BLOQUE 2. ÁLGEBRA LINEAL.
APLICACIONES LINEALES Y DIAGONALIZACIÓN(*)

- Aplicaciones lineales.
- Expresión matricial de una aplicación lineal.
- Diagonalización.

En contextos como Sistemas Dinámicos o procesos de cadenas de Markov es preciso conocer la forma general de una potencia o la exponencial de una matriz cuadrada ***A***, lo que no resulta fácil de obtener a no ser que la matriz sea diagonal o diagonal por bloques. Pero si ***A*** se puede expresar como producto de tres matrices en la forma ***A=PDP***$^{-1}$ con ***D*** diagonal

$$A = P\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}P^{-1}$$

entonces:

$$A^k = P\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}^k P^{-1} =$$

$$= P\begin{pmatrix} a^k & 0 & \cdots & 0 \\ 0 & b^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c^k \end{pmatrix}P^{-1}$$

En este tema se abordará la diagonalización de una matriz cuadrada ***A***, es decir será posible encontrar bajo determinadas circunstancias, si ***A=PDP***$^{-1}$ con ***D*** diagonal.

Para ello es necesario manejar los conceptos de aplicación lineal, matriz asociada a una aplicación lineal y vector y valor propio.

1. Aplicaciones lineales

Sea $f: IR^n \rightarrow IR^m$ una aplicación de IR^n en IR^m que a todo vector $\vec{x} \in IR^n$ asocia un vector $\vec{y} \in IR^m$. Se escribe entonces $\vec{y} = f(\vec{x})$.

Se dice que *f* es una *aplicación lineal* si se cumplen las dos condiciones siguientes:

$$f(\vec{x} + \vec{x}') = f(\vec{x}) + f(\vec{x}') \; \forall \vec{x}, \vec{x}' \in IR^n$$
$$f(\lambda \vec{x}) = \lambda f(\vec{x}) \; \forall \vec{x} \in IR^n, \forall \lambda \in IR$$

relaciones llamadas *axiomas de linealidad*.

De esta definición se deduce inmediatamente la siguiente caracterización de las aplicaciones lineales:

Es condición necesaria y suficiente para que una aplicación $f: IR^n \rightarrow IR^m$ sea una aplicación lineal que se cumpla:

$$\forall \vec{x}, \vec{x}' \in IR^n, \; \forall \alpha, \beta \in IR, \; f\left(\alpha \vec{x} + \beta \vec{x}'\right) = \alpha f\left(\vec{x}\right) + \beta f\left(\vec{x}'\right) \; .$$

La propiedad anterior se puede extender a una combinación lineal de cualquier número de sumandos, y se puede enunciar diciendo que una aplicación entre espacios vectoriales es una aplicación lineal si y sólo si transforma combinaciones lineales de vectores de IR^n en las combinaciones lineales de las imágenes en IR^m de esos vectores, con los mismos coeficientes.

Imagen de una aplicación lineal

Sea $f: IR^n \rightarrow IR^m$ una aplicación de IR^n en IR^m y sean $\vec{0}_{IR^n}$ y $\vec{0}_{IR^m}$ los elementos nulos de IR^n y IR^m respectivamente. Se cumplen las siguientes propiedades:

i) $f\left(\vec{0}_{IR^n}\right) = \vec{0}_{IR^m}$.

ii) $f\left(-\vec{x}\right) = -f\left(\vec{x}\right)$.

iii) La imagen por *f* de un subespacio vectorial de IR^n es un subespacio vectorial de IR^m. Esto es, si $U \subset IR^n$ es un subespacio vectorial de IR^n, entonces $f(U)$ es un subespacio vectorial de IR^m.
En particular, la *imagen de f, Im(f)* = $f(\mathrm{IR}^n)$ es un subespacio vectorial de IR^m.

$$Im(f) = \left\{\vec{y} = f(\vec{x}) \in IR^m \,/\, \vec{x} \in IR^n\right\}.$$

iv) Si $S = \left\{\vec{e}_1, \ldots, \vec{e}_r\right\}$ es un sistema de generadores de un subespacio vectorial U de IR^n, $f(S) = \left\{f\left(\vec{e}_1\right), \ldots, f\left(\vec{e}_r\right)\right\}$ es un sistema de generadores del subespacio vectorial $f(U)$ de IR^m.

Núcleo de una aplicación lineal

Sea $f : IR^n \rightarrow IR^m$ una aplicación de IR^n en IR^m. Se llama *núcleo de f* al conjunto, *Ker*(*f*), de los vectores de IR^n cuya imagen por *f* es el vector $\vec{0}_{IR^m}$:

$$Ker(f) = \left\{ \vec{x} \in IR^n \,/\, f(\vec{x}) = \vec{0}_{IR^m} \right\}.$$

El núcleo *Ker*(*f*) de una aplicación lineal nunca es vacío, ya que al menos contiene al vector $\vec{0}$ de IR^n. (Denotaremos con $\vec{0}$ tanto el vector nulo de IR^n como el de IR^m, distinguiéndose uno del otro por el contexto).

