**Universidad de Valladolid**

Facultad de Ciencias

# Trabajo Fin de Grado

Grado en Matemáticas

## Tensores y datos tensoriales

**Alumno: David Montalvo García**

**Tutor: Antonio Campillo López**

# Tensores y datos tensoriales

David Montalvo García

*Los encantos de esta ciencia sublime,*
*las matemáticas, sólo se revelan*
*a aquellas personas que tienen el valor*
*de profundizar en ella.*

***Carl Friedrich Gauss***

# Índice general

# Índice de notación

# Resumen

La primera parte del trabajo describe la noción de tensor, prestando especial atención al enfoque libre de coordenadas, sin dejar no obstante de lado su tratamiento clásico como datos de mediciones con respecto de un sistema de referencia (enfoque con coordenadas). Se incluye además un estudio detallado de los campos tensoriales, como extensión del concepto de tensor, tratando en especial los tensores y campos tensoriales métricos dada la importancia de los resultados que derivan de ellos. La segunda parte aborda el concepto de dato tensorial, desde el punto de vista de la generalización n-dimensional de las nociones de escalar, vector y matriz. A partir de esta generalización del enfoque con coordenadas, se tratarán sus operaciones prácticas actuales, junto con las descomposiciones tensoriales más relevantes en la actualidad. Estas descomposiciones se utilizan para generar algoritmos aplicables en numerosas áreas actuales, como en el estudio de señales, visión computacional, neurociencia, teoría de grafos, minería de datos y aprendizaje automático, entre otros. El trabajo trata de unificar y contrastar la terminología clásica y la práctica actual.

**Palabras clave:** tensor, producto tensorial, campo tensorial, métrica, datos tensoriales, descomposiciones tensoriales.

# Introducción

Los tensores están continuamente presentes en nuestras vidas, basta pensar en el tensor de presiones de la estancia en la que nos encontramos en este preciso instante, o en la imposibilidad de alinear las ruedas de nuestros vehículos si no fuera por el tensor de inercias. Pero sin duda, el hecho que marcó un importante hito en el desarrollo de los tensores fue su aplicación en la formulación de la teoría de la relatividad por Albert Einstein.

La palabra tensor fue utilizada por primera vez en 1846 por el matemático, físico y astrónomo irlandés, William Rowan Hamilton, en su libro "*Lectures on Quaternions*" relativo a la teoría de cuaternios que él mismo había desarrollado. Hamilton definió el tensor, dentro del marco de los cuaternios, como una cantidad positiva, o de forma más precisa, un escalar sin signo, que permite medir la magnitud de un cuaternio, del mismo modo que la longitud de un vector permite medir la magnitud del vector. En la actualidad este concepto se denomina valor absoluto del cuaternio, quedando desvinculado el término tensor de esta acepción. De hecho, el uso actual del término tensor fue introducido en 1898 por el físico alemán Woldemar Voigt, y comenzó asociándose a las magnitudes de esfuerzo y tensión en física.

El idea matemática de tensor tiene sus orígenes en el desarrollo de la geometría diferencial por Gauss, Riemann y Christoffel, quienes abrieron el camino para el desarrollo del cálculo tensorial. El cálculo tensorial, también llamado en sus orígenes cálculo absoluto, fue desarrollado en torno a 1890 por el matemático italiano Gregorio Ricci-Curbastro, quien en 1900, junto a su discípulo Tullio Levi-Civita, publicaron la clásica obra "*Méthodes de calcul différentiel absolu et leurs applications*", la cual establece las bases del cálculo tensorial.

Ya en el siglo XX, el cálculo tensorial tomó importante relevancia cuando Einstein aplicó todo el potencial que aportaban los tensores en la reformulación completa de la teoría general de la relatividad, que plasmó en diversos artículos publicados en 1915 y 1916. Estos avances provocaron un enorme progreso del cálculo tensorial, convirtiendo estas técnicas de cálculo en uno de los instrumentos más poderosos de toda la física moderna.

Son muchas las áreas de conocimiento sobre las que toma una importante trascendencia el cálculo tensorial, como por ejemplo la mecánica de medios continuos con tensores relevantes como el tensor de esfuerzos o el tensor de deformaciones; o la aplicación en el estudio del electromagnetismo, con el tensor electromagnético que permite derivar a partir de su definición las ecuaciones de Maxwell.

Las aplicaciones prácticas de los tensores van más allá de definiciones clásicas o reformulaciones de teorías físicas a lo largo de todo el siglo XX. Son muchas las aplicaciones en la actualidad, en áreas aplicadas como la neurociencia, la minería se datos o el aprendizaje automático.

¿Después de todo esto es posible seguir creyendo que los tensores no han sido una pieza primordial en la modelización del rumbo de nuestros avances científicos y de nuestro día a día?

En el avance del cálculo tensorial, existen dos tratamientos distintos del concepto de tensor. Por un lado, el enfoque clásico visualiza los tensores como generalizaciones n-dimensionales de los conceptos de escalar, vector y matriz. Estos tensores se interpretan como mediciones respecto de un determinado sistema de referencia, y los cambios de sistemas de referencias deben implicar cambios en las componentes siguiendo unas determinadas reglas o leyes de transformación. Este enfoque también se conoce como enfoque con coordenadas, y prioriza los datos sobre las estructuras matemáticas subyacentes.

En contraposición, el enfoque moderno o enfoque libre de coordenadas visualiza los tensores como objetos abstractos, construidos a partir de espacios vectoriales abstractos sobre los que se asientan las bases del álgebra multilineal. Este enfoque no trabaja con coordenadas hasta que no es necesario introducir bases del espacio vectorial sobre el que se definen. Dentro de este enfoque existen dos definiciones alternativas del concepto de tensor, por un lado como aplicaciones multilineales, y por otro como elementos de un nuevo espacio vectorial definido de forma axiomática y conocido como el producto tensorial de espacios vectoriales.

El enfoque moderno está mas estrechamente ligado al tratamiento que durante la historia los matemáticos han hecho del cálculo tensorial, dando importancia al objeto tensorial y dejando en un segundo plano sus componentes. Esto contrasta con el tratamiento clásico, predominante en la física aplicada, tomando mayor relevancia las componentes y estudiando los cambios de sistemas de referencia y los cambios de componentes.

La idea de tensor puede ser generalizada aún más desde el punto de vista de los campos tensoriales, generalizando a su vez la idea de campo escalar o campo vectorial. Estos campos se fundamentan en la asignación de un tensor a cada punto de una variedad diferenciable. Son muchos los casos en los que se utiliza el término tensor para hacer referencia realmente a un campo tensorial.

El trabajo se estructura en dos capítulos. El Capítulo 1 describe la teoría principal del álgebra multilineal, tratando los tensores tanto desde el punto de vista de las aplicaciones multilineales, como del espacio vectorial fruto del producto tensorial de espacios vectoriales, desarrollando para ello todos los resultados que permiten relacionar ambas definiciones. Se incluye además una mención al enfoque con coordenadas, que será generalizado posteriormente en el Capítulo 2, siendo necesario establecer las condiciones que definen los cambios de base en el espacio vectorial de partida sobre el que se estructuran los tensores. Asimismo, este primer capítulo ahonda en la construcción de los campos tensoriales, estudiando para ello todos los conceptos necesarios relativos a variedades diferenciables. Antes de terminar el capítulo se introduce una clase de tensores y campos tensoriales de notable relevancia, como son las métricas, profundizando en sus aplicaciones más trascendentes y que más valor aportan. Se cierra por último el capítulo con varios ejemplos clásicos de campos tensoriales en geometría.

El Capítulo 2, tal y como adelantábamos previamente, extiende el enfoque clásico de tensor para obtener estructuras de datos n-dimensionales que generalicen los conceptos de escalar, vector y matriz. Esta nueva acepción de tensor posee una gran riqueza de operaciones, tratando en el capítulo las más relevantes. La operación primordial desde el punto de vista del álgebra multilineal deriva del concepto de producto multilineal, y sus aplicaciones sobre esta teoría derivarán en la introducción del concepto de descomposición tensorial. En lo relativo a estas descomposiciones, se abordarán las dos más extendidas en la actualidad: la descomposición CP y la descomposición de Tucker. Para terminar el capítulo, de forma análoga al cierre del primer capítulo, se incluye un ejemplo de aplicación de estas descomposiciones relativo al área de ciencias de la computación.

# 1. Álgebra multilineal

A lo largo del trabajo $\mathbb{K}$ denotará un cuerpo conmutativo, típicamente $\mathbb{R}$ o $\mathbb{C}$, y $V(\mathbb{K})$ un espacio vectorial sobre $\mathbb{K}$. En aquellos casos donde no sea necesario indicar explícitamente el cuerpo, utilizaremos $V$ como notación de espacio vectorial. Además, salvo que se indique lo contrario, $V$ será un espacio vectorial de dimensión finita $n \in \mathbb{N}$.

## 1.1. Conceptos previos

En primer lugar haremos un repaso de aquellos conceptos que serán útiles a lo largo del estudio de tensores desde la perspectiva del álgebra multilineal.

**Definición 1.1.1** *Se define el dual algebraico de $V$, que denotaremos por $V^*$, como el conjunto de aplicaciones de $V$ en $\mathbb{K}$ lineales, es decir,*

$$V^* = \mathcal{L}(V, \mathbb{K}) := \{\phi : V \to \mathbb{K} : \phi \text{ es lienal}\}$$

*Cada elemento del dual $\phi \in V^*$ recibe el nombre de forma lineal.*

$V^*$ es un espacio vectorial sobre $\mathbb{K}$ de dimensión $n$ bajo las definiciones habituales de suma y producto por un escalar dadas para cada vector $\underline{x} \in V$:

- $(\phi + \psi)(\underline{x}) = \phi(\underline{x}) + \psi(\underline{x})$, $\forall \phi, \psi \in V^*$,
- $(a\phi)(\underline{x}) = a \cdot \phi(\underline{x})$, $\forall \phi \in V^*$, $\forall a \in \mathbb{K}$

Fijamos una base de $V$, $\{\underline{e}_1, ..., \underline{e}_n\} \subset V$. De esta forma, si $\underline{x} \in V$, sabemos que existen unos únicos $x^1, ..., x^n \in \mathbb{K}$ tales que $\underline{x} = \sum\limits_{i=1}^{n} x^i \underline{e}_i$. Por ello, aplicando la linealidad de los elementos $\phi \in V^*$, se tiene

$$\phi(\underline{x}) = \sum_{i=1}^{n} x^i \phi(\underline{e}_i) \in \mathbb{K}$$

**Proposición 1.1.2** *Sean $V$ un espacio vectorial de dimensión $n$ sobre $\mathbb{K}$ y $B = \{\underline{e}_1, ..., \underline{e}_n\}$ una base de $V$. Entonces existe una única base $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$ tal que, $\underline{e}^i(\underline{e}_j) = \delta^i_j$, para todos $i, j \in \{1, .., n\}$, donde $\delta^i_j$ representa la delta de Kronecker.*

**Definición 1.1.3** *Fijada $B = \{\underline{e}_1, ..., \underline{e}_n\}$ una base de $V$, la base $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$ dada en la proposición anterior se denomina base dual de $B$.*

Fijada por tanto la base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ de $V$ y su base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$, si $\underline{x} = \sum\limits_{i=1}^{n} x^i \underline{e}_i \in V$, se verifica

$$\underline{e}^j(\underline{x}) = \underline{e}^j\left(\sum_{i=1}^{n} x^i \underline{e}_i\right) = \sum_{i=1}^{n} x^i \underline{e}^j(\underline{e}_i) = \sum_{i=1}^{n} x^i \delta^j_i = x^j$$

Por tanto se tiene

$$\underline{x} = \sum_{j=1}^{n} \underline{e}^j(\underline{x})\, \underline{e}_j, \quad \forall \underline{x} \in V$$

De forma análoga, si $\underline{y} = \sum\limits_{i=1}^{n} y_i\, \underline{e}^i \in V^*$, entonces se verifica

$$\underline{y}(\underline{e}_j) = \sum_{i=1}^{n} y_i\, \underline{e}^i(\underline{e}_j) = \sum_{i=1}^{n} y_i\, \delta^i_j = y_j$$

Por tanto se tiene

$$\underline{y} = \sum_{j=1}^{n} \underline{y}(\underline{e}_j)\, \underline{e}^j, \quad \forall \underline{y} \in V^*$$

**Convenio de sumación de Einstein**

Si observamos las expresiones escritas hasta el momento, vemos que en todos los casos las sumas se producen con respecto a los índices que aparecen repetidos dos veces en lossumandos , y que figuran una vez como superíndices y otra como subíndices. Para abreviar la escritura, y sin pérdida de claridad, Einstein propuso omitir la sigma griega como símbolo de representación de la suma, $\Sigma$, y en su lugar introducir el siguiente criterio:

*Siempre que en un monomio aparezca dos veces el mismo índice, una como superíndice y otra como subíndice, se debe, a menos que se indique expresamente lo contrario, sumar los monomios obtenidos dando a este índice todos los valores posibles.*

## 1.2. Aplicaciones Multilineales y Tensores

**Definición 1.2.1** *Sean $V_1, V_2, ..., V_m, W$ espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$. Un operador $F : V_1 \times ... \times V_m \to W$ se denomina operador multilineal si es lineal en cada una de sus $m$ variables, esto es, si para cada $i = 1, ..., m$ y para todo $\underline{v}_i, \underline{w}_i \in V_i$, $\lambda \in \mathbb{K}$, se verifica*

$$F(\underline{v}_1, ..., \lambda\,\underline{v}_i + \underline{w}_i, ..., \underline{v}_m) = \lambda\,F(\underline{v}_1, ..., \underline{v}_i, ..., \underline{v}_m) + F(\underline{v}_1, ..., \underline{w}_i, ..., \underline{v}_m)$$

*Si $W = \mathbb{K}$, el operador multilineal $F : V_1 \times ... \times V_m \to \mathbb{K}$ recibe el nombre de aplicación multilineal. Si además se tiene que $V_1 = ... = V_m = V$, el operador $F : V \times ... \times V \to \mathbb{K}$ se denomina forma multilineal .*

**Observación 1.2.2** *En la literatura general, el término aplicación multilineal se utiliza también para denotar cualquier operador multilineal, independientemente de ser o no $\mathbb{K}$ el espacio de llegada, si bien en este texto mantendremos la terminología expuesta en 1.2.1 para facilitar la comprensión.*

Si $F$ es un operador multilineal, de la definición previa se deducen propiedades importantes, en particular las dos siguientes:

- $F(\underline{v}_1, ..., \underline{v}_m) = \underline{0}$ si $\underline{v}_i = \underline{0}$ para algún índice $i = 1, ..., m$.
- $F(\underline{v}_1, ..., \lambda\,\underline{v}_i + \mu\,\underline{w}_i, ..., \underline{v}_m) = \lambda\,F(\underline{v}_1, ..., \underline{v}_i, ..., \underline{v}_m) + \mu\,F(\underline{v}_1, ..., \underline{w}_i, ..., \underline{v}_m)$, para cada $i = 1, ..., m$ y para todo $\underline{v}_i, \underline{w}_i \in V_i$, $\lambda, \mu \in \mathbb{K}$.

**Definición 1.2.3** *Sea $V$ un espacio vectorial sobre $\mathbb{K}$ de dimensión finita $n$, y sean $r, s$ enteros no negativos. Un tensor $r$ veces covariante y $s$ veces contravariante, también llamado de tipo $(r, s)$, sobre $V$ es una aplicación multilineal*

$$\begin{aligned} T\colon V \times \overset{r}{\cdots} \times V \times V^* \times \overset{s}{\cdots} \times V^* &\to \mathbb{K} \\ (\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n) &\longmapsto T(\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n) \end{aligned}$$

*Por convención, definimos un tensor de tipo $(0, 0)$ como un elemento del cuerpo $\mathbb{K}$. Los tensores de tipo $(r, 0)$ se denominan tensores covariantes, y los tensores de tipo $(0, s)$ se denominan tensores contravariantes.*

Denotamos por $\mathcal{T}_r^s(V)$ el conjunto de tensores de tipo $(r, s)$ sobre $V$. Sobre este conjunto se definen de la forma natural las operaciones de suma y producto por escalares:

- Si $T, T' \in \mathcal{T}_r^s(V)$, para cualesquiera $\underline{x}_1, ..., \underline{x}_n \in V$ e $\underline{y}^1, ..., \underline{y}^n \in V^*$, se define la *suma* $T + T'$ como

$$(T + T')(\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n) = T(\underline{x}_1, ..., \underline{x}_n, y^1, ..., y^n) + T'(\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n)$$

- Si $T \in \mathcal{T}_r^s(V)$ y $\lambda \in \mathbb{K}$, de nuevo para cualesquiera $\underline{x}_1, ..., \underline{x}_n \in V$ e $\underline{y}^1, ..., \underline{y}^n \in V^*$, se define el *producto por un escalar* $\lambda \cdot T$ como

$$(\lambda \cdot T)(\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n) = \lambda \cdot T(\underline{x}_1, ..., \underline{x}_n, \underline{y}^1, ..., \underline{y}^n)$$

Es claro que para todo par de tensores $T, T' \in \mathcal{T}_r^s(V)$, $T + T' \in \mathcal{T}_r^s(V)$, y para todo tensor $T \in \mathcal{T}_r^s(V)$ y todo $\lambda \in \mathbb{K}$, $\lambda \cdot T \in \mathcal{T}_r^s(V)$. Además, consideramos como tensor nulo de $\mathcal{T}_r^s(V)$ el tensor $T \equiv 0$. A partir de estas observaciones se deriva de forma directa el siguiente resultado:

**Proposición 1.2.4** *Con las operaciones de suma y producto por escalares definidas previamente, $\mathcal{T}_r^s(V)$ es un espacio vectorial sobre $\mathbb{K}$.*

**Definición 1.2.5** *Dado un tensor $T \in \mathcal{T}_r^s(V)$, se define el orden de $T$ como el valor $r + s$.*

**Ejemplos 1.2.6**

1. *Un tensor de tipo $(1, 0)$ en $V$ se corresponde con una forma lineal de $V$ en $\mathbb{K}$. Por ello es claro que se tiene $\mathcal{T}_1^0(V) = \mathcal{L}(V, \mathbb{K}) = V^*$.*

2. *Un tensor de tipo $(0, 1)$ es una forma lineal del espacio $V^*$ en $\mathbb{K}$, luego $\mathcal{T}_0^1(V) = \mathcal{L}(V^*, \mathbb{K}) = V^{**}$, el bidual de $V$. Puesto que hemos considerado $V$ como un espacio vectorial de dimensión finita, sabemos que el Teorema de Reflexividad afirma que la aplicación*

$$\begin{aligned} i\colon V &\to V^{**} \\ \underline{x} &\mapsto i(\underline{x})\colon V^* \to \mathbb{K} \\ &\qquad\qquad \underline{y} \mapsto i(\underline{x})(\underline{y}) \coloneqq \underline{y}(\underline{x}) \end{aligned}$$

*es un isomorfismo entre $V$ y $V^{**}$, por lo que podemos identificar el espacio vectorial de los tensores de tipo $(0, 1)$ con el espacio $V$, considerando así $\mathcal{T}_0^1(V) = V$.*

3. *Los tensores de tipo $(1, 1)$ son aplicaciones bilineales de $V \times V^*$ en $\mathbb{K}$, lo que denotamos como $\mathcal{T}_1^1(V) = \mathcal{L}(V \times V^*, \mathbb{K})$. Un ejemplo particular de tensor $(1, 1)$ es la evaluación:*

$$\begin{aligned} ev : V \times V^* &\to \mathbb{K} \\ (\underline{x}, \underline{y}) &\mapsto ev(\underline{x}, \underline{y}) \coloneqq \underline{y}(\underline{x}) \end{aligned}$$

*En general, existe una correspondencia biunívoca entre los tensores de tipo $(1, 1)$ y los endomorfismos del espacio vectorial $V$. En efecto, al endomorfismo $f : V \to V$ le corresponde el tensor $(1, 1)$ dado por*

$$\begin{aligned} T : V \times V^* &\to \mathbb{K} \\ (\underline{x}, \underline{y}) &\mapsto T(\underline{x}, \underline{y}) \coloneqq \underline{y}\big(f(\underline{x})\big) \end{aligned}$$

*Recíprocamente, dado un tensor $T$ de tipo $(1,1)$, este se corresponde con el endomorfismo $f : V \to V^{**} = V$ tal que $f(\underline{x})$ es el vector de $V$ que se identifica con el vector de $V^{**}$ dado por $y \mapsto T(\underline{x}, \underline{y})$:*

$$\begin{aligned} i\big(f(\underline{x})\big) : V^* &\to \mathbb{K} \\ \underline{y} &\mapsto \underline{y}\big(f(\underline{x})\big) \coloneqq T(\underline{x}, \underline{y}) \end{aligned}$$

*donde $i\big(f(\underline{x})\big)$ denota la imagen de $f(\underline{x})$ por el isomorfismo $i$ entre $V$ y $V^{**}$ de identificación de ambos espacios visto en el ejemplo anterior.*

*En particular, al tensor ev de evaluación le corresponde el endomorfismo identidad. El espacio vectorial de los tensores $(1,1)$ queda identificado así con el espacio vectorial de los endomorfismos de $V$.*

4. *El producto interno cuando $\mathbb{K} = \mathbb{R}$, también conocido como producto escalar, representa un ejemplo de tensor de tipo $(2,0)$.*

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V &\to \mathbb{R} \\ (\underline{x}_1, \underline{x}_2) &\mapsto \langle \underline{x}_1, \underline{x}_2 \rangle \end{aligned}$$

*En general, para cualquier cuerpo $\mathbb{K}$, los tensores de tipo de $(2,0)$ son las aplicaciones bilineales $T : V \times V \to \mathbb{K}$, y para cualquier $r > 0$, los tensores de tipo $(r,0)$ son las aplicaciones multilineales $T : V \times \overset{r}{\cdots} \times V \to \mathbb{K}$. Los tensores $(r,0)$ reciben también el nombre de formas multilineales, en particular, son las formas lineales para $r = 1$ y las formas bilineales para $r = 2$.*

**Notación 1.2.7** *Denotaremos por $(V)^r \times (V^*)^s$ al espacio $V \times \overset{r}{\cdots} \times V \times V^* \times \overset{s}{\cdots} \times V^*$.*

Al igual que ocurría con las formas lineales, veamos que cada tensor $T \in \mathcal{T}_r^s(V)$ queda determinado por sus valores en una base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ y en su base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$.

**Proposición 1.2.8** *Sean $T \in \mathcal{T}_r^s(V)$, $\underline{x}_1, ..., \underline{x}_r \in V$ e $\underline{y}^1, ..., \underline{y}^s \in V^*$, y consideramos la base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ de $V$ y su base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$. Entonces existen unos únicos coeficientes $a_p^i$ y $b_i^q$ para cada $i \in \{1, ..., n\}$, $p \in \{1, ..., r\}$ y $q \in \{1, ..., s\}$ tales que*

$$T(\underline{x}_1, ..., \underline{x}_r, \underline{y}^1, ..., \underline{y}^s) = t_{i_1 \cdots i_r}^{j_1 \cdots j_s} \cdot a_1^{i_1} \cdots a_r^{i_r} \cdot b_{j_1}^1 \cdots b_{j_s}^s$$

*donde $t_{i_1 \cdots i_r}^{j_1 \cdots j_s} = T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s})$, con $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$. Estos coeficientes están unívocamente determinados: el coeficiente $a_p^i$ es la coordenada i-ésima del vector $\underline{x}_p$ en la base $B$, y el coeficiente $b_i^q$ es la i-ésima coordenada del vector $\underline{y}^q$ en la base dual $B^*$.*

<u>*Demostración*</u>.

A partir de las bases $B$ y $B^*$, sabemos que existen unos únicos coeficientes $a_p^i$ y $b_i^q$ para cada $i \in \{1, ..., n\}$, $p \in \{1, ..., r\}$ y $q \in \{1, ..., s\}$ tales que $\underline{x}_p = a_p^i \, \underline{e}_i$, $\forall p \in \{1, ..., r\}$, e $\underline{y}^q = b_j^q \, \underline{e}^j$, $\forall q \in \{1, ..., s\}$. Entonces

$$\begin{aligned} T(\underline{x}_1, ..., \underline{x}_r, \underline{y}^1, ..., \underline{y}^s) &= T(a_1^{i_1} \underline{e}_{i_1}, ..., a_r^{i_r} \underline{e}_{i_r}, b_{j_1}^1 \underline{e}^{j_1}, ..., b_{j_s}^s \underline{e}^{j_s}) \\ &= T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s}) \cdot a_1^{i_1} \cdots a_r^{i_r} \cdot b_{j_1}^1 \cdot b_{j_s}^s \\ &= t_{i_1 \cdots i_r}^{j_1 \cdots j_s} \cdot a_1^{i_1} \cdots a_r^{i_r} \cdot b_{j_1}^1 \cdots b_{j_s}^s \end{aligned}$$

□

**Definición 1.2.9** *Los coeficientes $t_{i_1 \cdots i_r}^{j_1 \cdots j_s}$ para $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$ determinados en 1.2.8 reciben el nombre de componentes de $T$ respecto de la base $B = \{\underline{e}_1, ..., \underline{e}_n\}$.*

Nótese que la base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$, asociada a $B$, juega un papel auxiliar, aunque fundamental, en esta noción de componentes tensoriales.

