Álgebra lineal y multilineal

Universidad Internacional de Valencia

Máster Universitario en Inteligencia Artificial

02MIAR | Matemáticas:
Matemáticas para la Inteligencia Artificial

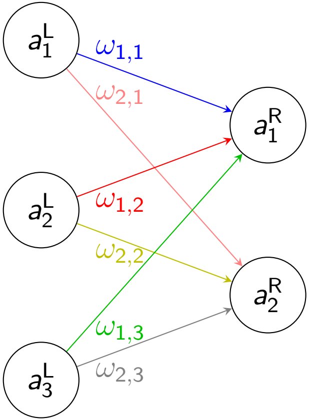
Profesor:
David Zorío Ventura

universidadviu.com

De: Planeta Formación y Universidades

Motivación

Las matrices constituyen un elemento importante en la descripción de redes neuronales.



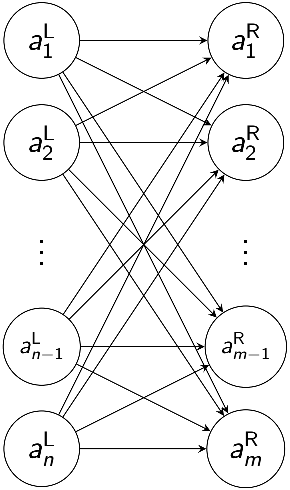
$$
\begin{aligned}
a_1^R &= f_1\left(\omega_{1,1}a_1^L + \omega_{1,2}a_2^L + \omega_{1,3}a_3^L + b_1\right),\\
a_2^R &= f_2\left(\omega_{2,1}a_1^L + \omega_{2,2}a_2^L + \omega_{2,3}a_3^L + b_2\right).
\end{aligned}
$$

$$
W = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \omega_{1,3} \\ \omega_{2,1} & \omega_{2,2} & \omega_{2,3} \end{pmatrix}, \quad a^L = \begin{pmatrix} a_1^L \\ a_2^L \\ a_3^L \end{pmatrix}.
$$

$$
\begin{aligned}
a_1^R &= f_1\left(W(1,:)a^L + b_1\right),\\
a_2^R &= f_2\left(W(2,:)a^L + b_2\right).
\end{aligned}
$$

Motivación

En general, para el caso de información de una capa de n neuronas que se transmite a otra capa de m neuronas:



$$a_i^R = f_i\left(\sum_{j=1}^{n} \omega_{i,j} a_j^L + b_i\right), \ 1 \leq i \leq m.$$

$$W = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,n-1} & \omega_{1,n} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,n-1} & \omega_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \omega_{m-1,1} & \omega_{m-1,2} & \cdots & \omega_{m-1,n-1} & \omega_{m-1,n} \\ \omega_{m,1} & \omega_{m,2} & \cdots & \omega_{m,n-1} & \omega_{m,n} \end{pmatrix},$$

$$a^L = \begin{pmatrix} a_1^L \\ a_2^L \\ \vdots \\ a_{n-1}^L \\ a_n^L \end{pmatrix} \to a_i^R = f_i\left(W(i,:)a^L + b_i\right), \ 1 \leq i \leq m.$$

Autovectores y autovalores

Universidad Internacional de Valencia

Definición

Sea $A \in \mathbb{R}^{n\times n}$ una matriz cuadrada. Un **vector propio** *o* **autovector** $v \in \mathbb{R}^{n\times 1}$, $v \neq \overrightarrow{\mathbf{0}}$, *asociado a A es aquel que cumple que $\exists\lambda \in \mathbb{R}$ tal que $Av = \lambda v$. El valor λ recibe el nombre de* **valor propio** *o* **autovalor** *de A asociado al autovector v.*

Ejemplo

Sea $A = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}$. $\lambda_1 = 2$ *y* $\lambda_2 = -3$ *son valores propios de A asociados a los vectores propios* $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ *y* $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, *resp.:*

$$\begin{aligned} Av_1 &= \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2v_1, \\ Av_2 &= \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \end{pmatrix} = -3\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -3v_2. \end{aligned}$$

Diagonalización

- ¿Qué interés tiene conocer los autovalores y autovectores?
- Supongamos que, dada $A \in \mathbb{R}^{n\times n}$, conseguimos encontrar n autovectores linealmente independientes, $\{v_i\}_{i=1}^n$ (base de $\mathbb{R}^n$), asociados respectivamente a n autovalores, $\{\lambda_i\}_{i=1}^n$.
- Por definición, se tiene $Av_i = \lambda_i v_i$ para $1 \leq i \leq n$.
- Denotamos $P \in \mathbb{R}^{n\times n}$ a la matriz formada por los vectores propios dispuestos en columna y D a la matriz diagonal formada por los valores propios:

$$P = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

- Entonces se cumple $AP = PD$. Equivalentemente: $P^{-1}AP = D$.
- $(AP)(:,j) = AP(:,j) = Av_j = \lambda_j v_j = PD(:,j) = (PD)(:,j)$.

