Índice general

Capítulo 3

Transformaciones lineales

El estudio de funciones (transformaciones) juega un rol fundamental en la ciencia y en la ingeniería. Una función es una operación que transforma un objeto en otro. Estudiaremos funciones entre espacios vectoriales, es decir, funciones que transforman vectores de un espacio vectorial en vectores de otro o del mismo espacio vectorial.

En este capítulo nos centraremos en una clase particular de funciones, llamadas *lineales*. Veremos que son simples y que hay una vasta teoría que nos permite comprenderlas profundamente. No es este el caso de las funciones en general, por esto mismo, las funciones lineales juegan un rol fundamental en la ingeniería, pues son práticamente las únicas que podemos representar totalmente en un computador, y las únicas que podemos "resolver" de manera directa. En algunos de los ejemplos en el capítulo comenzaremos a entender la utilidad de las transformaciones lineales.

El trabajo con las transformaciones lineales se simplifica si las representamos mediante matrices y esto podremos hacerlo porque conocemos bases de los espacios vectoriales con los que trabajaremos. Pero además, de nuevo gracias a los conocimientos que ya tenemos sobre bases y dimensión de espacios vectoriales, podremos formalizar el trabajo con matrices y comprenderemos mejor los métodos de solución de sistemas de ecuaciones lineales.

Comencemos entonces a entender qué es una transformación lineal.

Definición 3.1. Sean U y V espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$. Sea $T : U \longrightarrow V$ una función (transformación) entre U y V, es decir, T es una función que transforma vectores de U en vectores de V. T es una **transformación o aplicación lineal** si y solo si

1. $\forall u, w \in U \ : \ T(u + w) = T(u) + T(w)$,
2. $\forall \alpha \in \mathbb{K} \ : \ \forall u \in U \ : \ T(\alpha\, u) = \alpha T(u)$.

En otras palabras, una transformación lineal es una función que respeta la suma (T de una suma de vectores es la suma de los resultados de aplicar T a cada vector) y el producto por escalar. Nota que si, por ejemplo, U es la recta que pasa por el origen y tiene a $\vec{r} \in \mathbb{R}^3$ como director, es decir,

$$U = \langle \vec{r} \rangle$$

entonces una transformación lineal $T : U \to V$ es tal que $T(\alpha\vec{r}) = \alpha T(\vec{r})$, es decir, el resultado de aplicar T a vectores de U es un vector en

$$\langle T(\vec{r})\rangle \, .$$

Si $T(\vec{r}) = \vec{s}$, $\vec{s} \in \mathbb{R}^3$, la transformación lineal T modifica vectores en la recta U en vectores en otra recta, en la recta $\langle T(\vec{r})\rangle = \langle \vec{s}\rangle$.

Analicemos si las siguientes funciones entre espacios vectoriales son funciones lineales.

Ejemplo 3.2.

1. *La transformación $f : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $\forall\, x \in \mathbb{R} \ : \ f(x) = a$ con $a \in \mathbb{R}$ es lineal si y solo si $a = 0$.*
   Si $a \neq 0$, f no satisface ninguna de las dos propiedades necesarias para ser transformación lineal pues si $x, y, \alpha \in \mathbb{R}$, entonces

   $$\begin{aligned} f(x+y) = a &\neq f(x) + f(y) = 2a, \\ f(\alpha x) = a &\neq \alpha f(x) = \alpha a. \end{aligned}$$

   *En general la transformación $T : U \longrightarrow V$ tal que $\forall\, u \in U \ : \ T(u) = \theta_V$ es una transformación lineal, se denomina **transformación nula** y la denotaremos por Θ.*