**Observación 1.2.10** *La notación de índices desarrollada por los matemáticos italianos Tullio Levi-Civita y Gregorio Ricci-Curbastro, distingue entre subíndices y superíndices: los subíndices representan las componentes covariantes de los tensores, mientras los superíndices representan las componentes contravariantes.*

Veamos a continuación cómo se transforman las componentes de un tensor cuando se produce un cambio de base en el espacio vectorial $V$.

**Proposición 1.2.11** *Sea $B = \{\underline{e}_1, ..., \underline{e}_n\}$ una base de $V$ y $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ su base dual. Consideramos ahora $\widetilde{B} = \{\underline{\tilde{e}}_1, ..., \underline{\tilde{e}}_n\}$ otra base de $V$ y $\widetilde{B}^* = \{\underline{\tilde{e}}^1, ..., \underline{\tilde{e}}^n\}$ su respectiva base dual. Si las componentes de un tensor $T \in \mathcal{T}_r^s(V)$ en las bases $B$ y $\widetilde{B}$ son $t_{i_1 \cdots i_r}^{j_1 \cdots j_s} = T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s})$ y $\tilde{t}_{k_1 \cdots k_r}^{l_1 \cdots l_s} = T(\underline{\tilde{e}}_{k_1}, ..., \underline{\tilde{e}}_{k_r}, \underline{\tilde{e}}^{l_1}, ..., \underline{\tilde{e}}^{l_s})$ respectivamente, entonces*

$$\tilde{t}_{k_1 \cdots k_r}^{l_1 \cdots l_s} = p_{k_1}^{i_1} \cdots p_{k_r}^{i_r} \cdot (p')_{j_1}^{l_1} \cdots (p')_{j_s}^{l_s} \cdot t_{i_1 \cdots i_r}^{j_1 \cdots j_s}$$

*con $P = (p_k^i)$ la matriz de cambio de base de $\widetilde{B}$ a $B$ y $P^{-1} = \big((p')_j^l\big)$ su matriz inversa.*

<u>*Demostración*</u>.

Cada elemento de $\widetilde{B}$ puede expresarse de forma única como combinación lineal de elementos de $B$ a partir de la matriz de cambio de base $P$: $\underline{\tilde{e}}_k = p_k^i \, \underline{e}_i$. Sea ahora $Q$ la matriz de cambio de base de $\widetilde{B}^*$ a $B^*$. De nuevo cada elemento de $\widetilde{B}^*$ puede expresarse como combinación lineal de elementos de $B^*$ a partir de la matriz de cambio de base $Q$: $\underline{\tilde{e}}^{\,l} = q^l_j\, \underline{e}^j$. Podemos observar que, de hecho, se tiene que $Q = P^{-1}$:

$$\delta^l_k = \underline{\tilde{e}}^{\,l}(\underline{\tilde{e}}_k) = \underline{\tilde{e}}^{\,l}(p^i_k\, \underline{e}_i) = q^l_j\, \underline{e}^j(p^i_k\, \underline{e}_i) = q^l_j\, p^i_k\, \underline{e}^j(\underline{e}_i) = q^l_j\, p^i_k\, \delta^j_i = q^l_m\, p^m_k$$

Por tanto el cambio podemos escribirlo como $\underline{\tilde{e}}^{\,l} = (p')^l_j\, \underline{e}^j$. A partir de estos cambios se deduce ya fácilmente el cambio de coordenadas del tensor:

$$\begin{aligned}
\tilde{t}^{\,l_1\cdots l_s}_{k_1\cdots k_r} &= T(\underline{\tilde{e}}_{k_1}, ..., \underline{\tilde{e}}_{k_r}, \underline{\tilde{e}}^{\,l_1}, ..., \underline{\tilde{e}}^{\,l_s}) = T(p^{i_1}_{k_1}\underline{e}_{i_1}, ..., p^{i_r}_{k_r}\underline{e}_{i_r}, (p')^{l_1}_{j_1}\underline{e}^{j_1}, ..., (p')^{l_s}_{j_s}\underline{e}^{j_s}) \\
&= p^{i_1}_{k_1}\cdots p^{i_r}_{k_r}\cdot (p')^{l_1}_{j_1}\cdots (p')^{l_s}_{j_s}\cdot T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s}) \\
&= p^{i_1}_{k_1}\cdots p^{i_r}_{k_r}\cdot (p')^{l_1}_{j_1}\cdots (p')^{l_s}_{j_s}\cdot t^{j_1\cdots j_s}_{i_1\cdots i_r}
\end{aligned}$$

□

A partir de las reglas de cambio de base anteriores, se obtiene el resultado que rige el enfoque tensorial clásico basado en las coordenadas del tensor:

**Teorema 1.2.12** *Sea $V$ un espacio vectorial de dimensión $n$. Un tensor $T \in \mathcal{T}^s_r(V)$ representa el mismo dato que una asignación a cada base de $V$ de $n^{r+s}$ componentes $t^{j_1\cdots j_s}_{i_1\cdots i_r}$ para $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$ tales que las respectivas asignaciones a cada par de bases $B$ y $\widetilde{B}$ de $V$ satisfacen las fórmulas de cambio de base de 1.2.11. Esto es, si $t^{j_1\cdots j_s}_{i_1\cdots i_r}$ para $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$ es la asignación asociada a la base $B$ y $\tilde{t}^{\,l_1\cdots l_s}_{k_1\cdots k_r}$ para $k_1, ..., k_r, l_1, ..., l_s \in \{1, ..., n\}$ es la asignación asociada a $\widetilde{B}$, se debe cumplir para todo $k_1, ..., k_r, l_1, ..., l_s \in \{1, ..., n\}$ que*

$$\tilde{t}^{\,l_1\cdots l_s}_{k_1\cdots k_r} = p^{i_1}_{k_1}\cdots p^{i_r}_{k_r}\cdot (p')^{l_1}_{j_1}\cdots (p')^{l_s}_{j_s}\cdot t^{j_1\cdots j_s}_{i_1\cdots i_r}$$

*con $P = (p^i_k)$ la matriz de cambio de base de $\widetilde{B}$ a $B$ y $P^{-1} = \big((p')^l_j\big)$ su matriz inversa.*

<u>Demostración</u>.

Dado un tensor $T \in \mathcal{T}^s_r(V)$, la definición 1.2.9 muestra que dichas asignaciones existen, y la proposición 1.2.11 afirma que dichas asignaciones satisfacen las fórmulas de cambio de base.

Recíprocamente, dadas dos asignaciones $t^{j_1\cdots j_s}_{i_1\cdots i_r}$ para $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$ asociada a $B$ y $\tilde{t}^{\,l_1\cdots l_s}_{k_1\cdots k_r}$ para $k_1, ..., k_r, l_1, ..., l_s \in \{1, ..., n\}$ asociada a $\widetilde{B}$, verificando las fórmulas de cambio de base $\tilde{t}^{\,l_1\cdots l_s}_{k_1\cdots k_r} = p^{i_1}_{k_1}\cdots p^{i_r}_{k_r}\cdot (p')^{l_1}_{j_1}\cdots (p')^{l_s}_{j_s}\cdot t^{j_1\cdots j_s}_{i_1\cdots i_r}$, el tensor $T$ de tipo $(r, s)$ puede definirse con la asignación en cualquier selección de la base de $V$ y, entonces, las fórmulas del cambio de base muestran que dicho tensor no depende de la base elegida.

□

**Observación 1.2.13** *El teorema anterior sigue siendo válido si, en lugar de considerar las asignaciones a todas las bases de $V$, se consideran únicamente las asignaciones a las bases de cualquier conjunto no vacío de ellas, siempre y cuando se satisfagan las fórmulas de cambio de base entre dos cualesquiera de estas bases.*

## 1.3. Producto tensorial de espacios vectoriales

Dados dos espacios vectoriales $U$ y $V$ sobre $\mathbb{K}$, el espacio vectorial producto tensorial de ambos espacios, que denotaremos por $U \otimes V$, y la aplicación bilineal asociada, $m : U \times V \to U \otimes V$, quedan caracterizados salvo isomorfismos a partir de la siguiente propiedad universal:

**Teorema 1.3.1** *Dados $U$ y $V$ espacios vectoriales sobre $\mathbb{K}$, existe un único (salvo isomorfismo) espacio vectorial $T$ sobre $\mathbb{K}$ y un operador bilineal $m\colon U \times V \to T$ tales que $T = \langle Im(m) \rangle$ y para todo operador bilineal $b\colon U \times V \to W$ con llegada en un espacio vectorial $W$ arbitrario, existe un único operador lineal $l : T \to W$ tal que $l \circ m = b$.*

$$\begin{tikzcd} U \times V \arrow[r, "m"] \arrow[dr, "b"'] & T \arrow[d, dashed, "l"] \\ & W \end{tikzcd}$$

**Notación 1.3.2** *El espacio $T$ lo denotamos, salvo isomorfismo, como el espacio $U \otimes V$, y las imágenes por el operador bilineal $m$ las denotamos de la forma $m(u, v) = u \otimes v$.*

**Definición 1.3.3** *Definimos el producto tensorial de los espacios vectoriales $U$ y $V$ como el par $(U \otimes V, m)$, donde el espacio vectorial $U \otimes V$ y el operador bilineal $m : U \times V \to U \otimes V$ están caracterizados de forma única (salvo isomorfismo) por la propiedad universal 1.3.1.*

**Proposición 1.3.4** *Sean $U$ y $V$ espacios vectoriales sobre $\mathbb{K}$. Para todos $\underline{u}, \underline{u}_1, \underline{u}_2 \in U$, $\underline{v}, \underline{v}_1, \underline{v}_2 \in V$ y $\alpha, \beta \in \mathbb{K}$, se verifican las siguientes igualdades relativas al espacio $U \otimes V$:*

1. $(\alpha \underline{u}_1 + \beta \underline{u}_2) \otimes \underline{v} = \alpha(\underline{u}_1 \otimes \underline{v}) + \beta(\underline{u}_2 \otimes \underline{v})$

2. $\underline{u} \otimes (\alpha \underline{v}_1 + \beta \underline{v}_2) = \alpha(\underline{u} \otimes \underline{v}_1) + \beta(\underline{u} \otimes \underline{v}_2)$

3. $\alpha(\underline{u} \otimes \underline{v}) = (\alpha \underline{u}) \otimes \underline{v} = \underline{u} \otimes (\alpha \underline{v})$

*Demostración*.

Por la bilinealidad del operador $m : U \times V \to U \otimes V$ asociado al espacio $U \otimes V$, se tiene

1. $(\alpha \underline{u}_1 + \beta \underline{u}_2) \otimes \underline{v} = m\big((\alpha \underline{u}_1 + \beta \underline{u}_2), \underline{v}\big) = \alpha\, m(\underline{u}_1, \underline{v}) + \beta\, m(\underline{u}_2, \underline{v}) = \alpha(\underline{u}_1 \otimes \underline{v}) + \beta(\underline{u}_2 \otimes \underline{v})$

2. $\underline{u} \otimes (\alpha \underline{v}_1 + \beta \underline{v}_2) = m\big(\underline{u}, (\alpha \underline{v}_1 + \beta \underline{v}_2)\big) = \alpha\, m(\underline{u}, \underline{v}_1) + \beta\, m(\underline{u}, \underline{v}_2) = \alpha(\underline{u} \otimes \underline{v}_1) + \beta(\underline{u} \otimes \underline{v}_2)$

3. $\alpha(\underline{u} \otimes \underline{v}) = \alpha\, m(\underline{u}, \underline{v}) = m(\alpha\, \underline{u}, \underline{v}) = (\alpha \underline{u}) \otimes \underline{v}$ , $\alpha(\underline{u} \otimes \underline{v}) = \alpha\, m(\underline{u}, \underline{v}) = m(\underline{u}, \alpha\, \underline{v}) = \underline{u} \otimes (\alpha \underline{v})$

$\square$

**Proposición 1.3.5** *Sean $U$ y $V$ espacios vectoriales sobre $\mathbb{K}$. El producto tensorial de ambos espacios es simétrico, esto es, existe un isomorfismo canónico entre $U \otimes V$ y $V \otimes U$.*

Demostración.

Sea $m_1 : U \times V \to U \otimes V$ el operador bilineal asociado al producto tensorial $U \otimes V$, y $m_2 : V \times U \to V \otimes U$ el operador bilineal asociado al producto tensorial $V \otimes U$. Definimos el operador

$$\begin{aligned} b_1 : U \times V &\to V \otimes U \\ (\underline{u}, \underline{v}) &\mapsto b_1(\underline{u}, \underline{v}) \coloneqq m_2(\underline{v}, \underline{u}) = \underline{v} \otimes \underline{u} \end{aligned}$$

el cual es bilineal por serlo $m_2$. Por la propiedad universal del producto tensorial $U \otimes V$, existe un único operador lineal $l_1 : U \otimes V \to V \otimes U$ tal que $l_1 \circ m_1 = b_1$, y por tanto se verifica $l_1(\underline{u} \otimes \underline{v}) = \underline{v} \otimes \underline{u}$.

De forma análoga definimos el operador

$$\begin{aligned} b_2 : V \times U &\to U \otimes V \\ (\underline{v}, \underline{u}) &\mapsto b_2(\underline{v}, \underline{u}) \coloneqq m_1(\underline{u}, \underline{v}) = \underline{u} \otimes \underline{v} \end{aligned}$$

de nuevo bilineal por serlo $m_1$. Entonces por la propiedad universal del producto tensorial $V \otimes U$, existe un único operador lineal $l_2 : V \otimes U \to U \otimes V$ tal que $l_2 \circ m_2 = b_2$, luego $l_2(\underline{v} \otimes \underline{u}) = \underline{u} \otimes \underline{v}$. Se tiene así el siguiente diagrama de operadores lineales y bilineales:

$$\begin{tikzcd}
U \times V \arrow[r, "m_1"] \arrow[dr, "b_1"'] & U \otimes V \arrow[d, dashed, shift right, "l_1"'] & \\
& V \otimes U \arrow[u, dashed, shift right, "l_2"'] & V \times U \arrow[l, "m_2"] \arrow[ul, "b_2"']
\end{tikzcd}$$

El operador compuesto $l_2 \circ l_1 : U \otimes V \to U \otimes V$ verifica $(l_2 \circ l_1)\big(m_1(\underline{u}, \underline{v})\big) = (l_2 \circ l_1)(\underline{u} \otimes \underline{v}) = l_2(\underline{v} \otimes \underline{u}) = \underline{u} \otimes \underline{v} = m_1(\underline{u}, \underline{v})$ y por tanto $l_2 \circ l_1 \cong Id_{U \otimes V}$. De igual modo se razona con el operador $l_1 \circ l_2 : V \otimes U \to V \otimes U$ para concluir que $l_1 \circ l_2 \cong Id_{V \otimes U}$. Se concluye así que $l_1$ y $l_2$ son ambos isomorfismos y que a su vez son inversos entre sí, de donde se deduce $U \otimes V \cong V \otimes U$.

$\square$

**Proposición 1.3.6** *Sean $U$, $V$ y $W$ espacios vectoriales sobre $\mathbb{K}$. El producto tensorial de los espacios vectoriales es asociativo, esto es, existe un isomorfismo canónico entre $U \otimes (V \otimes W)$ y $(U \otimes V) \otimes W$.*

*Demostración*.

Para cada $\underline{w} \in W$ consideramos el operador $b_{\underline{w}} : U \times V \to U \otimes (V \otimes W)$ definido como $b_w(u,v) = \underline{u} \otimes (\underline{v} \otimes \underline{w})$. $b_{\underline{w}}$ es bilineal, y por la propiedad universal del producto tensorial $U \otimes V$ existe una única aplicación lineal $l_{\underline{w}} : U \otimes V \to U \otimes (V \otimes W)$ tal que $l_w(\underline{u} \otimes \underline{v}) = \underline{u} \otimes (\underline{v} \otimes \underline{w})$ para todo par $(\underline{u}, \underline{v}) \in U \times V$, o lo que es lo mismo, para todo elemento $\underline{u} \otimes \underline{v} \in U \otimes V$.

$$\begin{tikzcd} U \times V \arrow[r] \arrow[dr, "b_{\underline{w}}"'] & U \otimes V \arrow[d, dashed, "l_{\underline{w}}"] \\ & U \otimes (V \otimes W) \end{tikzcd}$$

Consideramos ahora el operador

$$\begin{aligned} b : (U \otimes V) \times W &\to U \otimes (V \otimes W) \\ (\underline{t}, \underline{w}) &\mapsto b(\underline{t}, \underline{w}) := l_{\underline{w}}(\underline{t}) \end{aligned}$$

El operador $b$ así definido es lineal en $\underline{t}$ para $\underline{w}$ fijo, puesto que $l_{\underline{w}}$ es lineal. Veamos que también es lineal en $\underline{w}$ para cada $\underline{t}$ fijo, esto es, veamos que se cumplen las igualdades $l_{\underline{w}_1+\underline{w}_2}(\underline{t}) = l_{\underline{w}_1}(\underline{t}) + l_{\underline{w}_2}(\underline{t})$ y $l_{\lambda\underline{w}}(\underline{t}) = \lambda l_{\underline{w}}(\underline{t})$ para todos $\underline{w}_1, \underline{w}_2 \in W$ y para todo $\lambda \in \mathbb{K}$. Es suficiente con probar estas igualdades cuando $\underline{t}$ es un elemento de la forma $\underline{u} \otimes \underline{v}$, puesto que ambos lados de las igualdades son lineales en $\underline{t}$ y $U \otimes V$ está constituido por combinaciones lineales de elementos de este tipo.

- $l_{\underline{w}_1+\underline{w}_2}(\underline{u} \otimes \underline{v}) = \underline{u} \otimes \big(\underline{v} \otimes (\underline{w}_1 + \underline{w}_2)\big) = \underline{u} \otimes \big((\underline{v} \otimes \underline{w}_1) + (\underline{v} \otimes \underline{w}_2)\big) = \big(\underline{u} \otimes (\underline{v} \otimes \underline{w}_1)\big) + \big(\underline{u} \otimes (\underline{v} \otimes \underline{w}_1)\big) = l_{\underline{w}_1}(\underline{u} \otimes \underline{v}) + l_{\underline{w}_2}(\underline{u} \otimes \underline{v})$
- $l_{\lambda\underline{w}}(\underline{u} \otimes \underline{v}) = \underline{u} \otimes (\underline{v} \otimes \lambda\underline{w}) = \underline{u} \otimes \lambda(\underline{v} \otimes \underline{w}) = \lambda\big(\underline{u} \otimes (\underline{v} \otimes \underline{w})\big) = \lambda \cdot l_{\underline{w}}(\underline{u} \otimes \underline{v})$

Como hemos visto que $b$ es bilineal, por la propiedad universal del producto tensorial del espacio $(U \otimes V) \otimes W$ existe un único operador lineal $l : (U \otimes V) \otimes W \to U \otimes (V \otimes W)$ tal que $l(t \otimes w) = l_w(t)$, luego $l((u \otimes v) \otimes w) = l_w(u \otimes v) = u \otimes (v \otimes w)$.

$$\begin{tikzcd} (U \otimes V) \times W \arrow[r] \arrow[dr, "b"'] & (U \otimes V) \otimes W \arrow[d, dashed, "l"] \\ & U \otimes (V \otimes W) \end{tikzcd}$$

De forma análoga se construye el operador lineal $\tilde{l} : U \otimes (V \otimes W) \to (U \otimes V) \otimes W$ de tal forma que $\tilde{l}(u \otimes (v \otimes w)) = (u \otimes v) \otimes w$. De esta forma se verifica que $(l \circ \tilde{l})\big(u \otimes (v \otimes w)\big) = u \otimes (v \otimes w)$ y $(\tilde{l} \circ l)\big((u \otimes v) \otimes w\big) = (u \otimes v) \otimes w$, quedando así probada la existencia del isomorfismo.

□

El producto tensorial de tres o más espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$ se define de forma análoga al producto tensorial de dos espacios vectoriales, a través de un resultado idéntico al expuesto en el Teorema 1.3.1, el cuál se expone a continuación en el Teorema 1.3.7. No obstante, el resultado 1.3.6 muestra que, aunque dispongamos de más de dos factores en el espacio vectorial producto tensorial, estos se pueden asociar en sucesivos pares de diversas formas, todas ellas isomorfas, de forma que el producto tensorial se expresará en función de productos tensoriales de dos factores.

**Teorema 1.3.7** *Dados $V_1, ..., V_h$ espacios vectoriales sobre $\mathbb{K}$, existe un único (salvo isomorfismo) espacio vectorial $T$ sobre $\mathbb{K}$ y un operador multilineal $m : V_1 \times \cdots \times V_h \to T$ tales que $T = \langle Im(m) \rangle$ y para todo operador multilineal $b : V_1 \times \cdots \times V_h \to W$ con llegada en un espacio vectorial $W$ arbitrario, existe un único operador lineal $l : T \to W$ tal que $l \circ m = b$.*

$$\begin{tikzcd}
V_1 \times \cdots \times V_h \arrow[r, "m"] \arrow[dr, "b"'] & T \arrow[d, dashed, "l"] \\
 & W
\end{tikzcd}$$

El espacio vectorial $T$ se denota por $V_1 \otimes \cdots \otimes V_h$. El par $(V_1 \otimes \cdots \otimes V_h, m)$ que satisfaga la propiedad universal 1.3.7 (único salvo isomorfismos) recibe el nombre de producto tensorial de los espacios vectoriales $V_1, ..., V_h$, y el operador $m$ se denomina operador tensorial. El vector $m(\underline{v}_1, ..., \underline{v}_h)$ de $V_1 \otimes \cdots \otimes V_h$, donde $\underline{v}_i \in V_i$ para $i \in \{1, ..., h\}$, se denota por $v_1 \otimes ... \otimes v_h$. Tal y como hemos visto en las proposiciones 1.3.5 y 1.3.6, salvo isomorfismo, el producto tensorial es conmutativo y asociativo, y lo mismo sucede con el operador tensorial.

**Proposición 1.3.8** *Sean $U$ y $V$ espacios vectoriales sobre $\mathbb{K}$, $B_U = \{\underline{u}_1, ..., \underline{u}_n\}$ una base de $U$ y $B_V = \{\underline{v}_1, ..., \underline{v}_m\}$ una base de $V$. Entonces $B = \left\{\underline{u}_i \otimes \underline{v}_j : i \in \{1, ..., n\}, j \in \{1, ..., m\}\right\}$ es una base del espacio $U \otimes V$.*

<u>*Demostración*</u>.

Sea $\underline{u} \otimes \underline{v} \in U \otimes V$. Sabemos que existen expresiones únicas de $\underline{u}$ y $\underline{v}$ como combinaciones lineales de los elementos de sus respectivas bases en los espacios $U$ y $V$, esto es, existen unos únicos coeficientes $\alpha^i$ para $i \in \{1, ..., n\}$ y $\beta^j$ para $j \in \{1, ..., m\}$ tales que $\underline{u} = \alpha^i \underline{u}_i$ y $\underline{v} = \beta^j \underline{v}_j$. Si $m : U \times V \to U \otimes V$ es el operador bilineal asociado al producto tensorial $U \otimes V$, se tiene

$$\underline{u} \otimes \underline{v} = m(\underline{u}, \underline{v}) = m(\alpha^i \underline{u}_i, \beta^j \underline{v}_j) = \alpha^i \beta^j m(\underline{u}_i, \underline{v}_j) = \alpha^i \beta^j (\underline{u}_i \otimes \underline{v}_j)$$

y por tanto $B$ tal y como lo hemos definido constituye un sistema de generadores de $U \otimes V$.

Veamos ahora que $B$ es linealmente independiente. Consideramos la combinación

$$\alpha^i\beta^j(\underline{u}_i \otimes \underline{v}_j) = 0 \;\Rightarrow\; (\alpha^i\underline{u}_i) \otimes (\beta^j\underline{v}_j) = 0 \tag{1.1}$$

Sea $B_U^* = \{\underline{u}^1, ..., \underline{u}^n\}$ la base dual de $B_U$ en $U$ y $B_V^* = \{\underline{v}^1, ..., \underline{v}^m\}$ la base dual de $B_V$ en $V$, y definimos para cada $(k, l) \in \{1, ..., n\} \times \{1, ..., m\}$ la aplicación

$$\begin{aligned} b_{kl} : U \times V &\to \mathbb{K} \\ (\underline{u}, \underline{v}) &\mapsto b_{kl}(\underline{u}, \underline{v}) := \underline{u}^k(\underline{u})\underline{v}^l(\underline{v}) \end{aligned}$$

Para cada $(k, l) \in \{1, ..., n\} \times \{1, ..., m\}$ es claro que la aplicación $b_{kl}$ es bilineal, por lo que aplicando la propiedad universal del espacio vectorial producto tensorial $U \otimes V$, sabemos que existe una única aplicación lineal $l_{kl} : U \otimes V \to \mathbb{K}$ tal que $l_{kl}(\underline{u} \otimes \underline{v}) = \underline{u}^k(\underline{u})\underline{v}^l(\underline{v})$. Aplicando para cada $(k, l) \in \{1, ..., n\} \times \{1, ..., m\}$ la forma lineal $l_{kl}$ en la igualdad 1.1, obtenemos

$$0 = l_{kl}(0) = l_{kl}\big((\alpha^i\underline{u}_i) \otimes (\beta^j\underline{v}_j)\big) = \underline{u}^k(\alpha^i\underline{u}_i)\underline{v}^l(\beta^j\underline{v}_j) = \alpha^i\,\underline{u}^k(\underline{u}_i)\beta^j\,\underline{v}^l(\underline{v}_j) = \alpha^k\beta^l$$

puesto que $\underline{u}^k(\underline{u}_i) = \delta_i^k$ y $\underline{v}^l(\underline{v}_j) = \delta_j^l$, lo que prueba que $B$ es un conjunto linealmente independiente, quedando entonces probado que es una base del espacio vectorial $U \otimes V$.