Se cumplen las siguientes propiedades:

i) El núcleo *Ker*(*f*) de toda aplicación lineal $f : IR^n \rightarrow IR^m$ es un subespacio vectorial de IR^n.

ii) Si $f : IR^n \rightarrow IR^m$ es una aplicación lineal del espacio vectorial IR^n en el espacio vectorial IR^m,

$$dim\ Ker(f) = dim\ \mathrm{IR}^n - dim\ Im(f) = n - dim\ Im(f)$$

o bien,

$$dim\ Ker(f) + dim\ Im(f) = n$$

A la dimensión de *Im*(*f*) se le suele llamar el *rango de la aplicación lineal f.*

2. Expresión matricial de una aplicación lineal

Sea $f : IR^n \rightarrow IR^m$ una aplicación lineal de IR^n en IR^m. Sean $B_1 = \left\{ \overrightarrow{u_1}, \ldots, \overrightarrow{u_n} \right\}$ una base de IR^n y $B_2 = \left\{ \overrightarrow{w_1}, \ldots, \overrightarrow{w_m} \right\}$ una base de IR^m. Por último consideremos las imágenes por *f* de los vectores de la base B_1:

$$f(\overrightarrow{u_1}) = \overrightarrow{a_1}, \ldots, f(\overrightarrow{u_n}) = \overrightarrow{a_n}$$

Estas imágenes son vectores de IR^m y, por tanto, son combinaciones lineales de los elementos de la base B_2.

$$\begin{array}{l} f(\overrightarrow{u_1}) = a_{11}\overrightarrow{w_1} + a_{12}\overrightarrow{w_2} + \cdots + a_{1m}\overrightarrow{w_m} \\ f(\overrightarrow{u_2}) = a_{21}\overrightarrow{w_1} + a_{22}\overrightarrow{w_2} + \cdots + a_{2m}\overrightarrow{w_m} \\ \vdots \\ f(\overrightarrow{u_n}) = a_{n1}\overrightarrow{w_1} + a_{n2}\overrightarrow{w_2} + \cdots + a_{nm}\overrightarrow{w_m} \end{array}$$

Cualquier vector $\vec{x} \in IR^n$ es combinación lineal de los vectores de la base B_1:

$$\vec{x} = x_1 \overrightarrow{u_1} + x_2 \overrightarrow{u_2} + \cdots + x_n \overrightarrow{u_n}$$

Por tanto:

$$\begin{aligned}
\vec{y} = f(\vec{x}) &= x_1 f(\overrightarrow{u_1}) + x_2 f(\overrightarrow{u_2}) + \cdots + x_n f(\overrightarrow{u_n}) = \\
&= x_1 (a_{11} \overrightarrow{w_1} + a_{12} \overrightarrow{w_2} + \cdots + a_{1m} \overrightarrow{w_m}) + \\
&\quad + x_2 (a_{21} \overrightarrow{w_1} + a_{22} \overrightarrow{w_2} + \cdots + a_{2m} \overrightarrow{w_m}) + \\
&\quad + \cdots\cdots\cdots + \\
&\quad + x_n (a_{n1} \overrightarrow{w_1} + a_{n2} \overrightarrow{w_2} + \cdots + a_{nm} \overrightarrow{w_m}) = \\
&= (x_1 a_{11} + x_2 a_{21} + \cdots + x_n a_{n1}) \overrightarrow{w_1} + \\
&\quad \cdots + (x_1 a_{1m} + x_2 a_{2m} + \cdots + x_n a_{nm}) \overrightarrow{w_m}
\end{aligned}$$

En consecuencia, las coordenadas de $\vec{y} = f(\vec{x}) \in IR^m$ respecto de B_2 son: $\vec{y} = (y_1, y_2, \cdots, y_m)$ con

$$\begin{aligned}
&y_1 = x_1 a_{11} + x_2 a_{21} + \cdots + x_n a_{n1} \\
&\cdots \\
&y_m = x_1 a_{1m} + x_2 a_{2m} + \cdots + x_n a_{nm}
\end{aligned}$$

Matricialmente:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o abreviadamente: ***Y=AX***.

A la matriz ***A*** se le llama *matriz asociada a la aplicación lineal f respecto de las bases B_1 y B_2 de IR^n y IR^m*. Esta matriz queda determinada por *f* de forma sencilla, ya que sus columnas son precisamente las coordenadas respecto de B_2 de las imágenes por *f* de los vectores de la base B_1.

