CONTENIDO DE LA PARTE DEL ÁLGEBRA LINEAL

Versión preliminar

D. Yakubovich (2018)

En algunas ocasiones, damos referencias a las páginas del libro E. Hernández, Álgebra y geometría, segunda edición, Addison-Wesley/UAM, 1994.

1. El repaso del espacio euclideo $\mathbb{R}^n$

En general identificamos cada punto $(x_1, x_2, \dots, x_n)^t \in \mathbb{R}^n$ con el vector $\vec{x}$ que va desde el origen a dicho punto: $\vec{x} = (x_1, x_2, \dots, x_n)^t$.

$$(x_1, x_2, \dots, x_n)^t = (y_1, y_2, \dots, y_n)^t \Leftrightarrow x_i = y_i, \ \forall i = 1, \dots, n.$$

Aquí ^{t} es el símbolo de transposición. En expresiones matriciales con vectores, los entendemos siempre como **vectores columna** (aunque a veces escribiremos por comodidad $(x_1, x_2, \dots, x_n)$ en vez de $(x_1, x_2, \dots, x_n)^t$). Se define la base canónica en $\mathbb{R}^n$ como el conjunto de vectores:

$$\vec{e}_1 = (1, 0, \dots, 0)^t, \ \ \vec{e}_2 = (0, 1, \dots, 0)^t, \ \ \dots, \qquad \vec{e}_n = (0, 0, \dots, 1)^t.$$

Operaciones básicas:

1. Suma:
$$(x_1, x_2, \dots, x_n)^t + (y_1, y_2, \dots, y_n)^t = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)^t$$

2. Producto por un escalar:
$$\lambda(x_1, x_2, \dots, x_n)^t = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)^t, \ \forall \lambda \in \mathbb{R}$$

Propiedades:

1. Asociativa:
$$(\mu\lambda)(x_1, x_2, \dots, x_n)^t = \mu(\lambda(x_1, x_2, \dots, x_n)^t), \ \forall \lambda, \mu \in \mathbb{R}$$

2. Distributiva 1:
$$(\mu + \lambda)(x_1, x_2, \dots, x_n)^t = \mu(x_1, x_2, \dots, x_n)^t + \lambda(x_1, x_2, \dots, x_n)^t, \ \forall \lambda, \mu \in \mathbb{R}$$

3. Distributiva 2:
$$\mu\Big((x_1, x_2, \dots, x_n)^t + (y_1, y_2, \dots, y_n)^t\Big) = \mu(x_1, x_2, \dots, x_n)^t + \mu(y_1, y_2, \dots, y_n)^t, \ \forall \mu \in \mathbb{R}$$

4. Elemento 0:

$$0 \cdot (x_1, x_2, \ldots, x_n)^t = (0, 0, \ldots, 0)^t, \qquad \lambda \cdot (0, 0, \ldots, 0)^t = (0, 0, \ldots, 0)^t$$

5. Neutro del producto:

$$1 \cdot (x_1, x_2, \ldots, x_n^t) = (x_1, x_2, \ldots, x_n)^t$$

(y otras)

PRODUCTOS ENTRE VECTORES: ESCALAR Y VECTORIAL

Producto escalar:

$$< \vec{x}, \vec{y} >=< (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) >= x_1\, y_1 + x_2\, y_2 + \cdots + x_n\, y_n = \sum_{i=1}^{n} x_i\, y_i.$$

Observación 1.1. *Recordemos que* $\langle \vec{x}, \vec{x} \rangle \geq 0,\ \langle \vec{x}, \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$.