2. *La transformación $g : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $\forall\, x \in \mathbb{R} \ : \ g(x) = ax + b$ con $a, b \in \mathbb{R}$ es lineal si y solo si $b = 0$.*
   Para que g sea lineal debe ocurrir que para todo par de números reales x, y, se cumpla la igualdad $g(x+y) = g(x) + g(y)$. Como

   $$g(x+y) = a(x+y) + b, \qquad g(x) + g(y) = ax + b + ay + b = a(x+y) + 2b,$$

   se cumple que $g(x+y) = g(x) + g(y)$ si y solo si $b = 0$. Si $b = 0$ y $\alpha \in \mathbb{R}$, g también respeta el producto por escalar pues

   $$g(\alpha x) = a(\alpha x), \qquad \alpha g(x) = \alpha(ax).$$

3. *La transformación $L : \mathbb{R}^2 \to \mathbb{R}$ dada por $L((x,y)^T) = 3x - 2y$ es lineal.*
   Utilicemos su definición para calcular el resultado de aplicar T a algunos vectores de $\mathbb{R}^2$.
   $L((1,-1)^T) = 3 \cdot 1 - 2 \cdot (-1) = 3 + 2 = 5$,
   $L((0,10)^T) = 3 \cdot 0 - 2 \cdot 10 = 0 - 20 = -20$,
   $L((3a, a+b)^T) = 3 \cdot 3a - 2(a+b) = 9a - 2a - 2b = 7a - 2b$.
   Demostremos que L es lineal. Si $u = (x,y)^T$ y $v = (a,b)^T$, entonces
   $L(u+v) = L((x+a, y+b)^T) = 3(x+a) - 2(y+b) = 3x + 3a - 2y - 2b = 3x - 2y + 3a - 2b = L(u) + L(v)$.
   Por otro lado, y si $\alpha \in \mathbb{R}$,
   $L(\alpha u) = L((\alpha x, \alpha y)^T) = 3\alpha x - 2\alpha y = \alpha(3x - 2y) = \alpha L(u)$.

4. *La transformación* $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *con*

$$L\left((x,y)^T\right) = (2y, 3, 4x)^T$$

no es lineal. Si $(x,y)^T, (u,v)^T \in \mathbb{R}^2$, *entonces*

$$L((x,y)^T) + L((u,v)^T) = (2y,3,4x)^T + (2v,3,4u)^T = (2(y+v),6,4(x+u))^T,$$

mientras que

$$L((x,y)^T + (u,v)^T) = L((x+u,y+v)^T) = (2(y+v),3,4(x+u))^T \neq L((x,y)^T) + L((u,v)^T).$$

5. *La transformación* $\frac{d}{dt}$ *que a cada* $f \in C^1([a,\, b])$ *le hace corresponder su derivada es una transformación lineal entre* $U = C^1([a,\, b])$ *y* $V = C([a,\, b])$.
Esto es cierto porque para cada par de funciones diferenciales f *y* g *se cumple que* $(f+g)' = f' + g'$ *y para todo escalar* $\alpha \in \mathbb{R}$ *y función diferenciable* f *se tiene que* $(\alpha f)' = \alpha f'$.

6. *La transformación* $I : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ *definida por* $I(p) = \displaystyle\int_0^1 p(x)\ dx$ *es lineal.*
En efecto, si $p, q \in \mathcal{P}(\mathbb{R})$ *y* $\lambda \in \mathbb{R}$,

$$\begin{aligned} I(p+q) &= \int_0^1 (p+q)(x)\ dx \\ &= \int_0^1 (p(x)+q(x))\ dx \\ &= \int_0^1 p(x)\ dx + \int_0^1 q(x)\ dx \\ &= I(p) + I(q), \end{aligned}$$

$$\begin{aligned} I(\lambda p) &= \int_0^1 (\lambda p)(x)\ dx \\ &= \int_0^1 \lambda p(x)\ dx \\ &= \lambda \int_0^1 p(x)\ dx \\ &= \lambda I(p). \end{aligned}$$