Ejemplo

Anteriormente vimos que los vectores propios de $A = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}$ *son* $v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ *y* $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, *asociados, respectivamente, a los autovalores* $\lambda_1 = 2$ *y* $\lambda_2 = -3$. *Definimos*

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$$

$$AP = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow \\ Av_1 & Av_2 \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 2 & -3 \end{pmatrix}$$

$$PD = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow \\ \lambda_1 v_1 & \lambda_2 v_2 \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 2 & -3 \end{pmatrix}$$

Diagonalización

Universidad Internacional de Valencia

Definición

Sea $A \in \mathbb{R}^{n\times n}$ *una matriz cuadrada. Diremos que* A *es* **diagonalizable** *si* $\exists P \in \mathbb{R}^{n\times n}$ *regular y* $D \in \mathbb{R}^{n\times n}$ *diagonal tales que*

$$P^{-1}AP = D.$$

- ¿Qué utilidad tiene diagonalizar una matriz?
- Cálculo del determinante:

$$\begin{aligned}\det(A) = \det\left(PDP^{-1}\right) &= \det(P)\det(D)\det\left(P^{-1}\right)\\ &= \det(P)\det(D)\frac{1}{\det(P)} = \det(D) = \prod_{i=1}^{n}\lambda_i.\end{aligned}$$

- Cálculo del rango:

$$\operatorname{rank}(A) = \left|\{\lambda_i \mid \lambda_i \neq 0,\ 1 \leq i \leq n\}\right|.$$

- Cálculo de potencias matriciales.
  Si $P^{-1}AP = D$, entonces multiplicando por P a la izquierda y P^{-1} a la derecha podemos obtener
$$A = PDP^{-1}.$$

  - $A^2 = A \cdot A = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDI_nDP^{-1} = P(D \cdot D)P^{-1} = PD^2P^{-1}.$
  - $A^3 = A^2 \cdot A = (PD^2P^{-1})(PDP^{-1}) = PD^2(P^{-1}P)DP^{-1} = PD^2I_nDP^{-1} = P(D^2D)P^{-1} = PD^3P^{-1}.$
  - En general, se tiene
$$A^k = PD^kP^{-1}, \text{ donde } D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

Diagonalización

- Si $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ es un valor propio de A y $v \in \mathbb{R}^{n\times 1}$ es un vector propio asociado a λ, entonces:
$$Av = \lambda v \leftrightarrow Av - \lambda v = \overrightarrow{\mathbf{0}} \leftrightarrow Av - \lambda I_n v = \overrightarrow{\mathbf{0}} \leftrightarrow (A - \lambda I_n)v = \overrightarrow{\mathbf{0}}.$$
- $(A - \lambda I_n)v = \overrightarrow{\mathbf{0}}$ es un sistema de ecuaciones lineal **homogéneo** (términos independientes nulos). Por tanto, siempre es compatible:
  - Si $\det(A - \lambda I_n) \neq 0$, entonces es compatible determinado y su única solución es $v = \overrightarrow{\mathbf{0}}$
  - Si $\det(A - \lambda I_n) = 0$, entonces es compatible indeterminado y tiene infinitas soluciones.
- Queremos encontrar vectores propios $v \neq \overrightarrow{\mathbf{0}}$, luego imponemos $\det(A - \lambda I_n) = 0$ (**ecuación característica**). Las raíces de esta ecuación con incógnita λ son los valores propios.
- Para cada valor propio obtenido $\{\lambda_i\}_{i=1}^{k}$ se obtiene el máximo número de vectores linealmente independientes como que verifican $(A - \lambda_i I_n)v = \overrightarrow{\mathbf{0}}$, $1 \leq i \leq k$.

