Tópicos en Teoría de Matrices

UNIVERSIDAD NACIONAL DE COLOMBIA
Sede Medellín
DEPARTAMENTO DE BIBLIOTECAS
Ciencias Agropecuarias y Ciencias

TOPICOS EN TEORIA DE MATRICES

ABRAHAM J. ASMAR CHARRIS

Profesor Asociado
Departamento de Matemáticas
Universidad Nacional de Colombia, Medellín

Universidad Nacional de Colombia
Sede Medellín
Facultad de Ciencias
1995

UNIVERSIDAD NACIONAL DE COLOMBIA
Sede Medellín
DEPARTAMENTO DE BIBLIOTECAS
Ciencias Agropecuarias y Ciencias

Tópicos en Teoría de Matrices
Primera Edición
1995

Prohibida la reproducción total o parcial de esta obra.
por cualquier medio, sin autorización escrita de la
Universidad Nacional de Colombia y del Autor.

UNAL-Medellín
6 4000 00062370 5

Impreso en Medellín
Esta obra se terminó de imprimir en Noviembre de 1995
En la Litografía Ecográficas
Se imprimieron 300 ejemplares

A mis padres Francisco y Josefina.

UNIVERSIDAD NACIONAL DE COLOMBIA
Sede Medellín
DEPARTAMENTO DE BIBLIOTECAS
Ciencias Agropecuarias y Ciencias

I
512.9434
A75

Contenido

UNIVERSIDAD NACIONAL DE COLOMBIA
BIBLIOTECA

Prólogo xi

Capítulo 1
Repaso y miscelánea 1

1.1 Matrices 1
Suma de matrices 4
Multiplicación por escalar 5
Producto de matrices 5
Inversas 9
Transpuesta, conjugada y transpuesta hermitiana 10
Matrices simétricas y hermitianas 11
Matrices y sistemas de ecuaciones lineales. Método de Eliminación de Gauss 12
Operaciones elementales y matrices elementales 17
Ejercicios 1.1 18

1.2 Espacios vectoriales 20
Subespacios 21
Subespacio generado 22
Subespacios asociados a una matriz 23
Dependencia e independencia lineal 26
Bases y dimensión 29
Independencia lineal e invertibilidad de matrices 30
Transformaciones lineales 31
Ejercicios 1.2 33

1.3 Rango de una matriz 34
Operaciones elementales y rango. Factorización rango 37
Relación entre $dim\,\mathcal{N}(\mathbf{A})$ y $\rho(\mathbf{A})$ 38
Ejercicios 1.3 41

1.4 Determinantes 42
Desarrollo de determinantes por cofactores 46
Adjunta e inversas 46
Ejercicios 1.4 48

Donación Cuello 7-11-76

UNIVERSIDAD NACIONAL DE COLOMBIA
Sede Medellín
36450
DEPARTAMENTO DE BIBLIOTECAS
Ciencias Agropecuarias y Ciencias

1.5 Sistemas de ecuaciones lineales. Resultados básicos 50
Ejercicios 1.5 52
1.6 Producto escalar y ortogonalidad 53
Ortogonalidad 55
Proyecciones ortogonales y proceso de ortogonalización de Gram-Schmidt . 56
Factorización QT 58
Subespacios ortogonales 61
Complemento ortogonal. Teorema fundamental del Algebra Lineal 64
Producto escalar estándar y norma euclidiana en $\mathbb{C}^n$ 67
Ejercicios 1.6 69

Capítulo 2
Valores propios, vectores propios y diagonalización 72

2.1 Coordenadas en F^n 72
Cambio de base (cambio de coordenadas) 76
Ejercicios 2.1 78
2.2 Matriz que representa una transformación lineal respecto a una base 79
Semejanza de matrices 83
Ejercicios 2.2 86
2.3 Problema de la diagonalización de una matriz 87
Valores y vectores propios de una matriz 89
Ejercicios 2.3 101
2.4 Matrices simétricas reales, matrices hermitianas y diagonalización. Lema de Schur. Teorema Espectral 103
Ejercicios 2.4 110
2.5 Matrices normales 111
Ejercicios 2.5 112
2.6 Descomposición en valores singulares 113
Ejercicios 2.6 117

