Álgebra

Universidad Internacional de Valencia

Máster Universitario en Inteligencia Artificial

02MIAR | Matemáticas:

Matemáticas para la Inteligencia Artificial

Profesor:

Víctor M. Campello

De:
Planeta Formación y Universidades

Definición

Un vector real v de dimensión n es una lista ordenada de n números reales:

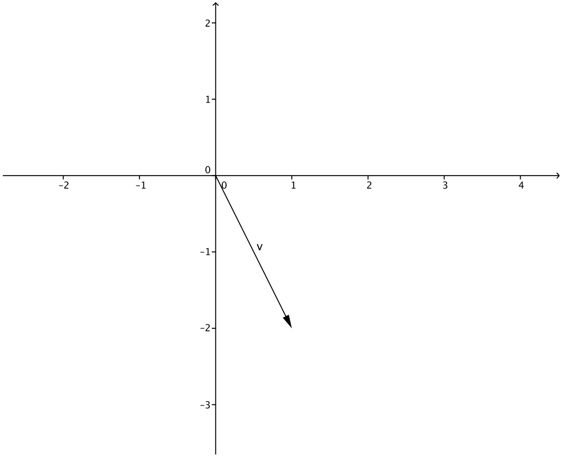
$$v = (v_1, v_2, \ldots, v_n), \quad v_i \in \mathbb{R}.$$

Se denota por $\mathbb{R}^n$ el conjunto formado por todos los vectores de dimensión n. Por tanto, podemos escribir $v \in \mathbb{R}^n$.

Ejemplos

1. $v = (1, -\frac{2}{3}) \in \mathbb{R}^2$.
2. $w = (\sqrt{2}, -1.2, \pi) \in \mathbb{R}^3$.

Los vectores en $\mathbb{R}^n$ se pueden interpretar de forma gráfica mediante un sistema de n coordenadas. En caso de tener una dimensión de $n = 3$ o inferior, éstos se pueden dibujar.



Definición

Dados dos vectores $v, w \in \mathbb{R}^n$, $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n)$, *podemos definir su suma como*

$$v + w = (v_1 + w_1, v_2 + w_2, \dots, v_n + w_n).$$

Definición

Dado un valor real $\lambda \in \mathbb{R}$ *(escalar) y un vector* $v \in \mathbb{R}^n$, $v = (v_1, v_2, \dots, v_n)$, *podemos definir su producto como*

$$\lambda \cdot v = (\lambda v_1, \lambda v_2, \dots, \lambda v_n).$$

La longitud de los vectores se puede medir a través de **normas**. Algunas normas vectoriales importantes son:

- Norma euclídea:

$$\|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

- Norma 1:

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| = |v_1| + |v_2| + \cdots + |v_n|.$$

- Norma del máximo:

$$\|v\|_\infty = \max_{1\leq i\leq n} |v_i|.$$

Ejemplo

Construimos un modelo de machine learning para predecir las temperaturas a partir de algunas variables de entrada (input). Supongamos que obtenemos un conjunto de n predicciones con el modelo considerado, almacenadas en un vector:

$$\hat{v} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n) \in \mathbb{R}^n.$$

Comparamos ahora el vector de predicciones, $\hat{v}$, con el vector de respuestas reales (observaciones), $v = (v_1, \ldots, v_n)$. Una forma de hacerlo es utilizar alguna norma sobre la diferencia de los dos vectores:

$$E = \frac{1}{n}\|\hat{v} - v\|.$$

Objetivo: encontrar un algoritmo que minimice E.