Cambio de bases en una aplicación lineal

Sea $f : IR^n \to IR^m$ una aplicación lineal de IR^n en IR^m que tiene como matriz asociada respecto a las bases B_1 *y* B_2 *de* IR^n *y* IR^m, respectivamente, la matriz $\boldsymbol{A}$, expresándose la aplicación lineal f en la forma matricial $\boldsymbol{Y = AX}$.

Sea B'_1 otra base de IR^n y P la matriz cambio de base B'_1 a B_1, es decir, que entre la matriz columna de coordenadas $\boldsymbol{X}$ de un vector $\vec{x} \in IR^n$ en B_1 y la de las coordenadas $\boldsymbol{X'}$ del mismo vector en B'_1 se tiene la relación $\boldsymbol{X = P\,X'}$.

Sea B'_2 otra base de IR^m y $\boldsymbol{Q}$ la matriz cambio de base B'_2 a B_2, de modo que la relación entre las coordenadas de un vector en B_2 y sus correspondientes en B'_2 es $\boldsymbol{Y = QY'}$.

Respecto de las bases B'_1 y B'_2 de IR^n y IR^m, la aplicación lineal f tendrá asociada otra matriz $\boldsymbol{A'}$, es decir, respecto de estas bases, la aplicación lineal f se expresará así: $\boldsymbol{Y' = A'\,X'}$.

Buscamos la relación que pueda existir entre $\boldsymbol{A}$ y $\boldsymbol{A'}$. Para ello, sustituyendo en la relación $\boldsymbol{Y = AX}$ las matrices de coordenadas $\boldsymbol{X = P\,X'}$ e $\boldsymbol{Y = QY'}$, se tiene:

$$\boldsymbol{QY'=APX'}$$

y puesto que $\boldsymbol{Q}$ es regular, multiplicando a la izquierda por $\boldsymbol{Q}^{-1}$, se tiene:

$$\boldsymbol{Y'=Q^{-1}APX'}$$

por lo que la nueva matriz asociada a f en las bases B'_1 y B'_2 es:

$$\boldsymbol{A'= Q^{-1}AP}$$

En el caso de que f sea una aplicación lineal de IR^n en sí mismo, si se efectúa un cambio de la base B_1 a la base B'_1 en este espacio, entonces será $\boldsymbol{Q = P}$ **y** la matriz asociada a la aplicación lineal f en la nueva base será:

$$\boldsymbol{A'= P^{-1}AP}$$

3. Diagonalización

Puesto que una matriz $\boldsymbol{A}$ está asociada a una aplicación lineal f, se plantea la pregunta de si es posible representar f mediante una matriz más sencilla, diagonal. Habrá que determinar, por tanto, si existen bases adecuadas en los espacios de partida y de llegada respecto de las cuales la matriz asociada sea diagonal.

En este apartado se abordará el problema de la diagonalización de una matriz partiendo de los conceptos de valor propio y subespacio de vectores propios correspondiente y teniendo en cuenta la interpretación de las columnas de una matriz asociada a una aplicación lineal.

Valores y vectores propios

Sea $f : IR^n \to IR^n$ una aplicación lineal. Se dice que un vector $\vec{x} \neq \vec{0} \in IR^n$ es un *vector propio* o *autovector* de *f* si existe $\lambda \in IR$ tal que:

$$f(\vec{x}) = \lambda \vec{x}$$

Al escalar λ, que si existe es único, se le conoce como *valor propio* o *autovalor* asociado al vector propio $\vec{x}$.

Cada vector propio está asociado a un único valor propio, sin embargo, en general existirán distintos vectores propios asociados a un mismo valor propio.

Sea $\boldsymbol{A}$ la matriz que determina *f* en una cierta base *B* de IR^n, sea $\boldsymbol{I}$ la matriz identidad de orden *n*, y sea $\boldsymbol{X}$ la matriz columna de las coordenadas del vector $\vec{x}$ en la base *B*. Entonces la condición de vector propio se escribe:

$$\boldsymbol{AX} = \lambda \boldsymbol{I} \boldsymbol{X}$$

o equivalentemente:

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$$

Esta ecuación matricial equivale a un sistema lineal homogéneo de *n* ecuaciones en las *n* incógnitas que corresponden a las *n* coordenadas de $\vec{x}$. Los vectores propios asociados a λ serán las soluciones no triviales de este sistema, que admite alguna de estas soluciones si y sólo si:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}_{\text{n}}| = 0$$

o en forma desarrollada:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0$$

El determinante anterior genera al desarrollarlo un polinomio en λ de grado *n*, que se denomina *polinomio característico* de la matriz $\boldsymbol{A}$, y la ecuación obtenida al igualarlo a 0, se llama la *ecuación característica* de la aplicación lineal f o de la matriz $\boldsymbol{A}$.

