ÁLGEBRA LINEAL I
NOTAS DE CLASE
UNIDAD 2

Abstract. Estas notas conciernen al álgebra de matrices y serán actualizadas conforme el material se cubre. Las notas no son substituto de la clase pues solo contienen un fragmento de la teoría discutida en el aula.

1. Álgebra de matrices

Definición 2.1: *(Matriz de $m \times n$.)* Sean m y n números naturales, entonces una *matriz* $\mathbf{A}$ de $m \times n$ sobre un campo $\mathbb{F}$ es un arreglo rectangular de mn números dispuestos en m renglones o filas y n columnas, en cuyo caso escribiremos $\mathbf{A} \in \mathbb{F}^{m\times n}$. Nos referiremos a m y n como las *dimensiones* de la matriz $\mathbf{A}$. En estas notas asumiremos que $\mathbb{F}$ es el campo de los números reales $\mathbb{R}$. Es usual referirse a los elementos del campo como *escalares*, lo cual también haremos aquí.

Observación 2.1: una matriz puede visualizarse de tres maneras distintas: (i) como un arreglo de números (definición 2.1), (ii) como una colección de renglones, (iii) como una colección de columnas. Entonces si $\mathbf{A} \in \mathbb{R}^{m\times n}$, las siguientes son las tres formas de visualizar esta matriz:

$$(i)\quad \mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \qquad (ii)\quad \mathbf{A} = [C_1 \ \cdots C_n]\,, \qquad (iii)\quad \mathbf{A} = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix},$$

en donde $a_{ij} \in \mathbb{R}$ para $i = 1\,, \ldots\,, m$ y $j = 1\,, \ldots\,, n$,

$$C_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}, \ j = 1, \ldots, n\,, \qquad R_i = [a_{i1} \ \cdots \ a_{in}]\,, \ i = 1, \ldots m\,.$$

Nos referiremos a los números a_{ij} como las *componentes* o *elementos* de la matriz $\mathbf{A}$.

Definición 2.2: *(Vectores columna y vectores renglón.)* Un *vector columna* X en $\mathbb{R}^n$ ($X \in \mathbb{R}^n$) es una matriz de $n \times 1$. Similarmente, un *vector renglón* Y en $\mathbb{R}^m$ ($Y \in \mathbb{R}^m$) es una matriz de $1 \times m$. Frecuentemente nos referiremos a las componentes de un vector como sus *coordenadas*.

Notación: entenderemos por el símbolo $(\mathbf{A})_{ij}$ como el elemento de $\mathbf{A}$ en el i-ésimo renglón y en la j-ésima columna de $\mathbf{A}$, es decir que si visualizamos $\mathbf{A}$ como en la forma (i) de la observación 2.1, entonces $(\mathbf{A})_{ij} = a_{ij}$.

Definición 2.3: *(Transpuesta de una matriz.)* Sea $\mathbf{A} \in \mathbb{R}^{m\times n}$, la matriz transpuesta de $\mathbf{A}$, denotada por $\mathbf{A^t}$ (léase "A transpuesta"), está definida como $\mathbf{A^t} \in \mathbb{R}^{n\times m}$ y $(\mathbf{A^t})_{ij} = a_{ji}$. Es decir que la matriz transpuesta se obtiene intercambiando renglones por columnas de manera ordenada. En particular, la transpuesta de un vector columna es un vector renglón y vice versa.

Definición 2.4: *(Multiplicación de un vector por un escalar.)* Sea $\alpha \in \mathbb{F}$ un elemento del campo $\mathbb{F}$ y $X = [x_1 \ \cdots \ x_n]$ un vector renglón con n componentes, entonces definimos la multiplicación de X por el escalar α de la siguiente manera,

$$\alpha X := [\alpha x_1 \ \ldots \ \alpha x_n]\,;$$

Date: February 16, 2010.

la multiplicación de un escalar por un vector columna se define de manera similar.

Definición 2.5: *(Suma de vectores.)* Sean X y Y vectores renglón ambos de dimensión n, $X = [x_1 \ \dots \ x_n]$, $Y = [y_1 \ \dots \ y_n]$. Definimos la suma de X y Y de la siguiente manera,

$$X + Y \ := \ [x_1 + y_1 \ \dots \ x_n + y_n]\,;$$

la suma de vectores columna se define de manera similar. Observe que la definición requiere sumar vectores de la misma dimensión y que la suma de dos vectores produce un tercer vector de la misma dimensión que los vectores sumandos.

