnas se corresponden con las coordenadas de los autovectores) que ya introdujimos en la Sección 3.1.2. Calculemos entonces, $\det(\mathbf{B} - \lambda \mathbf{I}_d)$. Tenemos:

$$\det(\mathbf{B} - \lambda \mathbf{I}_d) = \det(\Lambda^{-1} \mathbf{A} \Lambda - \lambda \mathbf{I}_d) = \det(\Lambda^{-1} \mathbf{A} \Lambda - \lambda \Lambda^{-1} \Lambda).$$

Tomando factor común, Λ^{-1} a izquierda, y Λ a derecha, tenemos:

$$\det(\mathbf{B} - \lambda \mathbf{I}_d) = \det[\Lambda^{-1} (\mathbf{A} - \lambda \mathbf{I}_d) \Lambda].$$

Entonces, $\det(\mathbf{A} - \lambda \mathbf{I}_d) = 0$ implica que $\det(\mathbf{B} - \lambda \mathbf{I}_d) = 0$. Luego, λ es también autovalor de \mathbf{B} .

R Si dos matrices son semejantes, es porque **representan al mismo operador lineal**, sólo que las matrices asociadas están escritas en bases diferentes.

9.2 El polinomio característico

La ecuación característica (Eq. 9.5) 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. Luego, como autovectores asociados a distintos autovalores son linealmente independientes, estos 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 9.1, que ya fue analizado.

9.2.1 Multiplicidad de las raíces

En el caso 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 9.2.1 — Multiplicidad geométrica. 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.

9.3 Operador Adjunto

Aplicando lo que hemos visto para operadores lineales en un espacio euclídeo, vamos a definir un operador llamado *adjunto*, y lo haremos a partir del siguiente teorema que garantiza su existencia y unicidad.

Teorema 9.3.1 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 vector \vec{v} , podemos afirmar que $f : V \rightarrow \mathbb{R}$ 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 de los vectores \vec{u} y \vec{v} 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:

$$\langle T(\vec{v}) | \vec{u} \rangle = v^\lambda [T^*]_\mu^\nu u^\mu g_{\lambda\nu} = v^\lambda [T^*]_\mu^\nu u^\mu \langle \mathbf{e}_\lambda | \mathbf{e}_\nu \rangle = \langle \vec{v} | T^*(\vec{u}) \rangle,$$

llegando 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 complejos, $\langle \vec{v} | \vec{u} \rangle = v^\lambda \bar{u}^\mu g_{\lambda\mu}$, lo que llevaría en un análisis completamente análogo, a la existencia del operador adjunto, cuya matriz asociada está dada por:

$$\mathbf{T}^* = \overline{\mathbf{T}}^t,$$

donde la barra indica conjugación compleja.

9.3.1 Matrices hermíticas. Matrices simétricas

Ahora consideraremos un caso particular de operadores adjuntos, los operadores hermíticos.

Definición 9.3.1 — Operador Hermítico. Vamos a definir como operador *hermítico* a todo operador que satisfice

$$T^* = T.$$

R Podemos notar que para el caso de espacios reales, los operadores hermíticos son simétricos. O en términos matriciales,

$$T^t = T.$$

Para matrices hermíticas, tenemos los siguientes resultados:

Teorema 9.3.2 — Autovalores de operadores hermíticos. Los operadores hermíticos tienen autovalores reales.

En efecto, consideremos el autovalor λ y su autovector \vec{v} :

$$\langle T(\vec{v})|\vec{v} \rangle = \langle \vec{v}|T^*(\vec{v}) \rangle = \langle \vec{v}|T(\vec{v}) \rangle \rightarrow \langle \lambda \vec{v}|\vec{v} \rangle = \langle \vec{v}|\lambda \vec{v} \rangle \text{ luego } \lambda \langle \vec{v}|\vec{v} \rangle = \overline{\lambda} \langle \vec{v}|\vec{v} \rangle.$$

Consecuentemente, $\lambda = \overline{\lambda}$, lo que implica que $\lambda \in \mathbb{R}$.

Otro resultado importante lo aporta el siguiente teorema:

Teorema 9.3.3 — Ortogonalidad. Los autovectores de operadores hermíticos asociados a distintos autovalores son ortogonales.