El producto escalar entre dos vectores $v, w \in \mathbb{R}^n$ se define como:

$$v \cdot w = \sum_{i=1}^{n} v_i w_i .$$

Ejemplo

Sean $v, w \in \mathbb{R}^3$, *con* $v = (2, -1, 0)$ *y* $w = (-3, -2, 1)$. *Entonces*

$$\begin{aligned} v \cdot w &= (2, -1, 0) \cdot (-3, -2, 1) \\ &= 2 \cdot (-3) + (-1) \cdot (-2) + 0 \cdot 1 = -6 + 2 + 0 = -4. \end{aligned}$$

Definición

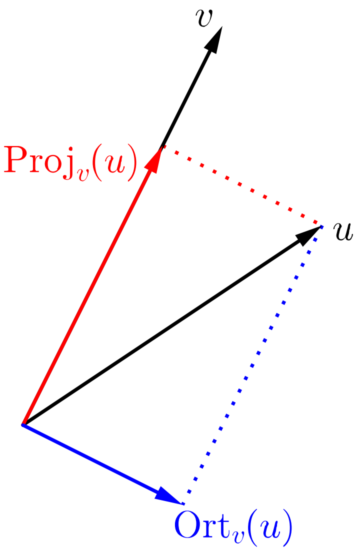
Sean $u, v \in \mathbb{R}^n$. Definimos la **proyección** *de u sobre v como*

$$\mathrm{Proj}_v\left(u\right) = \frac{u \cdot v}{v \cdot v} v.$$

Definición

Sean $u, v \in \mathbb{R}^n$. Definimos la **ortogonal** *de u sobre v como*

$$\mathrm{Ort}_v\left(u\right) = u - \mathrm{Proj}_v\left(u\right).$$



Definición

Diremos que dos vectores $u, v \in \mathbb{R}^n$ son **ortogonales** *si $u \cdot v = 0$.*

La definición de ortogonalidad entre dos vectores se corresponde con la noción geométrica de perpendicularidad.

Más concretamente,

$$u \cdot v = \|u\|_2 \|v\|_2 \cos(\widehat{uv}),$$

donde $\widehat{uv}$ es el ángulo que forman los vectores u y v.

Propiedades

Para $u, v, w \in \mathbb{R}^n$ *y* $a \in \mathbb{R}$*:*

- $u + v = v + u.$
- $u \cdot v = v \cdot u.$
- $u \cdot (v + w) = u \cdot v + u \cdot w.$
- $(a \cdot u) \cdot v = a \cdot (u \cdot v).$
- $v \cdot v = \|v\|_2^2.$
- $u \cdot v = 0$ *si y sólo si* u *y* v *son ortogonales.*
- $v \cdot u = v \cdot \mathrm{Proj}_v(u).$
- $\mathrm{Ort}_v(u) \cdot v = 0$

Definición

Dados $v_1, v_2, \ldots, v_k \in V$*, una* **combinación lineal** *de los* k *vectores anteriores es cualquier expresión de la forma*

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k,$$

con $a_1, a_2, \ldots, a_k \in K$*.*

Definición

Sean $v_1, v_2, \ldots, v_k \in V$*. Diremos que los vectores anteriores son* **linealmente independientes** *si se cumple que la ecuación*

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k = 0$$

tiene como única solución $a_1 = a_2 = \cdots = a_k = 0$*. En caso contrario, se dice que dichos vectores son* **linealmente dependientes***.*

Definición

*Una **matriz** de tamaño $m \times n$ sobre el cuerpo de los números reales $\mathbb{R}$ es un conjunto formado por $m \cdot n$ valores reales, $a_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$, distribuidos en m filas y n columnas de la siguiente forma:*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

*Escribimos $A \in \mathbb{R}^{m \times n}$, siendo $\mathbb{R}^{m \times n}$ el conjunto formado por todas las matrices $m \times n$. Asimismo, diremos que una matriz $A \in \mathbb{R}^{m \times n}$ es **cuadrada** si $m = n$, es decir, si tiene el mismo número de filas y de columnas.*