Definición 1.1. *Se define la longitud de un vector $\vec{x}$ como su norma o módulo y se denota por*

$$\|\vec{x}\| = \sqrt{< \vec{x}, \vec{x} >} = \sqrt{\sum_{i=1}^{n} x_i^2}\,.$$

Proposición 1.1.

$$< \vec{x}, \vec{y} >= \|\vec{x}\|\ \|\vec{y}\| \cos(\alpha).$$

Como consecuencia tenemos:

1. la desigualdad de Cauchy: $\quad \|\vec{x}\, \vec{y}\| \leq \|\vec{x}\|\ \|\vec{y}\|$
2. la desigualdad triangular: $\quad \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$
3. $\vec{x}, \vec{y}$ son vectores ortogonales ($\vec{x} \perp \vec{y}$) si y sólo si $< \vec{x}, \vec{y} >= 0$.

Definición 1.2. *La* **distancia** *entre dos puntos $\vec{x}$, $\vec{x}$ de $\mathbb{R}^n$ es la norma de su diferencia, es decir,*

$$\mathrm{dist}(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|.$$

La norma en $\mathbb{R}^n$ verifica propiedades similares al valor absoluto en $\mathbb{R}$:

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|\,, \qquad \big|\, \|\vec{x}\| - \|\vec{y}\| \,\big| \leq \|\vec{x} - \vec{y}\|\,.$$

De hecho, la norma es igual al valor absoluto si $n = 1$.

2. Sistemas lineales. El método de Gauss y la forma escalonada

p.10 - 21

Estudiamos un sistema lineal de la forma

$$\left\{ \begin{array}{ccccc} a_{11}x_1 + a_{12}x_1 + & \ldots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_1 + & \ldots & +a_{2n}x_n & = & b_2 \\ \ldots & \ldots & & \ldots & \ldots \\ a_{m1}x_1 + a_{m2}x_1 + & \ldots & +a_{mn}x_n & = & b_m \end{array} \right. \tag{1}$$

que en forma matricial se escribe como

$$A\vec{x} = \vec{b}. \tag{2}$$

Son m ecuaciones lineales con n incógnitas. La matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

es de tamaño $m \times n$, $\vec{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ es un vector de *incógnitas*, y $\vec{b} = (b_1, \dots, b_m)^t \in \mathbb{R}^m$ es el vector de *los datos*.

<u>Def:</u> La matriz $(A \,|\, b)$ ampliada del sistema. Tiene m filas y $n+1$ columnas.

<u>El método de Gauss</u> consiste en lo siguiente.

A la matriz ampliada del sistema se aplican sucesivamente *operaciones elementales con filas*. Son las siguientes.

1) Multiplicar una fila por un número real no nulo;

2) Intercambiar dos filas (es decir, dos ecuaciones);

3) Sumar o restar un múltipli de una ecuación a otra.

En cada paso, se obtiene un sistema equivalente al sistema inicial.

<u>Def:</u> Se dice que una matriz *está en forma escalonada* si

1) Cada peldaño tien altura 1;

2) Debajo de la escalera todos los elementos de la matriz son cero;

3) En cada esquina de un peldaño aparece el número 1;

4) Toda columna que contiene un pivote tiene todos los demás elementos nulos.

<u>Definiciones:</u>

- Sistema incompatible;
- Sistema compatible determinado;
- Sistema compatible indeterminado.

- Sistemas homogéneas (es decir, cuando el vector de los datos $\vec{b}$ es nulo).

p. 17 <u>Teorema de Rouche-Frobenius, 1^a forma</u> (en términos de pivotes de A, $(A \,|\, b)$ y del número n de incógnitas.

p. 35-49 <u>Def:</u> Funciones (o aplicaciones) lineales $f : \mathbb{R}^n \to \mathbb{R}^m$.

<u>Teorema</u>

1) Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal. Definimos $\vec{a}_j = f(\vec{e}_j) \in \mathbb{R}^m$, y sea $A = (\vec{a}_1 \,|\, \vec{a}_2 \,|\, \dots \,|\, \vec{a}_n)$ la matriz, cuyas columnas son $\vec{a}_1, \vec{a}_2, \vec{a}_n$. Entonces

$$f(\vec{x}) = A \cdot \vec{x} \quad \text{para todo } \vec{x} \in \mathbb{R}^n.$$

2) Recíprocamente, dada cualquier matriz $A \in M_{m\times n}(\mathbb{R})$, la aplicación $f(\vec{x}) = A \cdot \vec{x}$, $f : \mathbb{R}^n \to \mathbb{R}^m$ es lineal.