7. *Veamos un ejemplo donde el cuerpo sobre el que se definen los espacios vectoriales* U *y* V *es importante al momento de decidir si una transformación* $T : U \to V$ *es o no lineal.*
La transformación $T : \mathbb{C} \longrightarrow \mathbb{C}^4$ *tal que*

$$T\,(a + i\,b) = \begin{pmatrix} a \\ b \\ -b \\ a \end{pmatrix}$$

es una transformación lineal si consideramos $\mathbb{K} = \mathbb{R}$ *y no lo es si* $\mathbb{K} = \mathbb{C}$.
Sean $a + i\,b, x + i\,y \in \mathbb{C}$. *Entonces*

$$T(a+i\,b+x+i\,y) = \begin{pmatrix} a+x \\ b+y \\ -(b+y) \\ a+x \end{pmatrix} = \begin{pmatrix} a \\ b \\ -b \\ a \end{pmatrix} + \begin{pmatrix} x \\ y \\ -y \\ x \end{pmatrix} = T(a+i\,b) + T(x+i\,y).$$

Consideremos ahora a $\mathbb{C}$ *y* $\mathbb{C}^4$ *como espacios vectoriales sobre* $\mathbb{R}$.
Para que T *sea una transformación lineal debe cumplir además que para todo* $\alpha \in \mathbb{R}$, $a+i\,b \in \mathbb{C}$, $T(\alpha(a+i\,b)) = \alpha T(a+i\,b)$,

$$T(\alpha(a+i\,b)) = T(\alpha\,a + i\,\alpha\,b) = \begin{pmatrix} \alpha\,a \\ \alpha\,b \\ -\alpha\,b \\ \alpha\,a \end{pmatrix} = \alpha \begin{pmatrix} a \\ b \\ -b \\ a \end{pmatrix} = \alpha T(a+i\,b)$$

y, por tanto, T *es una transformación lineal si* $\mathbb{K} = \mathbb{R}$.
Consideremos a $\mathbb{C}$ *y* $\mathbb{C}^4$ *como espacios vectoriales complejos. Sea* $\alpha = u + i\,v \in \mathbb{C}$

$$T(\alpha(a+i\,b)) = T((u+i\,v)(a+i\,b)) = T((ua-vb)+i\,(ub+va)) = \begin{pmatrix} ua-vb \\ ub+va \\ -(ub+va) \\ ua-vb \end{pmatrix},$$

mientras que

$$\begin{aligned} \alpha T(a+i\,b) &= (u+i\,v)T(a+i\,b), \\ &= (u+i\,v)\begin{pmatrix} a \\ b \\ -b \\ a \end{pmatrix}, \\ &= \begin{pmatrix} au+i\,av \\ bu+i\,bv \\ -bu-i\,bv \\ au+i\,av \end{pmatrix} \neq T(\alpha(a+i\,b)). \end{aligned}$$

Por tanto, T *no es una transformación lineal si* $\mathbb{K} = \mathbb{C}$.

La transformación lineal del siguiente ejemplo es una función entre un espacio U *cualquiera y un s.e.v. de él.*

8. *Sean* U *un espacio vectorial sobre un cuerpo* $\mathbb{K}$ *y* $U_1, U_2 \subseteq U$ *subespacios vectoriales de* U *tales que* $U = U_1 \oplus U_2$. *Esto significa que para cada* $u \in U$ *existen un único vector* $u_1 \in U_1$ *y un único vector* $u_2 \in U_2$ *de modo que* $u = u_1 + u_2$. *Utilizaremos esta descomposición de* U *para definir las aplicaciones* $T : U \to U_1$ *y* $L : U \to U_2$,