$\square$

**Corolario 1.3.9** *Si $U$ es un espacio vectorial de dimensión $n$ y $V$ es un espacio vectorial de dimensión $m$, el producto tensorial $U \otimes V$ es un espacio vectorial de dimensión $n \cdot m$.*

**Proposición 1.3.10** *Sean $U$ y $V$ espacios vectoriales sobre $\mathbb{K}$. El espacio dual del producto tensorial de los espacios vectoriales es el producto tensorial de los espacios duales, esto es, existe un isomorfismo canónico entre $(U \otimes V)^*$ y $U^* \otimes V^*$.*

*Demostración.*

Sean $\underline{f} \in U^*$ y $\underline{g} \in V^*$ fijos, y consideramos la aplicación

$$\begin{aligned} b_{\underline{f}\,\underline{g}} : U \times V &\to \mathbb{K} \\ (\underline{u}, \underline{v}) &\mapsto b_{\underline{f}\,\underline{g}}(\underline{u}, \underline{v}) := \underline{f}(\underline{u})\underline{g}(\underline{v}) \end{aligned}$$

Es claro que $b_{\underline{f}\,\underline{g}}$ es bilineal gracias a la linealidad de $\underline{f}$ y $\underline{g}$. Si $m_1 : U \times V \to U \otimes V$ es el operador bilineal asociado al producto tensorial $U \otimes V$, por la propiedad universal se tiene que existe una única aplicación lineal $l_{\underline{f}\,\underline{g}} : U \otimes V \to \mathbb{K}$ tal que $l_{\underline{f}\,\underline{g}} \circ m_1 = b_{\underline{f}\,\underline{g}}$, esto es, $l_{\underline{f}\,\underline{g}}(\underline{u} \otimes \underline{v}) = \underline{f}(\underline{u})\underline{g}(\underline{v})$ para todo $\underline{u} \otimes \underline{v} \in U \otimes V$. Por tanto $l_{\underline{f}\,\underline{g}} \in (U \otimes V)^*$.

Definimos el operador

$$
\begin{aligned}
b : U^* \times V^* &\to (U \otimes V)^* \\
(\underline{f}, \underline{g}) &\mapsto b(\underline{f}, \underline{g}) \coloneqq l_{\underline{f}\,\underline{g}} : U \otimes V \to \mathbb{K} \\
&\qquad\qquad \underline{u} \otimes \underline{v} \mapsto l_{\underline{f}\,\underline{g}}(\underline{u} \otimes \underline{v}) = \underline{f}(\underline{u})\underline{g}(\underline{v})
\end{aligned}
$$

Veamos que $b$ es bilineal. Razonaremos únicamente con la primera componente, ya que el razonamiento es análogo con la segunda: $b(\alpha\underline{f}_1 + \beta\underline{f}_2, \underline{g})(\underline{u} \otimes \underline{v}) = (\alpha\underline{f}_1 + \beta\underline{f}_2)(\underline{u}) \cdot \underline{g}(\underline{v}) = \alpha\big(\underline{f}_1(\underline{u}) \cdot \underline{g}(\underline{v})\big) + \beta\big(\underline{f}_2(\underline{u}) \cdot \underline{g}(\underline{v})\big) = \alpha \cdot b(\underline{f}_1, \underline{g}) + \beta \cdot b(f_2, \underline{g})$.

Por ello, si $m_2 : U^* \times V^* \to U^* \otimes V^*$ es el operador bilineal asociado a $U^* \otimes V^*$, de nuevo por la propiedad universal del producto tensorial de espacios vectoriales existe un único operador lineal $l : U^* \otimes V^* \to (U \otimes V)^*$ tal que $l \circ m_2 = b$, esto es,

$$
l(\underline{f} \otimes \underline{g}) = b(\underline{f}, \underline{g}) = l_{\underline{f}\,\underline{g}} \quad \Rightarrow \quad l(\underline{f} \otimes \underline{g})(\underline{u} \otimes \underline{v}) = l_{\underline{f}\,\underline{g}}(\underline{u} \otimes \underline{v}) = \underline{f}(\underline{u})\underline{g}(\underline{v})
$$

Veamos a continuación que $l$ define un isomorfismo entre $U^* \otimes V^*$ y $(U \otimes V)^*$. Como hemos probado que $l$ es lineal, para probar que $l$ es inyectiva nos basta ver que $Ker(l) = \{\underline{0}\}$.

Sea $\underline{F} \in U^* \otimes V^*$ tal que $l(\underline{F}) \equiv 0$. Sean $\{\underline{f}^1, ..., \underline{f}^n\} \subset U^*$ una base de $U^*$ y $\{\underline{g}^1, ..., \underline{g}^m\} \subset V^*$ una base de $V^*$. Por 1.3.8 sabemos que existen unos únicos $\alpha_{ij} \in \mathbb{K}$ para todo $i \in \{1, ..., n\}$ y $j \in \{1, ..., m\}$ tales que $\underline{F} = \alpha_{ij}(\underline{f}^i \otimes \underline{g}^j)$. Por ello, para todos $\underline{u} \in U$ y $\underline{v} \in V$, se tiene

$$
0 = l(\underline{F})(\underline{u} \otimes \underline{v}) = l(\alpha_{ij}(\underline{f}^i \otimes \underline{g}^j))(\underline{u} \otimes \underline{v}) = \alpha_{ij}\big(l(\underline{f}^i \otimes \underline{g}^j)(\underline{u} \otimes \underline{v})\big) = \alpha_{ij}\underline{f}^i(\underline{u})\underline{g}^j(\underline{v})
$$

Como la igualdad es cierta para todo $\underline{v} \in V$, deducimos que se tiene $0 \equiv \alpha_{ij}\underline{f}^i(\underline{u})\underline{g}^j$, y como $\underline{g}^1, ..., \underline{g}^m$ es un conjunto linealmente independiente de elementos de $V^*$, se deduce que $\alpha_{ij}\underline{f}^i(\underline{u}) = 0$ para todo $j \in \{1, ..., m\}$. Aplicando el mismo razonamiento para todo $\underline{u} \in U$, deducimos que $0 \equiv \alpha_{ij}\underline{f}^i$, y por ser $\{\underline{f}^1, ..., \underline{f}^n\}$ un conjunto linealmente independiente, se tiene $\alpha_{ij} = 0$, $\forall\, i \in \{1, ..., n\}$, $j \in \{1, ..., m\}$, y por tanto $\underline{F} \equiv 0$, tal y como queríamos probar.

No falta por último probar que $l$ es sobreyectiva. Para ello sean $\{\underline{u}_1, ..., \underline{u}_n\} \subset U$ una base de $U$, $\{\underline{u}^1, ..., \underline{u}^n\} \subset U^*$ su base dual y $\{\underline{v}_1, ..., \underline{v}_m\} \subset V$ una base de $V$. Consideramos un elemento $\underline{h} \in (U \otimes V)^*$, y queremos probar que $\exists \underline{H} \in U^* \otimes V^*$ / $l(\underline{H}) \equiv \underline{h}$, para lo cual basta probar que las aplicaciones lineales $l(\underline{H})$ y $\underline{h}$ coinciden para todos los elementos de una base de $U \otimes V$.

Definimos $\underline{H} \in U^* \otimes V^*$ como $\underline{H} = \sum_{i=1}^{n} \underline{u}^i \otimes \underline{g}^i$ donde $\underline{g}_i \in V^*$ es tal que $\underline{g}^i(\underline{v}_k) = \underline{h}(\underline{u}_i \otimes \underline{v}_k)$ para todo $k \in \{1, ..., m\}$. Como $\big\{\underline{u}_j \otimes \underline{v}_k \ : \ j \in \{1, ..., n\}, k \in \{1, ..., m\}\big\}$ es base de $U \otimes V$, y además se tiene

$$l(\underline{H})(\underline{u}_j \otimes \underline{v}_k) = l\Big(\sum_{i=1}^{n} \underline{u}^i \otimes \underline{g}^i\Big)(\underline{u}_j \otimes \underline{v}_k) = \sum_{i=1}^{n} l(\underline{u}^i \otimes \underline{g}^i)(\underline{u}_j \otimes \underline{v}_k)$$
$$= \sum_{i=1}^{n} \underline{u}^i(\underline{u}_j)\underline{g}^i(\underline{v}_k) = \underline{g}^j(\underline{v}_k) = h(\underline{u}_j \otimes \underline{v}_k)$$

entonces $l(\underline{H}) \equiv \underline{h}$, y por tanto $l$ es también sobreyectiva.

$\square$

A partir de los resultados anteriores, el siguiente resultado muestra la relación entre el producto tensorial de espacios vectoriales y el espacio vectorial $\mathcal{T}_r^s(V)$. Antes de ello, definimos el siguiente convenio de notación:

**Notación 1.3.11** *Si $U$ y $V$ son dos espacios vectoriales, simplificaremos la notación del producto tensorial $U \otimes \overset{r}{\cdots} \otimes U \otimes V \otimes \overset{s}{\cdots} \otimes V$ utilizando en su lugar la abreviatura $(U)^r \otimes (V)^s$.*

**Corolario 1.3.12** *Dado un espacio vectorial $V$ sobre $\mathbb{K}$ existe un isomorfismo canónico entre el espacio vectorial $\mathcal{T}_r^s(V)$ de los tensores $r$ veces covariantes y $s$ veces contravariantes sobre $V$ y el producto tensorial $(V^*)^r \otimes (V)^s$.*

<u>*Demostración.*</u>

Hemos definido $\mathcal{T}_r^s(V)$ como el espacio vectorial de las aplicaciones multilineales de $(V)^r \times (V^*)^s$ en $\mathbb{K}$, lo que denotamos por $\mathcal{L}\big((V)^r \times (V^*)^s, \mathbb{K}\big)$. Por la propiedad universal extendida (Teorema 1.3.7), sabemos que existe, salvo isomorfismos, un único espacio vectorial $(V)^r \otimes (V^*)^s$ y un operador multilineal asociado $m : (V)^r \times (V^*)^s \to (V)^r \otimes (V^*)^s$ tales que para toda aplicación multilineal $b : (V)^r \times (V^*)^s \to \mathbb{K}$ existe una única aplicación lineal $l : (V)^r \otimes (V^*)^s \to \mathbb{K}$ tal que $l \circ m = b$.

$$\begin{tikzcd}
(V)^r \times (V^*)^s \arrow[r, "m"] \arrow[dr, "b"'] & (V)^r \otimes (V^*)^s \arrow[d, dashed, "l"] \\
& \mathbb{K}
\end{tikzcd}$$

Deducimos por tanto que se tiene $\mathcal{L}\big((V)^r \times (V^*)^s, \mathbb{K}\big) \cong \mathcal{L}\big((V)^r \otimes (V^*)^s, \mathbb{K}\big)$. Por último, aplicando 1.3.10, obtenemos

$$\mathcal{T}_r^s(V) \cong \mathcal{L}\big((V)^r \otimes (V^*)^s, \mathbb{K}\big) = \big((V)^r \otimes (V^*)^s\big)^* \cong (V^*)^r \otimes (V)^s$$

$\square$

A continuación se muestra la noción de producto tensorial de tensores, operación muy extendida en la práctica y que será de gran utilidad en la demostración del resultado 1.3.15.

**Definición 1.3.13** *Sean $T_1 \in \mathcal{T}_r^s(V)$ y $T_2 \in \mathcal{T}_{r'}^{s'}(V)$. Se define el producto tensorial de los tensores $T_1$ y $T_2$ como la forma multilineal $T_1 \otimes T_2 \in \mathcal{T}_{r+r'}^{s+s'}(V)$ dada por*

$$(T_1 \otimes T_2)(\underline{x}_1, ..., \underline{x}_{r+r'}, \underline{y}^1, ..., \underline{y}^{s+s'}) = T_1(\underline{x}_1, ..., \underline{x}_r, \underline{y}^1, ..., \underline{y}^s) \cdot T_2(\underline{x}_{r+1}, ..., \underline{x}_{r'}, \underline{y}^{s+1}, ..., \underline{y}^{s'})$$

**Proposición 1.3.14** *Dados $T_1, T_2, T_3 \in \mathcal{T}_r^s(V)$ y $\alpha, \beta \in \mathbb{K}$, se tiene*

1. $T_1 \otimes (T_2 \otimes T_3) = (T_1 \otimes T_2) \otimes T_3$.
2. $(\alpha T_1 + \beta T_2) \otimes T_3 = \alpha(T_1 \otimes T_3) + \beta(T_2 \otimes T_3)$.
3. $T_1 \otimes (\alpha T_2 \otimes \beta T_3) = \alpha(T_1 \otimes T_2) + \beta(T_1 \otimes T_3)$.

**Proposición 1.3.15** *Sea $B = \{\underline{e}_1, ..., \underline{e}_n\}$ una base de $V$ y sea la base $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$ su base dual. Entonces el conjunto*

$$B_r^s = \left\{\underline{e}^{i_1} \otimes \cdots \otimes \underline{e}^{i_r} \otimes \underline{e}_{j_1} \otimes \cdots \otimes \underline{e}_{j_s} \,:\, i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}\right\}$$

*es una base de $\mathcal{T}_r^s(V)$. De hecho, cada $T \in \mathcal{T}_r^s(V)$ se expresa como*

$$T = t_{i_1 \cdots i_r}^{j_1 \cdots j_s}\, \underline{e}^{i_1} \otimes \cdots \otimes \underline{e}^{i_r} \otimes \underline{e}_{j_1} \otimes \cdots \otimes \underline{e}_{j_s} \tag{1.2}$$

*donde la coordenada $t_{i_1 \cdots i_r}^{j_1 \cdots j_s}$ de $T$ en $B_r^s$ coincide con $T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s})$.*

<u>*Demostración*</u>.

Veamos en primer lugar la segunda parte. Para probar que la expresión (1.2) es cierta, veamos que las aplicaciones lineales de ambos lados de la igualdad coinciden para todo elemento $(\underline{x}_1, ..., \underline{x}_r, \underline{y}^1, ..., \underline{y}^s) \in (V)^r \times (V^*)^s$.

Por 1.2.8 sabemos que existen unos coeficientes únicos $a_p^i$ y $b_i^q$ para cada $i \in \{1, ..., n\}$, $p \in \{1, ..., r\}$ y $q \in \{1, ..., s\}$ tales que $\underline{x}_p = a_p^i \underline{e}_i$ e $\underline{y}^q = a_i^j \underline{e}^j$ para todo $p \in \{1, ..., r\}$ y $q \in \{1, ..., s\}$, y además

$$T(\underline{x}_1, ..., \underline{x}_r, \underline{y}^1, ..., \underline{y}^s) = t_{i_1 \cdots i_r}^{j_1 \cdots j_s} \cdot a_1^{i_1} \cdots a_r^{i_r} \cdot b_{j_1}^1 \cdots b_{j_s}^s$$

donde $t_{i_1 \cdots i_r}^{j_1 \cdots j_s} = T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s})$, con $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$. Evaluando el lado derecho de 1.2, utilizando para ello el hecho de que $\underline{e}_i \in \mathcal{T}_0^1(V)$ y $\underline{e}^i \in \mathcal{T}_1^0(V)$ para todo $i$ y la definición 1.3.13, tenemos

$$\begin{aligned}
&\left(t^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}\otimes\cdots\otimes\underline{e}^{i_r}\otimes\underline{e}_{j_1}\otimes\cdots\otimes\underline{e}_{j_s}\right)(\underline{x}_1,...,\underline{x}_r,\underline{y}^1,...,\underline{y}^s)\\
&\qquad= t^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}(\underline{x}_1)\cdots\underline{e}^{i_r}(\underline{x}_r)\cdot\underline{e}_{j_1}(\underline{y}^1)\cdots\underline{e}_{j_s}(\underline{y}^s)\\
&\qquad= t^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}(a^{k_1}_1\underline{e}_{k_1})\cdots\underline{e}^{i_r}(a^{k_r}_r\underline{e}_{k_r})\cdot\underline{e}_{j_1}(b^1_{l_1}\underline{e}^{l_1})\cdots\underline{e}_{j_s}(b^s_{l_s}\underline{e}^{s_1})\\
&\qquad= t^{\,j_1\cdots\, j_s}_{\,i_1\cdots\, i_r}\cdot a^{i_1}_1\cdots a^{i_r}_r\cdot b^1_{j_1}\cdots b^s_{j_s}
\end{aligned}$$

Por tanto se tiene que, para todo $(\underline{x}_1,...,\underline{x}_r,\underline{y}^1,...,\underline{y}^s)\in(V)^r\times(V^*)^s$,

$$T(\underline{x}_1,...,\underline{x}_r,\underline{y}^1,...,\underline{y}^s)=\left(t^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}\otimes\cdots\otimes\underline{e}^{i_r}\otimes\underline{e}_{j_1}\otimes\cdots\otimes\underline{e}_{j_s}\right)(\underline{x}_1,...,\underline{x}_r,\underline{y}^1,...,\underline{y}^s)$$
$$\Rightarrow T=t^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}\otimes\cdots\otimes\underline{e}^{i_r}\otimes\underline{e}_{j_1}\otimes\cdots\otimes\underline{e}_{j_s}$$

lo que prueba además que los elementos de $B^s_r$ generan el espacio $\mathcal{T}^s_r(V)$. Para probar que $B^s_r$ es base de dicho espacio nos falta ver su independencia lineal. Consideramos para ello una combinación lineal igualada al tensor nulo, esto es,

$$a^{j_1\cdots j_s}_{i_1\cdots i_r}\,\underline{e}^{i_1}\otimes\cdots\otimes\underline{e}^{i_r}\otimes\underline{e}_{j_1}\otimes\cdots\otimes\underline{e}_{j_s}=0$$

Evaluando este tensor en $(\underline{e}_{k_1},...,\underline{e}_{k_r},\underline{e}^{l_1},...,\underline{e}^{l_s})$ para cualesquiera $k_1,...,k_r,l_1,...,l_s\in\{1,...,n\}$, obtenemos

$$0=a^{j_1\cdots j_s}_{i_1\cdots i_r}\delta^{i_1}_{k_1}\cdots\delta^{i_r}_{k_r}\cdot\delta^{l_1}_{j_1}\cdots\delta^{l_s}_{j_s}=a^{l_1\cdots l_s}_{k_1\cdots k_r}$$

y por tanto los productos tensoriales de $B^s_r$ son linealmente independientes, a partir de lo cual se concluye que $B^s_r$ es base de $\mathcal{T}^s_r(V)$.

□

**Corolario 1.3.16** *Si $V$ es un espacio vectorial de dimensión finita $n$, entonces la dimensión del espacio vectorial $\mathcal{T}^s_r(V)$ es $n^{r+s}$.*

Nótese que, aunque en 1.3.15 hayamos probado que todo tensor de $\mathcal{T}^s_r(V)$ puede expresarse como una combinación lineal de productos tensoriales, esto no implica que todo tensor sea un producto tensorial.

**Definición 1.3.17** *Dado un espacio vectorial $V$, un tensor covariante $T\in\mathcal{T}^0_r(V)$ (resp. contravariante $T\in\mathcal{T}^s_0(V)$) es simétrico respecto de los índices $p$ y $q$ si las componentes del tensor en una base cualquiera verifican $t_{i_1,...,i_p,...,i_q,...,i_r}=t_{i_1,...,i_q,...,i_p,...,i_r}$ (resp. $t^{j_1,...,j_p,...,j_q,...,j_s}=t^{j_1,...,j_q,...,j_p,...,j_s}$).*
*Si además se verifica que $t_{i_1,i_2,...,i_r}=t_{i_{\sigma(1)},i_{\sigma(2)},...,i_{\sigma(r)}}$ (resp. $t^{j_1,j_2,...,j_s}=t^{j_{\sigma(1)},j_{\sigma(2)},...,j_{\sigma(r)}}$) para toda permutación $\sigma$ del conjunto $\{1,...,r\}$ (resp. $\{1,...,s\}$), el tensor se denomina simétrico.*

Estas definiciones se extienden sin dificultad a tensores mixtos, tomando parejas de índices ambos covariantes o ambos contravariantes.

## 1.4. Campos tensoriales

Al igual que los tensores generalizan los conceptos de escalar y vector, un campo tensorial es una generalización de los campos escalares y vectoriales, asignando a cada punto del espacio un tensor.

### 1.4.1. Variedades diferenciables

Comenzaremos el estudio de los campos tensoriales analizando los conceptos y resultados necesarios para la definición de campo tensorial. En esta sección nos limitaremos únicamente a enunciar los resultados, centrando la atención en el contenido de los mismos y dejando al lector la profundización en ellos (véase [8], [2]).

**Definición 1.4.1** *Sea $M$ un espacio topológico Hausdorff verificando el segundo axioma de numerabilidad.*

- *Una carta n-dimensional de $M$ se define como una terna $(U, V, x)$ donde $U$ es un abierto de $M$, $V$ es un abierto de $\mathbb{R}^n$ y $x : U \to V$ es un homeomorfismo.*
- *Dadas dos cartas $(U_1, V_1, x_1)$ y $(U_2, V_2, x_2)$ de la misma dimensión y tales que $U_1 \cap U_2 \neq \emptyset$, se define el cambio de carta de $(U_1, V_1, x_1)$ a $(U_2, V_2, x_2)$ como el homeomorfismo*

$$x_2 \circ x_1 : x_1(U_1 \cap U_2) \to x_2(U_1 \cap U_2)$$

- *Se define atlas de la estructura $M$ a toda familia de cartas n-dimensionales $\{(U_i, V_i, x_i)\}_{i \in I}$ tal que $M = \bigcup_{i \in I} U_i$ y todos los cambios de carta son diferenciables (de clase $\mathcal{C}^\infty$).*

**Definición 1.4.2** *Diremos que un espacio topológico Hausdorff que verifica el segundo axioma de numerabilidad, $M$, tiene estructura de variedad diferenciable n-dimensional, si existe un atlas asociado a dicha estructura formado por cartas n-dimensionales.*

**Definición 1.4.3** *Sea $M$ una variedad diferenciable. Una aplicación $F : M \to \mathbb{R}^m$ es diferenciable en $p \in M$ (de clase $\mathcal{C}^\infty$) si existe una carta local en $p$, $(U, V, x)$, tal que la expresión local de $F$ en la carta, $F \circ x^{-1} : V \to \mathbb{R}^m$, es diferenciable en $x(p)$.*

*Diremos que la aplicación $F : M \to \mathbb{R}^m$ es diferenciable en un abierto $W$ de $M$ si es diferenciable en todo $p \in W$.*

Un ejemplo de aplicaciones diferenciables son las funciones coordenadas de una carta:

**Ejemplo 1.4.4** *Dada una variedad diferenciable $M$ y una carta $(U, V, x)$ de $M$, se define para cada $i \in \{1, ..., n\}$ la función coordenada $x^i : U \to \mathbb{R}$ asociada a la carta, dada por $x^i = pr_i \circ x$, donde $pr_i$ representa la proyección i-ésima de $V \subset \mathbb{R}^n$ en $\mathbb{R}$.*

$$\begin{tikzcd} U \arrow[r, "x"] \arrow[dr, "x^i"'] & V \arrow[d, "pr_i"] \\ & \mathbb{R} \end{tikzcd}$$

*Esta función es diferenciable en $U$ puesto que su expresión local en la carta $(U, V, x)$ viene dada por $(x^i)^{-1} \circ x = pr_i$.*

**Definición 1.4.5** *Sea $M$ una variedad diferenciable y $(U, V, x)$ una carta de $M$. El abierto $U$ de $M$ junto con las funciones coordenadas de la carta $x^1, ..., x^n$ constituyen un abierto coordenado de la variedad diferenciable $M$, el cual denotamos por $(U; x^1, ..., x^n)$. La aplicación*

$$\begin{aligned} (x^1, ..., x^n) : U &\to \mathbb{R}^n \\ p &\mapsto \big(x^1(p), ..., x^n(p)\big) \end{aligned}$$

*establece un difeomorfismo de $U$ en $\mathbb{R}^n$, y los elementos $x^1(p), ..., x^n(p)$ se denominan coordenadas locales de $p$ en $(U, V, x)$.*

El siguiente paso consiste en definir el concepto de germen de función, a partir del cual podremos derivar la definición de espacio tangente a una variedad.

**Definición 1.4.6** *Sea $M$ una variedad diferenciable y sea $p \in M$. Consideramos el conjunto de todos los pares $(U, h)$ donde $U$ es abierto de $M$ con $p \in U$, y $h : U \to \mathbb{R}$ es diferenciable. Definimos sobre este conjunto la relación de equivalencia:*

$$(U_1, h_1) \sim (U_2, h_2) \iff \exists\, W \text{ abierto } / \;\; W \subset U_1 \cap U_2, \;\; p \in W \text{ y } h_1|_W = h_2|_W$$

*Cada clase de equivalencia por la relación anterior recibe el nombre de germen de función diferenciable en $p$.*

Denotamos por $[(U, h)]$ la clase de equivalencia del par $(U, h)$. En aquellos casos en los que no sea necesario conocer el abierto $U$ sobre el que se define la correspondiente función $h$, simplificamos la notación denotando por $[h]_p$ la clase de equivalencia de $(U, h)$, o incluso únicamente por $[h]$ o $h$ sobreentendiendo en cada contexto el punto $p \in M$ sobre el que se trabaja.