Ejemplo

- *Sea* $A = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}$.
- *Ec. caract.:* $\det(A - \lambda I_2) = 0 \rightarrow \det \begin{pmatrix} 7-\lambda & -10 \\ 5 & -8-\lambda \end{pmatrix} = 0 \rightarrow \lambda^2 + \lambda - 6 = 0 \rightarrow$ $\lambda_1 = -3,\ \lambda_2 = 2.$
- *Vectores propios:*
  - $(A - \lambda_1 I_2)v = \overrightarrow{\mathbf{0}} \rightarrow \begin{pmatrix} 10 & -10 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}, \alpha \in \mathbb{R}.$
    $\alpha = 1 \rightarrow v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$
  - $(A - \lambda_2 I_2)v = \overrightarrow{\mathbf{0}} \rightarrow \begin{pmatrix} 5 & -10 \\ 5 & -10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2\alpha \\ \alpha \end{pmatrix}, \alpha \in \mathbb{R}.$
    $\alpha = 1 \rightarrow v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$

Otras descomposiciones matriciales

- No todas las matrices son diagonalizables en $\mathbb{R}$.
- Si A es cuadrada y no diagonalizable $\rightarrow$ **Forma canónica de Jordan.**
- Si A es rectangular $\rightarrow$ **Factorización SVD** (*singular value decomposition*).
- Otras factorizaciones útiles:
  - Factorización LU (método de eliminación de Gauss).
  - Factorización de Cholesky (mínimos cuadrados, Montecarlo...).
  - Factorización QR (mínimos cuadrados).

Definición

Una matriz $A \in \mathbb{R}^{n\times n}$ es simétrica si $A' = A$

Teorema

Si $A \in \mathbb{R}^{n\times n}$ es simétrica, entonces A es diagonalizable.

Matrices simétricas

Definición

Una matriz simétrica $A \in \mathbb{R}^{n\times n}$ *es* **definida positiva** *(respectivamente* **semidefinida positiva***) si* $\forall v \in \mathbb{R}^{n\times 1}$, $v \neq \overrightarrow{\mathbf{0}}$, $v'Av > 0$ *(respectivamente* $v'Av \geq 0$*). Por otra parte,* A *es* **(semi)definida negativa** *si* $-A$ *es (semi)definida positiva.*

Ejemplos

1. $I_n \in \mathbb{R}^{n\times n}$ *es definida positiva* $\forall n \in \mathbb{N}$. *En efecto, dado* $v \in \mathbb{R}^{n\times 1}$, $v \neq \overrightarrow{\mathbf{0}}$, $v'I_nv = v'v = \|v\|_2^2 > 0$.
2. $0_n \in \mathbb{R}^{n\times n}$ *es semidefinida positiva* $\forall n \in \mathbb{N}$. *En efecto, dado* $v \in \mathbb{R}^{n\times 1}$, $v \neq \overrightarrow{\mathbf{0}}$, $v'0_nv = 0$.
3. $A = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ *no es (semi)definida positiva, pues, por ejemplo, tomando* $v' = (1,1)$, *se tiene* $v'Av = -6 < 0$.

Matrices simétricas

Definición

Sea $A \in \mathbb{R}^{n\times n}$ *y* $k \in \{1, 2, \dots, n\}$*. El* **menor principal dominante** *de orden* k *asociado a* A*, al que denotaremos por* A_k*, es* $A_k = \det\left(A(1:k, 1:k)\right)$*; es decir, el determinante de la submatriz formada por las primeras* k *filas y columnas.*

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ *una matriz simétrica. Los enunciados siguientes son equivalentes:*

- A *es definida positiva (respectivamente, semidefinida positiva).*
- $A_k > 0$, $\forall k \in \{1, 2, \dots, n\}$ *(respectivamente,* $A_k \geq 0$*).*
- $\forall \lambda \in \mathbb{R}$ *valor propio de* A*,* $\lambda > 0$ *(respectivamente,* $\lambda \geq 0$*).*

Álgebra tensorial

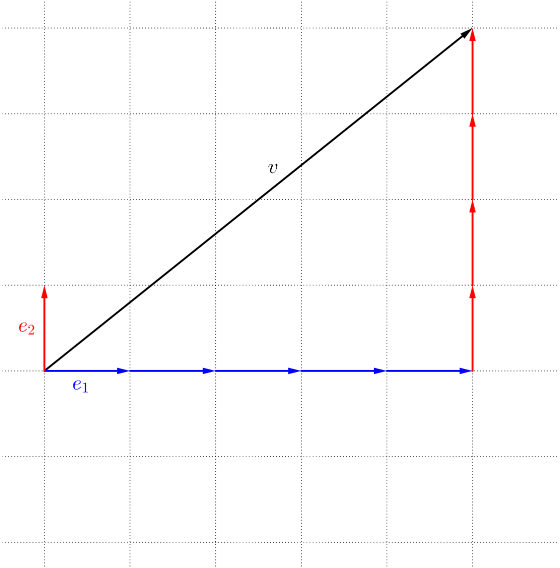
Definición

Un **tensor** *es un objeto invariante con respecto a un cambio de coordenadas.*

Ejemplos

1. *Los* **escalares** *en* $\mathbb{R}$ *son tensores.*
2. *Los* **vectores** v *de un espacio vectorial* V *son tensores.*
3. *Los elementos del conjunto* V^*, *formado por las aplicaciones lienales de la forma* $V \to \mathbb{R}$, *llamados* **covectores**, *son tensores.*
4. *Las aplicaciones lineales entre dos espacios vectoriales* V *y* W, *de la forma* $V \to W$, *tienen representación como tensor.*
5. *En consecuencia, las matrices de* $\mathbb{R}^{m\times n}$ *tienen representación como tensor.*
6. *Tensores métricos (ecuaciones de campo de Einstein).*