En efecto, sean \vec{v} y \vec{u} autovectores asociados a los autovalores λ y ξ , respectivamente. Calculemos,

$$\langle T(\vec{v})|\vec{u} \rangle = \langle \vec{v}|T^*(\vec{u}) \rangle = \langle \vec{v}|T(\vec{u}) \rangle \rightarrow \langle \lambda \vec{v}|\vec{u} \rangle = \langle \vec{v}|\xi \vec{u} \rangle,$$

luego:

$$\lambda \langle \vec{v}|\vec{u} \rangle = \overline{\xi} \langle \vec{v}|\vec{u} \rangle = \xi \langle \vec{v}|\vec{u} \rangle \rightarrow (\lambda - \xi) \langle \vec{v}|\vec{u} \rangle = 0,$$

y como los autovalores son distintos, tenemos que \vec{v} y \vec{u} deben ser ortogonales.

Finalmente, podemos afirmar lo siguiente:

Teorema 9.3.4 — Diagonalización. Los operadores hermíticos son diagonalizables. En términos de matrices reales, tenemos que **toda matriz simétrica es diagonalizable**^a.

^aLa demostración de este teorema tiene una complejidad estrictamente técnica, y la dejamos para que sea leída en los textos de la bibliografía recomendada.

R 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.

9.3.2 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.

En cambio, si tenemos raíces múltiples para el polinomio característico, podremos operar de la siguiente manera:

- 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, e.g., de Gram–Schmidt.

9.4 Operadores unitarios

Un ejemplo final de operador.

Definición 9.4.1 — Operador unitario. Un operador se dice *unitario* si preserva el 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 9.4.1 — Inverso. Un operador unitario tiene por inverso, su adjunto.

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 .

R Para el caso de operadores sobre espacios euclídeos en los reales, tendremos que el inverso de un operador adjunto es su transpuesto.

▪ **Ejemplo 9.2 — Rotación en \mathbb{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:

$$R_\theta^{-1} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = R'_\theta = R_{(-\theta)},$$

que no solo coincide con la transpuesta, sino que también indica que una rotación de ángulo $-\theta$ es la operación inversa a rotar θ . ■

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 \mathbf{T} , entonces al realizar un

cambio de coordenadas ortogonales cuya matriz es \mathbf{Q} , la matriz asociada en la nueva base será:

$$\mathbf{T}' = \mathbf{Q}' \mathbf{T} \mathbf{Q}.$$

Matrices relacionadas a través de esta relación se las denomina [ortogonalmente semejantes](#).

10. Formas canónicas

Bibliografía recomendada para el capítulo: [A78; F71; L73; R73; S07].

10.1 Forma canónica de Jordan

Hemos visto en el capítulo anterior 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 del polinomio característico–, la matriz es diagonalizable (Sección 9.2.1). Si estos números no coinciden, es porque **la multiplicidad geométrica es menor a la multiplicidad algebraica**.

Veamos qué se puede hacer en estos casos para expresar al operador (en su representación matricial) en una forma canónica que, por supuesto, ya no será diagonal.

10.1.1 Matriz de bloque de Jordan

Para un escalar dado, λ , se define la **matriz de bloques de Jordan**, $\mathbf{B}_k(\lambda)$, a la matriz:

$$\mathbf{B}_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} = \lambda \mathbf{I}^{k \times k} + \mathbf{N}_k$$

donde:

$$\mathbf{N}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Podemos notar que $(\mathbf{N}_k)^k = 0$ (i.e. la k -veces aplicación sucesiva de \mathbf{N}_k da como resultado el valor nulo) y se denomina **nilpotente**.

R En general, toda matriz triangular con ceros en la diagonal es nilpotente.

10.1.2 Matriz de Jordan

Una matriz de Jordan, está compuesta por matrices de bloque de Jordan de la forma:

$$\mathbf{J} = \begin{bmatrix} \mathbf{B}_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & \mathbf{B}_2(\lambda_1) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathbf{B}_\ell(\lambda_\ell) \end{bmatrix}.$$

R Una matriz diagonal es una matriz de Jordan. En efecto, está compuesta por matrices de bloques de Jordan de 1×1 .

■ **Ejemplo 10.1** Sea, por ejemplo:

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 7 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} & 0 \\ 0 & 7 \end{bmatrix}.$$

Es decir, \mathbf{A} está compuesta por dos matrices de bloque de Jordan, la primera de 2×2 , y la segunda de 1×1 . ■

■ **Ejemplo 10.2** Otro ejemplo, sea la siguiente matriz:

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} & 0 & 0 & 0 \\ 0 & [3] & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} & 0 \\ 0 & 0 & 0 & \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix} \end{bmatrix}.$$

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 el que damos en el siguiente ejemplo.