**Proposición 1.4.7** *Sobre el conjunto de gérmenes de funciones diferenciables en $p$, definimos las operaciones*

- $[(U_1, h_1)] + [(U_2, h_2)] = [(U_1 \cap U_2, h_1 + h_2)]$
- $[(U_1, h_1)] \cdot [(U_2, h_2)] = [(U_1 \cap U_2, h_1 h_2)]$

*Estas operaciones no dependen de los representantes elegidos, y dotan al conjunto cociente de estructura de anillo conmutativo y unitario.*

**Definición 1.4.8** *Dada una variedad diferenciable $M$, el conjunto de gérmenes de funciones diferenciables en $p \in M$ dotado de las operaciones anteriores recibe el nombre de anillo de gérmenes de funciones diferenciables en $p$, y se denota por $\mathcal{O}_{M,p}$.*

En adelante, cuando no haya confusión sobre la variedad diferenciable $M$ a la que pertenece $p$, para simplificar denotaremos $\mathcal{O}_p$ en lugar de $\mathcal{O}_{M,p}$.

Es claro que podemos ver $\mathbb{R}$ como el subanillo de $\mathcal{O}_p$ formado por los gérmenes de las funciones constantes, y por tanto $\mathcal{O}_p$ tiene estructura de $\mathbb{R}$-álgebra, considerando para ello el producto $\alpha[h]_p = [\alpha\, h]_p$ para todo $\alpha \in \mathbb{R}$. En particular, $\mathcal{O}_p$ tiene estructura de $\mathbb{R}$-espacio vectorial.

**Proposición 1.4.9** *El anillo $\mathcal{O}_p$ de gérmenes de funciones diferenciables en $p$ es local, es decir, posee un único ideal maximal, concretamente el ideal $m_p = \{[h]_p \in \mathcal{O}_p \,:\, h(p) = 0\}$.*

**Definición 1.4.10** *Llamaremos derivación de $M$ centrada en $p$ a toda aplicación $\delta : \mathcal{O}_p \to \mathbb{R}$ tal que*

1. *$\delta$ es $\mathbb{R}$-lineal, esto es, $\delta(\lambda[h]_p + \mu[g]_p) = \lambda\, \delta([h]_p) + \mu\, \delta([g]_p)$, $\forall\, [h]_p, [g]_p \in \mathcal{O}_p$ y $\lambda, \mu \in \mathbb{R}$.*
2. *$\delta$ verifica la regla de Leibniz: $\delta([h]_p[g]_p) = h(p)\, \delta([g]_p) + g(p)\, \delta([h]_p)$, $\forall\, [h]_p, [g]_p \in \mathcal{O}_p$.*

**Proposición 1.4.11** *El conjunto de todas las derivaciones de $M$ centradas en $p$, con la suma y producto por escalares definidos del modo habitual, es un espacio vectorial sobre $\mathbb{R}$.*

**Definición 1.4.12** *Dada una variedad diferenciable $M$ de dimensión $n$, el espacio tangente a $M$ en $p$, denotado por $T_p(M)$ , es el $\mathbb{R}$-espacio vectorial de las derivaciones centradas en $p$. Además, se define el espacio vectorial cotangente $T_p(M)^*$ , como el espacio dual de $T_p(M)$.*

Volviendo al resultado 1.4.9, consideramos ahora el ideal $m_p^2$ generado por los productos $[h_1]_p[h_2]_p$ con $[h_1]_p, [h_2]_p \in m_p$. Tanto $m_p$ como $m_p^2$ son espacios vectoriales sobre $\mathbb{R}$, con $m_p^2 \subset m_p$, y tomamos ahora el espacio vectorial cociente $m_p/m_p^2$ sobre $\mathbb{R}$. Denotaremos por $\overline{[g]}_p$ la clase módulo $m_p^2$ de $[g]_p$, definida como $\overline{[g]}_p \coloneqq \{[g]_p + [h]_p \,:\, [h]_p \in m_p^2\}$. A partir de este espacio obtenemos el siguiente resultado:

**Teorema 1.4.13** *Dada una variedad diferenciable $M$, y un elemento $p \in M$, existe un isomorfismo canónico entre el espacio cotangente, $T_p(M)^*$, y el espacio cociente ${}^{m_p}\!/_{m_p^2}$.*

*Demostración.*

Comenzamos la prueba viendo que $T_p(M)$ es canónicamente isomorfo a $\left({}^{m_p}\!/_{m_p^2}\right)^*$. En primer lugar, tomamos $\delta \in T_p(M)$, una derivación centrada en $p$, la cual verifica:

1. $\delta(\lambda) = 0$ para todo $\lambda \in \mathbb{R}$ (germen constante).
2. $\delta([h]_p) = 0$ si $[h]_p \in m_p^2$ :

   Un elemento $[h]_p$ de $m_p^2$ es de la forma $[h]_p = \sum\limits_{i=1}^{r} [f_i]_p\, [g_i]_p$, con $[f_i]_p\, [g_i]_p \in m_p$, luego $f_i(p) = g_i(p) = 0$, para $i \in \{1, ..., r\}$. Entonces

$$\delta([h]_p) = \delta\Bigg( \sum_{i=1}^{r} [f_i]_p\, [g_i]_p \Bigg) = \sum_{i=1}^{r} \Big( f_i(p)\, \delta([g_i]_p) + g_i(p)\, \delta([f_i]_p) \Big) = 0$$

La derivación $\delta$ induce la aplicación lineal

$$\begin{aligned} D_\delta : {}^{m_p}\!/_{m_p^2} &\to \mathbb{R} \\ \overline{[g]}_p &\mapsto D_\delta(\overline{[g]}_p) := \delta([g]_p) \end{aligned}$$

la cual es lineal por la linealidad de $\delta$, y está bien definida gracias a la propiedad (2) anterior: dados $[g_1]_p \neq [g_2]_p \in m_p$ con $\overline{[g_1]}_p = \overline{[g_2]}_p$, se tiene que $\overline{[g_1]}_p - \overline{[g_2]}_p \in m_p^2$, luego

$$D(\overline{[g_1]}_p) - D(\overline{[g_2]}_p) = D_\delta(\overline{[g_1]}_p - \overline{[g_2]}_p) = \delta(\overline{[g_1]}_p - \overline{[g_2]}_p) = 0 \;\Rightarrow\; D_\delta(\overline{[g_1]}_p) = D_\delta(\overline{[g_2]}_p)$$

Obtenemos así el operador lineal $\phi_1$ entre los espacios $T_p(M)$ y $\left({}^{m_p}\!/_{m_p^2}\right)^*$:

$$\begin{aligned} \phi_1 \colon T_p(M) &\to \left({}^{m_p}\!/_{m_p^2}\right)^* \\ \delta \quad &\longmapsto \quad D_\delta \; : \; {}^{m_p}\!/_{m_p^2} \to \mathbb{R} \\ &\qquad\qquad\quad \overline{[g]}_p \mapsto D_\delta(\overline{[g]}_p) := \delta([g]_p) \end{aligned}$$

Recíprocamente, si $D : {}^{m_p}\!/_{m_p^2} \to \mathbb{R}$ es una aplicación lineal dada, esta induce la aplicación

$$\begin{aligned} \delta_D : \mathcal{O}_p &\to \mathbb{R} \\ [g]_p &\mapsto \delta_D([g]_p) := D\big(\overline{[g - g(p)]}_p\big) \end{aligned}$$

donde $\overline{[g-g(p)]}_p$ es la clase módulo $m_p^2$ de $[g]_p - g(p) = [g-g(p)]_p$. Es fácil ver que $\delta_D$ es una derivación centrada en $p$, y por tanto, $\delta_D \in T_p(M)$. Obtenemos así el operador lineal $\phi_2$:

$$\begin{aligned} \phi_2 \colon \left( {}^{m_p}\!/\!_{m_p^2} \right)^* &\to T_p(M) \\ D \quad &\longmapsto \quad \delta_D \colon \mathcal{O}_p \to \mathbb{R} \\ & \qquad\qquad [g]_p \mapsto \delta_D([g]_p) := D(\overline{[g-g(p)]}_p) \end{aligned}$$

Observemos por último que se tiene $\phi_2 \circ \phi_1 \cong Id_{T_p(M)}$ y $\phi_1 \circ \phi_2 \cong Id_{\left( {}^{m_p}\!/\!_{m_p^2} \right)^*}$ para probar de esta forma que $T_p(M) \cong \left( {}^{m_p}\!/\!_{m_p^2} \right)^*$.

- Consideramos la aplicación $\phi_2 \circ \phi_1 : T_p(M) \to T_p(M)$. Para todo $\delta \in T_p(M)$ y todo $[g]_p \in \mathcal{O}_p$, se tiene

$$\begin{aligned} (\phi_2 \circ \phi_1)(\delta([g]_p)) &= \delta_{D_\delta}([g]_p) = D_\delta(\overline{[g-g(p)]}_p) = \delta([g-g(p)]_p) = \delta([g]_p - [g(p)]_p) \\ &= \delta([g]_p) - \delta(g(p)) = \delta([g]_p) - 0 = \delta([g]_p) \quad \Rightarrow \quad \phi_2 \circ \phi_1 \cong Id_{T_p(M)} \end{aligned}$$

- Consideramos ahora la aplicación $\phi_1 \circ \phi_2 : \left( {}^{m_p}\!/\!_{m_p^2} \right)^* \to \left( {}^{m_p}\!/\!_{m_p^2} \right)^*$. Para todo $D \in \left( {}^{m_p}\!/\!_{m_p^2} \right)^*$ y para todo $\overline{[g]}_p \in {}^{m_p}\!/\!_{m_p^2}$ (es decir, para todo $[g]_p \in m_p$ con $\overline{[g]}_p = [g]_p + m_p^2$), se tiene:

$$(\phi_1 \circ \phi_2)(D(\overline{[g]}_p)) = D_{\delta_D}(\overline{[g]}_p) = \delta_D([g]_p) = D(\overline{[g-g(p)]}_p) = D(\overline{[g]}_p) \text{ [puesto que } [g]_p \in m_p]$$

$$\Rightarrow \quad \phi_1 \circ \phi_2 \cong Id_{\left( {}^{m_p}\!/\!_{m_p^2} \right)^*}$$

Como el bidual de un espacio sabemos que es canónicamente isomorfo al espacio original, obtenemos que el espacio ${}^{m_p}\!/\!_{m_p^2}$ es canónicamente isomorfo al espacio cotangente $T_p(M)^*$.

$\square$

Siguiendo con la notación del teorema, dado un germen $[f]_p \in \mathcal{O}_p$, se tiene $[f - f(p)]_p \in m_p$, y llamamos *diferencial de $f$ en $p$* a la clase $\overline{[f-f(p)]_p} \in m_p/m_p^2$, que denotamos como $df|_p$, o simplemente $df$ cuando no dé lugar a confusión.
Se define la derivación $\frac{\partial}{\partial x^n}\big|_p : \mathcal{O}_p \to \mathbb{R}$ por $\frac{\partial}{\partial x^n}\big|_p([f]_p) := \frac{\partial f \circ x^{-1}}{\partial x^n}\big(x(p)\big)$ para todo $[f]_p \in \mathcal{O}_p$.

**Proposición 1.4.14** *Si $(U; x^1, ..., x^n)$ es un abierto coordenado asociado a la variedad diferenciable $M$, y $x^1(p), ..., x^n(p)$ son las coordenadas locales en $p \in U$, el conjunto $\left\{ \frac{\partial}{\partial x^1}\big|_p, ..., \frac{\partial}{\partial x^n}\big|_p \right\}$ es base del espacio tangente $T_p(M)$, y el conjunto $\{ dx^1\big|_p, ..., dx^n|_p \}$ es la correspondiente base dual del espacio $T_p(M)^*$, y por tanto se verifica $dx^i\left( \frac{\partial}{\partial x^i}\big|_p \right)\Big|_p = \delta_j^i$, para todos $i, j \in \{1, ..., n\}$.*

**Corolario 1.4.15** *Dada una variedad diferenciable $M$ y un punto $p \in M$, los espacios $T_p(M)$ y $T_p(M)^*$ son espacios vectoriales de dimensión $n$.*

**Notación 1.4.16** *Si $(U; x^1, ..., x^n)$ es un abierto coordenado asociado a la variedad diferenciable $M$, la asignación a cada $q \in U$ de $dx^i\big|_q$ se denota por $dx^i$. En el caso de las derivadas parciales, simplificaremos la notación anteriormente utilizada, pasando a denotar la asignación a cada $q \in U$ de $\frac{\partial}{\partial x^i}\big|_q$ por $\partial_i$.*

A efectos del convenio de sumación de Einstein, para la notación $\frac{\partial}{\partial x^i}$, el índice $i$ se considera subíndice. De esta forma, un vector tangente de $T_p(M)$ se puede denotar de forma única como $a^i \frac{\partial}{\partial x^i}$, o alternativamente como $a^i\, \partial_i$, y un vector cotangente de $T_p(M)^*$ se puede denotar de forma única como $b_i\, dx^i$.

**Proposición 1.4.17** *Sea $(U; x^1, ..., x^n)$ un abierto coordenado asociado a la variedad diferenciable $M$. Dada una función $g : W \to \mathbb{R}$ diferenciable, si $U \subset W$, es decir, si $g$ está definida en el abierto coordenado, se tiene*

$$dg|_U = \frac{\partial g}{\partial x^i}\, dx^i$$

*donde $dg|_U$ denota la restricción de $dg$ a $U$.*

### 1.4.2. Campos tensoriales sobre variedades diferenciables

Existen dos enfoques a los campos tensoriales sobre variedades diferenciables, que resultan ser equivalentes para variedades de dimensión finita. En este trabajo introduciremos los campos tensoriales utilizando aplicaciones multilineales asociadas al espacio vectorial tangente $T_p(M)$, y asignando "de forma diferenciable" a cada punto de la variedad un tensor sobre dicho espacio.

El segundo enfoque posible parte de la teoría clásica de los $\mathcal{C}^\infty$-módulos $\mathcal{T}(M)$ y $\mathcal{T}^*(M)$ de campos vectoriales y 1-formas respectivamente, y a partir de ellos define los campos tensoriales como el $\mathcal{C}^\infty$-módulo de las aplicaciones $\big(\mathcal{T}(M)\big)^r \times \big(\mathcal{T}^*(M)\big)^s \to \mathcal{C}^\infty$ que son $\mathcal{C}^\infty$-multilineales.

Para nuestro caso a estudio consideramos una variedad diferenciable $M$ y un elemento $p \in M$, y tomamos el espacio vectorial $\mathcal{T}_r^s\big(T_p(M)\big)$ de los tensores de tipo $(r, s)$ sobre el espacio tangente a $M$ en $p$, formado por las aplicaciones multilineales

$$T : T_p(M) \times \overset{r}{\cdots} \times T_p(M) \times T_p(M)^* \times \overset{s}{\cdots} \times T_p(M)^* \longrightarrow \mathbb{R}$$

Consideramos ahora el conjunto

$$
\begin{aligned}
T_r^s(M) &:= \bigsqcup_{p\in M} \mathcal{T}_r^s\big(T_p(M)\big) \quad \text{[unión disjunta]}\\
&\equiv \bigcup_{p\in M} \big\{(p,T) : T \in \mathcal{T}_r^s\big(T_p(M)\big)\,\big\} = \big\{(p,T) : p\in M, T\in \mathcal{T}_r^s\big(T_p(M)\big)\big\}
\end{aligned}
$$

**Proposición 1.4.18** *El conjunto $T_r^s(M)$ sobre una variedad diferenciable $M$ admite estructura natural de variedad diferenciable de dimensión $n + n^{r+s}$, tal que la aplicación de proyección*

$$
\begin{aligned}
\pi : T_r^s(M) &\to M\\
(p,T) &\mapsto p
\end{aligned}
$$

*es diferenciable y de rango $n^{r+s}$ en todos los puntos.*

**Definición 1.4.19** *Dada $M$ una variedad diferenciable de dimensión $n$, se define el espacio total del fibrado tensorial de tipo $(r,s)$ sobre $M$ como la variedad diferenciable*

$$
T_r^s(M) := \bigsqcup_{p\in M} \mathcal{T}_r^s\big(T_p(M)\big)
$$

*El fibrado tensorial de tipo $(r,s)$ sobre la variedad diferenciable $M$ se define como el par formado por el espacio total del fibrado tensorial, $T_r^s(M)$ y la aplicación diferenciable de proyección, $\pi$.*

Aunque los fibrados tensoriales consisten también en la aplicación de proyección, en lo que sigue pondremos énfasis en el espacio total del fibrado puesto que, al tomar coordenadas locales, son suficientes estas coordenadas junto con el espacio total para describir los campos tensoriales y trabajar con ellos.

A partir del concepto de fibrado tensorial (cuyas propiedades y resultados exceden los objetivos de este trabajo y por ello nos limitamos a hacer uso únicamente de su definición) obtenemos la noción de campo tensorial:

**Definición 1.4.20** *Un campo tensorial de tipo $(r,s)$ sobre un abierto $W \subset M$, es una función $T : W \to T_r^s(M)$ tal que $T(p) \in \mathcal{T}_r^s\big(T_p(M)\big)$ para todo $p \in W$.*

**Ejemplos 1.4.21**

1. *Si $r = s = 0$, un campo tensorial de tipo $(0,0)$ asigna un escalar a cada elemento $p \in W \subset M$, por ello estos campos tensoriales son simplemente funciones reales $T : W \to \mathbb{R}$.*

2. *Si $r = 0$ y $s = 1$, los campos tensoriales de tipo $(0,1)$ se denominan campos vectoriales. Los campos vectoriales asignan a cada $p \in W \subset M$ un elemento del espacio tangente $T_p(M)$.*

3. *Si $r=1$ y $s=0$, los campos tensoriales de tipo $(1,0)$ reciben el nombre de formas diferenciales (de clase $\mathcal{C}^\infty$), o 1-formas . Las 1-formas asignan a cada $p \in W \subset M$ un elemento del espacio cotangente $T_p(M)^*$.*

Veamos ahora cómo definir la diferenciabilidad de los campos tensoriales, para lo cual partimos de los casos concretos de diferenciabilidad de campo vectorial y 1-forma:

**Definición 1.4.22** *Sea $M$ una variedad diferenciable, y $W \subseteq W$ un abierto no vacío.*

- *Un campo vectorial $D : W \to T_0^1(M)$ se dice diferenciable (de clase $\mathcal{C}^\infty$) si para toda función $f \in \mathcal{C}^\infty(W)$, se cumple que la función*

$$\begin{aligned} D(f) : W &\to \mathbb{R} \\ p &\mapsto D(p)(f) \end{aligned}$$

  *es diferenciable. El conjunto de todos los campos vectoriales diferenciables en $W$ se denota por $\mathcal{T}(W)$.*

- *Una 1-forma $\theta : W \to T_1^0(M)$ se dice diferenciable (de clase $\mathcal{C}^\infty$) si para todo campo vectorial diferenciable $D \in \mathcal{T}(W)$, se cumple que la función*

$$\begin{aligned} \theta(D) : W &\to \mathbb{R} \\ p &\mapsto \theta(p)\big(D(p)\big) \end{aligned}$$

  *es diferenciable. El conjunto de las 1-formas diferenciables en $W$ se denota por $\mathcal{T}^*(W)$.*

- *Un campo tensorial $T : W \to T_r^s(M)$ se dice diferenciable (de clase $\mathcal{C}^\infty$) si para todos los campos vectoriales $D_1, ..., D_r \in \mathcal{T}(W)$ y todos las 1-formas $\theta^1, ..., \theta^s \in \mathcal{T}^*(W)$, se cumple que la función*

$$\begin{aligned} T(D_1, ..., D_r, \theta^1, ..., \theta^s) : W &\to \mathbb{R} \\ p &\mapsto T(p)\big(D_1(p), ..., D_r(p), \theta^1(p), ..., \theta^s(p)\big) \end{aligned}$$

  *es diferenciable. El conjunto de los campos tensoriales diferenciables de tipo $(r,s)$ sobre el abierto $W \subseteq M$, se denota por $\mathcal{T}_r^s(W)$.*

Fijado el par $(r,s)$, sobre el conjunto $\mathcal{T}_r^s(M)$ podemos definir del modo habitual las operaciones de suma y producto por funciones $f \in \mathcal{C}^\infty$:

- Dados $T_1, T_2 \in \mathcal{T}_r^s(M)$, se define para todo $p \in M$ la suma $(T_1+T_2)(p) = T_1(p) + T_2(p)$.

- Dados $T \in \mathcal{T}_r^s(M)$ y $f \in \mathcal{C}^\infty$, se define para cada $p \in M$ el producto de un campo tensorial por un elemento de $\mathcal{C}^\infty$ como $(f \cdot T)(p) = f(p) \cdot T(p)$.

Ambas operaciones están bien definidas dada la estructura de espacio vectorial del espacio $\mathcal{T}_r^s\big(T_p(M)\big)$. Se deduce así el siguiente resultado:

**Proposición 1.4.23** *Dada una variedad diferenciable $M$, el conjunto $\mathcal{T}_r^s(M)$ posee estructura de módulo sobre el anillo de funciones diferenciables $\mathcal{C}^\infty(M)$. Además, dado que todo abierto no vacío $W$ de $M$ es también una variedad diferenciable (con la estructura inducida de $M$), se deduce que el conjunto $T_r^s(W)$ es también un módulo sobre el anillo de funciones diferenciables sobre $W$, $\mathcal{C}^\infty(W)$.*

Las respectivas estructuras de módulo sobre los conjuntos de campos tensoriales son compatibles con las restricciones (los campos tensoriales definidos sobre $M$ se restringen a campos tensoriales definidos sobre $W$). A continuación, mostraremos cómo cuando tomamos $W = U$ un abierto coordenado, gracias a las coordenadas, los módulos de campos tensoriales definidos sobre $U$ tienen bases (es decir, son módulos libres) y, por tanto, se pueden manejar cómo los espacios vectoriales tomando como escalares a las funciones de $\mathcal{C}^\infty(U)$.

**Definición 1.4.24** *Sean $(U; x^1, ..., x^n)$ un abierto coordenado asociado a la variedad diferenciable $M$ de dimensión $n$, y $T$ un campo tensorial de tipo $(r, s)$. Se definen las componentes de $T$ con respecto al abierto coordenado $(U; x^1, ..., x^n)$, como las funciones*

$$
\begin{aligned}
h_{i_1\cdots i_r}^{j_1\cdots j_s} = T\big(\partial_{i_1}, ..., \partial_{i_r}, dx^{j_1}, ..., dx^{j_s}\big) : U &\to \mathbb{R} \\
p &\mapsto T(p)\big(\partial_{i_1}|_p, ..., \partial_{i_r}|_p, dx^{j_1}|_p, ..., dx^{j_s}|_p\big)
\end{aligned}
$$

*para todo $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$.*

A partir de la definición 1.3.13 relativa al producto tensorial de dos tensores definidos sobre el mismo espacio vectorial, definimos el producto tensorial de campos tensoriales:

**Definición 1.4.25** *Sea $T_1 \in \mathcal{T}_r^s\big(T_p(M)\big)$ un campo tensorial definido en el abierto no vacío $W_1 \subseteq M$, y $T_2 \in \mathcal{T}_{r'}^{s'}\big(T_p(M)\big)$ un campo tensorial definido sobre el abierto no vacío $W_2 \subseteq M$, tales que $W = W_1 \cap W_2$ es no vacío. Se define el producto tensorial de $T_1$ y $T_2$ como el campo tensorial $T_1 \otimes T_2 \in \mathcal{T}_{r+r'}^{s+s'}\big(T_p(M)\big)$ definido sobre el abierto $W$, dado por*

$$(T_1 \otimes T_2)(p) = (T_1|_W)(p) \otimes (T_2|_W)(p)$$

*donde $T_1|_W$ y $T_2|_W$ denotan las restricciones a $W$ de ambos campos tensoriales.*

**Proposición 1.4.26** *Si $(U; x^1, ..., x^n)$ es un abierto coordenado asociado a la variedad diferenciable $M$, sobre los puntos de $p \in U$, un campo tensorial $T$ de tipo $(r, s)$ se expresa de la forma*

$$T = h_{i_1 \cdots i_r}^{j_1 \cdots j_s} \, dx^{i_1} \otimes \cdots \otimes dx^{i_r} \otimes \partial_{j_1} \otimes \cdots \otimes \partial_{j_s}$$

*con $h_{i_1 \cdots i_r}^{j_1 \cdots j_s}$ las componentes de $T$ en $(U; x^1, ..., x^n)$, $dx^{i_1}, ..., dx^{i_r} \in \mathcal{T}^*(U)$ y $\partial_{j_1}, ..., \partial_{j_s} \in \mathcal{T}(U)$ para todo $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$.*

Veamos por último cómo relacionar la diferenciabilidad local de los campos tensoriales con la diferenciabilidad de las componentes del campo tensorial.

**Proposición 1.4.27** *Sean $(U; x^1, ..., x^n)$ un abierto coordenado asociado a la variedad diferenciable $M$, y $T$ un campo tensorial de tipo $(r, s)$ definido en un abierto no vacío $W \subseteq M$ con $U \subseteq W$. Entonces $T \in \mathcal{C}^\infty(U)$ si y solo si $h_{i_1 \cdots i_r}^{j_1 \cdots j_s} = T\big(\partial_{i_1}, ..., \partial_{i_r}, dx^{j_1}, ..., dx^{j_s}\big) \in \mathcal{C}^\infty(U)$ para todos $i_1, ..., i_r, j_1, ..., j_s \in \{1, ..., n\}$.*

## 1.5. Métricas

Los tensores y campos tensoriales métricos, también conocidos como métricas tanto en el contexto vectorial como diferencial, constituyen una clase particularmente importante de tensores y campos tensoriales desde el punto de vista de las aplicaciones prácticas. Mas allá de su utilidad para la formulación de aspectos puramente métricos (distancias, ángulos, áreas...) en esta sección daremos forma a estos tensores y campos tensoriales métricos y nos centraremos

- desde el punto de vista vectorial, en la posibilidad de establecer un isomorfismo canónico entre un espacio y su dual en presencia de un tensor métrico, lo que derivará en el estudio de la manipulación de índices de tensores, y
- desde el punto de vista diferencial, en el estudio de las condiciones a imponer para asegurar la existencia de campos tensoriales métricos sobre variedades diferenciables.

