Álgebra Lineal (R211 - CE9) 2024

3.1 Los cuatro espacios fundamentales

Importante: sólo por este párrafo trabajaremos exclusivamente con ev de dimensión finita.

Además, para no errarle, consideraremos todos los vectores como columnas. Si en algún momento no dan las dimensiones, tener en cuenta que se toma el transpuesto.

Revisaremos muchos conceptos que conocemos hasta ahora, los relacionaremos e introduciremos algunos nuevos para darle una nueva mirada a todo lo aprendido.

Para fijar ideas, consideraremos:

- V F-ev con $dim_F V = n$, $B_V = \{v_1, \ldots, v_n\}$ base de V. Sabemos que existe un isomorfismo $V \simeq F^n$ tal que $v \to [v]_{B_V} = (\alpha_1, \ldots, \alpha_n)$ si $v = \alpha_1 v_1 + \ldots \alpha_n v_n$.
- W F-ev con $dim_F W = m$, $B_W = \{w_1, \ldots, w_m\}$ base de W. Sabemos que existe un isomorfismo $W \simeq F^m$ tal que $w \to [w]_{B_W} = (\beta_1, \ldots, \beta_m)$ si $w = \beta_1 w_1 + \ldots \beta_m w_m$.
- Como $L(V, W) \simeq F^{m\times n}$, si $T \in L(V, W)$ tenemos que $[T]_{B_V B_W} \in F^{m\times n}$ es tal que

$$[T]_{B_V B_W}[v]_{B_V} = [Tv]_{B_W}.$$

Escribamos ahora

$$\begin{aligned} A =& [T]_{B_V B_W} = (a_{ij}) \in \mathbb{F}^{m\times n}, \\ \overline{x} =& [v]_{B_V} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n, \text{ y} \\ b =& [Tv]_{B_W} = (\beta_1, \ldots, \beta_m) \in \mathbb{F}^m. \end{aligned}$$

Entonces,

- Para hallar $ker(T)$ (o equivalentemente, chequear si T es mono), resolvemos $[T]_{B_V B_W}[v]_{B_V} = (0, \ldots, 0)^t$, es decir,

  $$A\overline{x} = \overline{0}.$$

  Y para esto aplicamos la resolución de sistemas homogéneos que aprendimos antes.
- Para calcular $Im(T)$ (y de paso chequear si T es epi), buscamos todos los $b = (\beta_1, \ldots, \beta_m) \in \mathbb{F}^m$ tal que $[T]_{B_V B_W}[v]_{B_V} = (\beta_1, \ldots, \beta_m)^t$ tiene solución, es decir, tal que

  $$A\overline{x} = b$$

  tiene solución, y para esto aplicamos la resolución de sistemas generales que aprendimos antes.

El espacio columna de A

Recordemos un poquito cómo construimos la notación matricial de un sistema de ecuaciones y revisemos más detenidamente lo que mencionamos recién. Consideremos el siguiente sistema de ecuaciones lineales con m ecuaciones y n incógnitas:

$$(\mathrm{S}) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

Si ponemos $A = (a_{ij})$, $\overline{x} = (x_1, \ldots, x_n)^t$ y $b = (b_1, \ldots, b_m)^t$, el sistema es $A\overline{x} = b$. Pero esto no es lo que queremos observar. Miremos nuevamente:

$$\begin{array}{|c|cc|c|ccc|c|cc|c|}
a_{11} & x_1 & + & a_{12} & x_2 & + \ \ldots & + & a_{1n} & x_n & = & b_1, \\
a_{21} & x_1 & + & a_{22} & x_2 & + \ \ldots & + & a_{2n} & x_n & = & b_2, \\
\vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots & \vdots & & \vdots \\
a_{m1} & x_1 & + & a_{m2} & x_2 & + \ \ldots & + & a_{mn} & x_n & = & b_m. \\
\hline
\uparrow & \uparrow & & \uparrow & \uparrow & & & \uparrow & \uparrow & & \uparrow \\
col_1(A) & x_1 & + & col_2(A) & x_2 & + \ \ldots & + & col_n(A) & x_n & = & b.
\end{array}$$

Es decir, b es una combinación lineal de las columnas de A. Y el conjunto de todas las combinaciones lineales de las columnas de A es el span de las columnas de A.