■ **Ejemplo 10.3** Sea: $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} [3] & 0 \\ 0 & \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \end{bmatrix}$,

donde nótese que la matriz está compuesta por dos bloques de Jordan. ■

10.1.3 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. En el caso del Ejemplo 10.3, la matriz posee un único autovalor, el 3, el cual posee una multiplicidad geométrica igual a 2, dando como resultado la matriz presentada.

Cuando tenemos una matriz diagonalizable, lo que procuramos son los autovectores que son base del espacio y 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 10.1.1 Sea \mathbf{A} una matriz cuadrada. Entonces siempre existe una matriz Λ tal que:

$$\Lambda^{-1} \mathbf{A} \Lambda = \mathbf{J},$$

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.

Procedimiento para la obtención del bloque de Jordan de una matriz 2×2

Sea \mathbf{A} una matriz 2×2 cuyo único autovalor es λ con multiplicidad algebraica igual 2, pero que el subespacio asociado a λ tenga sólo dimensión 1. Sea \vec{v}_1 un autovector asociado a λ . Consideremos, entonces, un vector \vec{v}_2 definido a través de la relación:

$$(\mathbf{A} - \lambda \mathbf{I})\vec{v}_2 = \vec{v}_1.$$

Debemos demostrar que:

1. \vec{v}_2 existe,
2. \vec{v}_2 y \vec{v}_1 son linealmente independientes,
3. la matriz asociada en la base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$, es una matriz de Jordan.

1. Para demostrar la existencia de \vec{v}_2 , tomemos un vector \vec{u} que **no sea autovector de \mathbf{A}** . Sea \vec{w} definido a través de $\vec{w} = (\mathbf{A} - \lambda \mathbf{I})\vec{u}$. Notemos entonces que \vec{w} no es nulo, ya que \vec{u} no es un autovector. Calculemos $(\mathbf{A} - \lambda \mathbf{I})\vec{w}$:

$$(\mathbf{A} - \lambda \mathbf{I})\vec{w} = (\mathbf{A} - \lambda \mathbf{I})(\mathbf{A} - \lambda \mathbf{I})\vec{u} = (\mathbf{A} - \lambda \mathbf{I})^2\vec{u}.$$

Ahora bien, el polinomio característico asociado a \mathbf{A} será entonces:

$$p(x) = (x - \lambda)^2.$$

El Teorema de Cayley–Hamilton establece que el polinomio característico evaluado en el operador debe dar como resultado el operador nulo, luego:

$$p(\mathbf{A}) = \mathbf{0},$$

con lo que demostramos que $(\mathbf{A} - \lambda \mathbf{I})\vec{w} = \mathbf{0}$ o lo que es lo mismo que \vec{w} es un autovector. Entonces, si \vec{w} es un autovector, podemos definir un vector \vec{u} que cumpla con $\vec{w} = (\mathbf{A} - \lambda \mathbf{I})\vec{u}$, lo que prueba la primera parte de la demostración, si asociamos $\vec{w} \rightarrow \vec{v}_1$ y $\vec{u} \rightarrow \vec{v}_2$.

2. Para la segunda parte, supongamos:

$$c_1 \odot \vec{v}_1 \oplus c_2 \odot \vec{v}_2 = \vec{0}.$$

Apliquemos el operador $(\mathbf{A} - \lambda \mathbf{I})$,

$$(\mathbf{A} - \lambda \mathbf{I})(c_1 \odot \vec{v}_1 \oplus c_2 \odot \vec{v}_2) = \vec{0},$$

entonces,

$$c_1 \odot (\mathbf{A} - \lambda \mathbf{I})\vec{v}_1 \oplus c_2 \odot (\mathbf{A} - \lambda \mathbf{I})\vec{v}_2 = \vec{0}.$$

Como \vec{v}_1 es un autovector, $(\mathbf{A} - \lambda \mathbf{I})\vec{v}_1 = \vec{0}$ y además, $(\mathbf{A} - \lambda \mathbf{I})\vec{v}_2 = \vec{v}_1$ entonces,

$$(\mathbf{A} - \lambda \mathbf{I})(c_1 \odot \vec{v}_1 \oplus c_2 \odot \vec{v}_2) = c_1 \odot \vec{0} \oplus c_2 \odot \vec{v}_1 = \vec{0},$$

con lo que $c_2 = 0$ y entonces, fuerza a que $c_1 = 0$. En otras palabras, los vectores son linealmente independientes.