Sean $A, B \in \mathbb{R}^{m\times n}$. Podemos definir la **suma** y la **resta** $A \pm B$ como:

$$A \pm B = \begin{pmatrix} a_{1,1} \pm b_{1,1} & a_{1,2} \pm b_{1,2} & \cdots & a_{1,n} \pm b_{1,n} \\ a_{2,1} \pm b_{2,1} & a_{2,2} \pm b_{2,2} & \cdots & a_{2,n} \pm b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} \pm b_{m,1} & a_{m,2} \pm b_{m,2} & \cdots & a_{m,n} \pm b_{m,n} \end{pmatrix}$$

Asimismo, dado un escalar $\lambda \in \mathbb{R}$, definimos el **producto** $\lambda \cdot A$ como

$$\lambda A = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \cdots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \cdots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \cdots & \lambda a_{m,n} \end{pmatrix}$$

Por último, si $A \in \mathbb{R}^{m \times s}$ y $B \in \mathbb{R}^{s \times n}$, entonces definimos el **producto matricial** $A \cdot B \in \mathbb{R}^{m \times n}$ como

$$A \cdot B = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,n} \end{pmatrix}$$

donde $c_{i,j} = \sum_{k=1}^{s} a_{i,k} b_{k,j} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \cdots + a_{i,s} b_{s,j}$.

Nota

En general, el producto de matrices **NO** *es conmutativo. Por ejemplo:*

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -6 & 7 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ -3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -7 & 8 \end{pmatrix}$$

Definición

Sea $A \in \mathbb{R}^{n\times n}$. Diremos que A es **invertible** *o* **regular** *si $\exists B \in \mathbb{R}^{n\times n}$ tal que $A \cdot B = B \cdot A = I_n$. La matriz B recibe el nombre de* **matriz inversa** *y se denota por A^{-1}.*

Definición

Sea $A \in \mathbb{R}^{n\times n}$ *una matriz cuadrada de la forma*

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.$$

Definimos el **determinante** *de* A *de forma recurrente como sigue:*

- *Si* $n = 1$, $\det(A) = a_{1,1}$.
- *Para* $n > 1$,

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1,j} \det(A_{1,j}),$$

donde $A_{i,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ *es la matriz resultante de eliminar en* A *la fila* i *y la columna* j.

Teorema

El determinante de una matriz A puede calcularse de cualquiera de las siguientes formas:

- *Desarrollo por la fila i, $1 \leq i \leq n$:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}).$$

- *Desarrollo por la columna j, $1 \leq j \leq n$:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}).$$

donde $A_{i,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ es la matriz resultante de eliminar en A la fila i y la columna j.

Teorema

$\forall A, B \in \mathbb{R}^{n\times n}$, se cumple

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Teorema

$A \in \mathbb{R}^{n\times n}$ es regular si y sólo si $\det(A) \neq 0$.

Definición

El **rango** *de una matriz $A \in \mathbb{R}^{m\times n}$ es el número máximo de columnas que, como vectores, son linealmente independientes entre sí. Se denota por* $\operatorname{rank}(A)$.

Teorema

$A \in \mathbb{R}^{n\times n}$ es regular si y sólo si $\operatorname{rank}(A) = n$.

Definición

Sea $A \in \mathbb{R}^{m\times n}$*. La* **traspuesta** *de la matriz* A *se define como la matriz* $A' \in \mathbb{R}^{n\times m}$ *tal que* $A'(i,j) = A(j,i)$, $1 \leq i \leq n$, $1 \leq j \leq m$.

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ *regular. Entonces la inversa de* A, A^{-1}, *viene dada por*

$$A^{-1} = \frac{1}{\det(A)} \cdot \operatorname{adj}(A),$$

donde $\operatorname{adj}(A)$ *es la* **matriz adjunta** *de* A*, dada por* $\operatorname{adj}(A) = C'$, *con* $C(i,j) = (-1)^{i+j}|A_{i,j}|$, $1 \leq i,j \leq n$, *con* $|A_{i,j}| = \det(A_{i,j})$.