Así, para calcular $Im(T)$ (y de paso chequear si T es epi), buscamos todos los $b \in \mathbb{F}^m$ tales que $A\overline{x} = b$ tiene solución, esto es todos los $b \in < \{col_1(A), \ldots, col_n(A)\} >$. Esto último nos dice que $Im(T) =< \{col_1(A), \ldots, col_n(A)\} >$. Se los tomé en el parcial, se acuerdan?

MORALEJA: El espacio generado por las columnas de A es RE importante. Tiene nombre y todo:

Definición 1 *Dada una matriz $A \in \mathbb{F}^{m\times n}$ llamamos espacio columna de A al conjunto*

$$C(A) :=< \{col_1(A), \ldots, col_n(A)\} > .$$

$C(A)$ es un sev de $\mathbb{F}^m$.

Con esta definición sigue que $A\overline{x} = b$ es compatible sii $b \in C(A)$.

El espacio nulo de A

Con la notación anterior, para hallar $ker(T)$ (o equivalentemente, chequear si T es mono), tenemos que ver si $A\overline{x} = \overline{0}$ tiene sólo la solución trivial, es decir, ver si las columnas de A son li. Podemos decir algunas cosas más:

- El sistema $A\overline{x} = \overline{0}$ siempre tiene solución, siempre es compatible. En términos de transformaciones lineales esto nos dice que $T\overline{0}_V = \overline{0}_W$.

- Si A es cuadrada e invertible, la única solución es la trivial. En términos de transformaciones lineales, tenemos que $dimV = dimW$, T mono y luego iso.

- Si A no es invertible (aún en el caso que A sea cuadrada, hay soluciones no triviales, lo que implica que T no será monomorfismo.

Definición 2 *Dada una matriz $A \in \mathbb{F}^{m\times n}$ llamamos espacio nulo de A al conjunto*

$$N(A) = \{\overline{x} \in \mathbb{F}^n : A\overline{x} = \overline{0}\}.$$

$N(A)$ es un sev de $\mathbb{F}^n$.

Así, $\overline{x} \in N(A)$ sii $v \in ker(T)$ donde $[v]_{B_V} = \overline{x}$.

Escalonando A

En resumen, hemos definido dos espacios asociados a una matriz $C(A)$ y $N(A)$ que tienen interpretación a nivel de transformaciones lineales como los espacios imagen y núcleo asociados a ella. Todo siempre que tengamos bases fijas para dominio y codominio. El siguiente cuadro es el que tenemos que tener en mente:

$$\begin{array}{ccccccc} \mathbb{F}^n & \longleftrightarrow & V, & \qquad\qquad & \mathbb{F}^m & \longleftrightarrow & W, \\ & & & & & & \\ N(A) & \longleftrightarrow & ker(T), & & C(A) & \longleftrightarrow & Im(T). \end{array}$$

Pero más aún, como ya hemos visto, para resolver sistemas recurrimos a escalonar estas matrices. Veamos cómo se traduce ésto en términos de transformaciones lineales.

Volvamos a nuestro sistema y recordemos las operaciones entre ecuaciones:

$$\text{(S)} \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

- ESCALAMIENTO (la llamamos Tipo 1): multiplicar la i-ésima ecuación por un escalar no nulo $\alpha \in \mathbb{F}$:

$$\text{(S')} \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \vdots \\ \alpha a_{i1}x_1 + \alpha a_{i2}x_2 + \cdots + \alpha a_{in}x_n = \alpha b_i, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

- ELIMINACIÓN (la llamamos Tipo 2): sumar a la i-ésima ecuación un múltiplo por un escalar no nulo $\alpha \in \mathbb{F}$ de la k-ésima ecuación:

$$
\text{(S')} \begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\
\vdots \\
(a_{i1} + \alpha a_{k1})x_1 + (a_{i2} + \alpha a_{k2})x_2 + \cdots + (a_{in} + \alpha a_{kn})x_n = (b_i + \alpha b_k), \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.
\end{cases}
$$