3. Finalmente, obtengamos la matriz del operador, pero en la base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$. Dado que $\mathbf{A}\vec{v}_1 = \lambda \odot \vec{v}_1$, al escribir la matriz asociada en la base elegida tendremos que la primera columna será $\begin{bmatrix} \lambda \\ 0 \end{bmatrix}$. Por otro lado, a partir de la expresión que establece que \vec{v}_2 se obtiene de: $(\mathbf{A} - \lambda \mathbf{I})\vec{v}_2 = \vec{v}_1$, tenemos que $\mathbf{A}\vec{v}_2 = \vec{v}_1 \oplus \lambda \odot \vec{v}_2$, con lo que la segunda columna será $\begin{bmatrix} 1 \\ \lambda \end{bmatrix}$ por lo que la matriz asociada en esta base será:

$$\mathbf{A}_{\mathcal{B}} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

esto es, la forma de Jordan.

10.1.4 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 (s_i) es menor que la algebraica (d_i), i.e. sea λ_1 tal que $s_1 < d_1$. 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 (Sección 10.1.3).

En el caso en que $d_1 = 3$ los posibles valores para la multiplicidad geométrica serán 1 ó 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 sea 1 ó 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 algebraica).

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:

$$(\mathbf{A} - \lambda_1 \mathbf{I})\vec{v}_2^{(\lambda_1)} = \vec{v}_1^{(\lambda_1)} \text{ y } (\mathbf{A} - \lambda_1 \mathbf{I})\vec{v}_3^{(\lambda_1)} = \vec{v}_2^{(\lambda_1)}.$$

Ejercicio 10.1 Se deja como ejercicio demostrar que los tres vectores son linealmente independientes. ■

Entonces, para este subespacio, el bloque de Jordan será:

$$\mathbf{B}_{\lambda_1} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

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:

- Dejar $\vec{v}_1^{(\lambda_1)}$ y construir el $\vec{v}_3^{(\lambda_1)}$ a partir de $(\mathbf{A} - \lambda_1 \mathbf{I})\vec{v}_3^{(\lambda_1)} = \vec{v}_2^{(\lambda_1)}$. 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:

$$\mathbf{B}_{\lambda_1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

- El otro caso posible sería 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:

$$\mathbf{B}_{\lambda_1} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

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 extender 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 algebraica d , con $d > 3$. La multiplicidad geométrica que nos interesará será 1, 2, 3 hasta $d - 1$, ya que si coincide con la multiplicidad algebraica, carece de interés, dado que 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 coincida 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:

$$\mathbf{B}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

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:

$$\mathbf{B}_\lambda = \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

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:

$$\mathbf{B}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

De esta manera, siempre habrá dos bloques, pero en función de cómo se extienda la base tendremos diferentes tamaños para los mismos.