$$T(u) = u_1, \quad L(u) = u_2,$$

es decir $T(u)$ es la parte de u en U_1, mientras que $L(u)$ es su parte en U_2. T es la proyección de U sobre U_1 (a lo largo de U_2) y L es la proyección de U sobre U_2 (a lo largo de U_1). Demostremos que T es una función lineal. Consideremos $u, w \in U$. Como $U = U_1 \oplus U_2$, existen únicos $u_1, w_1 \in U_1$ y $u_2, w_2 \in U_2$ de modo que $u = u_1 + u_2$ y $w = w_1 + w_2$. Como $u+w = (u_1+u_2)+(w_1+w_2) = (u_1+w_1)+(u_2+w_2)$ con $u_1+w_1 \in U_1$ y $u_2+w_2 \in U_2$ y U_1+U_2 es directa, los vectores u_1+w_1 y u_2+w_2 son los únicos vectores en U_1 y U_2 que permiten escribir a $u+w$ como suma de un vector en U_1 y uno en U_2. Entonces $T(u+w) = u_1+w_1 = T(u)+T(w)$. Por otro lado, si $\alpha \in \mathbb{K}$, Como $\alpha u = \alpha(u_1 + u_2) = \alpha u_1 + \alpha u_2$ con $\alpha u_1 \in U_1$ y $\alpha u_2 \in U_2$ y $U_1 + U_2$ es directa, los vectores αu_1 y αu_2 son los únicos vectores en U_1 y U_2 que permiten escribir a αu como suma de un vector en U_1 y uno en U_2. Entonces $T(\alpha u) = \alpha u_1 = \alpha T(u)$.

En el ejemplo anterior definimos la transformación nula entre $\mathbb{K}$-espacios vectoriales U y V.

Otra transformación (o función) importante es la **función identidad** entre un $\mathbb{K}$-e.v. V y él mismo. Esta transformación se representa mediante $\mathrm{id}\colon V \to V$ y es tal que para cada $u \in V$, $\mathrm{id}(u) = u$. Lo invitamos a demostrar que ella es, efectivamente, una función lineal.

Veamos algunos ejemplos de transformaciones que no son lineales.

Ejemplo 3.3.

- *$F : \mathbb{R}^2 \to \mathbb{R}^2$ definida por*

$$F\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+1 \\ y+1 \end{pmatrix}$$

no es lineal. Veamóslo: Si $u = \begin{pmatrix} x \\ y \end{pmatrix}$ y $v = \begin{pmatrix} a \\ b \end{pmatrix}$ son vectores cualesquiera de $\mathbb{R}^2$, entonces

$$F(u+v) = \begin{pmatrix} x+a+1 \\ y+b+1 \end{pmatrix},$$

mientras que

$$F(u) + F(v) = \begin{pmatrix} x+1 \\ y+1 \end{pmatrix} + \begin{pmatrix} a+1 \\ b+1 \end{pmatrix} = \begin{pmatrix} x+a+2 \\ y+b+2 \end{pmatrix} \neq F(u+v).$$

- *La transformación $E : \mathbb{R}^3 \to \mathbb{R}^2$ definida por*

$$E\left((x,y,z)^T\right) = (xy, xz, zx)^T$$

no es lineal. Veamos por qué. Sean $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ y $\alpha \in \mathbb{R}$, entonces

$$E(\alpha u) = \left(\alpha^2 xy, \alpha^2 xz, \alpha^2 zx\right)^T = \alpha^2 (xy, xz, zx)^T = \alpha^2 E(u)$$

El vector $E(\alpha u) = \alpha E(u)$ solamente si $\alpha = 1$ o $\alpha = 0$.

Dos aplicaciones lineales entre e.v. U y V son iguales si y solo si resultan los mismos valores al ser evaluadas en los mismos vectores.

Definición 3.4. Dos aplicaciones lineales $T : U \longrightarrow V$, $L : U \longrightarrow V$ con U, V espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$ son **iguales** si y solo si

$$\forall\, u \in U \ : \ T(u) = L(u).$$

En el siguiente lema se demuestran propiedades importantes en el trabajo con las transformaciones lineales, éstas no son propiedades que pueda tener cualquier función entre e.v., sino que son propiedades que solo podemos garantizar las cumplan las aplicaciones lineales.