### 1.5.1. Métricas en el contexto vectorial

**Definición 1.5.1** *Sea $V$ un espacio vectorial sobre $\mathbb{R}$ de dimensión finita $n$.*

- *Una métrica semi-euclídea es un tensor $T \in \mathcal{T}_2^0(V)$ de tipo $(2, 0)$ simétrico y no degenerado.*
- *Una métrica euclídea es una métrica semi-euclídea $T \in \mathcal{T}_2^0(V)$ cuya forma cuadrática definida a partir de la forma bilineal simétrica es definida positiva.*

Un tensor $T$ de tipo $(2,0)$ es una forma bilineal $g : V \times V \to \mathbb{R}$. Si tomamos una base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ de $V$ y su respectiva base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$, por 1.3.15 sabemos que $B_2^0 = \left\{e^i \otimes e^j \,:\, i, j \in \{1, ..., 2\}\right\}$ es base de $\mathcal{T}_2^0(V)$, y podemos expresar la forma bilineal como

$$g = g_{ij}\, \underline{e}^i \otimes \underline{e}^j$$

con $g_{ij} = g(\underline{e}_i, \underline{e}_j) \in \mathbb{R}$, para $i, j \in \{1, ..., n\}$.

Si consideramos la matriz $(g_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, el tensor $g$ es una métrica semi-euclídea si la matriz $(g_{ij})$ es simétrica y tiene determinante distinto de 0, y es una métrica euclídea si además todos los autovalores de $(g_{ij})$ son positivos (sabemos que son reales por ser la matriz simétrica).

Llamaremos de forma genérica tensor métrico a la métrica $g \in \mathcal{T}_2^0(V)$, independientemente si se trata de una métrica semi-euclídea o euclídea.

La noción de signatura de un tensor métrico, extensible para cualquier forma cuadrática real definida a partir de una forma bilineal simétrica real sobre un espacio vectorial de dimensión finita, permitirá dar una definición alternativa de métrica euclídea:

**Definición 1.5.2** *Se define la signatura de un tensor métrico $g$ como el número de autovalores positivos, negativos y nulos de la matriz simétrica $(g_{ij})$ con respecto a una determinada base.*

El Criterio de Sylvester asegura que la signatura del tensor no depende de la elección de la base. Por ello, denotaremos por $(n_+, n_-)$ la dupla compuesta por el número de autovalores positivos y negativos del tensor métrico (el número de autovalores idénticamente nulos se deriva de los datos anteriores a partir de la dimensión del espacio). De esta forma, podemos definir una métrica euclídea sobre un espacio vectorial de dimensión $n$ como una métrica semi-euclídea con signatura $(n, 0)$.

**Definición 1.5.3** *Sea $V$ un espacio vectorial sobre $\mathbb{R}$ de dimensión $n$. Si dotamos a $V$ de una métrica euclídea (resp. semi-euclídea) $g$, se dice que $(V, g)$ (o simplemente $V$ si no hay confusión) es un espacio euclídeo (resp. semi-euclídeo).*

**Observación 1.5.4** *Para el caso de las métricas euclídeas, remarcar que estas se corresponden con el conocido producto escalar en $\mathbb{R}$, determinado por aplicaciones*

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V &\to \mathbb{R} \\ (\underline{x}_1, \underline{x}_2) &\mapsto \langle \underline{x}_1, \underline{x}_2 \rangle \end{aligned}$$

*bilineales, simétricas y definidas positivas.*

**Proposición 1.5.5** *Sea $V$ un espacio vectorial y $g$ un tensor métrico sobre $V$. El operador*

$$
\begin{aligned}
\phi \colon V &\longrightarrow V^* \\
\underline{v} &\longmapsto T_{\underline{v}} \colon V \longrightarrow \mathbb{R} \\
&\qquad\qquad \underline{w} \longmapsto T_{\underline{v}}(\underline{w}) \coloneqq g(\underline{v}, \underline{w})
\end{aligned}
\tag{1.3}
$$

*define un isomorfismo entre los espacios $V$ y $V^*$.*

Demostración.

Es claro que $T_{\underline{v}} \in V^*$ gracias a la bilinealidad de $g$. Queremos ver que $\phi$ es un isomorfismo, para ello comenzamos viendo que es inyectivo. Sean $\underline{v}_1, \underline{v}_2 \in V$ tales que $T_{\underline{v}_1} \equiv T_{\underline{v}_2}$, lo que equivale a que $g(\underline{v}_1, \underline{w}) = g(\underline{v}_2, \underline{w})$ para todo $\underline{w} \in V$. Por tanto se tiene

$$g(\underline{v}_1, \underline{w}) - g(\underline{v}_2, \underline{w}) = g(\underline{v}_1 - \underline{v}_2, \underline{w}) = 0, \text{ para todo } \underline{w} \in V$$

y por ser $g$ no degenerado se tiene $\underline{v}_1 - \underline{v}_2 = 0$, $\underline{v}_1 = \underline{v}_1$, lo que prueba la inyectividad de $\phi$.

Veamos ahora que la transformación en sobreyectiva. Para ello, tomamos una aplicación $T \in V^*$, y queremos probar que existe $\underline{v} \in V$ tal que $T \equiv T_{\underline{v}}$. Consideramos para ello una base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ de $V$ y su base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$, a partir de las cuales se tienen las expresiones $T = a_i\, \underline{e}^i$ y $\underline{v} = b^i\, \underline{e}_i$.

Para ver cuáles han de ser los coeficientes en la base $B$ del vector $\underline{v}$ buscado, suponemos $T \equiv T_{\underline{v}}$, que equivale a que $T$ y $T_{\underline{v}}$ coincidan en los elementos de la base $B$:

- $T(\underline{e}_i) = a_j\, \underline{e}^j(\underline{e}_i) = a_i$
- $T_{\underline{v}}(\underline{e}_i) = g(\underline{v}, \underline{e}_i) = g(b^j\, \underline{e}_j, \underline{e}_i) = b^j\, g(\underline{e}_j, \underline{e}_i) = d^j\, g_{ij}$

y por tanto se tiene $a_i = b^j\, g_{ij}$. Utilizando la matriz inversa de $(g_{ij})$, que existe por ser $g$ no degenerado, y que denotamos por $(g^{ij})$, se tiene que $b^j = a_i\, g^{ij}$. De esta forma, tomando el vector $v = a_i\, g^{ij}\, \underline{e}_j$, se verifica $T \equiv T_{\underline{v}}$, luego queda probada así la sobreyectividad de $\phi$, y en consecuencia $\phi$ define un isomorfismo entre $V$ y $V^*$.

□

Observamos que la prueba anterior no tiene dificultad alguna, sin embargo muestra cuál es el procedimiento para transformar un tensor contravariante $\underline{v} \in \mathcal{T}^1_0(V)$ en un tensor covariante $T_{\underline{v}} \in \mathcal{T}^0_1(V)$ en presencia de una métrica, y viceversa. La idea básica vemos que radica en multiplicar los coeficientes del tensor por los coeficientes $g_{ij}$ o $g^{ij}$, según el objetivo buscado. A partir de esta idea se deduce el siguiente resultado:

**Proposición 1.5.6** *Sean $V$ un espacio vectorial de dimensión $n$ con base $B = \{\underline{e}_1, ..., \underline{e}_n\}$ y base dual $B^* = \{\underline{e}^1, ..., \underline{e}^n\}$ de $V^*$, y $g$ un tensor métrico sobre $V$. Dada la base $B_r^s$ definida en 1.3.15 de $\mathcal{T}_r^s(V)$, consideramos un tensor $T \in \mathcal{T}_r^s(V)$ con componentes en la base $B_r^s$*

$$t_{i_1 \cdots i_r}^{j_1 \cdots j_s} = T(\underline{e}_{i_1}, ..., \underline{e}_{i_r}, \underline{e}^{j_1}, ..., \underline{e}^{j_s})$$

*Dado $p \in \{1, ..., r\}$, el isomorfismo (1.3) entre $V$ y $V^*$ inducido por $g$ nos permite "subir el índice del lugar $p$" y definir un nuevo tensor $\widetilde{T} \in \mathcal{T}_{r-1}^{s+1}(V)$ con componentes en la base $B_{r-1}^{s+1}$:*

$$\tilde{t}_{i_1 \cdots \widehat{i}_p \cdots i_r}^{j_1 \cdots j_s \, j_{s+1}} = t_{i_1 \cdots i_p \cdots i_r}^{j_1 \cdots j_s} \, g^{i_p \, j_{s+1}}$$

*donde el símbolo $\widehat{\ }$ indica que el índice ha sido omitido. De forma análoga, dado $q \in \{1, ..., s\}$, el isomorfismo (1.3) inducido por $g$ nos permite "bajar el índice del lugar $q$" y definir un nuevo tensor $\overline{T} \in \mathcal{T}_{r+1}^{s-1}(V)$ con componentes en la base $B_{r+1}^{s-1}$*

$$\bar{t}_{i_1 \cdots i_r \, i_{r+1}}^{j_1 \cdots \widehat{j}_q \cdots j_s \, j_{s+1}} = t_{i_1 \cdots i_r}^{j_1 \cdots j_q \cdots j_s} \, g_{j_q \, i_{r+1}}$$

Como consecuencia obtenemos que podemos identificar los tensores de tipo $(r, s)$ con los tensores de tipo $(0, r + s)$, y con los tensores de tipo $(r + s, 0)$, etc. sin más que aplicar la transformación con las matrices $(g_{ij})$ o $(g^{ij})$ de forma alternada tantas veces como sean necesarias.

**Corolario 1.5.7** *Dado un espacio vectorial $V$ y un tensor métrico $g$ asociado al espacio, todo tensor $T \in \mathcal{T}_r^s(V)$ se asocia con $r + s$ tensores de orden $r + s$. En particular, todo tensor tiene asociado un tensor covariante $T : V \times \overset{r+s}{\cdots} \times V \to \mathbb{R}$*

### 1.5.2. Métricas en el contexto diferencial

De forma análoga a lo estudiado para tensores en el contexto vectorial, los conceptos y resultados estudiados se pueden extender al contexto diferencial.

**Definición 1.5.8** *Sea $M$ una variedad diferenciable de dimensión $n$ y $W \subseteq M$ un abierto no vacío de la variedad diferenciable.*

- *Una métrica semi-riemanniana sobre $W$ es un campo tensorial diferenciable $T \in \mathcal{T}_2^0(W)$ de tipo $(2, 0)$ tal que, para todo $p \in W$, $T(p) \in \mathcal{T}_2^0\big(T_p(M)\big)$ es una métrica semi-euclídea.*
- *Una métrica riemanniana es una métrica semi-riemanniana $T \in \mathcal{T}_2^0(W)$ tal que para cada punto $p \in W$ el tensor $T(p) \in \mathcal{T}_2^0\big(T_p(M)\big)$ es una métrica euclídea.*
- *Una métrica de Lorentz en una métrica semi-riemanniana $T \in \mathcal{T}_2^0(W)$ tal que la signatura del tensor $T(p)$ en cada punto $p \in W$ es $(1, n - 1)$.*

**Observación 1.5.9** *Si el abierto $W \subseteq M$ es conexo, la signatura del tensor métrico $T(p)$ es constante para todo $p \in W$.*

**Definición 1.5.10** *Una variedad diferenciable $M$ dotada de una métrica semi-riemanniana, riemanniana o de Lorentz, se denomina respectivamente variedad semi-riemanniana, riemanniana o de Lorentz.*

**Ejemplo 1.5.11** *Como curiosidad, un caso particular utilizado en relatividad general es la variedad de Lorentz de dimensión 4 utilizada para modelar el espacio-tiempo. En ella, si denotamos por $g$ la métrica de Lorentz asociada a la variedad, los vectores tangentes en cada punto de la variedad se clasifican en tres tipos: vectores temporales o de tipo tiempo (si $g(u,u) < 0$), vectores nulos, lumínicos o de tipo luz (si $g(u,u) = 0$) y vectores espaciales o de tipo espacio (si $g(u,u) > 0$).*

Dado un abierto no vacío $W \subseteq M$, el resultado 1.5.5 aplicado para cada $p \in W$ sobre el espacio vectorial $V = T_p(M)$, establece un isomorfismo entre el espacio tangente a la variedad $M$ en $p$, y el espacio cotangente $T_p(M)^*$ en $p$, en presencia de una métrica $g \in \mathcal{T}_2^0(W)$, bien sea una métrica semi-riemanniana, riemanniana o de Lorentz.

Del mismo modo que hemos procedido con las métricas, llamaremos campo tensorial métrico a toda métrica en el contexto diferencial, bien sea semi-riemanniana, riemanniana o de Lorentz. De esta forma, un campo tensorial métrico $g$ definido sobre el abierto no vacío $W \subseteq M$, asigna a cada punto $p \in W$ un tensor métrico $g(p)$ (que denotaremos por $g_p$ para simplificar) del espacio tangente a $M$ en $p$

$$g_p : T_p(M) \times T_p(M) \longrightarrow \mathbb{R}$$

de modo que varíe de forma diferenciable con $p$, lo cual se traduce en que la aplicación

$$\begin{aligned} g_p(D_1, D_2) : W &\to \mathbb{R} \\ p &\mapsto g(p)\big(D_1(p), D_2(p)\big) \end{aligned}$$

sea diferenciable para todo par de campos vectoriales diferenciables $D_1, D_2 \in \mathcal{T}(W)$.

Consideramos ahora un abierto coordenado $(U; x^1, ..., x^n)$ sobre una variedad $M$, y un campo tensorial métrico $g$ definido sobre $U \subseteq M$. Sea la base $\left\{ \frac{\partial}{\partial x^1}\big|_p, ..., \frac{\partial}{\partial x^n}\big|_p \right\}$ de $T_p(M)$ formada por campos vectoriales, y la respectiva base dual $\{ dx^1\big|_p, ..., dx^n|_p\}$ de $T_p(M)^*$ formada por 1-formas. Consideramos la representación como en 1.4.26 del campo tensorial métrico

$$g = g_{ij}\, dx^i \otimes dx^j$$

donde $g_{ij} = g\big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\big) : U \to \mathbb{R}$ son funciones tales que $g_{ij}(p) = g(p)\big(\frac{\partial}{\partial x^i}\big|_p, \frac{\partial}{\partial x^j}\big|_p\big) \in \mathbb{R}$. Como $g$ es diferenciable, por 1.4.27 sabemos que $g_{ij}$ es diferenciable en $U$, y además, por definición de métrica, para cada $p \in U$ la matriz $(g_{ij}(p))$ es simétrica y no degenerada.

Los componentes del campo tensorial métrico dependen de la elección del abierto coordenado. Veamos por ello cómo se transforman las componentes bajo un cambio de coordenadas. Consideramos otro abierto coordenado $(V; y^1, ..., y^n)$ tal que $p \in U \cap V$. Entonces, el campo tensorial métrico determinará una nueva matriz $g_{kl} = g\big(\frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l}\big)$. Esta nueva matriz de funciones se relaciona con la matriz original mediante la aplicación de la regla de la cadena

$$\frac{\partial}{\partial y^k} = \frac{\partial x^i}{\partial y^k}\frac{\partial}{\partial x^i}$$

por lo que la relación ente las expresiones locales de $g$ con respecto a $(U; x^1, ..., x^n)$ y $(V; y^1, ..., y^n)$ viene dada por

$$g_{kl} = \frac{\partial x^i}{\partial y^k}\,\frac{\partial x^j}{\partial y^l}\,g_{ij}$$

En términos de matrices, si denotamos por $(g_{ij})^{[x]}$ la matriz de los coeficientes del campo tensorial métrico asociada al abierto coordenado $(U; x^1, ..., x^n)$, y por $(g_{ij})^{[y]}$ la matriz asociada al abierto coordenado $(V; y^1, ..., y^n)$, ambas matrices se relacionan mediante la igualdad matricial

$$(g_{ij})^{[x]} = \left((Dy)^{-1}\right)^T (g_{ij})^{[y]}\,(Dy)^{-1}$$

donde $Dy$ denota la matriz Jacobiana del cambio de coordenadas.

Llegamos así a la conclusión que la existencia de campos tensoriales métricos definidos sobre un abierto coordenado $(U; x^1, ..., x^n)$ está garantizada, pero no ocurre lo mismo en el caso general, cuando tratamos de definir un campo tensorial métrico sobre un abierto $W$ de $M$, incluyendo así el caso de toda la variedad. Acabamos de ver que en el caso de disponer de dos abiertos coordenados, es necesario imponer condiciones adicionales para garantizar que no existen incompatibilidades en los cambios de coordenadas entre cartas.

Veamos a continuación qué condiciones son las que garantizan la existencia de métricas riemannianas sobre variedades diferenciables. Para ello, comenzamos viendo unos conceptos previos que serán de utilidad en la formulación del resultado buscado:

**Definición 1.5.12**

- *Un recubrimiento de un conjunto $X$ es una colección de subconjuntos de $X$ cuya unión contiene a $X$, esto es, una familia indexada $U = \{U_\alpha : \alpha \in A\}$ de subconjuntos de $X$ es un recubrimiento de $X$ si $X \subseteq \bigcup\limits_{\alpha\in A} U_\alpha$.*

- *Un recubrimiento de un espacio topológico $X$ es abierto si todos sus elementos son conjuntos abiertos.*
- *Un recubrimiento $V = \{V_\beta : \beta \in B\}$ de $X$ es un refinamiento del recubrimiento $U = \{U_\alpha : \alpha \in A\}$ de $X$ si para todo $V_\beta$ en $V$ existe un $U_\alpha$ de $U$ tal que $V_\beta \subseteq U_\alpha$.*
- *Un recubrimiento por abiertos de $X$ es localmente finito si todo punto del espacio posee un entorno que interseca únicamente a un número finito de elementos del recubrimiento, esto es, si dado un recubrimiento $U = \{U_\alpha : \alpha \in A\}$ de $X$, para todo $x \in X$ existe un entorno $V_x$ de $x$ tal que el conjunto $\{\alpha \in A : U_\alpha \cap V_x \neq \emptyset\}$ es finito.*
- *Un espacio topológico $X$ se dice que es paracompacto si todo recubrimiento abierto admite un refinamiento localmente finito.*

La principal característica de los espacios $X$ paracompactos que son a su vez espacios de Haussdorf es que admiten particiones de la unidad subordinadas a cualquier recubrimiento por abiertos. Esto significa que, dado un recubrimiento abierto del espacio $X$, entonces existen una colección de funciones continuas en $X$ con valores en el intervalo unidad $[0, 1]$ tales que:

- Para toda función $f : X \to \mathbb{R}$ de la colección, existe un conjunto abierto $U$ del recubrimiento tal que el soporte de $f$ está contenido en $U$.
- Para todo punto $x \in X$, existe un entorno $V_x$ de $x$ tal que todas salvo un número finito de ellas son idénticamente nulas en $V$ y la suma de las funciones no nulas es 1 en $V$.

A partir de estos conceptos y resultados previos, obtenemos el siguiente resultado:

**Proposición 1.5.13** *Toda variedad diferenciable paracompacta $M$ puede ser dotada de una métrica riemanniana.*

<u>*Demostración*</u>.

Sea $M = \bigcup\limits_{\alpha \in U} U_\alpha$ un recubrimiento de $M$ por abiertos coordenados $\{(U_\alpha, \varphi_\alpha)\}_\alpha$. Para cada $\alpha$, consideramos la métrica riemanniana $g_\alpha$ en $U_\alpha$ cuya expresión local $\big((g_\alpha)_{ij}\big)$ es la matriz identidad. Sea $\{\rho_\alpha\}_\alpha$ una partición diferenciable de la unidad subordinada al recubrimiento $\{U_\alpha\}_\alpha$, y definimos $g = \sum\limits_{\alpha \in A} \rho_\alpha g_\alpha$.

Como la familia de soportes de $\rho_\alpha$ es localmente finita, $g$ será localmente finita, y por tanto $g$ está bien definida. Además, es claro que $g$ será diferenciable, bilineal y simétrica en cada punto. Como $\rho_\alpha \geq 0$ para todo $\alpha \in A$, y $\sum\limits_{\alpha \in A} \rho_\alpha = 1$, también se deriva que $g$ es definida positiva, y por tanto es una métrica riemanniana sobre $M$.

□

## 1.6. Ejemplos

Veamos seguidamente algunos ejemplos notables de campos tensoriales que aparecen en distintas áreas de conocimiento dentro de las matemáticas y la física.

### 1.6.1. Primera Forma Fundamental

Consideramos una superficie diferenciable $S$ en $\mathbb{R}^3$, la cual sabemos que constituye una variedad diferenciable de dimensión 2. Sobre esta variedad definimos la primera forma fundamental de $S$ como el campo tensorial métrico $\mathrm{I} : S \to T_2^0(S)$ tal que, para cada $p \in S$, el tensor métrico $\mathrm{I}(p) = \mathrm{I}_p \in \mathcal{T}_2^0\big(T_p(S)\big)$ se define como la restricción al espacio $T_p(S)$ del producto escalar en $\mathbb{R}^3$ (métrica euclídea en $\mathbb{R}^3$), obteniendo así el tensor métrico

$$\begin{aligned}\mathrm{I}_p : T_p(S) \times T_p(S) &\to \mathbb{R} \\ (v, w) &\mapsto \mathrm{I}_p(v, w) \coloneqq v \cdot w\end{aligned}$$

Veamos cómo representar de forma local las componentes del campo tensorial métrico. Dado un punto $p \in S$, consideramos una carta $\big(U, V, x^{-1}\big)$ tal que $p \in U$. $S$ está contenida, como subvariedad diferenciable de dimensión 2, en la variedad diferenciable $\mathbb{R}^3$, y la inmersión vendrá dada localmente por las ecuaciones paramétricas diferenciables

$$x^1 = x^1(u^1, u^2), \quad x^2 = x^2(u^1, u^2), \quad x^3 = x^3(u^1, u^2)$$

En $\mathbb{R}^3$, las derivaciones parciales $\left\{ \frac{\partial}{\partial x^1}\big|_p, \frac{\partial}{\partial x^2}\big|_p, \frac{\partial}{\partial x^3}\big|_p \right\}$ forman una base ortonormal para la métrica euclídea en $\mathbb{R}^3$ en cada uno de los puntos $p \in \mathbb{R}^3$. Por tanto se tiene que

$$\left.\frac{\partial}{\partial x^k}\right|_p \cdot \left.\frac{\partial}{\partial x^h}\right|_p = \delta_{kh}$$

Por otro lado, en los puntos $p \in S$, las derivaciones parciales $\left\{ \frac{\partial}{\partial u^1}\big|_p, \frac{\partial}{\partial u^2}\big|_p \right\}$ son elementos linealmente independientes del espacio $T_p(S)$. Por la inmersión de $S$ en $\mathbb{R}^3$, el espacio $T_p(S)$ será un subespacio bidimensional del espacio $T_p(\mathbb{R}^3)$ de dimensión 3. Por ello, aplicando la regla de la cadena a las ecuaciones $x^1$, $x^2$ y $x^3$ que definen la inmersión, obtenemos las coordenadas de $\frac{\partial}{\partial u^1}\big|_p$ y $\frac{\partial}{\partial u^2}\big|_p$ en la base $\left\{ \frac{\partial}{\partial x^1}\big|_p, \frac{\partial}{\partial x^2}\big|_p, \frac{\partial}{\partial x^3}\big|_p \right\}$:

$$\left.\frac{\partial}{\partial u^i}\right|_p = \left.\frac{\partial x^1}{\partial u^i}\right|_p \left.\frac{\partial}{\partial x^1}\right|_p + \left.\frac{\partial x^2}{\partial u^i}\right|_p \left.\frac{\partial}{\partial x^2}\right|_p + \left.\frac{\partial x^3}{\partial u^i}\right|_p \left.\frac{\partial}{\partial x^3}\right|_p = \left.\frac{\partial x^k}{\partial u^i}\right|_p \left.\frac{\partial}{\partial x^k}\right|_p, \quad i = 1, 2$$

donde $\left.\frac{\partial x^k}{\partial u^i}\right|_p$ se calcula como la parcial con respecto a $u^i$ de la función $x^k$ en $(u^1, u^2) = x^{-1}(p)$:

$$\left.\frac{\partial x^k}{\partial u^i}\right|_p = \left.\frac{\partial}{\partial u^i}\right|_p (x^k) = \frac{\partial x^k}{\partial u^i}(u^1, u^2)$$

De este modo, las componentes $g_{ij}$ de I en esta carta son

$$g_{ij} = g\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = \frac{\partial}{\partial u^i} \cdot \frac{\partial}{\partial u^j}$$

y utilizando el hecho de que las derivaciones parciales $\frac{\partial}{\partial x^i}$ son una base ortonormal, se tiene

$$g_{ij} = \frac{\partial x^1}{\partial u^i}\frac{\partial x^1}{\partial u^j} + \frac{\partial x^2}{\partial u^i}\frac{\partial x^2}{\partial u^j} + \frac{\partial x^3}{\partial u^i}\frac{\partial x^3}{\partial u^j} = \sum_{k=1}^{3} \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j}$$

**Observación 1.6.1** *De hecho, esta expresión es la utilizada habitualmente, que no requiere del uso de derivaciones dado que habitualmente se trabaja con vectores $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ y el espacio $T_p(\mathbb{R}^3)$ se identifica con el espacio vectorial habitual $\mathbb{R}^3$ dotado con la métrica euclídea. Realmente, si tomamos los vectores*

$$r_{u^i} = r_{u^i}(u^1, u^2) = \left(\frac{\partial x^1}{\partial u^i}, \frac{\partial x^2}{\partial u^i}, \frac{\partial x^3}{\partial u^i}\right)\Bigg|_{(u^1,u^2)}, \quad i = 1, 2$$

*y definimos $g_{ij} = r_{u^i} \cdot r_{u^j}$, tomando el espacio $T_p(S)$ como el subespacio vectorial de $\mathbb{R}^3$ generado por los vectores $r_{u^1}$ y $r_{u^2}$ tangentes a $S$ en $p$, estos coinciden con las derivaciones $\frac{\partial}{\partial u^1}$ y $\frac{\partial}{\partial u^2}$, y por tanto ambas definiciones de los coeficientes $g_{ij}$ coinciden.*

De forma habitual, los coeficientes de la primera forma fundamental se denotan siguiendo la nomenclatura $E(p) = g_{11}(p)$, $F(p) = g_{12}(p) = g_{21}(p)$ y $G(p) = g_{11}(p)$, obteniendo así la matriz asociada al tensor métrico

$$(g_{ij})|_p = \begin{pmatrix} E(p) & F(p) \\ F(p) & G(p) \end{pmatrix}$$

Dados $v = v^i \left.\frac{\partial}{\partial u^i}\right|_p$, $w = w^i \left.\frac{\partial}{\partial u^i}\right|_p \in T_p(S)$, se tiene

$$\mathrm{I}_p(v, w) = \mathrm{I}_p\left(v^i \left.\frac{\partial}{\partial u^i}\right|_p, w^i \left.\frac{\partial}{\partial u^i}\right|_p\right) = g_{ij}(p)\, v^i\, w^i = \begin{pmatrix} v^1 & v^2 \end{pmatrix} \begin{pmatrix} E(p) & F(p) \\ F(p) & G(p) \end{pmatrix} \begin{pmatrix} w^1 \\ w^2 \end{pmatrix}$$

$(S, \mathrm{I})$ constituye una variedad riemanniana con campo tensorial métrico I. Gracias a la primera forma fundamental, podemos estimar longitudes de curvas definidas sobre la superficie y ángulos de intersección entre curvas.