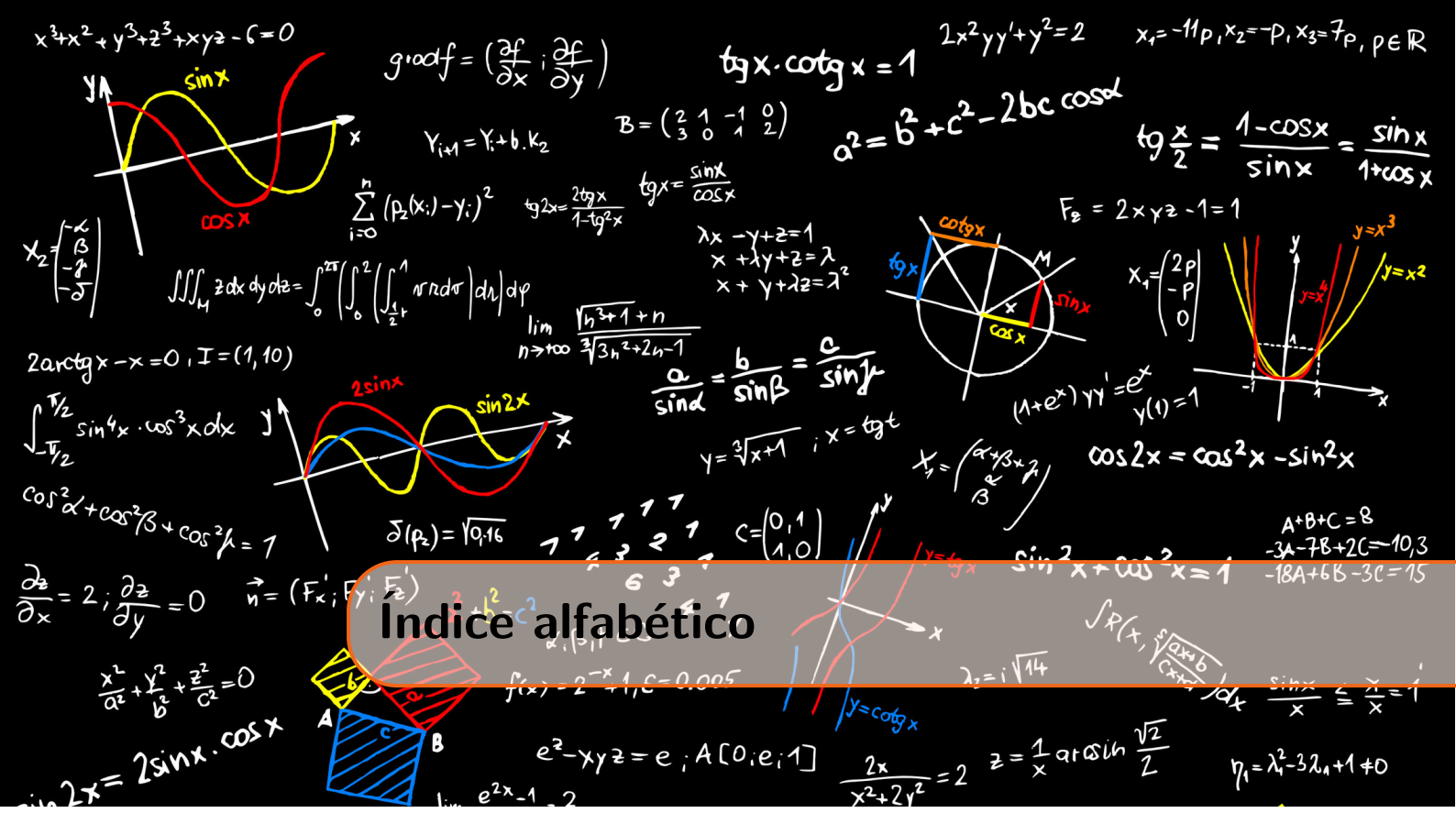
En resumen, la cantidad de bloques coincidirá con la multiplicidad algebraica del autovalor.



Bibliografía

Libros

- [A78] Maltsev A. *Fundamentos de Álgebra Lineal*. 3.^a edición. Editorial MIR, 1978 (véase páginas 7, 15, 23, 31, 43, 69, 77, 89).
- [B85] Schutz B. *A First Course in General Relativity*. Ed. Cambridge University Press, 1985 (véase páginas 51, 66).
- [D76] Kay D. *Tensor Calculus*. Editorial Mc Graw Hill, 1976 (véase página 51).
- [F71] Kreider D.; Kuller R.; Ostberg D.; Perkins F. *Introducción al Análisis Lineal*. 1.^a edición. Volumen 1,2. Editorial Fondo Educativo Interamericano, 1971 (véase páginas 7, 15, 23, 31, 43, 69, 77, 89).
- [L73] Santaló L. *Vectores y Tensores*. Eudeba, 1973 (véase páginas 7, 11, 15, 23, 31, 38, 43, 51, 55, 69, 77, 89).
- [R73] Hoffman K.; Kunze R. *Álgebra Lineal*. Editorial Prentice Hall, 1973 (véase páginas 7, 15, 23, 31, 43, 69, 77, 89).
- [S68] Bishop R.; Goldberg S. *Tensor Analysis on Manifolds*. Editorial Dover, 1968 (véase página 51).
- [S07] Grossman S. *Álgebra Lineal*. 6.^a edición. Editorial Mc Graw Hill, 2007 (véase páginas 7, 15, 23, 31, 43, 69, 77, 80, 89).