Lema 3.5. Sean U y V e.v. sobre un mismo cuerpo $\mathbb{K}$ y sea $T : U \longrightarrow V$ una transformación lineal. Entonces se cumple que

1. T transforma al vector nulo de U en el vector nulo de V, es decir, $T(\theta_U) = \theta_V$.
2. El resultado de aplicar T al inverso aditivo de $u \in U$ es el inverso aditivo de $T(u)$.
3. Si $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ y $u_1, u_2, \ldots, u_n \in U$, entonces

$$T(\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n) = \alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots + \alpha_n T(u_n).$$

Demostración. Demostremos cada una de las propiedades anteriores.

1. Como para todo $u \in U$ se cumple que $0u = \theta_U$, entonces $T(\theta_U) = T(0u) = 0T(u) = \theta_V$.
2. Dado que en todo espacio vectorial se cumple que el inverso aditivo de un vector u es el vector $(-1)u$ se tiene que $T(-u) = T((-1)u) = (-1)T(u) = -T(u)$.
3. Esta propiedad puede demostrarse por inducción en el número de vectores en la suma tomando como base de inducción las propiedades que tiene T por ser una aplicación lineal.
   Una consecuencia de esta propiedad es la siguiente: si $\{u_1, u_2, \ldots, u_m\}$ es una base de U y $T : U \to W$ es una transformación lineal entre U y W, entonces, si conocemos los valores de $T(u_1), T(u_2), \ldots, T(u_m)$ podemos determinar $T(u)$ para cada $u \in U$ pues

$$u \in U \Rightarrow \exists\, \beta_1, \ldots, \beta_m \in \mathbb{K} \ : \ u = \sum_{i=1}^{m} \beta_i u_i \Rightarrow$$

$$T(u) = T\left(\sum_{i=1}^{m} \beta_i u_i\right) = \sum_{i=1}^{m} \beta_i \underbrace{T(u_i)}_{\text{conocidos}} .$$

   Si $\{u_1, \ldots, u_m\}$ es una base de U, entonces una transformación lineal $T : U \longrightarrow V$ está unívocamente determinada por sus valores en la base de U.
   Dos aplicaciones lineales $T : U \longrightarrow V,\ L : U \longrightarrow V$ son iguales si y solo si dan el mismo resultado al aplicarse a los elementos en una base de U.

□

Observación 3.6. *Los dos primeros resultados del lema anterior nos permiten identificar si una transformación* ***no*** *es lineal.*

*Si $T : U \to V$ no es tal que $T(\theta_U) = \theta_V$, entonces T no es lineal. La primera transformación en el ejemplo 3.3 transforma al vector nulo de $\mathbb{R}^2$ en el vector $\begin{pmatrix}1\\1\end{pmatrix}$. Podemos concluir entonces que ella **no** es una transformación lineal.*

Nota que la segunda de las transformaciones anteriores sí transforma al nulo de $\mathbb{R}^3$ en el nulo de $\mathbb{R}^2$. Sin embargo, ella no es lineal. Ella no convierte al inverso aditivo de cualquier $u \in \mathbb{R}^3$ en el inverso aditivo de $E(u)$. Por ejemplo,

$$E\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) = \begin{pmatrix}1\\1\\1\end{pmatrix},$$

y

$$E\left(-\begin{pmatrix}1\\1\\1\end{pmatrix}\right) \neq -E\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right).$$

*Sin embargo, que una transformación T convierta al nulo del espacio de partida en el nulo del espacio de llegada y al inverso aditivo de un vector u en el inverso aditivo de $T(u)$ no significa que ella sea lineal. Nota que $\sin : \mathbb{R} \to \mathbb{R}$ satisface estas dos condiciones, pero **no** es una transformación lineal.*

Veamos ahora un ejemplo en el que utilizamos el hecho de que una transformación lineal T se define completamente a través de los resultados de aplicar T a una base del espacio de partida.

Ejemplo 3.7. *Sean $T, L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ aplicaciones lineales tales que*

$$T((1,0,0)^T) = (1,1,1)^T,\ T((0,1,0)^T) = (0,2,1)^T,\ T((0,0,1)^T) = (0,1,2)^T$$

y

$$L((1,0,1)^T) = (1,2,3)^T,\ L((0,1,1)^T) = (0,3,3)^T,\ L((1,1,1)^T) = (1,3,3)^T.$$

¿Están T y L bien definidas?