### 1.6.2. Segunda Forma Fundamental

Siguiendo con la notación desarrollada para la primera forma fundamental, veamos cómo definir la segunda forma fundamental asociada una superficie $S$ de $\mathbb{R}^3$. Para ello necesitamos en primer lugar suponer que la superficie está orientada (cómo vamos a trabajar de forma local, toda carta $(U, V, x^{-1})$ induce una orientación en $U$, que será con la que trabajaremos).

Tomamos las coordenadas locales $u^1$ y $u^2$ compatibles con la orientación, y definimos el vector normal inducido por la carta en cada punto $p \in U$ en coordenadas locales como

$$n = n(u^1, u^2) = \frac{r_{u^1}(u^1, u^2) \times r_{u^2}(u^1, u^2)}{\|r_{u^1}(u^1, u^2) \times r_{u^2}(u^1, u^2)\|}$$

tomando para ello los vectores $r_{u^1}$ y $r_{u^1}$ definidos como en 1.6.1, con $\times$ el producto vectorial de ambos vectores. El campo $n$ así definido para cada punto de $U$ es diferenciable, y define así los vectores $n_{u^1}$ y $n_{u^2}$ de $\mathbb{R}^3$.

Se define la segunda forma fundamental asociada a la superficie $S$ de $\mathbb{R}^3$ como el campo tensorial $\mathrm{II} : S \to T_2^0(S)$ tal que, para cada $p \in S$, el tensor $\mathrm{II}_p \in \mathcal{T}_2^0\big(T_p(S)\big)$ viene dado por sus coordenadas locales en la carta $(U, V, x^{-1})$ por $b_{ij} = -n_{u^i} \cdot r_{u^j}$, que son funciones reales diferenciables definidas sobre $U$.

El tensor $\mathrm{II}_p$ verifica $b_{ij} = b_{ji}$, luego es un tensor simétrico, y no depende de las coordenadas locales elegidas, esto es, se comporta bien ante cambios de base en $T_p(S)$ dado por cambios de coordenadas locales.

Por razones históricas las componentes de la segunda forma fundamental se designan por $L(p) = b_{11}(p)$, $M(p) = b_{12}(p) = b_{21}(p)$ y $N(p) = b_{11}(p)$, obteniendo así la matriz

$$(b_{ij})|_p = \begin{pmatrix} L(p) & M(p) \\ M(p) & N(p) \end{pmatrix}$$

La segunda forma fundamental II permite principalmente estudiar la curvatura de la superficie y clasificar sus puntos.

# 2. Datos tensoriales

Tal y como mencionábamos en la introducción, existen dos formas de enfocar la teoría tensorial. Mientras el primer capítulo establece el enfoque moderno basado en la estructuración abstracta de los tensores a partir de aplicaciones multilineales y productos tensoriales sobre espacios vectoriales (enfoque libre de coordenadas), en este segundo capítulo trataremos el enfoque clásico, visualizando los tensores como generalizaciones n-dimensionales de los conceptos de escalar, vector o matriz (enfoque con coordenadas).

La teoría clásica establece las componentes de los tensores como mediciones respecto de un sistema de referencia. Estas componentes se interpretan como las componentes del tensor en una determinada base. Si cambia el sistema de referencia, las nuevas componentes deben cambiar siguiendo unas determinadas leyes impuestas por el cambio de base (véase el resultado 1.2.12). En este capítulo extenderemos este enfoque clásico, y trataremos los tensores como componentes de almacenamiento de datos, manteniendo fijo el sistema de referencia. De esta forma, podremos definir el tensor del siguiente modo:

**Definición 2.0.1** *Sean $V_1, ..., V_N$ espacios vectoriales de dimensiones respectivas $I_1, ..., I_N$ definidos sobre el cuerpo $\mathbb{K}$. Se define tensor de orden $N$ como un elemento del espacio vectorial $V_1 \otimes \cdots \otimes V_N$. Por convención definimos los tensores de orden $0$ como los elementos de $\mathbb{K}$.*

Observamos que esta definición difiere notablemente del resultado 1.3.12. Con esta definición alternativa podremos representar estructuras de datos de forma que la dimensión de cada espacio no tenga porqué coincidir. Esto nos permite acercarnos de un modo más realista al concepto de arreglo (o array) n-dimensional, sin perder de vista el enfoque matemático de la definición.

Dada una base de cada uno de los espacios vectoriales, podemos considerar la base del espacio tensorial producto como el conjunto de los productos tensoriales de los elementos de cada base (véase 1.3.6 y 1.3.8), esto es, si $B_i = \{\underline{v}_1^i, \underline{v}_2^i, ..., \underline{v}_{I_i}^i\}$ es una base de $V_i$ para cada $i = \{1, ..., N\}$, el conjunto

$$B = \left\{\underline{v}_{i_1}^1 \otimes \underline{v}_{i_2}^2 \otimes \cdots \otimes \underline{v}_{i_N}^N \ : \ 1 \leq i_j \leq I_j, \ j \in \{1, ..., n\}\right\}$$

es base del espacio $V_1 \otimes \cdots \otimes V_N$. De esta forma, cada tensor podrá expresarse de forma única en la forma

$$T = t^{\,i_1\,i_2\cdots i_N}\, \underline{v}^1_{i_1} \otimes \underline{v}^2_{i_2} \otimes \cdots \otimes \underline{v}^N_{i_N}$$

donde los elementos $t^{\,i_1\cdots i_N} \in \mathbb{K}$ se denominan las componentes del tensor respecto de la base $B$.

Mediante esta expresión, es claro que podemos identificar cada elemento del espacio tensorial $V_1 \otimes \cdots \otimes V_N$ con un elemento del espacio $\mathbb{K}^{I_1\times\cdots\times I_N} := \mathbb{K}^{I_1} \otimes \cdots \otimes \mathbb{K}^{I_N}$, expresado en la forma

$$t^{\,i_1\,i_2\cdots i_N}\, \underline{e}^1_{i_1} \otimes \underline{e}^2_{i_2} \otimes \cdots \otimes \underline{e}^N_{i_N}$$

donde $\underline{e}^k_j$ se corresponde con el vector $j$-ésimo de la base estándar de $\mathbb{K}^{I_k}$.

Dado que de aquí en adelante no será necesario expresar explícitamente el tensor a partir de la base anterior y por tanto no se hará uso del convenio de sumación de Einstein, utilizaremos subíndices en lugar de superíndices para denotar las componentes del tensor. Este criterio está respaldado dado que los espacios vectoriales reales son susceptibles de ser provistos de métricas euclídeas. En el resto del capítulo trabajaremos con tensores definidos sobre $\mathbb{R}$-espacios vectoriales, y simplificaremos su uso tratándolos como elementos de $\mathbb{R}^{I_1\times\cdots\times I_N}$, si bien los resultados presentados podrán extenderse sin dificultad al caso general.

A partir de esta definición alternativa de tensor es claro que hemos obtenido una generalización de los conceptos de escalar, vector y matriz, puesto que estos se corresponden con tensores de orden 0, 1 y 2, respectivamente.

## 2.1. Operaciones con tensores

Veamos a continuación las operaciones básicas habituales entre tensores, para lo cual necesitamos en primer lugar definir una serie de conceptos básicos que nos permitirán definir con mayor facilidad estas operaciones.

Dado un tensor, podemos crear subtensores, definidos como subconjuntos de componentes del tensor. Estos subtensores se obtienen dejando fijos un determinado conjunto de índices, permitiendo variar únicamente los índices restantes. Utilizaremos el carácter ":" para denotar aquellos índices cuyos valores pueden variar en todo su dominio. Los dos casos más importantes de subtensores son los siguientes:

**Definición 2.1.1** *Sea $T \in \mathbb{R}^{I_1\times\cdots\times I_N}$ un tensor de orden $N$.*

- *Una fibra asociada al modo k-ésimo (también llamada fibra k-modal) del tensor $T$, es un subtensor de orden 1 obtenido al dejar fijos todos los índices menos el índice k-ésimo. Se denota por $T_{i_1,\ldots,i_{k-1},:,i_{k+1},\ldots,i_N}$.*

- *Una sección (slice) del tensor $T$ es un subtensor de orden 2 (matriz) obtenido al fijar todos los índices menos dos de ellos.*

En una matriz $A \in \mathbb{R}^{I \times J}$ (tensor de orden 2), las columnas son las fibras 1-modales (segundo índice fijo, $A_{:,j}$), y las filas son las fibras 2-modales (primer índice fijo $A_{i,:}$).

En un tensor $T \in \mathbb{R}^{I \times J \times K}$ de orden 3, las fibras 1-modales ($T_{:,j,k}$) se denominan *columnas*, las fibras 2-modales ($T_{i,:,k}$) se denominan *filas*, y las fibras 3-modales ($T_{i,j,:}$) se denominan *tubos*. A su vez, las secciones del tensor de orden 3, en función del índice fijado, se denominan secciones horizontales (fijando el primer índice $T_{i,:,:}$), verticales (fijando el segundo índice $T_{:,j,:}$) y frontales (fijando el tercer índice $T_{:,:,k}$). En la Figura 2.1 se muestra un ejemplo gráfico de estos conceptos sobre un tensor de orden 3.

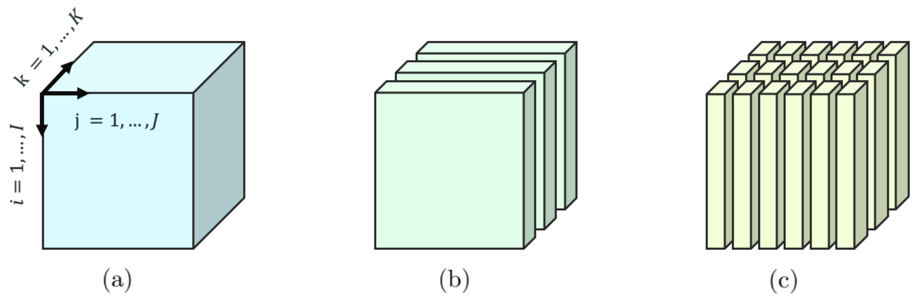


**Figura 2.1:** Representación gráfica de (a) un tensor $T \in \mathbb{R}^{I \times J \times K}$ de orden 3, con componentes $t_{i,j,k}$, (b) las secciones frontales $T_{:,:,k}$, y (c) las fibras 1-modales (columnas) $T_{:,j,k}$

**Definición 2.1.2** *Un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ es diagonal si $x_{i_1,\ldots,i_N} \neq 0$ sólo si $i_1 = \cdots = i_N$.*

Una generalización del producto tensorial de tensores (1.3.13) permite trasladar esta operación desde el punto de vista de las componentes de un tensor en una base:

**Definición 2.1.3** *Dados dos tensores $T^{(1)} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ y $T^{(2)} \in \mathbb{R}^{J_1 \times \ldots \times J_M}$, se define el producto tensorial (o exterior) de $T^{(1)}$ y $T^{(2)}$, y se denota $T^{(1)} \circ T^{(2)}$, como el tensor $C \in \mathbb{R}^{I_1 \times \ldots \times I_N \times J_1 \times \ldots \times J_M}$ de orden $N + M$ con componentes $c_{i_1,\ldots,i_N,j_1,\ldots,j_N} = t^{(1)}_{i_1,\ldots,i_N}\, t^{(2)}_{j_1,\ldots,j_M}$.*

Nótese que utilizamos el símbolo $\circ$ para denotar el producto tensorial en lugar de $\otimes$, tal y como hicimos en el primer capítulo. Esto se debe a que más adelante definiremos el producto de Kronecker, el cuál utiliza también el símbolo $\otimes$ para su representación, permitiendo así distinguir fácilmente ambos productos sin necesidad de añadir subíndices al operador.

A partir de este producto de tensores, se deriva un nuevo tipo de tensor, conocido como tensor de rango uno, que será un tensor clave en el estudio de las descomposiciones de tensores.

**Definición 2.1.4** *Un tensor $T \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ se denomina de rango uno si puede escribirse como producto tensorial de $N$ tensores de orden 1, esto es, si existen $N$ vectores $v^{(1)}, v^{(2)}, \ldots, v^{(N)}$ con $v^{(k)} \in \mathbb{R}^{I_k}$ para cada $k \in \{1, \ldots, N\}$, tales que $T = v^{(1)} \circ v^{(2)} \circ \cdots \circ v^{(N)}$, y por consiguiente las componentes del vector se escriben de la forma $t_{i_1, i_2, \ldots, i_N} = v^{(1)}_{i_1} v^{(2)}_{i_2} \cdots v^{(N)}_{i_N}$, donde $v^{(k)}_j$ denota la componente $j$-ésima del vector $v^{(k)}$.*

El siguiente concepto a desarrollar se basa en la representación de un tensor en forma matricial. Este proceso, que llamaremos despliegue matricial del tensor (del término inglés *unfolding*), se basa en la reordenación de los elementos de un tensor de orden $N$ para construir una matriz. Consideraremos únicamente los despliegue n-modales, que serán los despliegues de utilidad para la definición del producto n-modal.

**Definición 2.1.5** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \ldots \times I_N}$, su matriz de despliegue asociada al modo $n$-ésimo se define como la matriz $T_{(n)} \in \mathbb{R}^{I_n \times (I_1 \cdots I_{n-1} I_{n+1} \cdots I_N)}$ con componentes $t_{i_n, j}$ donde el índice $j$ que determina la columna de la matriz viene determinado por el multi-índice $j = \overline{i_1, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N}$ definido por*

$$\overline{i_1, i_2, \ldots, i_N} = i_1 + (i_2 - 1)I_1 + (i_3 - 1)I_1 I_2 + \cdots (i_N - 1)I_1 \cdots I_{N-1}$$

Nótese que las columnas de la matriz de despliegue asociado al modo $n$-ésimo son las fibras $n$-modales del tensor $T$, dispuestas en el orden establecido por el multi-indice $j$. El criterio de ordenación establecido se denomina convenio *little-endian*, y se corresponde con el orden lexicográfico. En algunos contextos, se utilizan otros criterios de ordenación, como es el caso del convenio de ordenación *big-endian*, definido a partir del multi-índice

$$\overline{i_1, i_2, \ldots, i_N} = i_N + (i_{N-1} - 1)I_N + (i_{N-2} - 1)I_N I_{N-1} + \cdots (i_1 - 1)I_N \cdots I_2$$

En la Figura 2.2 se muestra una representación gráfica de un ejemplo de despliegue aplicado sobre un tensor $T \in \mathbb{R}^{I \times J \times K}$ de orden 3.

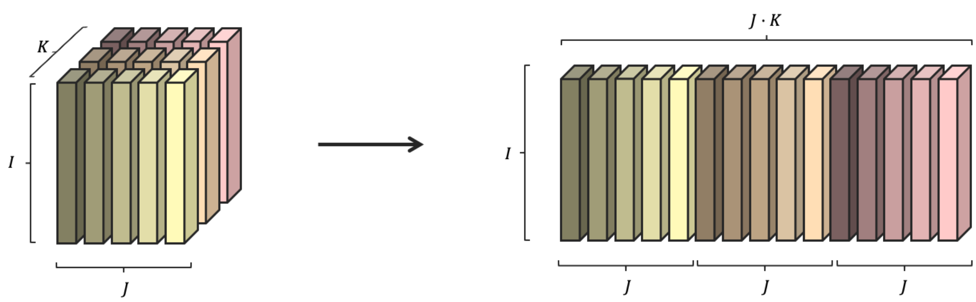


**Figura 2.2:** Representación gráfica de la matriz de despliegue $T_{(1)} \in \mathbb{R}^{I \times (JK)}$ asociada al primer modo de un tensor $T \in \mathbb{R}^{I \times J \times K}$ de orden 3, utilizando el convenio *little-endian* sobre las fibras 1-modales

Si nos centramos ahora en analizar el despliegue de un tensor desde el punto de vista del espacio vectorial a partir del cual hemos definido los tensores, vemos que existe un isomorfismo de espacios vectoriales sobre el que se construye este despliegue.

Si denotamos por $S_N$ el grupo simétrico de $N$ elementos, para cualquier permutación $\pi \in S_N$ sabemos que existe un isomorfismo canónico entre los espacios tensoriales $V_1 \otimes V_2 \otimes \cdots \otimes V_N$ y $V_{\pi(1)} \otimes V_{\pi(2)} \otimes \cdots \otimes V_{\pi(N)}$ (véase 1.3.5). Los paréntesis suelen omitirse sobre estos productos tensoriales de espacios debido al isomorfismo canónico existente entre $V_i \otimes (V_j \otimes V_k)$ y $(V_i \otimes V_j) \otimes V_k$ (véase 1.3.6). No obstante podemos utilizarlos para enfatizar un agrupamiento particular de los factores. Tomamos entonces el agrupamiento

$$(V_{\pi(1)} \otimes \cdots \otimes V_{\pi(r_1)}) \otimes \cdots (V_{\pi(r_{j-1}+1)} \otimes \cdots \otimes V_{\pi(r_j)}) \otimes \cdots \otimes (V_{\pi(r_{\ell-1}+1)} \otimes \cdots \otimes V_{\pi(r_\ell)})$$

de $\ell$ grupos con $r_j - r_{j-1}$ espacios en el grupo $j$-ésimo, definiendo para ello $r_0 = 0$ y $r_l = N$.

Fijadas bases de los $N$ espacios vectoriales, sabemos que

$$V_{\pi(r_{j-1}+1)} \otimes \cdots \otimes V_{\pi(r_j)} \;\cong\; \mathbb{K}^{I_{\pi(r_{j-1}+1)}} \otimes \cdots \otimes \mathbb{K}^{I_{\pi(r_j)}}$$

Si denotamos por $[d]$ el conjunto $\{1, ..., d\}$, dada una aplicación $\mu : [n_1] \times \cdots \times [n_d] \to [n_1 \cdots n_d]$ biyectiva, se tiene el isomorfismo

$$\begin{array}{rccc} \phi: & \mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_d} & \longrightarrow & \mathbb{K}^{n_1 \cdots n_d} \\ & t^{\,j_1 \cdots j_d}\, \underline{e}^1_{j_1} \otimes \cdots \otimes \underline{e}^d_{j_d} & \longmapsto & t^{\mu(j_1,\ldots,j_d)}\, \underline{e}_{\mu(j_1,\ldots,j_d)} \end{array} \tag{2.1}$$

Definimos ahora $s_j = \{\pi(r_{j-1}+1), \pi(r_{j-1}+2), ..., \pi(r_j)\}$ para cada $j \in \{1, ..., \ell\}$, y consideramos una aplicación biyectiva $\mu_j : [I_{\pi(r_{j-1}+1)}] \times \cdots \times [I_{\pi(r_j)}] \to [N_{s_j}]$ con $N_{s_j} := \prod\limits_{k \in s_j} I_k$.

Tal y como acabamos de ver con la definición del isomorfismo (2.1), haciendo uso de $\mu_j$ podemos establecer un isomorfismo entre $\mathbb{K}^{I_{\pi(r_{j-1}+1)}} \otimes \cdots \otimes \mathbb{K}^{I_{\pi(r_j)}}$ y $\mathbb{K}^{N_{s_j}}$.

Mediante este procedimiento, dados $\ell$ grupos de índices $s_1, ..., s_\ell$ disjuntos, se tiene que

$$V_1 \otimes \cdots \otimes V_N \cong \mathbb{K}^{N_{s_l}} \otimes \cdots \otimes \mathbb{K}^{N_{s_\ell}}$$

y el nuevo tensor de orden $\ell$, obtenido tras aplicar el isomorfismo entre ambos espacios, se denomina $(s_1, ..., s_\ell)$-aplanamiento del tensor $T$, y lo denotaremos por $T_{(s_1,\ldots,s_\ell)}$.

Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, el despliegue asociado al modo $k$-ésimo se corresponde con el aplanamiento asociado a los conjuntos de índices $s_1 = \{k\}$ y $s_2 = \{1, 2, ..., k-1, k+1, ..., N\}$, y se denota por $T_{(k)} = T_{(s_1,s_2)}$.

Otro caso particular de aplanamiento de una matriz consiste en tomar un único conjunto $s$ compuesto por todos los índices del tensor. De esta forma, podemos transformar un tensor en un vector, y la técnica de aplanamiento se conoce como vectorización. Para ello bastará con fijar un orden de reordenación de los elementos consistente, como por ejemplo fijando el convenio *little-endian* correspondiente al orden lexicográfico:

**Definición 2.1.6** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, se define la vectorización del tensor $T$, y se denota como $vec(T)$, como el tensor de orden 1 (vector) $vec(T) \in \mathbb{R}^{I_1 \cdots I_N}$ con componentes $c_j = t_{\overline{i_1,\ldots,i_N}}$, utilizando el multi-índice fijado por el convenio de ordenación little-endian.*

En la Figura 2.3 se muestra un ejemplo de representación gráfica de la vectorización de un tensor de orden 3, aplicando el criterio de ordenación *little-endian* tal y como se ha establecido en la definición.

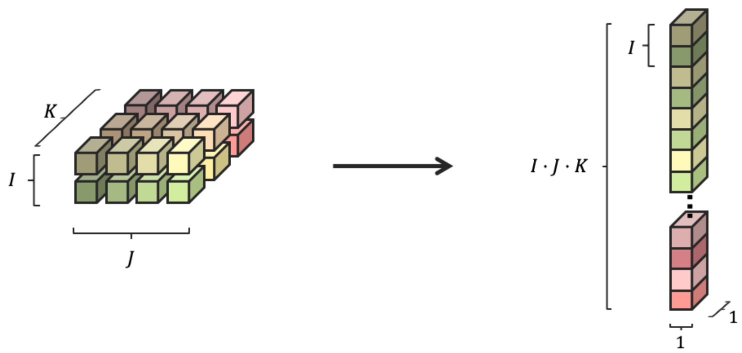


**Figura 2.3:** Representación gráfica del vector $vec(T) \in \mathbb{R}^{I \cdot J \cdot K}$ obtenido tras la vectorización del tensor $T \in \mathbb{R}^{I \times J \times K}$ aplicando el convenio de ordenación *little-endian*. El ejemplo muestra de hecho la vectorización de un tenor $T \in \mathbb{R}^{2 \times 4 \times 3}$, obteniendo el vector $vec(T) \in \mathbb{R}^{24}$

A continuación se definen una nueva operación relativa a los tensores, la cual, tal y como veremos más adelante, está ligada al concepto previo de despliegue.