Definición

Diremos que una aplicación $f : \mathbb{R}^n \to \mathbb{R}^m$ *es una* **aplicación lineal** *si* $\forall u_1, u_2 \in \mathbb{R}^n$ *y* $\forall a, b \in \mathbb{R}$ *se cumple*

$$f(au_1 + bu_2) = af(u_1) + bf(u_2).$$

Ejemplos

1. *La aplicación* $f : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por* $f(x, y, z) = (x + y, 2z - x)$ *es lineal.*
2. *La aplicación* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $f(x, y) = (x, x^2 - y)$ *no es lineal.*
3. *La aplicación* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $f(x, y) = (x - y, x + 1)$ *no es lineal.*

Teorema

Toda aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^m$ *tiene una matriz que la representa, que además es única. Concretamente,* $\exists! M_f \in \mathbb{R}^{m \times n}$ *tal que* $\forall v \in \mathbb{R}^n$ *puede escribirse:*

$$f(v) = M_f v.$$

Ejemplo

Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ *dada por*

$$f(x, y, z) = \begin{pmatrix} x + y \\ y - z \\ 2x + y + z \\ 3z \end{pmatrix} \to M_f = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Teorema

Sea $f : \mathbb{R}^n \to \mathbb{R}^s$ y $g : \mathbb{R}^s \to \mathbb{R}^m$ aplicaciones lineales, con M_f y M_g las matrices que las representan, respectivamente. Entonces $g \circ f : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal, cuya matriz que la representa viene dada por $M_{g \circ f} = M_g M_f$.

Ejercicios

1. *Comprueba que el teorema se cumple para las dos ejemplos anteriores.*
2. *Probar el teorema para dos aplicaciones lineales cualesquiera. Es decir, probar lo siguiente:*
   - 2.1 *$g \circ f$ es una aplicación lineal si f y g lo son.*
   - 2.2 *La matriz que representa $g \circ f$, $M_{g \circ f}$, puede escribirse como el producto de las matrices que representan g y f, esto es, $M_g M_f$.*

Ejemplo

Un ejemplo clásico de aplicaciones lineales son las rotaciones. En el caso de $\mathbb{R}^2$ *vienen dadas por* $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ *definidas como:*

$$f_\theta(x, y) = \begin{pmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{pmatrix} \to M_{f_\theta} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

donde θ *representa el ángulo de rotación (en sentido antihorario). Por tanto, esta aplicación envía un vector* $(x, y) \in \mathbb{R}^2$ *a su correspondiente versión rotada* θ *grados:* $\big(\cos(\theta)x - \sin(\theta)y, \sin(\theta)x + \cos(\theta)y\big)$. *Asimismo, y como ya sabemos, podemos expresar la transformación matricialmente:*

$$f_\theta(x, y) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Ejemplo

Por ejemplo, la rotación de 30° *(o* $\pi/6$ *rad) viene dada por:*

$$f_{\pi/6} \equiv \begin{pmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Por otra parte, la rotación de 60° *(o* $\pi/3$ *rad) viene dada por:*

$$f_{\pi/3} \equiv \begin{pmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Por tanto, si componemos sendas rotaciones queda:

$$f_{\pi/3} \circ f_{\pi/6} \equiv \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Ejemplo

Cabe esperar que el resultado sea una rotación de 90° *(o* $\pi/2$ *rad) con respecto al original:*

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

cumpliéndose $(x, y) \cdot (-y, x) = -xy + yx = 0$, *luego son ortogonales.*

Ejercicio

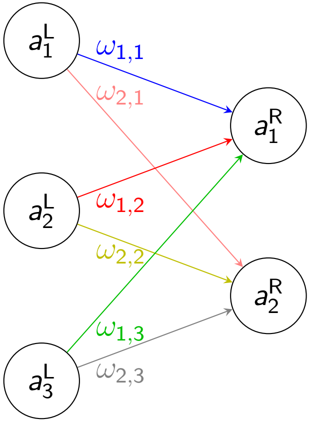
Utilizando las identidades trigonométricas que procedan, demuestra los siguientes enunciados:

(a) $\|f_\theta(v)\|_2 = \|v\|_2 \ \forall v \in \mathbb{R}^2, \forall \theta \in \mathbb{R}.$

(b) $f_\beta \circ f_\alpha = f_{\alpha+\beta} \ \forall \alpha, \beta \in \mathbb{R}.$

Motivación

Las matrices (y las aplicaciones lineales que representan) constituyen un elemento importante en la descripción de redes neuronales.