Los conjuntos $\mathcal{B}_1 = \{(1,0,0)^T, (0,1,0)^T, (0,0,1)^T\}$ y $\mathcal{B}_2 = \{(1,0,1)^T, (0,1,1)^T, (1,1,1)^T\}$ son bases de $\mathbb{R}^3$, por tanto, T y L están completamente definidas.

Encontremos la expresión para $T((x,y,z)^T)$ para cualquier elemento $(x,y,z)^T \in \mathbb{R}^3$. Sea $(x,y,z)^T \in \mathbb{R}^3$, entonces

$$\begin{aligned}
T((x,y,z)^T) &= T(x(1,0,0)^T + y(0,1,0)^T + z(0,0,1)^T),\\
&= xT((1,0,0)^T) + yT((0,1,0)^T) + zT((0,0,1)^T),\\
&= x(1,1,1)^T + y(0,2,1)^T + z(0,1,2)^T\\
&= (x, x+2y+z, x+y+2z)^T.
\end{aligned}$$

es decir:

$$T((x,y,z)^T) = (x, x+2y+z, x+y+2z)^T.$$

Ahora es sencillo calcular T en los vectores de la base de $\mathbb{R}^3$ que se utilizó para definir a L. Los valores de T en los elementos de $\mathcal{B}_2$,

$$\begin{aligned}
T((1,0,1)^T) &= (1,2,3)^T = L((1,0,1)^T),\\
T((0,1,1)^T) &= (0,3,3)^T = L((0,1,1)^T),\\
T((1,1,1)^T) &= (1,4,4)^T \neq L((1,1,1)^T),
\end{aligned}$$

por tanto, T y L no son iguales pues no resultan en los mismos vectores al aplicarse a la misma base del espacio de partida.

Asegurémonos de esto. Calculemos $L((x,y,z)^T)$. Sea $(x,y,z)^T \in \mathbb{R}^3$, entonces existen α, β y $\gamma \in \mathbb{R}$ tales que:

$$\begin{aligned}
(x,y,z)^T &= \alpha(1,0,1)^T + \beta(0,1,1)^T + \gamma(1,1,1)^T,\\
&= (\alpha+\gamma, \beta+\gamma, \alpha+\beta+\gamma)^T
\end{aligned}$$

y por tanto se tiene el siguiente sistema de ecuaciones:

$$\begin{aligned}
x &= \alpha+\gamma,\\
y &= \beta+\gamma,\\
z &= \alpha+\beta+\gamma.
\end{aligned}$$

De la primera ecuación obtenemos $\alpha = x-\gamma$ y de la segunda, $\beta = y-\gamma$. Reemplazando α y β en la tercera ecuación se obtiene que $\gamma = x+y-z$ y, por tanto, $\alpha = z-y$ y $\beta = z-x$. En consecuencia todo vector $(x,y,z)^T$ es la siguiente combinación lineal de los elementos de $\mathcal{B}_2$

$$(x,y,z)^T = (z-y)(1,0,1)^T + (z-x)(0,1,1)^T + (x+y-z)(1,1,1)^T,$$

luego:

$$\begin{aligned}
L((x,y,z)^T) &= L((z-y)(1,0,1)^T + (z-x)(0,1,1)^T + (x+y-z)(1,1,1)^T),\\
&= (z-y)L((1,0,1)^T) + (z-x)L((0,1,1)^T) + (x+y-z)L((1,1,1)^T),\\
&= (z-y)(1,2,3)^T + (z-x)(0,3,3)^T + (x+y-z)(1,3,3)^T\\
&= (x, 2z-y, 3z)^T.
\end{aligned}$$

es decir:

$$L((x,y,z)^T) = (x, 2z-y, 3z)^T.$$

Es evidente que T y L no son iguales.