Definición 2.6: *(Multiplicación de una matriz por un vector.)*
(i) *Multiplicación por la derecha:* sea $\mathbf{A} = [C_1 \ \cdots \ C_n]$ una matriz de $m \times n$ y $X = [x_1 \ \cdots \ x_n]^t$ un vector columna. Entonces definiremos el producto $\mathbf{A}X$ de la siguiente manera:

$$\mathbf{A}X \ = \ x_1 C_1 + \cdots + x_n C_n\,.$$

Observe que la definición de multiplicación por la derecha requiere de que el vector X tenga tantas coordenadas como columnas tenga la matriz $\mathbf{A}$ y que el producto $\mathbf{A}X$ es también un vector columna con tantas componentes como renglones tenga la matriz $\mathbf{A}$, es decir, $\mathbf{A}X \in \mathbb{R}^m$.
(i) *Multiplicación por la izquierda:* sea

$$\mathbf{A} = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}$$

una matriz de $m \times n$ y $Y = [y_1 \ \cdots \ y_m]$ un vector renglón. Entonces definiremos el producto $Y\mathbf{A}$ de la siguiente manera:

$$Y\mathbf{A} \ = \ y_1 R_1 + \cdots + y_m R_m\,.$$

Observe que la definición de multiplicación por la izquierda requiere de que el vector Y tenga tantas coordenadas como renglones tenga la matriz $\mathbf{A}$ y que el producto $Y\mathbf{A}$ es también un vector renglón con tantas componentes como columnas tenga la matriz $\mathbf{A}$, es decir, $Y\mathbf{A} \in \mathbb{R}^n$.

Lema 2.1: *(Otra interpretación de la multiplicación derecha e izquierda de matrices por vectores.)*
Sean $\mathbf{A} \in \mathbb{R}^{m\times n}$, X un vector columna en R^n y Y un vector renglón en $\mathbb{R}^m$ de la siguiente forma,

$$\mathbf{A} = [C_1\,, \, \dots \,, \, C_n] = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}, \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad Y = [y_1\,, \, \dots \,, \, y_m]\,.$$

Entonces se tiene que

$$\textbf{(a)} \quad \mathbf{A}X = \begin{bmatrix} R_1 X \\ R_2 X \\ \vdots \\ R_m X \end{bmatrix}, \qquad \textbf{(b)} \quad Y\mathbf{A} = [YC_1 \ \ YC_2 \ \ \dots \ \ YC_n]\,.$$

Observe que $R_i X$, $i = 1\,, \dots\,, m$, y YC_j, $j = 1\,, \dots\,, n$, son escalares.

Definición 2.8: *(matrices diagonales)*
Sea $\mathbf{D} \in \mathbb{R}^{n\times n}$ definida de la siguiente manera:

$$(1) \qquad (\mathbf{D})_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ d_i \in \mathbb{R} & \text{si } i = j \end{cases}\,;$$

es decir, $\mathbf{D}$ es tal que todas sus componentes son cero, excepto, quizás, las componentes sobre la diagonal principal (en donde la diagonal principal de $\mathbf{D}$ está constituida por todos sus elementos $(\mathbf{D})_{ii}$; también hablaremos de diagonal principal cuando la matriz no es cuadrada). Entonces decimos que $\mathbf{D}$ es una *matriz diagonal.*

Lema 2.2: sea $\mathbf{D} \in \mathbb{R}^{n\times n}$ una matriz diagonal de tamaño n como en (1), y sean $\mathbf{A} \in \mathbb{R}^{m\times n}$ y $\mathbf{B} \in \mathbb{R}^{n\times p}$ a las cuales visualizaremos como arreglos columna y renglón, respectivamente, es decir,

$$\mathbf{A} = [A_1 \cdots A_n] \qquad \text{y} \qquad \mathbf{B} = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix},$$

en donde $A_j = [a_{1j} \cdots a_{mj}]$, $j = 1, \ldots, n$, y $B_i = [b_{i1} \cdots b_{ip}]$, $i = 1, \ldots n$. Entonces

$$\mathbf{AD} = [d_1A_1 \cdots d_nA_n] \qquad \text{y} \qquad \mathbf{DB} = \begin{bmatrix} d_1B_1 \\ \vdots \\ d_nB_n \end{bmatrix}.$$

Definición 2.9: sea $\mathbf{I_n} \in \mathbb{R}^{n\times n}$ una matriz cuadrada de tamaño n definida como

$$(\mathbf{I_n}) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

Llamamos a $\mathbf{I_n}$ la *matriz identidad de tamaño n.*

Corolario 2.1:
(a) Para cualesquier matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ y $\mathbf{B} \in \mathbb{R}^{n\times p}$ se tiene que

$$\mathbf{AI_n} = \mathbf{A} \qquad \text{y} \qquad \mathbf{I_nB} = \mathbf{B}. \tag{2}$$

En particular, si $\mathbf{C} \in \mathbb{R}^{n}$ entonces

$$\mathbf{CI_n} = \mathbf{I_nC} = \mathbf{C}. \tag{3}$$

(b) $\mathbf{I_n}$ es la única matriz en $\mathbb{R}^{n\times n}$ que satisface las propiedades (2) y (3).