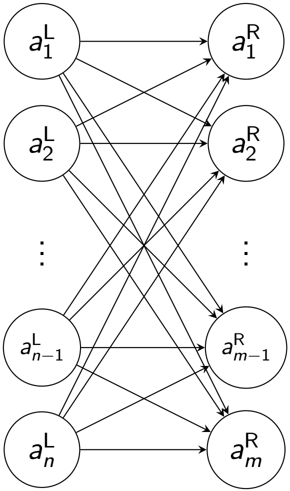
$$
\begin{aligned}
a_1^R &= f_1\left(\omega_{1,1}a_1^L + \omega_{1,2}a_2^L + \omega_{1,3}a_3^L + b_1\right),\\
a_2^R &= f_2\left(\omega_{2,1}a_1^L + \omega_{2,2}a_2^L + \omega_{2,3}a_3^L + b_2\right).
\end{aligned}
$$

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

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

Motivación

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



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

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

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

Autovectores y autovalores

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

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:

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

Diagonalización

- 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
  $$A^k = PD^kP^{-1}, \text{ donde } D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

- 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 $(A - \lambda_i I_n)v = \overrightarrow{\mathbf{0}}$, $1 \leq i \leq k$.

Universidad Internacional de Valencia

Ejemplo

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

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*) $\rightarrow$ PCA.
- 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. Además, $\exists P \in \mathbb{R}^{n\times n}$ ortogonal $(P' = P^{-1})$ tal que $P'AP = D$.

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 $k \in \{1, 2, \ldots, n\}$. El **menor principal dominante** *de orden k asociado a A, al que denotaremos por A_k, es $A_k = \det(A(1:k, 1:k))$; 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).*
- *$A_k > 0$, $\forall k \in \{1, 2, \ldots, n\}$ (respectivamente, $A_k \geq 0$).*
- *$\forall \lambda \in \mathbb{R}$ valor propio de A, $\lambda > 0$ (respectivamente, $\lambda \geq 0$).*

Álgebra tensorial

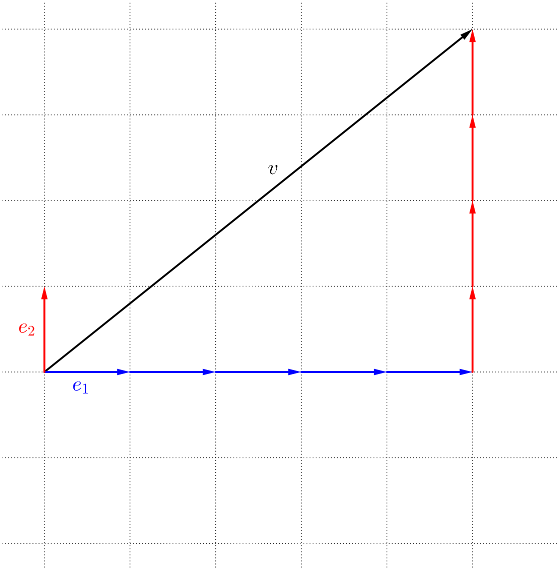
Definición

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 lineales 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