<u>Def:</u> Funciones inyectivas, sobreyectivas, biyectivas. Ejemplos.

<u>Def:</u> El dominio y la imagen de una función.

Def: La composición de dos funciones.

Teorema: 1) La composición de dos funciones inyectivas es inyectiva;
2) La composición de dos funciones sobreyectivas es sobreyectiva;
3) La composición de dos funciones biyectivas es biyectiva.
Def: La aplicación identidad $\mathrm{id} : A \to A$, donde A es un conjunto.

Def: Función inversa.
Def: Restricción de una función f a un subconjunto de su dominio de definición.

- Las funciones arcsen, arccos, arctg como inversas a restricciones de funciones $\mathrm{sen}, \cos, \mathrm{tg}$.
- La matriz identidad I_n como la matriz de la aplicación identidad $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$.
- La multiplicación de matrices y su relación con la composición de las correspondientes aplicaciones lineales.
- Operaciones con aplicaciones lineales y operaciones con matrices. Propiedades.

Def: Sistemas de vectores linealmente independientes. La relación con aplicaciones lineales inyectivas.

Teorema: Una familia de vectores $(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$ en $\mathbb{R}^m$ es linealmente independiente si y solo si la forma escalonada $\bar{A}$ de la matriz $A = (\vec{a}_1 \,|\, \vec{a}_2 \,|\, \ldots \,|\, \vec{a}_n)$, cuyas columnas son estos vectores, tiene pivote en cada columna (es decir, tiene n columnas).

p. 24-25 Def: 1) Por el rango de una familia de vectores $(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$ se se entiende el mayor número de ellos, que son linealmente independientes.
2) Por el rango de una matriz se entiende el rango de sus vectores columna.
Teorema: El rango de una matriz es igual al número de pivotes de su forma escalonada.

Teorema de Rouche-Frobenius, 2^a forma (en términos de los rangos de las matrices A y $(A \,|\, b)$ y del número n de incógnitas del sistema lineal).
Propiedad: Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es biyectiva, entonces $m = n$.
p. 50-60 Def.: Una matriz A $n \times n$ es *invertible* si la correspondiente aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^n$ es biyectiva.
En este caso, la aplicación inversa $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ es también lineal. La matriz inversa A^{-1} se define como la matriz de f^{-1}. Como $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$, se obtiene que

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

Teorema: Sea A una matriz $n \times n$. Entonces las siguientes propiedades son equivalentes:
1) A es invertible;
2) $r(A) = n$;
3) Las columnas de A son linealmente independientes.

Matrices elementales y su relación con operaciones elementales con filas.
Hay 3 tipos de operaciones elementales con filas, y les corresponden 3 tipos de matrices elementales. Son cuadradas e invertibles.

Método de Gauss de inversión de una matriz cuadrada.
Matriz transpuesta. Propiedades.

3. Determinantes

Solo se definen determinantes para matrices cuadradas.

Def Una permutación de n elementos es una biyección del conjunto $\{1, 2, \dots, n\}$ sobre sí mismo.

Para definir una permutación $\alpha : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$, basta escribir la sucesión $[\alpha(1), \alpha(2), \dots, \alpha(n)]$ de todos los valores de α. Hay $n!$ permutaciones de n elementos.

Ejemplo: Por la permutación $\alpha = [312]$ se entiende la función $\alpha : \{1, 2, 3\} \to \{1, 2, 3\}$, definida por $\alpha(1) = 3, \alpha(2) = 1, \alpha(3) = 2$. Hay $3! = 6$ permutaciones de 3 elementos.

Notación: Denotamos por S_n el conjunto de todas las permutaciones de n elementos.