Definición 2.10: decimos que $\mathbf{A} \in \mathbb{R}^{n\times n}$ es *invertible* si existe una matriz $\mathbf{B}$ (que depende de $\mathbf{A}$) tal que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I_n}. \tag{4}$$

Lema 2.3: dada $\mathbf{A} \in \mathbb{R}^{n\times n}$, si existe una matriz $\mathbf{B}$ como en (4), entonces esta matriz es la única en $\mathbb{R}^{n\times n}$ que satisface dicha propiedad.

Notación: dado el lema anterior, a la matriz $\mathbf{B}$ se la denota de manera especial como $\mathbf{A^{-1}}$.

Lema 2.4: sea $\mathbf{A} \in \mathbb{R}^{n\times n}$ tal que $\mathbf{A}X = b$ tiene solución única para algún $b \in \mathbb{R}^n$ dado. Entonces, para cualquier otro vector $b' \in \mathbb{R}^n$, el sistema $\mathbf{A}X = b'$ también tiene solución única.
<u>Prueba</u>: Sea X_s la única solución de la ecuación $\mathbf{A}X = b$ (es decir, suponga que X_s es el único vector que satisface que $\mathbf{A}X_s = b$). Si, por el contrario, pudiéramos encontrar un vector $b' \in \mathbb{R}^n$ tal que el sistema $\mathbf{A}X = b'$ no tuviera solución única, esto quiere decir que podemos encontrar dos vectores, X_1' y X_2', tales que $X_1' \neq X_2'$ y $\mathbf{A}X_i' = b'$, $i = 1, 2$. Sea $Y' = X_1' - X_2'$ y observe que $Y' \neq 0$ y que $\mathbf{A}Y' = 0$. Dado que $Y' \neq 0$, entonces $X_s \neq X_s + Y'$, mas aún, $\mathbf{A}(X_s + Y') = \mathbf{A}X_s + \mathbf{A}Y' = b + 0 = b$; es decir, el vector $X_s + Y'$ es otra solución al sistema $\mathbf{A}X = b$, lo cual contradice la suposición original de que dicho sistema tiene solución única. Por lo tanto, para cualquier vector $b' \in \mathbb{R}^n$, el sistema $\mathbf{A}X = b'$ debe tener solución única.

Nota: recordemos que el método de eliminación de Gauss, o bien el método de eliminación de Gauss-Jordan, consiste en escribir un sistema de ecuaciones lineales en forma escalonada, o bien escalonada reducida, mediante el uso de operaciones elementales. Tales operaciones elementales se definieron como: *escalamiento* (multiplicar una ecuación por un escalar), *intercambio* (intercambiar dos ecuaciones), *reemplazo* (sumar a una ecuación un múltiplo de otra ecuación). Ahora bien, cuando definimos la matriz aumentada de un sistema de ecuaciones lineales, vimos que el realizar operaciones elementales sobre las ecuaciones del sistema es equivalente a realizar esas mismas operaciones elementales sobre la matriz aumentada del sistema. La consecuencia mas importante de dicha observación fué que las operaciones elementales sobre los renglones de la matriz aumentada se pueden interpretar como la multiplicación por la izquierda de la matriz aumentada por matrices que llamamos *elementales*.

Defінción 2.11: decimos que $\mathbf{E} \in \mathbb{R}^{n\times n}$ es una *matriz elemental* si para cualquier matriz $\mathbf{A} \in \mathbb{R}^{n\times p}$, la matriz producto $\mathbf{EA}$ se obtiene de la matriz $\mathbf{A}$ después de sumarle a un renglón de $\mathbf{A}$ un múltiplo de otro renglón.

Lema 2.5: sea $\mathbf{E}$ la matriz elemental tal que la matriz producto $\mathbf{EA}$ se obtiene de la matriz $\mathbf{A}$ mediante sumarle al renglón ℓ-ésimo, α veces el renglón k-ésimo. Dado que el único uso que le daremos a las matrices elementales es en conexión con la eliminación Gaussiana, podemos suponer que $\ell > k$. Entonces

$$(\mathbf{E})_{ij} = \begin{cases} 1 & \text{si } i = j\,, \\ \alpha & \text{si } i = \ell \text{ y } j = k\,, \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Definición 2.12: sea $\mathbf{T} \in \mathbb{R}^{m\times n}$. Decimos que
(a) $\mathbf{T}$ es una matriz *triangular superior* si $(\mathbf{T})_{ij} = 0$ para toda $i > j$; es decir, si los elementos por debajo de la diagonal principal son todos cero.
(b) $\mathbf{T}$ es una mariz *triangular inferior* si $(\mathbf{T})_{ij} = 0$ para toda $i < j$; es decir, si los elementos por arriba de la diagonal principal son todos cero.