Indice alfabético

A

- Análisis tensorial 53
- Análisis tensorial: cálculo operacional . 53
- Análisis tensorial: derivación covariante 57
- Análisis tensorial: geometría Riemanniana 68
- Análisis tensorial: introducción..... 53
- Análisis tensorial: operadores diferenciales 61
- Análisis tensorial: teoría de curvas 66
- Autovalores y autovectores..... 79
- Autovalores y autovectores: definiciones 79
- Autovalores y Autovectores: Operador adjunto..... 86
- Autovalores y autovectores: operadores unitarios 88
- Autovalores y autovectores: polinomio característico..... 85

C

- Cálculo operacional: derivación de tensores 54
- Cálculo operacional: multiplicación de tensores 53

D

- Derivación covariante: símbolos de Christoffel 57
- Derivación covariante: tensor dos veces contravariante..... 60
- Derivación covariante: tensor dos veces covariante 61
- Derivación covariante: vector contravariante 58
- Derivación covariante: vector covariante 59
- Derivación de tensores: coordenadas curvilíneas..... 55

E

- Espacios euclídeos y espacios métricos 45
- Espacios euclídeos y espacios métricos: 1-forma asociada 51
- Espacios Euclídeos y Espacios Métricos: Axiomas de producto interno. 45
- Espacios euclídeos y espacios métricos: construcción de una base ortogonal 48, 49
- Espacios euclídeos y espacios métricos: el tensor métrico..... 50
- Espacios euclídeos y espacios métricos: longitud de arco..... 50
- Espacios Euclídeos y Espacios Métricos: Norma o módulo 46

Espacios euclídeos y espacios métricos: ortogonalidad	47
Espacios euclídeos y espacios métricos: producto interno	45
Espacios vectoriales	9
Espacios vectoriales: base y coordenadas	11
Espacios vectoriales: convención de Einstein	14
Espacios vectoriales: introducción	9
Espacios vectoriales: invarianza y representación	13

F

Formas bilineales y multilineales. Tensores	33
Formas bilineales y multilineales: Álgebra de bilineales	34
Formas bilineales y multilineales: aspectos metodológicos	43
Formas bilineales y multilineales: bilineales	33
Formas bilineales y multilineales: cambio de coordenadas	38
Formas bilineales y multilineales: componentes	40
Formas bilineales y multilineales: componentes y coordenadas	41
Formas bilineales y multilineales: coordenadas tensores	35
Formas bilineales y multilineales: covarianza y contravarianza	38
Formas bilineales y multilineales: formas cuadráticas	43
Formas bilineales y multilineales: producto tensorial de 1-formas	34
Formas bilineales y multilineales: tensores cartesianos	37
Formas canónicas	91
Formas canónicas: construcción de bloques de Jordan	94
Formas canónicas: Jordan	91
Formas canónicas: Matriz de bloque de Jordan	91
Formas canónicas: matriz de Jordan ..	92
Formas canónicas: multiplicidad	92
Formas Lineales y Espacio Dual	25

Formas lineales y espacio dual: 1-formas	25
Formas lineales y espacio dual: aspectos metodológicos	30
Formas lineales y espacio dual: cambio de coordenadas	31
Formas lineales y espacio dual: coordenadas	29
Formas lineales y espacio dual: el doble dual	28
Formas lineales y espacio dual: resumen cambios	31

G

Geometría Riemanniana: escalar de Ricci	69
Geometría Riemanniana: tensor de curvatura de Riemann	69
Geometría Riemanniana: tensor de Ricci	69

O

Operador adjunto: aspecto metodológico para la diagonalización	87
Operador adjunto: matrices hermíticas y simétricas	86
Operadores diferenciales: divergencia ..	63
Operadores diferenciales: gradiente	62
Operadores diferenciales: laplaciano ..	65
Operadores diferenciales: rotor	63

P

Polinomio característico: multiplicidad de las raíces	85
Prefacio	7

T

Tensores en Física	71
Tensores en Física: elasticidad	74
Tensores en Física: energía cinética	71
Tensores en Física: energía cinética de un cuerpo rígido	72

Tensores en Física: energía cinética de una partícula	71
Tensores en Física: tensor momento cua- drupolar	75
Teoría de curvas: campo de tensores sobre curvas.....	66
Teoría de curvas: geodésicas	68
Teoría de curvas: transporte paralelo ..	67
Transformaciones lineales	17
Transformaciones lineales: cambio de base y coordenadas	20
Transformaciones lineales: espacio.....	22
Transformaciones lineales: introducción	17
Transformaciones lineales: matriz asociada	18