- INTERCAMBIO (la llamamos Tipo 3): intercambiar la i-ésima ecuación por la k-ésima:

$$
\text{(S')} \begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\
\vdots \\
\alpha a_{k1}x_1 + \alpha a_{k2}x_2 + \cdots + \alpha a_{kn}x_n = \alpha b_k, \\
\vdots \\
\alpha a_{i1}x_1 + \alpha a_{i2}x_2 + \cdots + \alpha a_{in}x_n = \alpha b_i, \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.
\end{cases}
$$

Efectuar cualquiera de estas operaciones produce un sistema de ecuaciones equivalentes (recordar que esto significa que el conjunto solución es el mismo).

A nivel matricial, esto lo implementamos a partir de las operaciones elementales de filas. Dado un sistema (S) $A\overline{x} = b$, sabemos que las operaciones elementales por fila e pueden describirse a partir de matrices elementales E de modo que el sistema (S') $e(A)x = e(b)$ es equivalente. Para hallar las matrices E simplemente efectuamos la operación a la matriz identidad: $E = e(Id_m)$. Tenemos así que:

- ESCALAMIENTO:

$$
E = \begin{array}{r} \\ \\ \\ \text{fila } i \\ \\ \\ \end{array} \begin{pmatrix}
1 & 0 & \cdots & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \alpha & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & \cdots & 1
\end{pmatrix}.
$$

- ELIMINACIÓN:

$$E = \begin{array}{cc} & \begin{matrix} & & & \text{col i} & & \text{col } k & & \end{matrix} \\ \begin{matrix} \\ \\ \\ \text{fila } i \\ \\ \text{fila } k \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \end{array}.$$

- INTERCAMBIO:

$$E = \begin{array}{cc} & \begin{matrix} & & & \text{col i} & & \text{col } k & & \end{matrix} \\ \begin{matrix} \\ \\ \\ \text{fila } i \\ \\ \text{fila } k \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} \end{array}.$$

Todas estas matrices E son invertibles. Entonces si definimos $S \in L(W)$ tal que $[S]_{B_W} = E$, resulta que S es un automorfismo de W.

Si partimos de una matriz A y **usamos sólo operaciones de eliminación**, podemos llevarla a una matriz U en forma de escalón (algoritmo de eliminación de Gauss), luego aplicamos más operaciones de eliminación para llevarla a una matriz R en forma de escalón reducida (aquí es cuando el algoritmo se llama de Gauss-Jordan).

Supongamos que llevamos a cabo la eliminación de *Gauss* efectuando las operaciones de eliminación dadas por las matrices $E_1, \dots, E_k$. Llamemos directamente $E = E_1 \dots E_k$ y sea S el automorfismo resultante. Tenemos que $U = EA$, luego hemos obtenido una nueva transformación lineal $T' = S \circ T \in L(V, W)$.

Las matrices A, U y R son de tamaño $m \times n$. Hemos visto que la cantidad de pivots que obtenemos en U más la cantidad de variables libres nos dan la cantidad de columnas de A. Es decir, ya conocíamos el teorema de la dimensión!! En efecto:

- Las columnas de U que contienen pivots son li. Esto se interpreta fácilmente pensando que cada columna de U que contiene un pivot aporta una componente más, de modo que será li con las anteriores. La cantidad de columnas pivots es entonces la dimensión del espacio $C(U)$, que deberá ser la misma dimensión que $C(A)$ puesto que los conjuntos solución de ambos sistemas son iguales.

- Más aún, como S es un automorfismo, las columnas de A correspondientes a las columnas de U que contienen pivots, también son li.

- Ahora bien, tenemos que $ran(T) = dim(Im(T)) = dim(C(A)) = dim(C(U)) = \#$pivots.

- Además, la cantidad de variables libres que tenemos nos indica la dimensión del kernel de T, pues para hallarlo debíamos resolver el sistema homogéneo $A\overline{x} = \overline{0}$. Es decir, $\#$variables libres $= nul(T)$.

- Finalmente, puesto que $n = \#$columnas de $A = \#$pivots $+ \#$variables libres, resulta que $dimV = ran(T) + nul(T)$.