Observe que la matriz elemental de la definición 2.11 es triangular inferior.

Definición 2.13: sea $\mathbf{P} \in \mathbb{R}^m$ tal que para cualquier $\mathbf{A} \in \mathbb{R}^{m\times n}$, la matriz producto $\mathbf{PA}$ se obtiene de la matriz $\mathbf{A}$ mediante el intercambio de los *renglones* k y ℓ. Entonces decimos que $\mathbf{P}$ es una *matriz de permutación*.

Lema 2.6: Sea $\mathbf{P} \in \mathbb{R}^m$ la matriz de permutación de la definición anterior. Entonces,
(a) $\mathbf{P}$ se obtiene de la matriz identidad $\mathbf{I_m}$ intercambiando los renglones k con ℓ de esta última.
(b) $\mathbf{P^2} = \mathbf{I_m}$ ($\mathbf{P}$ es su propia inversa).
(c) Para cualquier $\mathbf{B} \in \mathbb{R}^{p\times m}$, la matriz producto $\mathbf{BP}$ se obtiene de la matriz $\mathbf{B}$ mediante el intercambio de las *columnas* k y ℓ.

Proposición 2.1: considere el sistema cuadrado de n ecuaciones lineales en n incógnitas, escrito en forma matricial,

$$\mathbf{A}X = b\,, \tag{5}$$

en donde $\mathbf{A} \in \mathbb{R}^{n\times n}$ y $X = [x_1\ \ldots\ x_n]^t$. Entonces tenemos que (5) tiene solución única si y solamente si $\mathbf{A}$ es invertible.
<u>Prueba</u>:
(i) Supongamos que $\mathbf{A}$ es invertible y procedamos a demostrar que el sistema (5) tiene solución única. Supongamos que existen dos soluciones, X_1 y X_2, de (5); es decir, $\mathbf{A}X_i = b$, $i = 1, 2$; queremos demostrar que estas soluciones son de hecho la misma (i.e., que $X_1 = X_2$). Dado que $\mathbf{A^{-1}}$ existe por hipótesis, entonces se tiene que $\mathbf{A^{-1}}(\mathbf{A}X_i) = \mathbf{A^{-1}}b$, para $i = 1, 2$; es decir, usando la ley asociativa del producto de matrices, que $X_i = \mathbf{A^{-1}}b$, $i = 1, 2$. Ahora observe que el lado derecho de esta última igualdad no depende de i, por lo tanto $X_1 = X_2$, que es lo que se quería demostrar.
(ii) Supongamos ahora que el sistema $\mathbf{A}X = b$ tiene solución única; es decir, que existe un único vector X_s, tal que satisface (5). Queremos demostrar que entonces existe una matriz $\mathbf{B}$ que satisface (4). Consideremos la matrix aumentada del sistema (5), $(\mathbf{A}|b)$. Dado que el sistema tiene solución única, podemos aplicar el método de eliminación de Gauss-Jordan para encontrar la solución, X_s; esto lo que significa es que, mediante el uso de operaciones elementales por renglón, podemos reescribir la matriz aumentada de la siguiente manera, $(\mathbf{I_n}|X_s)$. Lo anterior es equivalente a afirmar la existencia de una sucesión finita de *matrices elementales*, $\mathbf{E_i}$, $i = 1, \ldots, k$, y, posiblemente, una matriz de permutación $\mathbf{P}$, tales que

$$\mathbf{E_k E_{k-1}} \cdots \mathbf{E_1 P}(\mathbf{A}|b) = (\mathbf{U}|X_s)\,,$$

es decir, tales que,

$$\mathbf{E_k E_{k-1}} \cdots \mathbf{E_1 P A} = \mathbf{I_n}\,,$$
$$\mathbf{E_k E_{k-1}} \cdots \mathbf{E_1 P}b = \mathbf{X_s}\,.$$