Def: El número de inversiones $\mathrm{Inv}(\alpha)$ de una permutación α.

Ejemplos: $\mathrm{Inv}([1, 2, 3, 4]) = 0$, $\mathrm{Inv}([1, 3, 2, 4]) = 1$, $\mathrm{Inv}([4, 3, 2, 1]) = 6$

Def: Dada una permutación $\alpha \in S_n$, su *signatura* $\mathrm{Sig}(\alpha)$ se define como

$$\mathrm{Sig}(\alpha) = (-1)^{\mathrm{Inv}(\alpha)}.$$

Se dice que α es *par* si $\mathrm{Inv}(\alpha)$ es par (es decir, $\mathrm{Sig}(\alpha) = 1$). Se dice que α es *inpar* si $\mathrm{Inv}(\alpha)$ es inpar. En este caso, $\mathrm{Sig}(\alpha) = -1$.

La relación con determinantes en los ejemplos de matrices 2×2 y 3×3.

p. 123 La Definición del determinante:

$$|A| = \det(A) = \sum_{\alpha \in S_n} \mathrm{Sig}(\alpha) a_{1\alpha(1)} a_{2\alpha(2)} \dots a_{n\alpha(n)}.$$

Notación: Si A es una matriz de tamaño $n \times n$, $A = (\vec{v}_1 \,|\, \vec{v}_2 \,|\, \dots \,|\, \vec{v}_n)$, usamos también la notación

$$\det A = \det(\vec{v}_1 \,|\, \vec{v}_2 \,|\, \dots \,|\, \vec{v}_n).$$

Propiedad: Sea $\alpha \in S_n$ una permutación arbitaria de n elementos y sea $\beta \in S_n$ una transposición, lo que quiere decir que hay unos índices $\ell \neq m$, $\ell, m \in \{1, 2, \dots, n\}$, tales que β permuta estos dos elementos: $\beta(\ell) = m$, $\beta(m) = \ell$, y $\beta(k) = k$ para los demás números entre 1 y n. Entonces

$$\mathrm{Sig}(\beta \circ \alpha) = -\mathrm{Sig}(\alpha).$$

p. 72 Propiedades de determinantes (el comportamiento cuando se intercambian 2 columnas, cuando una columna se multiplica por una constante, descomposición por $k$ésima columna, linearidad en cada columna, etc.)

Teorema: Para dos matrices A, B de tamaño $n \times n$,

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

Cálculo de determinantes con el método de Gauss (es decir, mediante la reducción a la forma escalonada).

Def: *La matriz de cofactores* $\mathrm{cof}(A)$ de una matriz cuadrada A.

Teorema: Si $A \in M_{n\times n}(\mathbb{R})$ y $C = \mathrm{cof}(A)$, entonces

$$AC^t = \det(A) \cdot I_n.$$

Corolario: Si A es invertible, entonces

$$A^{-1} = \frac{1}{|A|}\big(\mathrm{cof}(A)\big)^t.$$

La regla de Cramer para la resolución de sistemas cuadrados determinados.

Binomio de Newton.

4. Sistemas generadores y bases

Consideramos una familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ de n vectores en $\mathbb{R}^m$.

Def 1: Ya hemos definido antes, qué familias de vectores son *linealmente independientes.*

Criterio (para ser l.i.): Sea $A = (\vec{v}_1 \,|\, \vec{v}_2 \,|\, \dots \,|\, \vec{v}_n)$ la matriz, cuyas columnas son los vectores de nuestra familia. Sea $\overline{A}$ la forma escalonada de A. Entonces son equivalentes las propiedades:

1) La familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ es linealmente independiente;

2) La aplicación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$, que corresponde a la matriz A, es inyectiva;

3) $\overline{A}$ tiene pivote en cada columna (es decir, tiene n pivotes).