Capítulo 3
Formas cuadráticas reales 118

3.1 Formas cuadráticas en $\mathbb{R}^n$ 119
Ejercicios 3.1 121

3.2 Diagonalización de una forma cuadrática 121
Diagonalización de una forma cuadrática mediante una transformación ortogonal 123
Método de Lagrange 130
Diagonalización de una forma cuadrática usando operaciones elementales . 132
Ley de inercia de las formas cuadráticas 134
Valores extremos de formas cuadráticas 136
Ejercicios 3.2 139
3.3 Formas cuadráticas y matrices definidas 140
Ejercicios 3.3 150

Capítulo 4
Proyecciones ortogonales y mínimos cuadrados 152

4.1 Proyección ortogonal de un vector sobre un subespacio 153
Matrices de proyección ortogonal 155
Ejercicios 4.1 163
4.2 Descomposición espectral de una matriz simétrica real 165
Ejercicios 4.2 169
4.3 Problemas lineales de mínimos cuadrados 170
Ejercicios 4.3 180
4.4 Algo de Cálculo Diferencial 182
Ejercicios 4.4 185

Capítulo 5
Inversa generalizada e inversas condicionales 187

5.1 Inversa generalizada 187
Condiciones de Penrose 195
Sistemas de ecuaciones lineales y la inversa generalizada 199
Ejercicios 5.1 200
5.2 Inversas condicionales 201
Tres tipos básicos de c-inversas 204
Ejercicios 5.2 206

Capítulo 6
Normas de vectores y matrices 209

6.1 Normas de vectores en F^n. Equivalencia de normas 210

Ejercicios 6.1 217
6.2 Normas matriciales 218
Norma matricial inducida por una norma vectorial 221
Ejercicios 6.2 225
6.3 Aplicaciones. Resultados básicos 227
Ejercicios 6.1 231

Referencias 233
Símbolos, notaciones y abreviaturas 235
Indice 237

Prólogo

UNIVERSIDAD NACIONAL
BIBLIOTECA CENTRAL

Si bien es cierto que existe un gran número de textos elementales y no elementales sobre Algebra Lineal, también lo es que en ocasiones se hace difícil hallar uno que se adecúe a los requerimientos y necesidades de un determinado curso. Justamente esta dificultad encontré al encargarme, en el primer semestre de 1993, de la asignatura Algebra Matricial Aplicada de la Especialización en Estadística de mi Facultad, lo que me obligó a preparar un material que hiciera las veces de texto guía para el curso. Algunos estudiantes, como también algunos colegas, me animaron a pensar en la posibilidad de publicar ese material. Pues bien, este texto es el resultado de aquellas primeras notas, revisadas y ampliadas durante los semestres siguientes en los que las usé en el curso de Algebra Matricial ya mencionado y también en el de Algebra Lineal para los Posgrados en Matemáticas, incluyendo el que dicté en la Universidad de Córdoba durante el primer semestre de 1994, dirigido a estudiantes de la Especialización en Matemáticas que la Universidad Nacional adelanta en esa Universidad.

Los estudiantes contribuyeron de manera importante, con sus sugerencias y comentarios, a la versión que aquí se presenta.

La asignatura Algebra Matricial a la que me he referido está dirigida principalmente a estudiantes de la Especialización o Maestría en Estadística, y el curso debe proporcionarles el lenguaje, los conceptos y resultados del Álgebra Matricial de uso en áreas como Modelos Lineales y Análisis Multivariante. La selección de tópicos que aparecen en este material obedece a esta necesidad. Pero, sin ninguna duda, tales temas interesan también a quienes adelantan una Especialización o Maestría en Matemáticas; es más - y no creo exagerar - tales conocimientos deberían hacer parte de la formación básica en pregrado para Matemáticos, Licenciados en Matemáticas e Ingenieros.

El primer capítulo es introductorio. En él repasamos rápidamente conceptos, resultados y procedimientos que corresponden, por lo general, a un primer curso en Algebra Lineal. Se ha incluido este capítulo, pensando en aquellos profesionales que llegan a los programas de posgrado, luego de un largo período de inactividad en lo que a Algebra Lineal se refiere.

En el capítulo 2 se trata el problema de establecer condiciones necesarias y/o suficientes para que una transformación lineal de F^m en F^n ($F = \mathbb{R}$ o $F = \mathbb{C}$) se pueda representar mediante una matriz diagonal. Los conceptos de valor propio y vector propio de una matriz se introducen a partir de este problema. Se destacan en este capítulo tres resultados: el Lema de Schur, el Teorema espectral y la Descomposición en valores singulares.