Sistemas de coordenadas



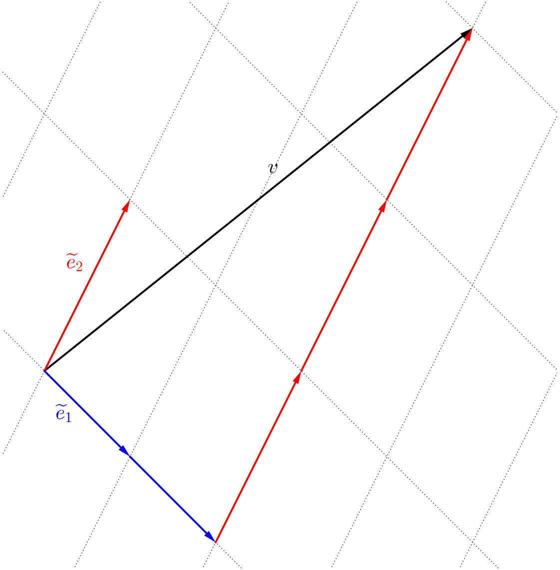
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\mathcal{B} = \{e_1, e_2\}$, donde:

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- $v = {\color{blue}5} e_1 + {\color{red}4} e_2$
- Coordenadas de v en la base $\mathcal{B}$:

$$\begin{bmatrix} {\color{blue}5} \\ {\color{red}4} \end{bmatrix}_{\mathcal{B}}$$

Sistemas de coordenadas



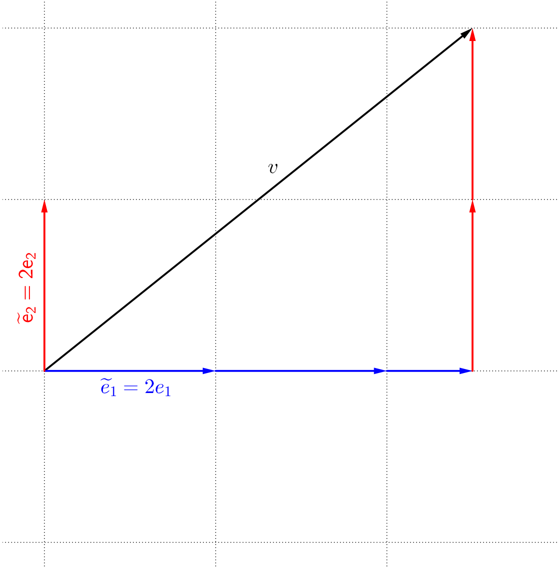
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\widetilde{\mathcal{B}} = \{\widetilde{e}_1, \widetilde{e}_2\}$, donde:

$$\widetilde{e}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \widetilde{e}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

- $v = {\color{blue}2}\widetilde{e}_1 + {\color{red}3}\widetilde{e}_2$
- Coordenadas de v en la base $\widetilde{\mathcal{B}}$:

$$\begin{bmatrix} {\color{blue}2} \\ {\color{red}3} \end{bmatrix}_{\widetilde{\mathcal{B}}}$$

Sistemas de coordenadas



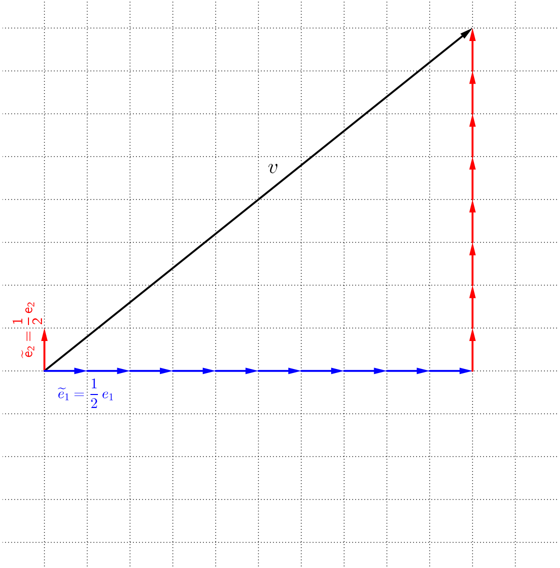
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\widetilde{\mathcal{B}} = \{\widetilde{e}_1 = 2e_1, \widetilde{e}_2 = 2e_1\}$, donde:

$$\widetilde{e}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \ \widetilde{e}_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