**Definición 2.1.7** *Dados un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, un vector $b \in \mathbb{R}^{I_n}$ y una matriz $A \in \mathbb{R}^{J \times I_n}$, se definen el producto $n$-modal del tensor $T$ por el vector $b$ como el tensor $C = T \bar{\times}_n b \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$ con componentes $c_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N} = \sum_{i_n} t_{i_1,\ldots,i_n,\ldots,i_N}\, b_{i_n}$; y el producto $n$-modal del tensor $T$ por la matriz $A$ como el tensor $D = T \times_n A \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ con componentes $d_{i_1,\ldots,i_{n-1},j,i_{n+1},\ldots,i_N} = \sum_{i_n} t_{i_1,\ldots,i_n,\ldots,i_N}\, a_{j,i_n}$*

De la definición de producto $n$-modal de un tensor $T$ y una matriz $A$, podemos observar cómo sus cálculos se reducen al producto de cada fibra $n$-modal de $T$ por la matriz $A$. De esta observación se deduce el siguiente resultado:

**Proposición 2.1.8** *Dados un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ y una matriz $A \in \mathbb{R}^{J \times I_n}$, el producto n-modal $D = T \times_n A$ equivale al cálculo despliegue tensorial del producto matricial de $A$ y $T_{(n)}$, esto es,*

$$D = T \times_n A \iff D_{(n)} = A\,T_{(n)}$$

En la Figura 2.4 podemos observar, mediante una representación gráfica, la equivalencia anterior asociada al producto 1-modal de un tensor $T$ de orden 3.

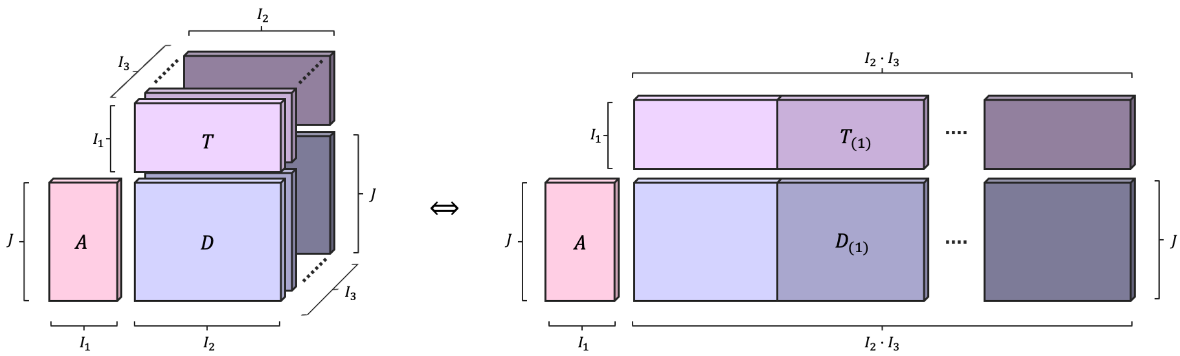


**Figura 2.4:** Representación gráfica del producto 1-modal de un tensor $T \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ de orden 3 y una matriz $A \in \mathbb{R}^{J \times I_1}$, obteniendo un tensor $D = T \times_n A \in \mathbb{R}^{J \times I_2 \times I_3}$. Esto equivale al ecuación de multiplicación matricial $D_{(1)} = A\,T_{(1)}$

El producto $k$-modal de un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ y una matriz $A \in \mathbb{R}^{J \times I_k}$ se relaciona, en términos del espacio vectorial $V_1 \otimes \cdots \otimes V_N$, con un cambio de base del espacio $V_k$:

Sea $B_k = \{\underline{e}_1^k, ..., \underline{e}_{I_k}^k\}$ una base de $V_k$, y consideramos una nueva base $\widetilde{B}_k = \{\underline{\tilde{e}}_1^k, ..., \underline{\tilde{e}}_{I_k}^k\}$ de $V_k$. Cada elemento $\underline{e}_i^k$ de $B_k$ puede expresarse de forma única como combinación lineal de los elementos de $\widetilde{B}_k$, esto es, $\underline{e}_i^k = p_i^j\, \underline{\tilde{e}}_j^k$, donde $p_i^j$ representan las componentes de la matriz de cambio de base de $B_k$ a $\widetilde{B}_k$. Se tiene así

$$\begin{aligned} t^{\,i_1 \cdots i_k \cdots i_N}\, \underline{e}_{i_1}^1 \otimes \cdots \otimes \underline{e}_{i_k}^k \otimes \cdots \otimes \underline{e}_{i_N}^N &= t^{\,i_1 \cdots i_k \cdots i_N}\, \underline{e}_{i_1}^1 \otimes \cdots \otimes \left(p_{i_k}^j\, \underline{\tilde{e}}_j^k\right) \otimes \cdots \otimes \underline{e}_{i_N}^N \\ &= t^{\,i_1 \cdots i_k \cdots i_N}\, p_{i_k}^j\, \underline{e}_{i_1}^1 \otimes \cdots \otimes \underline{\tilde{e}}_j^k \otimes \cdots \otimes \underline{e}_{i_N}^N \end{aligned}$$

Por tanto, si denotamos por $P_k = (p_i^j)$ la matriz de cambio de base en el espacio $V_k$, utilizando la notación por componentes, dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_k \times \cdots \times I_N}$, la representación del tensor en la nueva base vendrá dada por el cambio $\widetilde{T} = T \times_n P_k$, y sus componentes serán

$$\tilde{t}_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_N} = \sum_{i_k} t_{\,i_1,\ldots,i_k,\ldots,i_N}\, p_{i_k}^j$$

donde el índice $j$ varía en el mismo conjunto que $i_k$, puesto que la matriz de cambio de base será cuadrada, con dimensión $I_k$.

Para modos distintos de un tensor, el orden de los productos por matrices asociados a estos modos es irrelevante:

**Proposición 2.1.9** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ y las matrices $A \in \mathbb{R}^{J_n \times I_n}$ y $B \in \mathbb{R}^{J_m \times I_m}$, con $n \neq m$, entonces*

$$T \times_n A \times_m B = (T \times_n A) \times_m B = (T \times_m B) \times_n A$$

*Demostración.*

Es claro que se tiene $(T \times_n A) \times_m B,\ (T \times_m B) \times_n A \in \mathbb{R}^{I_1 \times \cdots \times J_n \times \cdots \times J_m \times \cdots \times I_N}$ (suponiendo que $n < m$, el otro caso es análogo). Veamos que las componentes de ambos tensores coinciden:

$$\begin{aligned}
\big((T \times_n A) \times_m B\big)_{i_1,\ldots,j_n,\ldots,j_m,\ldots,i_N} &= \sum_{i_m} (T \times_n A)_{i_1,\ldots,i_n,\ldots,i_m,\ldots,i_N}\, b_{j_m,i_m} \\
&= \sum_{i_n,i_m} t_{i_1,\ldots,j_n,\ldots,i_m,\ldots,i_N}\, a_{j_n,i_n}\, b_{j_m,i_m} \\
&= \sum_{i_n} (T \times_m B)_{i_1,\ldots,i_n,\ldots,j_m,\ldots,i_N}\, b_{j_n,i_n} \\
&= \big((T \times_m B) \times_n A\big)_{i_1,\ldots,j_n,\ldots,j_m,\ldots,i_N}
\end{aligned}$$

□

No ocurre lo mismo con el producto $n$-modal de un tensor y una sucesión de vectores, puesto que el orden de las operaciones y de los resultados intermedios cambia. Si $T \in \mathbb{R}^{I_1 \times \cdots \times I_n \times \cdots \times I_m \times \cdots \times I_N}$, $a \in \mathbb{R}^{I_n}$ y $b \in \mathbb{R}^{I_m}$ con $n < m$, se verifica

$$T \,\bar{\times}_n\, a \,\bar{\times}_m\, b = (T \,\bar{\times}_n\, a) \,\bar{\times}_{m-1}\, b = (T \,\bar{\times}_m\, b) \,\bar{\times}_n\, a$$

Si los modos sobre los que se aplica el producto $n$-modal por matrices coinciden, para matrices de dimensión adecuada se tiene el siguiente resultado, que muestra que bajo estos supuestos el orden de aplicación de los productos $n$-modales sí es relevante.

**Proposición 2.1.10** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ y las matrices $A \in \mathbb{R}^{J_n \times I_n}$ y $B \in \mathbb{R}^{K_n \times J_n}$, entonces*

$$T \times_n A \times_n B = T \times_n (B \cdot A)$$

*Demostración.*

Es claro que se tiene $T \times_n A \times_n B,\ T \times_n (B \cdot A) \in \mathbb{R}^{I_1 \times \cdots \times K_n \times \cdots \times I_N}$. Veamos que las componentes de ambos tensores coinciden:

$$\begin{aligned}
(T \times_n A \times_n B)_{i_1,\dots,k_n,\dots,i_N} &= \sum_{i_1,j_n} t_{i_1,\dots,i_n,\dots,i_N}\, a_{j_n,i_n}\, b_{k_n,j_n} \\
&= \sum_{i_n} t_{i_1,\dots,i_n,\dots,i_N} \cdot \left( \sum_{j_n} a_{j_n,i_n}\, b_{k_n,j_n} \right) \\
&= \sum_{i_n} t_{i_1,\dots,i_n,\dots,i_N} \cdot (A \cdot B)_{k_n,i_n} \\
&= \big(T \times_n (B \cdot A)\big)_{i_1,\dots,k_n,\dots,i_N}
\end{aligned}$$

□

A partir de estas propiedades, si tomamos todos los modos de un tensor, obtenemos un producto multilineal de tensores por un conjunto de matrices,

**Definición 2.1.11** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ y un conjunto de matrices $B^{(1)}, ..., B^{(N)}$ con $B^{(k)} \in \mathbb{R}^{J_k \times I_k}$ para cada $k \in \{1, ..., N\}$, se define el producto multilineal del tensor $T$ y las matrices $B^{(1)}, ..., B^{(N)}$, que denotaremos por $[\![T; B^{(1)}, ..., B^{(N)}]\!]$, como el producto*

$$[\![\, T;\, B^{(1)}, ...,\, B^{(N)} \,]\!] = T\, \times_1\, B^{(1)}\, \times_2\, B^{(2)} \cdots\, \times_N\, B^{(N)}$$

Esta operación se corresponde con el concepto de producto o multiplicación multilineal, enmarcado dentro del álgebra tensorial. Para su definición algebraica debemos recordar que, tal y como hemos definido el producto tensorial de espacios vectoriales a partir a la propiedad universal 1.3.7, sabemos que todo elemento del espacio tensorial puede expresarse como combinación lineal de productos tensoriales de vectores de cada uno de los espacios, esto es, dado $T \in V_1 \otimes \cdots \otimes V_d$, este puede expresarse en la forma $T = t^{i_1,\dots,i_d}\, v^1_{i_1} \otimes \cdots \otimes v^d_{i_d}$ con $v^k_{i_k} \in V_k$ para todo $k \in \{1, ..., d\}$.

**Definición 2.1.12** *Sean $V_1, ..., V_N$ espacios vectoriales sobre $\mathbb{K}$ de dimensión finita. Dada una colección de operadores lineales $A_k : V_k \to W_k$ para $k \in \{1, ..., n\}$, se define la multiplicación multilineal asociada a los operadores $(A_1, ..., A_N)$ como el operador derivado del producto tensorial de los operadores $A_1, ..., A_N$, esto es,*

$$\begin{aligned}
A_1 \otimes \cdots \otimes A_N : V_1 \otimes \cdots \otimes V_N \;&\longrightarrow\; W_1 \otimes \cdots \otimes W_N \\
t^{i_1,\dots,i_d}\, v^1_{i_1} \otimes \cdots \otimes v^d_{i_d} \;&\mapsto\; t^{i_1,\dots,i_d}\, A_1(v^1_{i_1}) \otimes \cdots \otimes A_N(v^d_{i_d})
\end{aligned}$$

Tomando bases de los espacios $V_k$ y $W_k$ para todo $k \in \{1, .., N\}$, las aplicaciones lineales pueden representarse en forma matricial a partir de las coordenadas en la representación en las bases, de forma que para cada $k$ consideramos la matriz $\widehat{A}_k = \big(a^{(k)}_{j,i}\big) \in \mathbb{K}^{J_k \times I_k}$ con $J_k$ la dimensión de $W_k$ e $I_k$ la dimensión de $V_k$.

Si consideramos ahora el tensor como un elemento de $\mathbb{K}^{I_1 \times \cdots \times I_N}$, representado como $T = t^{\,i_1\, i_2 \cdots i_N}\, \underline{e}^1_{i_1} \otimes \underline{e}^2_{i_2} \otimes \cdots \otimes \underline{e}^N_{i_N}$, con $\underline{e}^k_j$ el vector $j$-ésimo de la base estándar de $\mathbb{K}^{I_k}$, el producto multilineal de $A_1, ..., A_k$ aplicado al tensor $T$ queda de la forma

$$\begin{aligned}(A_1 \otimes \cdots \otimes A_N)(T) &= (A_1 \otimes \cdots \otimes A_N)(t^{\,i_1\, i_2 \cdots i_N}\, \underline{e}^1_{i_1} \otimes \underline{e}^2_{i_2} \otimes \cdots \otimes \underline{e}^N_{i_N}) \\ &= t^{\,i_1\, i_2 \cdots i_N} \big(\widehat{A}_1\, \underline{e}^1_{i_1}\big) \otimes \big(\widehat{A}_2\, \underline{e}^2_{i_2}\big) \otimes \cdots \otimes \big(\widehat{A}_N\, \underline{e}^N_{i_N}\big) \\ &= t^{\,i_1\, i_2 \cdots i_N} a^{(1)}_{j_1,i_1}\, a^{(2)}_{j_2,i_2} \cdots a^{(N)}_{j_N,i_N} \underline{e}^1_{i_1} \otimes \underline{e}^2_{i_2} \otimes \cdots \otimes \underline{e}^N_{i_N}\end{aligned}$$

Obtenemos de esta forma un nuevo tensor $(A_1 \otimes \cdots \otimes A_N)(T) \in \mathbb{K}^{J_1 \times J_2 \times \cdots \times J_N}$. Observamos cómo este producto tensorial se corresponde con el producto multilineal desde el punto de vista de las coordenadas del tensor.

Si tomamos el operador $Id_k : V_k \to V_k$ para cada $k \in \{1, ..., n-1, n+1, ..., N\}$, y tomamos un operador lineal $A_n : V_n \to W_n$, el correspondiente producto multilineal desde el punto de vista del análisis de las coordenadas del tensor se corresponde con el producto $n$-modal de un tensor y una matriz. De forma similar podemos establecer la analogía entre este producto multilineal así definido y el producto $n$-modal de un tensor por un vector, sin más que tomar $W_n = \mathbb{K}$.

Por último desarrollaremos dos productos matriciales que serán de gran utilidad en la simplificación de las descomposiciones tensoriales que trabajaremos en la siguiente sección.

**Definición 2.1.13**

- *Dadas dos matrices $A \in \mathbb{R}^{I \times J}$ y $B \in \mathbb{R}^{K \times L}$, se define el producto de Kronecker de $A$ y $B$, como la matriz $A \otimes B \in \mathbb{R}^{(I \cdot K) \times (J \cdot L)}$ con componentes $(A \otimes B)_{\overline{i,\,k},\,\overline{j,\,l}} = a_{i,j}\, b_{k,l}$, donde $\overline{i,\,k} = k + (i-1)K$ y $\overline{j,\,l} = l + (j-1)L$.*

- *Dadas dos matrices $A \in \mathbb{R}^{I \times K}$ y $B \in \mathbb{R}^{J \times K}$, se define el producto de Khatri-Rao de $A$ y $B$, como la matriz $C = A \odot B \in \mathbb{R}^{(I \cdot J) \times K}$ con columnas $c_{:,j} = a_{:,j} \otimes b_{:,j}$, con $a_{:,j}$, $b_{:,j}$ las columnas de las matrices $A$ y $B$. respectivamente.*

- *Dadas dos matrices $A, B \in \mathbb{R}^{I \times J}$ se define el producto de Hadamard de $A$ y $B$, como la matriz $A * B \in \mathbb{R}^{I \times J}$ con componentes $(A * B)_{i,j} = a_{i,j}\, b_{i,j}$.*

## 2.2. Descomposiciones tensoriales

El objetivo principal de las descomposiciones de tensores estándar es la factorización del tensor en factores que puedan ser interpretables desde el punto de vista de la contextualización de los datos tensoriales. Estas descomposiciones permiten establecer correspondencias entre los distintos espacios que componen el tensor. En esta sección estudiaremos las dos descomposiciones tensoriales por excelencia, la descomposición CP y la descomposición de Tucker, junto con sus aplicaciones prácticas en amplias áreas de conocimiento.

Tanto la descomposición CP como la descomposición de Tucker se consideran generalizaciones de la descomposición en valores singulares (SVD, de sus siglas en ingles *Singular Value Decomposition*) de matrices, y del análisis de componentes principales (PCA, de sus siglas en ingles *Principal Component Analysis*).

### 2.2.1. Descomposición CP

En esta sección nos centraremos en la descomposición CP, cuyas siglas derivan de los términos CANDECOMP (Canonical Decomposition) y PARAFAC (Parallel Factors Decomposition).

Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, la descomposición CP factoriza el tensor como suma de tensores de rango uno, esto es,

$$T \cong \sum_{r=1}^{R} \lambda_r \, a_r^{(1)} \circ \cdots \circ a_r^{(N)}$$

con $a_r^{(k)} \in \mathbb{R}^{I_k}$ para cada $k \in \{1, ..., N\}$ y $\lambda_r \in \mathbb{R}$, para todo $r \in \{1, ..., R\}$. A partir de estos datos, construimos las matrices $A^{(k)} = \left[\, a_1^{(k)} \,|\, a_2^{(k)} \,|\, \cdots \,|\, a_R^{(k)} \,\right] \in \mathbb{R}^{I_k \times R}$, denominadas matrices factor, definidas por columnas a partir de los vectores que constituyen los componentes de los tensores de rango uno. Podemos de esta forma reescribir la descomposición como

$$T \cong \sum_{r=1}^{R} \lambda_r \, a_r^{(1)} \circ a_r^{(2)} \circ \cdots \circ a_r^{(N)} = \Lambda \times_1 A^{(1)} \times_2 A^{(2)} \cdots \times_N A^{(N)} = [\![ \Lambda;\, A^{(1)}, A^{(2)}, ..., \, A^{(N)} ]\!]$$

con $\Lambda \in \mathbb{R}^{R \times \cdots \times R}$ un tensor diagonal de orden $R$ con valores en la diagonal $i_{r,\ldots,r} = \lambda_r$ (véase la Figura 2.5 para ver este esquema de descomposición con un tensor de orden 3).

Utilizando el producto matricial de Khatri-Rao y tomando la matriz $\Gamma = diag(\lambda)$, la versión matricial de la descomposición CP asociada al modo $n$-ésimo viene dada por la identificación

$$T_{(n)} \cong A^{(n)} \, \Gamma \left( A^{(N)} \odot \cdots \odot A^{(n+1)} \odot A^{(n-1)} \odot \cdots \odot A^{(1)} \right)^T \tag{2.2}$$

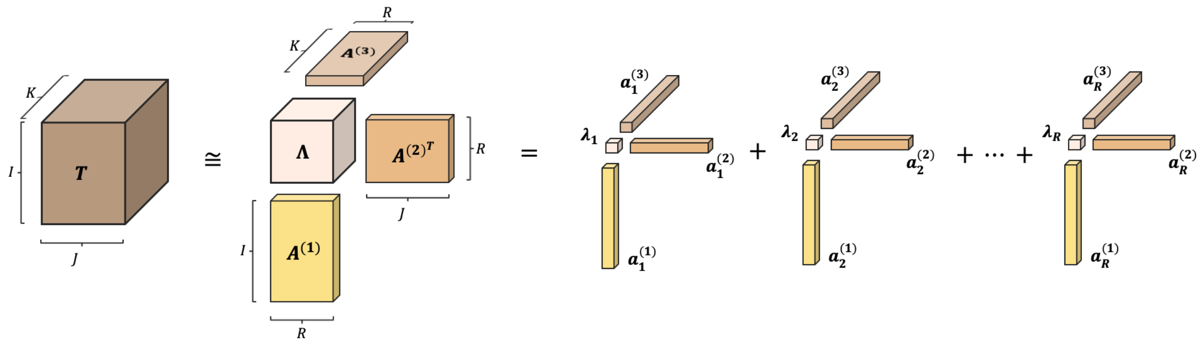


**Figura 2.5:** Representación gráfica del esquema de descomposición CP, para un tensor $T \in \mathbb{R}^{I\times J\times K}$ de orden 3, en la forma $T \cong [\![\, \Lambda;\, A^{(1)}, A^{(2)}, A^{(3)} \,]\!] = \sum_{r=1}^{R} \lambda_r\, a_r^{(1)} \circ a_r^{(2)} \circ a_r^{(3)}$

**Definición 2.2.1** *Dado un tensor $T \in \mathbb{R}^{I_1\times\cdots\times I_N}$ de orden $N$, el rango del tensor, representado por $R = rank(T)$, es el mínimo número de tensores de rango uno que permiten expresar $T$ como combinación lineal de los mismo*

En otras palabras, el rango de un tensor es el mínimo número de componentes necesarias en la descomposición CP para que esta sea exacta. Una descomposición CP exacta con $R = rank(T)$ componentes se denomina descomposición de rango $R$ de $T$.

**Observación 2.2.2** *La definición de rango de un tensor es análoga a la definición de rango de una matriz, aunque existen diferencias notables en sus propiedades. Por ejemplo, a diferencia de lo que ocurre con las matrices, no existe un algoritmo directo para determinar el rango de un tensor dado, y de hecho su cálculo es un problema NP-complejo. Otra peculiaridad del rango del tensor viene asociada a los términos de rango máximo (el mayor rango alcanzable) y de rango típico (cualquier rango que ocurre con probabilidad mayor que cero sobre un conjunto con medida de Lebesgue positiva). Sobre la colección de matrices de dimensión $I \times J$, los rangos máximo y típico coinciden, con valor $min\{I, J\}$. En cambio en el caso de los tensores ambos rangos pueden no coincidir.*

Veamos a continuación bajo qué supuestos podemos garantizar la unicidad de la descomposición de rango $R$ de un tensor. Sea así un tensor $T \in \mathbb{R}^{I_1\times\cdots\times I_N}$ de rango $R$, esto es,

$$T = \sum_{r=1}^{R} a_r^{(1)} \circ a_r^{(2)} \circ \cdots \circ a_r^{(N)} = [\![\, A^{(1)}, \ldots, A^{(N)} \,]\!]$$

Consideramos que el tensor $\Lambda$ de normalización de los factores se toma inmerso dentro de una de las matrices factor. De esta forma fijamos el siguiente convenio de notación:

**Notación 2.2.3** *Sean $A^{(1)}, ..., A^{(N)}$ matrices tales que $A^{(k)} \in \mathbb{R}^{J_k \times I_k}$ para cada $k \in \{1, ..., N\}$, y consideramos el tensor identidad $I \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ con componentes $i_{i_1,...,i_N} = 1$ si $i_1 = ... = i_N$ y $0$ en el resto de casos. Entonces entonces denotaremos por $[\![ A^{(1)}, \dots, A^{(N)} ]\!]$ el producto multilineal del tensor $I$ por las matrices $A^{(1)}, \dots, A^{(N)}$, esto es,*

$$[\![ A^{(1)}, \dots, A^{(N)} ]\!] \equiv [\![ I;\, A^{(1)}, \dots, A^{(N)} ]\!]$$

Cuando hablamos de unicidad en la descomposición hacemos referencia a la existencia de una única combinación de tensores de rango uno cuya suma permite recrear el tensor $T$ original, con las excepciones de los cambios de escala y permutaciones:

- Cambio de escala:

$$T = \sum_{r=1}^{R} (\alpha_r^{(1)}\, a_r^{(1)}) \circ (\alpha_r^{(2)}\, a_r^{(2)}) \circ \cdots \circ (\alpha_r^{(N)}\, a_r^{(N)}), \quad \text{con} \quad \prod_{k=1}^{N} \alpha_r^{(k)} = 1, \quad \forall\, k \in \{1, ..., N\}$$

- Permutación:

$$T = [\![ A^{(1)}, \dots, A^{(N)} ]\!] = [\![ A^{(1)}\, \Pi, \dots, A^{(N)}\, \Pi ]\!] \;\; \text{para toda matriz de permutación } \Pi \in \mathbb{R}^{R \times R}$$

El resultado más general y conocido acerca de la unicidad de la descomposición de rango $R$ viene asociado al concepto de $k$-rango.

**Definición 2.2.4** *Dada una matriz $A \in \mathbb{R}^{I \times J}$, el k-rango de $A$, denotado $k_A$, se define como el máximo $k$ tal que $k$ columnas cualesquiera de $A$ son linealmente independientes.*

**Teorema 2.2.5** *Sea $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ un tensor de orden $N$ y rango $R$, y supongamos que su descomposición de rango $R$ es*

$$T = \sum_{r=1}^{R} a_r^{(1)} \circ a_r^{(2)} \circ \cdots \circ a_r^{(N)} = [\![ A^{(1)}, \dots, A^{(N)} ]\!]$$

*Entonces se verifica:*

*(a) Es condición suficiente para que la descomposición de rango $R$ sea única que*

$$\sum_{n=1}^{N} k_{A^{(n)}} \geq 2R + (N - 1)$$

*(b) En los casos $R = 2$ y $R = 3$, la condición (a) es también una condición necesaria, aunque esto no es cierto de forma general para $R > 3$.*

*(c) Es condición necesaria y suficiente para que la descomposición de rango R sea única que se verifique la condición*

$$\min_{1\leq n\leq N}\left\{rank\Big(A^{(N)}\odot\cdots\odot A^{(n+1)}\odot A^{(n-1)}\odot\cdots\odot A^{(1)}\Big)\right\}=R$$

Las demostraciones sobre estas condiciones de unicidad no se añaden en este trabajo dado que exceden los objetivos del mismo, no obstante pueden consultarse en [13], [14] y [10].