Sea
$$\mathbf{B} := \mathbf{E_k E_{k-1}} \cdots \mathbf{E_1 P}\,,$$
entonces $\mathbf{BA} = \mathbf{I_n}$. Queda por demostrar que $\mathbf{AB} = \mathbf{I_n}$, esto lo hacemos de la siguiente manera:
$$\begin{aligned} \mathbf{BA} &= \mathbf{I_n} \\ \mathbf{A}(\mathbf{BA}) &= \mathbf{AI_n} \\ (\mathbf{AB})\mathbf{A} &= \mathbf{A} \\ (\mathbf{AB} - \mathbf{I_n})\mathbf{A} &= \mathbf{0_n}\,, \end{aligned} \tag{6}$$
en donde $\mathbf{0_n}$ es la matriz cuadrada nula de tamaño n. Ahora, por el lema 2.4, dado que $\mathbf{A}X = b$ tiene solución única, entonces también el sistema $\mathbf{A}X = e_i$ tiene solución única, en donde e_i es el vector columna cuyas componentes son todas cero excepto la i-ésima; llame a ξ_i la solución de dicha ecuación (es decir $\mathbf{A}\xi_i = e_i$). Multiplicando entonces la última ecuación en (6) por la derecha por ξ_i llegamos a que,
$$(\mathbf{AB} - \mathbf{I_n})e_i = 0\,.$$
El lado izquierdo de la ecuación de arriba es la columna i-ésima de la matriz $\mathbf{AB} - \mathbf{I_n}$, dado que i es arbitrario, concluimos que todas las columnas de la matriz $\mathbf{AB} - \mathbf{I_n}$ son cero, es decir $\mathbf{AB} - \mathbf{I_n} = \mathbf{0_n}$, por lo tanto $\mathbf{AB} = \mathbf{I_n}$, como se quería demostrar. Concluimos entonces que $\mathbf{AB} = \mathbf{BA} = \mathbf{I_n}$, lo cual implica que $\mathbf{B} = \mathbf{A^{-1}}$, como se había afirmado.

Corolario 2.2: supongamos que el sistema $\mathbf{A}X = b$ es consistente y que tiene un número infinito de soluciones. El sistema podría no ser cuadrado. Entonces,
(a) $\mathbf{A}$ no es invertible.
(b) Sea Y una solución entre el número infinito de soluciones que satisfacen $AX = b$. Entonces cualquier otra solución es de la forma $X = Y + X'$, en donde X' es solución del sistema homogéneo asociado, es decir, $\mathbf{A}X' = 0$.

Nota: la parte (b) del corolario anterior, se puede demostrar sin necesidad de los resultados anteriores, simplemente haciendo uso de la teoría ya cubierta de sistemas de ecuaciones lineales. En efecto, si $\mathbf{A}X = b$ tiene un número infinito de soluciones, entonces sabemos que no todas la variables son pivotales. Supongamos que $x_{i_1}, \ldots, x_{i_k}$ son las variables parámetro, entonces se puede demostrar mediante eliminación de Gauss-Jordan que la solución del sistema se puede escribir de la siguiente forma,
$$X = V + x_{i_1}H_1 + x_{i_2}H_2 + \cdots + x_{i_k}H_k\,,$$
en donde V satisface $\mathbf{A}V = b$ (solución particular), y H_j, $j = 1, \ldots, k$, es un vector columna tal que $\mathbf{A}H_j = 0$ y cuyas componentes correspondientes a los lugares en donde se encuentran las otras variables parámetro distintas de x_{i_j} son cero.

Definición 2.14: supongamos que el sistema $\mathbf{A}X = b$ es consistente y tiene un número infinito de soluciones.
(a) A cualquiera de entre el número infinito de soluciones del sistema no homogéneo, $\mathbf{A}X = b$, la llamaremos *solución particular* y la representaremos por X_p.
(b) Representaremos por X_h a la solución general del sistema homogéneo asociado, $\mathbf{A}X = 0$.
(c) Llamamos a $X_g = X_p + X_h$ la *solución general* del sistema, en donde X_p y X_h son como se definieron en los incisos anteriores.

De la parte (ii) de la demostración de la proposición 2.1 se desprende la siguiente

Proposición 2.2: (*Descomposición LU*)
Sea $\mathbf{A} \in \mathbb{R}^{m\times n}$. Entonces existe una matriz de permutación $\mathbf{P} \in \mathbb{R}^{m\times m}$, una matriz triangular inferior invertible $\mathbf{L} \in \mathbb{R}^{m\times m}$ con diagonal principal de unos, y una matriz triangular superior $\mathbf{U} \in \mathbb{R}^{m\times n}$ tales que $\mathbf{PA} = \mathbf{LU}$.

Jorge Viveros. Centro de Investigación en Matemáticas, Universidad Autónoma del Estado de Hidalgo. Hgo, México.

E-mail address: jviveros@uaeh.edu.mx