Definición 3 *Dada una matriz $A \in \mathbb{F}^{m\times n}$ llamamos rango de A al entero $ran(A) = dim(C(A))$. Si $r = ran(A)$, entonces $dim(N(A)) = n - r$.*

Los espacios que faltan

Definición 4 *Dada una matriz $A \in \mathbb{F}^{m\times n}$ llamamos:*

- *espacio fila de A al espacio generado por las filas de A, que es lo mismo que el espacio columna de A^t. El espacio fila $C(A^t)$ es un sev de $\mathbb{F}^n$,*
- *espacio nulo a izquierda de A al espacio nulo de A^t. El espacio nulo a izquierda $N(A)^t$ es un sev de $\mathbb{F}^m$.*

El espacio nulo a izquierda recibe su nombre de la siguiente observación: si planteamos la ecuación matricial $A^t\overline{y} = \overline{0}$ que lo describe, y la transponemos, tenemos que $\overline{y}^tA = \overline{0}^t$, es decir, estamos resolviendo un sistema homogéneo para el cual la incógnita multiplica a izquierda a la matriz A, es decir, la anula a izquierda.

Las dimensiones resultan: $dim(C(A^t)) = dim(C(A)) = ran(A) = r$, $dim(N(A)) = n-r$, $dim(N(A^t)) = m - r$.

Veamos algunos ejemplos:

Ejemplos 1 *Calculemos los cuatro espacios fundamentales (y sus dimensiones) asociados a la matriz A dada por:*

1.
$$A = \begin{pmatrix} 2 & 0 & 3 & -1 \\ 1 & -2 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}.$$

Escalonando esta matriz aplicando sólo operaciones de eliminación obtenemos una matriz U como la siguiente (recordar que la forma de escalón de una matriz no es única, sí lo es la reducida):

$$U = \begin{pmatrix} 2 & 0 & 3 & -1 \\ 0 & -2 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{5}{4} & \frac{3}{4} \end{pmatrix}.$$

U tiene pivots en las columnas 1, 2 y 3, que son li. Luego, el espacio columna de A está generado por esas mismas tres columnas que son li. Pero más aún, dado que $C(A)$ es sev de $\mathbb{R}^3$, si tenemos tres generadores li (una base), debe ser todo $\mathbb{R}^3$. De paso, sigue que el rango de A es 3, y decimos que el rango es total.

Para hallar $N(A)$ debemos resolver $Ax = 0$, pero ya que tenemos U, sigamos reduciendo utilizando operaciones de eliminación hasta llegar a R, la forma de escalón reducida:

$$R = \begin{pmatrix} 1 & 0 & 0 & -\frac{7}{5} \\ 0 & 1 & 0 & -\frac{7}{5} \\ 0 & 0 & 1 & \frac{3}{5} \end{pmatrix},$$

de donde vemos claramente que $N(A) = \langle (\frac{7}{5}, \frac{7}{5}, -\frac{3}{5}, 1) \rangle$, además $dimN(A) = 4 - 3 = 1$.

Para hallar $C(A^t)$, puesto que $dimC(A^t) = dimC(A) = ranA = 3$, el espacio fila de A está generado por las tres filas de A: $C(A^t) = \langle \{(2, 0, 3, -1), (1, -2, 1, 0), (-1, 1, 0, 1)\} \rangle$.

Para hallar $N(A^t)$ observamos que $dimN(A^t) = 3 - 3 = 0$, luego sigue fácilmente que $N(A^t) = \{\overline{0}\}$.

2.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ -1 & 0 & 1 \end{pmatrix}.$$

Eliminando obtenemos

$$U = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Y la FER resulta

$$R = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

U tiene pivots en las primera y segunda columnas, luego $C(A)$ tiene por base sus columnas primera y segunda, de modo que $ranA = 2$: $C(A) = \langle \{(1, 0, -1), (1, 2, 0)\} \rangle$. De R obtenemos rápidamente $N(A) = \langle (1, -1, 1) \rangle$, que ya sabíamos que tiene dimensión 1. El espacio fila está generado por las dos primeras filas de A y el nulo a izquierda hay que calcularlo como ejercicio, también sabemos que tiene dimensión 1.