Los vectores propios cumplen las siguientes propiedades:

i) Si $\vec{x}$ es un vector propio de la aplicación lineal *f*, asociado al valor propio λ, también lo es el vector $\mu\vec{x}$, asociado al mismo valor propio, donde μ es un escalar cualquiera no nulo.

ii) Si $\vec{x}$ y $\vec{x}'$ son dos vectores propios asociados al mismo valor propio λ, toda combinación lineal de dichos vectores también es un vector propio asociado al mismo autovalor puesto que:
$$f\left(\mu\vec{x}+\mu'\vec{x}'\right)=\mu f\left(\vec{x}\right)+\mu' f\left(\vec{x}'\right)=\mu\lambda\vec{x}+\mu'\lambda\vec{x}'=\lambda\left(\mu\vec{x}+\mu'\vec{x}'\right)$$

Como consecuencia de *i)* y *ii)*, el conjunto de los vectores propios asociados a un valor propio, constituye un subespacio vectorial de IR^n, llamado *subespacio propio* asociado a dicho valor propio.

iii) El polinomio característico de una matriz $\boldsymbol{A}$ asociada a una aplicación lineal *f* de IR^n en IR^n, resulta invariante ante un cambio de base en IR^n.

iv) Si $\overrightarrow{x_1}, \overrightarrow{x_2},\ldots,\overrightarrow{x_h}$ son *h* vectores propios asociados a *h* valores propios distintos $\lambda_1, \lambda_2,\ldots,\lambda_h$, entonces $\left\{\overrightarrow{x_1}, \overrightarrow{x_2},\ldots,\overrightarrow{x_h}\right\}$ es un sistema libre.

Diagonalización de una matriz cuadrada

- Sea $\boldsymbol{A}$ la matriz asociada a $f: IR^n \to IR^n$ respecto a una base *B* de IR^n. Supongamos que se obtienen *n* **autovalores, todos reales y distintos** y sean $\overrightarrow{x_1}, \overrightarrow{x_2},\ldots,\overrightarrow{x_n}$ autovectores asociados a dichos valores propios. Estos vectores forman una base *B*' de IR^n, respecto de la cual f tiene una matriz asociada $\boldsymbol{A}$' cuyas columnas son las coordenadas de $f\left(\overrightarrow{x_1}\right)=\lambda_1\overrightarrow{x_1},\ldots,f\left(\overrightarrow{x_n}\right)=\lambda_n\overrightarrow{x_n}$ referidas a la base *B*', por lo que $\boldsymbol{A}$' será una matriz diagonal

$$\boldsymbol{A}'=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

  Se dice entonces que la matriz $\boldsymbol{A}$ es *diagonalizable*, ya que existe una matriz diagonal $\boldsymbol{A}$' que representa la misma aplicación lineal *f* en cierta base. Si denotamos con $\boldsymbol{P}$ la matriz de cambio de base de *B*' a *B*, se tiene $\boldsymbol{A}' = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

  Si *B* es la base canónica de IR^n, las columnas de $\boldsymbol{P}$ son los n vectores propios $\overrightarrow{x_1}, \overrightarrow{x_2},\ldots,\overrightarrow{x_n}$.

- Una matriz $\boldsymbol{A}$ cuadrada de orden n es diagonalizable si y sólo si todos los valores propios λ_i (i = 1…, h) de $\boldsymbol{A}$ son reales, y para todo λ_i con orden de multiplicidad α_i, el rango de la matriz $\boldsymbol{A} - \lambda_i \boldsymbol{I}$ es igual a $n - \alpha_i$.

Bibliografía

- Barbolla, R. y Sanz, P. (1998). *Álgebra lineal y teoría de matrices*. Ed. Prentice Hall.
- Caballero, R. E., Calderón, S. y Galache, T. P. (2000). *Matemáticas aplicadas a la economía y a la empresa. 434 ejercicios resueltos y comentados*. Ed. Pirámide.
- Grossman, S. I. (1997). *Álgebra lineal*. Ed. McGraw-Hill.
- Hernández, E. (1999). *Álgebra y geometría*. Ed. Addison-Wesley/U.A.M.
- Kolman, B. (1999). *Álgebra lineal con aplicaciones y Matlab*. Ed. Prentice Hill.
- Martínez Salas. (1992). *Elementos de matemáticas*. Ed. Lex Nova.
- Sanz, P., Vázquez, F. J. y Ortega, P. (1998). *Álgebra lineal. Cuestiones, ejercicios y tratamiento en Derive*. Ed. Prentice Hall.

(*) En este tema se ha utilizado como fuente, tanto para la estructuración de los contenidos como en la inclusión de algunos párrafos, el siguiente libro: *Álgebra lineal y teoría de matrices*. R. Barbolla y P. Sanz. Ed. Prentice Hall. (1998).