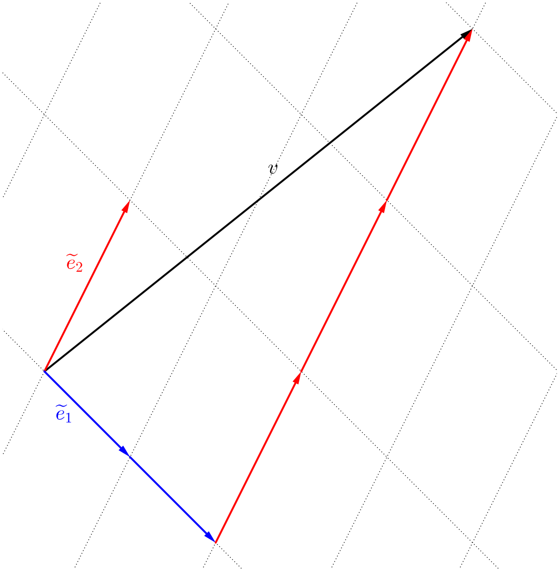
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\mathcal{B} = \{e_1, e_2\}$, donde:

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

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

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

Sistemas de coordenadas



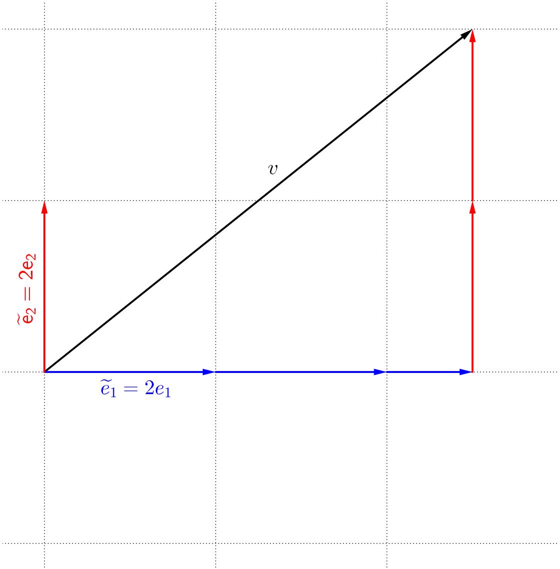
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\widetilde{\mathcal{B}} = \{\widetilde{e}_1, \widetilde{e}_2\}$, donde:

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

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

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

Sistemas de coordenadas



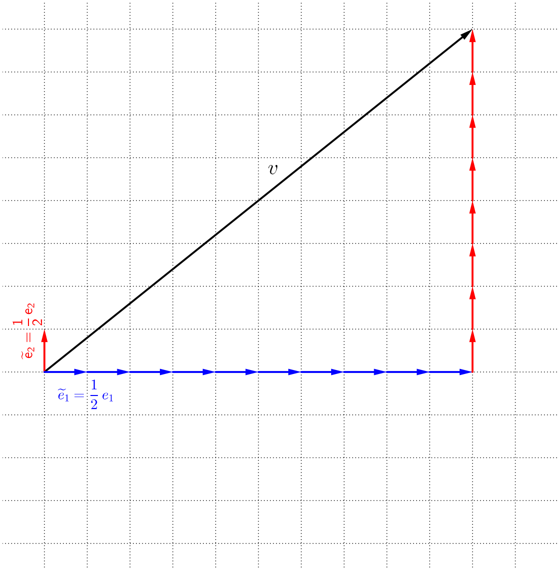
- Vector: $v = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$
- Base $\widetilde{\mathcal{B}} = \{\widetilde{e}_1 = 2e_1, \widetilde{e}_2 = 2e_1\}$, donde:

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

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

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

Sistemas de coordenadas



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

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

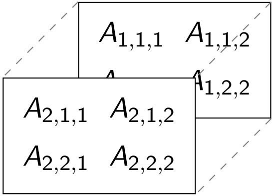
Sistemas de coordenadas

- 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, $T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}$ 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:* 2×132300.
3. *Imagen monocroma, con* 600 *píxeles de anchura y* 800 *píxeles de altura.*
   - *Rango:* 2.
   - *Dimensión:* 600×800.

Arrays multidimensionales

Universidad Internacional de Valencia

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:

victormanuel.campello@campusviu.es