Una vez fijamos el número de componentes de la descomposición, debemos definir un algoritmo que permita computar la descomposición CP. Nos centraremos a continuación en el algoritmo más popular y extendido en este contexto: el algoritmo de Mínimos Cuadrados Alternados (ALS, por sus siglas en inglés *Alternating Least Squares*).

Consideramos un tensor $T\in\mathbb{R}^{I_1\times\cdots\times I_N}$ de orden $N$. El objetivo a tratar radica en encontrar la descomposición CP con $R$ componentes que mejor aproxima al tensor $T$, esto es,

$$\min_{\widetilde{T}}\|T-\widetilde{T}\|\quad \text{con}\quad \widetilde{T}=\sum_{r=1}^{R}a_r^{(1)}\circ a_r^{(2)}\circ\cdots\circ a_r^{(N)}=[\![\,A^{(1)},\ldots,A^{(N)}\,]\!] \tag{2.3}$$

para lo cual consideramos la norma tensorial definida como la raíz cuadrada de la suma de todos los elementos del tensor al cuadrado, esto es,

$$\|T\|=\sqrt{\sum_{i_1,\ldots,i_N}t^2_{i_1,\ldots,i_N}}$$

(Esta norma tensorial es la análoga a la norma de Frobenius para el caso matricial).

El algoritmo ALS minimiza el error en 2.3 optimizando individualmente cada matriz factor $A^{(k)}$, mientras mantiene fijas el resto de matrices. Tomando las ventajas que aporta la simplificación matricial 2.2, el algoritmo actualiza las matrices factor en la forma

$$V=\big((A^{(1)})^T A^{(1)}*\cdots*(A^{(n-1)})^T A^{(n-1)}*(A^{(n+1)})^T A^{(n+1)}*\cdots*(A^{(N)})^T A^{(N)}\big)$$
$$A^{(k)}=T_{(n)}\big(A^{(N)}\odot\cdots\odot A^{(n+1)}\odot A^{(n-1)}\odot\cdots\odot A^{(1)}\big)V^\dagger$$

donde $V\in\mathbb{R}^{R\times R}$ y $V^\dagger$ denota la pseudo-inversa de Moore-Penrose de la matriz $V$. La pseudo-inversa de Moore-Penrose de una matriz, no necesariamente cuadrada, se debe recordar que se calcula como $V^\dagger=(V^T V)^{-1}V^T$.

En el Pseudocódigo 2.1 se muestra el algoritmo ALS, para el cual se supone fijado el criterio de inicialización de las matrices $A^{(k)}$ iniciales. Esta inicialización podrá llevarse a cabo de forma aleatoria, o siguiendo algún criterio preestablecido. El algoritmo se repite de forma iterada hasta alcanzar alguna condición de parada, definida a partir de criterios relativos al orden de mejora en la función objetivo, en las modificaciones de las matrices factor y/o al establecimiento del número máximo de iteraciones.

**procedure** *CP-ALS(*$T, R$*)*
  **initialize:** $A^{(n)} \in \mathbb{R}^{I_n \times R}$, $n \in \{1, ..., N\}$
  **repeat**
    **for** $n = 1, ..., N$ **do**
      $V \leftarrow (A^{(1)})^T A^{(1)} * \cdots * (A^{(n-1)})^T A^{(n-1)} * (A^{(n+1)})^T A^{(n+1)} * \cdots * (A^{(N)})^T A^{(N)}$
      $A^{(n)} \leftarrow T_{(n)}\big(A^{(N)} \odot \cdots \odot A^{(n+1)} \odot A^{(n-1)} \odot \cdots \odot A^{(1)}\big)V^{\dagger}$
      normalizar las columnas de $A^{(n)}$ y almacenar las normas en $\lambda$ (opcional)
    **end**
  **until** *condición de parada*
  return $\Lambda, A^{(1)}, A^{(2)}, ..., A^{(N)}$
**end**

**Pseudocódigo 2.1.** Algoritmo ALS para computar la descomposición CP con $R$ componentes de un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$

### 2.2.2. Descomposición de Tucker

La segunda descomposición que trataremos en este trabajo es la descomposición de Tucker, cuyo nombre se debe al matemático Ledyard R. Tucker, que en 1963 (véase [15]) introdujo esta descomposición.

La descomposición de Tucker descompone el tensor en un tensor núcleo multiplicado a lo largo de cada modo por una matriz. Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, la descomposición de Tucker puede expresarse como

$$T \cong \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} g_{r_1, r_2 ..., r_N} \left(a_{r_1}^{(1)} \circ a_{r_2}^{(2)} \circ \cdots \circ a_{r_N}^{(N)}\right)$$

con $a_{r_k}^{(k)} \in \mathbb{R}^{I_k}$ para cada $r_k \in \{1, ..., R_k\}$ y $k \in \{1, ..., N\}$. Construimos a partir de estos vectores las matrices factor $A^{(k)} = \left[\, a_1^{(k)} \,|\, a_2^{(k)} \,|\, \cdots \,|\, a_{R_k}^{(k)} \,\right] \in \mathbb{R}^{I_k \times R_k}$. De esta forma podemos reescribir la factorización de Tucker de $T$ como

$$
\begin{aligned}
T &\cong \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} g_{r_1,r_2\ldots,r_N} \left(a_{r_1}^{(1)} \circ a_{r_2}^{(2)} \circ \cdots \circ a_{r_N}^{(N)}\right) \\
&= G \times_1 A^{(1)} \times_2 A^{(2)} \cdots \times_N A^{(N)} = [\![ G;\, A^{(1)}, A^{(2)}, ..., A^{(N)} ]\!]
\end{aligned}
\tag{2.4}
$$

con $G \in \mathbb{R}^{R_1\times\cdots\times R_N}$ el tensor de componentes $g_{r_1,\ldots,r_N}$, denominado tensor núcleo (véase la Figura 2.6). Las entradas del tensor núcleo muestran el nivel de interacción entre las diferentes componentes de la descomposición.

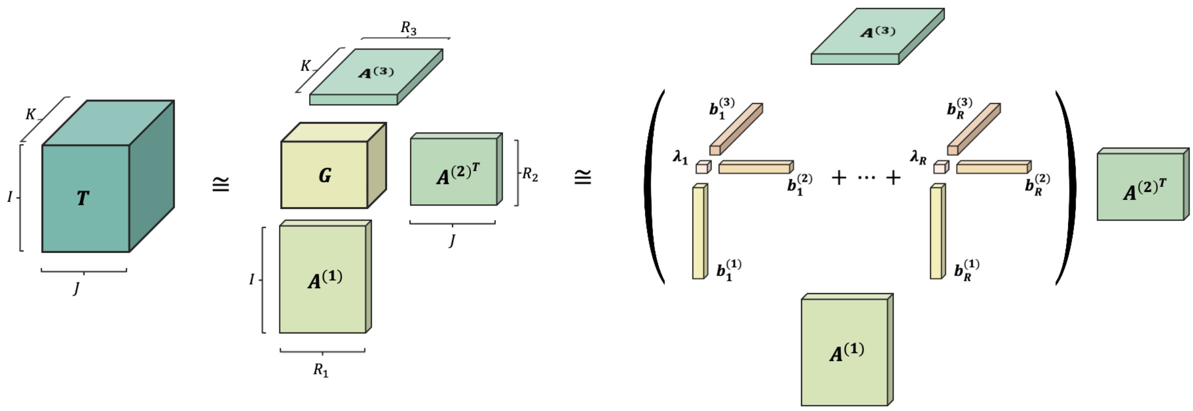


**Figura 2.6:** Representación gráfica de la descomposición de Tucker, para un tensor $T \in \mathbb{R}^{I\times J\times K}$ de orden 3, en la forma $T \cong [\![ G;\, A^{(1)}, A^{(2)}, A^{(3)} ]\!] = [\![ [\![ \Lambda;\, B^{(1)}, B^{(2)}, B^{(3)} ]\!];\, A^{(1)}, A^{(2)}, A^{(3)} ]\!]$. En la segunda expresión se realiza la descomposición CP del tensor núcleo $G$, de utilidad en determinados contextos

Por componentes, la descomposición de Tucker de $T \in \mathbb{R}^{I_1\times\cdots\times I_N}$ es

$$
t_{i_1,i_2,\ldots,i_N} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} g_{r_1,r_2\ldots,r_N}\; a_{i_1\,r_1}^{(1)}\; a_{i_2\,r_2}^{(2)}\; \cdots\; a_{i_N\,r_N}^{(N)}
$$

para $i_k \in \{1, ..., I_k\}$ con $k \in \{1, ..., N\}$. Utilizando el producto matricial de Kronecker, la versión matricial asociada al modo $n$-ésimo de la descomposición de Tucker viene dada por la identificación

$$
T_{(n)} \cong A^{(n)}\, G_{(n)} \left(A^{(N)} \otimes \cdots \otimes A^{(n+1)} \otimes A^{(n-1)} \otimes \cdots \otimes A^{(1)}\right)^T \tag{2.5}
$$

**Definición 2.2.6** *Dado un tensor $T \in \mathbb{R}^{I_1\times\cdots\times I_N}$ de orden $N$, el rango $n$-modal del tensor, denotado por $\mathrm{rank}_n(T)$, es el rango, en el sentido matricial, de la matriz de despliegue $T_{(n)}$.*

En otras palabras, el rango $n$-modal de un tensor es la dimensión del espacio vectorial generado por las fibras $n$-modales (véase la Figura 2.2). Si tomamos en la descomposición de Tucker los valores $R_n = rank_n(T)$ para cada $n \in \{1, ..., N\}$, se dice que el tensor $T$ es un tensor de rangos modales $(R_1, ..., R_N)$, o simplemente de rangos $(R_1, ..., R_N)$. Es importante no confundir el rango $n$-modal con el rango del tensor (véase 2.2.1). Es claro que $rank_n(T) \leq I_n$ para cada $n \in \{1, ..., N\}$, y se verifica también el siguiente resultado:

**Teorema 2.2.7** *Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, se verifica que, para cada $k \in \{1, ..., N\}$, $rank_k(T) \leq rank(T)$.*

*Demostración.*

Sea $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de rango $R$, luego existe una descomposición CP en la forma

$$T = \sum_{r=1}^{R} a^{(1),r} \circ a^{(2),r} \circ \cdots \circ a^{(N),r}$$

con $a^{(k),r} \in \mathbb{R}^{I_k}$ para cada $k \in \{1, ..., N\}$ y $r \in \{1, ..., R\}$. Fijamos un $k \in \{1, ..., N\}$. Consideramos el tensor $T^r = a^{(1),r} \circ a^{(2),r} \circ \cdots \circ a^{(N),r}$, con componentes

$$t^r_{i_1,...,i_{k-1},i_k,i_{k+1},...,i_N} = a^{(1),r}_{i_1} \cdots a^{(k-1),r}_{i_{k-1}} \cdot a^{(k),r}_{i_k} \cdot a^{(k+1),r}_{i_{k+1}} \cdots a^{(N),r}_{i_N}$$

Las fibras $k$-modales del tensor $T^r$ son los vectores

$$t^r_{i_1,...,i_{k-1},:,i_{k+1},...,i_N} = a^{(1),r}_{i_1} \cdots a^{(k-1),r}_{i_{k-1}} \cdot a^{(k),r} \cdot a^{(k+1),r}_{i_{k+1}} \cdots a^{(N),r}_{i_N}$$

y por tanto la matriz de despliegue asociada al modo $k$-ésimo de la matriz $T^r$ viene dada por

$$T^r_{(k)} = \left[ c^{(k),r}_1 \cdot a^{(k),r} \mid c^{(k),r}_2 \cdot a^{(k),r} \mid \cdots \mid c^{(k),r}_s \cdot a^{(k),r} \mid \cdots \mid c^{(k),r}_M \cdot a^{(k),r} \right] \in \mathbb{R}^{I_k \times M}$$

con $s = \overline{i_1, ..., i_{k-1}, i_{k+1}, ..., i_N}$ siguiendo el criterio *little-endian*, $M = I_1 \cdots I_{k-1} \cdot I_{k+1} \cdot I_N$, y

$$c^{(k),r}_s = c^{(k),r}_{\overline{i_1,...,i_{k-1},i_{k+1},...,i_N}} = a^{(1),r}_{i_1} \cdots a^{(k-1),r}_{i_{k-1}} \cdot a^{(k+1),r}_{i_{k+1}} \cdots a^{(N),r}_{i_N}$$

Por tanto se verifica que

$$T_{(k)} = \sum_{r=1}^{R} T^r_{(k)}$$

De esta forma, la $s$-ésima columna de la matriz de despliegue $k$-modal $T_{(k)}$ viene dada por

$$c^{(k),1}_s \cdot a^{(k),1} + c^{(k),2}_s \cdot a^{(k),2} + \cdots + c^{(k),R}_s \cdot a^{(k),R} \in \langle \{a^{(k),1}, a^{(k),2}, \ldots, a^{(k),R}\} \rangle$$

Por tanto es claro que la dimensión del espacio generado por las columnas de la matriz $T_{(k)}$ es a lo sumo $R$, quedando así probado que $rank_k(T) \leq R$ para cada $k \in \{1, ..., N\}$.

□

Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, se puede encontrar fácilmente, tal y como se muestra a continuación, una descomposición de Tucker de rango $(R_1, ..., R_N)$ con $R_n = rank_n(T)$, para cada $n \in \{1, \ldots, N\}$. Sin embargo, si en alguno de los modos tomamos $R_n \leq rank_n(T)$, el proceso de computo de una descomposición de Tucker es más costoso y al mismo tiempo más inexacto. En este segundo caso, haremos referencia a dicha descomposición como descomposición de Tucker truncada.

En cuanto a la unicidad, la descomposición de Tucker de un tensor $T$ como en 2.4 no es única. Para ello, basta considerar matrices $B^{(1)}, \ldots, B^{(N)}$ no singulares, con $B^{(k)} \in \mathbb{R}^{R_k \times R_k}$, y de esta forma es fácil comprobar que se verifica

$$[\![ G;\, A^{(1)}, ..., A^{(N)} ]\!] = \left[\!\!\left[ [\![ G;\, B^{(1)}, ..., B^{(N)} ]\!];\, A^{(1)}\big(B^{(1)}\big)^{-1}, ..., A^{(N)}\big(B^{(N)}\big)^{-1} \right]\!\!\right]$$

En otras palabras, podemos modificar el tensor núcleo $G$ sin afectar al ajuste siempre y cuando apliquemos las modificaciones inversas sobre las matrices factor $A^{(1)}, ..., A^{(N)}$. Esto nos permitirá tomar aquellas transformaciones que permitan simplificar en la medida de lo posible la estructura del tensor núcleo, tratando de obtener un tensor con el mayor número de componentes nulas posibles.

Tucker introdujo, junto con la descomposición, varios métodos para su cómputo. El primero de ellos, y el más extendido en la actualidad, es el conocido como método de descomposición en valores singulares de orden superior o HOSVD (de sus siglas en inglés, High-Order Singular Value Decomposition). Este método será de utilidad, y además fácil de aplicar, en aquellos casos en los que se busque una descomposición de Tucker de orden $(R_1, ..., R_N)$ con $R_n = rank_n(T)$ para cada $n \in \{1, ..., N\}$.

Veamos a continuación una descripción del algoritmo HOSVD para el cómputo de la descomposición de Tucker (véase además el Pseudocódigo 2.2). Dado un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, para cada $k \in \{1, ..., N\}$ se considera la descomposición en valores singulares de la matriz de despliegue $k$-modal, esto es, $T_{(k)} = U_k \, \Sigma_k \, V_k^T$ con $U_k \in \mathbb{R}^{I_k \times I_k}$ y $V_k \in \mathbb{R}^{M \times M}$ ortogonales y $\Sigma_k \in \mathbb{R}^{I_k \times M}$ una matriz formada por los valores singulares de $T_{(k)}$ en su diagonal principal ordenados de mayor a menor ($M = I_1 \cdots I_{k-1} I_{k+1} \cdots I_N$). Por las propiedades del producto multilineal, dada $I \in \mathbb{R}^{I_k \times I_k}$ la matriz identidad, se tiene

$$\begin{aligned} T &= T \times_1 I \times_2 I \cdots \times_N I \\ &= T \times_1 U_1\, U_1^T \times_2 U_2\, U_2^T \cdots \times_N U_N\, U_N^T \\ &= \left(T \times_1 U_1^T \times_2 U_2^T \cdots \times_N U_N^T\right) \times_1 U_1 \times_2 U_2 \cdots \times_N U_N \end{aligned}$$

Tomando el tensor núcleo $G = T \times_1 U^{(1)^T} \times_2 U^{(2)^T} \cdots \times_N U^{(N)^T}$, se obtiene la siguiente descomposición de Tucker:

$$T = G \times_1 U_1 \cdots \times_N U_N = [\![\, G;\, U_1, ..., U_N ]\!] = \Big[\!\!\Big[\, [\![\, T;\, U_1^T, ..., U_N^T\, ]\!];\, U_1, ..., U_N \Big]\!\!\Big]$$

**procedure** *HOSVD($T, R_1,...,R_N$)*
  **for** $n = 1, ..., N$ **do**
    $A^{(n)} \leftarrow$ matriz $U_n$ de la descompocisión en valores singulares $T_{(n)} = U_n\, \Sigma_n\, V_n^T$
  **end**
  $G \leftarrow T \times_1 A^{(1)^T} \times_2 A^{(2)^T} \cdots \times_N A^{(N)^T}$
  return $G, A^{(1)}, A^{(2)}, ..., A^{(N)}$
**end**

**Pseudocódigo 2.2.** Algoritmo HOSVD para computar la descomposición de Tucker de rango $(R_1, \ldots, R_N)$ de un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$

Cuando para algún modo se verifica $R_n \leq rank_n(T)$, la descomposición HOSVD se denomina descomposición HOSVD truncada. Tal y como se menciona anteriormente, esta descomposición truncada no es óptima en términos de obtener el mejor ajuste medido como la norma de la diferencia entre el tensor y la aproximación. Sin embargo, sí resulta útil cómo método de partida para el siguiente algoritmo a estudio.

El algoritmo de iteración ortogonal de orden superior (HOOI, de las siglas en ingles *High-Order Orthogonal Iteration*) es esencialmente un algoritmo ALS que utiliza la salida del algoritmo HOSVD como punto de entrada para inicializar las matrices factor de la descomposición (véase el Pseudocódigo 2.3 con los diferentes pasos del algoritmo). Si consideramos un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$, el problema de optimización que tratamos de resolver es

$$\min_{\widetilde{T}} \|T - \widetilde{T}\| \quad \text{con} \quad \widetilde{T} = [\![\, G\,; A^{(1)}, \ldots, A^{(N)}\, ]\!] \tag{2.6}$$

con $G \in \mathbb{R}^{R_1 \times \cdots \times R_N}$ y $A^{(n)} \in \mathbb{R}^{I_n \times R_n}$ y ortogonal por columnas, para cada $n \in \{1, ..., N\}$. El tensor núcleo debe satisfacer

$$G = T \times_1 A^{(1)^T} \cdots \times_N A^{(N)^T} = [\![\, G;\, A^{(1)^T}, ..., A^{(N)^T} ]\!]$$

**procedure** *HOOI*($T, R_1,...,R_N$)

  **initialize:** $A^{(n)} \in \mathbb{R}^{I_n \times R_n}$, $n \in \{1, ..., N\}$ utilizando HOSVD

  **repeat**

    **for** $n = 1, ..., N$ **do**

      $V \leftarrow T \times_1 A^{(1)^T} \cdots \times_{n-1} A^{(n-1)^T} \times_{n+1} A^{(n+1)^T} \cdots \times_N A^{(N)^T}$

      $A^{(n)} \leftarrow R_n$ vectores singualres por la izquierda de $V_{(n)}$

    **end**

  **until** *condición de parada*

  $G \leftarrow T \times_1 A^{(1)^T} \times_2 A^{(2)^T} \cdots \times_N A^{(N)^T}$

  return $\Lambda, A^{(1)}, A^{(2)}, ..., A^{(N)}$

**end**

**Pseudocódigo 2.3.** Algoritmo HOOI para computar la descomposición de Tucker de rango $(R_1, ..., R_N)$ de un tensor $T \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ de orden $N$

### 2.2.3. Aplicaciones prácticas

La descomposiciones tensoriales se utilizan de forma aplicada en distintas áreas de conocimiento, como neurociencia, procesamiento de señales, visión computacional, minería de datos o aprendizaje automático.

Dado un tensor es claro que, tras descomponerlo, se puede obtener una nueva representación que puede ser manejada con un menor número de componentes. Esta técnica de compresión de tensores puede aplicarse en distintos campos, permitiendo trabajar con estructuras de datos de ordenes de magnitud inferior, reduciendo así el coste computacional de muchos algoritmos. Un ejemplo muy sencillo en el que se puede sacar partido de esta reducción de componentes de la nueva descomposición podemos encontrarlo en la compresión de imágenes. Por este motivo, dada la sencillez del método, se muestra a continuación un ejemplo de compresión de imágenes con el fin de poner en práctica los algoritmos de descomposición estudiados.

Consideramos una imagen digital en formato PNG con un modelo de color RGB. De esta forma la imagen puede ser tratada como una matriz de píxeles, y a su vez cada uno de estos píxeles puede tratarse como un vector de 3 componentes. Por todo ello estas imágenes pueden tratarse como tensores de orden 3. Para el ejemplo se considera una imagen de un tamaño de $256 \times 256$ píxeles, obteniendo así tensores de dimensión $256 \times 256 \times 3$ (Figura 2.7).

Utilizando el algoritmo HOOI para obtener descomposiciones que aproximen el tensor original, podemos obtener descomposiciones de Tucker para un rango fijo $(R_1, R_2, R_3)$. En la Figura

2.7 se muestra el resultado de la imagen obtenida tras reconstruir el tensor a partir de descomposiciones de Tucker con rangos $(128, 128, 3)$, $(64, 64, 3)$, $(32, 32, 3)$ y $(16, 16, 3)$. Para cada descomposición se analiza el porcentaje de compresión obtenido al descomponer el correspondiente tensor, y el error relativo cometido medido en la norma de Frobenius tensorial.

| Imagen original | Reconstrucción | Reconstrucción | Reconstrucción | Reconstrucción |
|---|---|---|---|---|
| Rango $= (256, 256, 3)$ | Rango $= (128, 128, 3)$ | Rango $= (64, 64, 3)$ | Rango $= (32, 32, 3)$ | Rango $= (16, 16, 3)$ |
| Compresión $= 0\,\%$ | Compresión $= 41{,}66\,\%$ | Compresión $= 77{,}07\,\%$ | Compresión $= 90{,}10\,\%$ | Compresión $= 95{,}44\,\%$ |
| Error rel. $= 0\,\%$ | Error rel. $= 2{,}29\,\%$ | Error rel. $= 6{,}24\,\%$ | Error rel. $= 10{,}16\,\%$ | Error rel. $= 14{,}32\,\%$ |

**Figura 2.7:** Compresión de imágenes mediante descomposiciones de Tucker

El tensor original se compone de un total de $256 \times 256 \times 3 = 196698$ elementos, mientras que, por ejemplo, la descomposición de rango $(128, 128, 3)$ se compone del tensor núcleo de dimensión $128 \times 128 \times 3$ (49152 componentes), dos matrices de dimensión $256 \times 128$ (32768 componentes) y una matriz de $3 \times 3$ (9 componentes), lo que hacen un total de 114697 elementos, permitiendo una compresión del $41{,}66\,\%$, con un error relativo cometido del $2,29\,\%$. De forma análoga se estudian los tres casos restantes.

# Índice alfabético

# Bibliografía


[1] Z. Ahsan, *Tensors: mathematics of differential geometry and relativity*, PHI Learning, 2015.

[2] R. L. Bishop and S. I. Goldberg, *Tensor analysis on manifolds*, Dover, 1980.

[3] A. Cichocki, *Era of big data processing: A new approach via tensor networks and tensor decompositions*, arXiv preprint arXiv:1403.2048 (2014).

[4] W. H. Greub, *Multilinear algebra*, Springer Berlin Heidelberg, 1967.

[5] T. G. Kolda, *Multilinear operators for higher-order decompositions.*, Sandia National Laboratories (2006).

[6] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM review **51** (2009), no. 3, 455–500.

[7] J. M. Lee, *Riemannian manifolds: an introduction to curvature*, Vol. 176, Springer Science & Business Media, 2006.

[8] J. M. Lee, *Smooth manifolds*, Springer, 2013.

[9] A. Lichnerowicz, *Elementos de cálculo tensorial*, Aguilar, 1972.

[10] X. Liu and N. D. Sidiropoulos, *Cramér-rao lower bounds for low-rank decomposition of multidimensional arrays*, IEEE Trans. on Signal Process. **49** (2001), 2074–2086.

[11] S. Rabanser, O. Shchur, and S. Günnemann, *Introduction to tensor decompositions and their applications in machine learning*, arXiv preprint arXiv:1711.10781 (2017).

[12] R Rios, *Métodos de descomposición en valores singulares para tensores*, Facultad de Ciencias de la Universidad de Nacional de Colombia (2009).

[13] N. D. Sidiropoulos and R. Bro, *On the uniqueness of multilinear decomposition of n-way arrays*, J. Chemometrics **14** (2000), no. 3, 229–239.

[14] J. M. F. Ten Berge and N. D. Sidiropoulos, *On uniqueness in CANDECOMP/PARAFAC*, Psychometrika **67** (2002), 399–409.

[15] L. R. Tucker, *Implications of factor analysis of three-way matrices for measurement of change*, in Problems in Measuring Change, University of Wisconsin Press (1963), 122–137.