Factorización LU

Consideremos en este párrafo una matriz cuadrada A. Usando sólo operaciones de eliminación para obtener ceros debajo de la diagonal resulta en una matriz U, como hemos visto: $U = E_1 \dots E_k A$. Como cada E_k es invertible, tenemos definida la matriz $L = (E_1 \dots E_k)^{-1} = E_k^{-1} \dots E_1^{-1}$, de modo tal que $A = LU$. La particularidad de estas matrices es que L es triangular inferior y U es triangular superior.

- U es triangular superior por construcción.

- L es triangular inferior puesto que es el producto de triangulares inferiores: en efecto, cada uno de sus factores es una matriz triangular inferior pues es la inversa de una matriz elemental que es triangular inferior (al ser la matriz asociada a la eliminación de un elemento debajo de la diagonal).

Esta factorización de A como producto de una matriz triangular inferior por una matriz triangular superior se conoce como **factorización LU** y es muy conocida. Una de sus múltiples aplicaciones es para resolver sistemas: si queremos resolver $A\overline{x} = b$, resolvemos primero $L\overline{y} = b$ y luego $U\overline{x} = \overline{y}$. Es claramente más sencillo resolver un sistema cuya matriz asociada es triangular.

Ortogonalidad en $\mathbb{R}^n$

Recordemos la definición de producto escalar en $\mathbb{R}^n$: para $u, v \in \mathbb{R}^n$, con $u = (u_1, \ldots, u_n)^t$, $v = (v_1, \ldots, v_n)^t$ (y consideraremos todo el tiempo vectores columna, salvo que en algún momento debamos hacer alguna transposición para que las dimensiones no nos traigan problemas), se define

$$u \times v = u_1 v_1 + \cdots + u_n v_n.$$

De acuerdo a la notación matricial con la que contamos, podemos escribir de forma similar $u \times v = u^t v$. Observemos que $u \times v$ es un escalar real.

Decimos que u y v son ortogonales y denotamos $u \perp v$ si $u \times v = 0$.

El único vector que es ortogonal a todos es el vector nulo: $\overline{0} \perp v$ para todo $v \in \mathbb{R}^n$. Además, el vector nulo es el único ortogonal a si mismo: si $v \perp v$ entonces $v = \overline{0}$.

Si U y V son subconjuntos de $\mathbb{R}^n$, decimos que son ortogonales y escribimos $U \perp V$ si $u \perp v$ para todo $u \in U$ y todo $v \in V$.

En $\mathbb{R}^3$, tenemos que $\overline{i} \perp \overline{j} \perp \overline{k} \perp \overline{i}$. Y por ejemplo, $pl(xy) \perp eje(z)$.

Aplicando estos conceptos ahora a los espacios fundamentales podemos probar el siguiente resultado para matrices reales:

Proposición 1 *Sea $A \in \mathbb{R}^{m \times n}$. Entonces:*

1. *$N(A) \perp C(A^t)$ y $C(A) \perp N(A^t)$.*
2. *Más aún, $N(A) \overset{\perp}{\oplus} C(A^t) = \mathbb{R}^n$ y $N(A^t) \overset{\perp}{\oplus} C(A)$.*

Demostración:

1. Veamos primero que $N(A) \perp C(A^t)$. Sean $x \in N(A)$ e $y \in C(A^t)$. Luego $Ax = 0$ y existe x' tal que $A^t x' = y$. Entonces, $x^t y = x^t A^t y = (Ax)^t y = 0^t y = 0$. Que $C(A) \perp N(A^t)$ es análogo (EJERCICIO).

2. Veamos ahora que $N(A) \overset{\perp}{\oplus} C(A^t) = \mathbb{R}^n$. Para esto, veamos primero que $N(A) \cap C(A^t) = \{\overline{0}\}$. En efecto, si $v \in N(A) \cap C(A^t)$ entonces, por el item anterior, $v \perp v$. Luego debe ser $v = \overline{0}$. La proposición sigue de observar que $dim N(A) = n - r$ y $dim C(A^t) = r$. La otra descomposición en suma directa ortogonal queda como EJERCICIO.

$\square$