- $v = 2.5\widetilde{e}_1 + 2\widetilde{e}_2$
- Coordenadas de v en la base $\widetilde{\mathcal{B}}$:

$$\begin{bmatrix} 2.5 \\ 2 \end{bmatrix}_{\widetilde{\mathcal{B}}} = \frac{1}{2}\begin{bmatrix} 5 \\ 4 \end{bmatrix}_{\mathcal{B}}$$

Sistemas de coordenadas



- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\widetilde{\mathcal{B}} = \{\widetilde{e}_1 = \frac{1}{2} e_1, \widetilde{e}_2 = \frac{1}{2} e_1\}$, donde:

$$\widetilde{e}_1 = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \; \widetilde{e}_2 = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$$

- $v = {\color{blue}10}\widetilde{e}_1 + {\color{red}8}\widetilde{e}_2$
- Coordenadas de v en la base $\widetilde{\mathcal{B}}$:

$$\begin{bmatrix} {\color{blue}10} \\ {\color{red}8} \end{bmatrix}_{\widetilde{\mathcal{B}}} = 2 \begin{bmatrix} {\color{blue}5} \\ {\color{red}4} \end{bmatrix}_{\mathcal{B}}$$

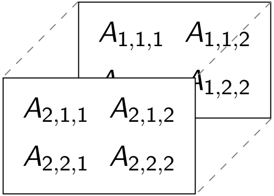
Sistemas de coordenadas

- Diremos que una componente de un tensor es **contravariante** si ésta varía en proporción inversa con respecto a un cambio de coordenadas.
- **Ejemplo:** coordenadas de un vector con respecto a una base.
- Diremos que una componente de un tensor es **covariante** si ésta varía en proporción directa con respecto a un cambio de coordenadas.
- **Ejemplo:** coordenadas de un vector v formadas a partir de su producto escalar con los elementos de la base $\mathcal{B} = \{b_i\}_{i=1}^{n}$:

$$\begin{bmatrix} v \cdot b_1 \\ \vdots \\ v \cdot b_n \end{bmatrix}$$

Arrays multidimensionales

- Los arrays multidimensionales son casos particulares de objetos que tienen representación de tensor.
- El **rango** de un tensor en forma de array representa las dimensiones espaciales en términos de la disposición de sus entradas.
  - Un escalar $x \in \mathbb{R}$ tiene rango 0.
  - Un vector $v \in \mathbb{R}^n$ tiene rango 1 (y dimensión n).
  - Una matriz $A \in \mathbb{R}^{n \times n}$ tiene rango 2 (y dimensión n).
  - Un array del tipo $A \in \mathbb{R}^{n \times n \times n}$ tiene rango 3 (y dimensión n).



- En general, $T^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ es un tensor de tipo (p, q) (p componentes contravariantes y q componentes covariantes), con rango $p + q$.

Arrays multidimensionales

Ejemplos

1. *Audio PCM Mono de* 2 *segundos, con frecuencia de muestreo de* 48 *kHz* (48000 *muestras por segundo).*
   - *Rango:* 1.
   - *Dimensión:* 96000.
2. *Audio PCM Stereo de* 3 *segundos, con frecuencia de muestreo* 44.1 *kHz* (44100 *muestras por segundo).*
   - *Rango:* 2.
   - *Dimensión:* 2×132200.
3. *Imagen monocroma, con* 600 *píxeles de anchura y* 800 *píxeles de altura.*
   - *Rango:* 2.
   - *Dimensión:* 600×800.

Arrays multidimensionales

Ejemplos

4. *Imagen RGB, con* 2000 *píxeles de anchura y* 1200 *píxeles de altura.*
   - *Rango:* 3.
   - *Dimensión:* $3 \times 2000 \times 1200$.
5. *Vídeo monocromo de* 1 *minuto, a* 30 *fps (frames por segundo), con resolución 720p:*
   - *Rango:* 3.
   - *Dimensión:* $1800 \times 1280 \times 720$.
6. *Vídeo RGB de* 40 *segundos, a* 60 *fps (frames por segundo), con resolución 1080p:*
   - *Rango:* 4.
   - *Dimensión:* $2400 \times 3 \times 1920 \times 1080$.

¡Muchas gracias!

Universidad Internacional de Valencia

Contacto:

david.zorio@campusviu.es