Def 2: Una familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ de n vectores en $\mathbb{R}^m$ se dice que es un *sistema generador* si todo vector $\vec{w}$ en $\mathbb{R}^m$ tiene al menos una representación

$$\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n.$$

Criterio (para ser un sistema generador): Definimos las matriz A y $\overline{A}$ como antes. Entonces son equivalentes las propiedades:

1) La familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ es un sistema generador;

2) La aplicación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$, que corresponde a la matriz A, es sobreyectiva;

3) $\overline{A}$ tiene pivote en cada fila (es decir, tiene m pivotes).

Def 3: Una familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ de n vectores en $\mathbb{R}^m$ se dice que es *una base* de $\mathbb{R}^m$ si es linealmente independiente y un sistema generador.

Aplicando los dos criterios anteriores, obtenemos lo siguiente.

Criterio (de una base): Son equivalentes las propiecades:

1) La familia $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ es una base;

2) La aplicación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$, que corresponde a la matriz A, es biyectiva;

3) $\overline{A}$ tiene pivote en cada fila y en cada fila.

Corolario Una familia de m vectores en $\mathbb{R}^n$ puede ser una base sólo en el caso cuando $m = n$.

Por supuesto, si $m = n$, esto no implica todavía que nuestra familia es una base. La correspondiente matriz A tiene que ser invertible.

Dada una base $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)$ de m vectores en $\mathbb{R}^m$, todo vector $\vec{w}$ en $\mathbb{R}_m$ tiene una única representación

$$\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_m\vec{v}_m.$$

Se dice que la sucesión de m números $(x_1, x_2, \dots, x_m)$ son *coordenadas del vector $\vec{w}$ en la base $\beta = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)$.*

p. 179 **El producto vectorial en $\mathbb{R}^3$** se puede definir como un determinante simbólico:

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \stackrel{\text{def}}{=} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \vec{e}_3 , \qquad \vec{x}, \vec{y} \in \mathbb{R}^3 .$$

Observación 4.1. *El vector $\vec{x} \times \vec{y}$ verifica:*

1. $\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \, \|\vec{y}\| \sin(\alpha) =$ *área del paralelogramo formado por ambos vectores*

2. *$\vec{x} \times \vec{y}$ es un vector perpendicular a cada uno de ellos en el sentido del avance del tornillo.*

Más propiedades del producto vectorial. Relación con los determinantes 3×3.

Bases ortonormales

Def.: Dados n vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ en $\mathbb{R}^n$, se dice que son *una base ortonormal* si $vecv_j$ es ortogonal a $vecv_k$ para todo par de índices $j \neq k$ y $\|\vec{v}_j\| = 1$ para todo j.

En otras palabras, se tiene que cumplir

$$< \vec{v}_j, \vec{v}_k >= \delta_{jk},$$

donde

$$\delta_{jk} = \begin{cases} 0 & \text{si } j \neq k, \\ 1 & \text{si } j = k. \end{cases}$$

Propiedades:

1) Toda base ortonormal es una base;

2) Las coordenadas de un vector $\vec{w}$ en una base ortonormal $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ se expresan mediante la simple fórmula $x_j =< \vec{w}, \vec{v}_j >$. En otras palabras,

$$\vec{w} = \sum_{k=1}^{n} x_k \vec{v}_k \qquad \Longrightarrow \qquad x_j =< \vec{w}, \vec{v}_j >,\ j = 1, 2, \dots, n.$$

3) Dados dos vectores cualesquiera $\vec{a} = \sum_{k=1}^{n} \alpha_k \vec{v}_k$, $\vec{b} = \sum_{k=1}^{n} \beta_k \vec{v}_k$, su producto escalar se expresa como

$$< \vec{a}, \vec{b} >= \sum_{k=1}^{n} \alpha_k \beta_k.$$

Distintas definiciones del rango de una matriz

Teorema Sea A cualquier matriz rectangular.

1) El rango de A coincide con el tamaño máximo de su menor no nulo.

2) El rango de A por columnas (que es cómo lo hemos definido) coincide con el rango de A por filas (el mayor número de filas linealmente independientes).