El capítulo 3 es para las formas cuadráticas en $\mathbb{R}^n$. El problema fundamental es el de la

UNIVERSIDAD NACIONAL
BIBLIOTECA CENTRAL

reducción a una suma de cuadrados mediante una transformación lineal real invertible, de las variables; se dan distintos métodos para la reducción. Un resultado a destacar es la Ley de Inercia. También tratamos en este capítulo el tema de las matrices reales (formas cuadráticas en $\mathbb{R}^n$) definidas.

El capítulo 4 está dedicado al concepto de proyección ortogonal de un vector sobre un subespacio de $\mathbb{R}^n$, y al tema asociado de las soluciones, según mínimos cuadrados, de sistemas de ecuaciones lineales inconsistentes. La sección 4.2, en la que se trata la descomposición espectral de una matriz simétrica real, es un complemento al capítulo 2, particularmente al teorema espectral.

En el capítulo 5 se aborda el tema de la inversa generalizada para una matriz real, concepto que se introduce a partir de la noción de solución según mínimos cuadrados, de longitud mínima, para sistemas de ecuaciones lineales. Este concepto, como el de inversas condicionales (otros tipos de inversas), es útil en áreas como Modelos Lineales.

El capítulo 6 es una introducción al concepto de Norma (medida de la "magnitud" o "tamaño") para vectores en $\mathbb{R}^n$ o en $\mathbb{C}^n$ y para matrices cuadradas. En los capítulos anteriores hemos hecho únicamente "Algebra", en éste hacemos un poco de "Análisis", preocupándonos ahora por la noción de "cercanía" entre vectores y entre matrices. Este tema es esencial en lo que se conoce como Algebra Lineal Numérica.

He procurado, a lo largo del texto, combinar una exposición lo más simple posible con un desarrollo teórico riguroso. Los ejercicios propuestos, en su gran mayoría, complementan la teoría expuesta; por lo general son de fácil solución si se tiene claridad conceptual.

Mis agradecimientos especiales para el profesor Arturo Jessie M., Director del Departamento de Matemáticas, quien apoyó desde su inicio el proyecto de publicación de este texto. Igualmente agradezco a la Facultad de Ciencias, a sus Posgrados en Matemáticas y Estadística, y a CINDEC-Medellín, por su patrocinio a esta publicación.

Debo también expresar mi gratitud a los profesores Carlos Mejía S. y Juan Carlos Correa M., quienes gentilmente se tomaron el trabajo de revisar todo el material. Muchas rectificaciones y mejoras las debo a sus valiosas sugerencias y acertadas observaciones.

Finalmente doy las gracias a los estudiantes Jorge H. Betancur y Diego F. Chávez quienes se encargaron de la transcripción del material utilizando el procesador matemático $\TeX$, como también a la profesora Luz Helena Muñoz S., quien amablemente facilitó los formatos para el levantamiento de este texto.

Abraham Asmar Charris.

Capítulo 1

REPASO Y MISCELANEA

Este capítulo es una recopilación de conceptos, procedimientos y resultados básicos que, por lo general, forman parte de un primer curso en Algebra Lineal. Se advierte que una gran parte de esos resultados aparecen aquí sin prueba.

El propósito fundamental de este material es servir como pre-requisito para los siguientes capítulos.

Referencias:

[*Noble-Daniel*], cap. 1, 3, 4 y 5.
[*Strang*], cap. 1, 2 y 4.
[*Searle*], cap. 1, 2, 3, 4, 6 y 7.
[*Lipschutz*], cap. 1, 2, 3, 4, 5 y 8.
[*Res-Fran-Mu*], cap. 1, 2 y 4.
[*Kreyszig*], cap. 7.

1.1 Matrices

A lo largo de este texto, F denotará o bien el cuerpo de los números reales $\mathbb{R}$, o bien el cuerpo de los números complejos $\mathbb{C}$. A menudo nos referiremos a los elementos de F como **escalares**.

Llamaremos **matriz sobre** F, o simplemente **matriz**, cualquier arreglo u ordenación rectangular como el siguiente

$$\mathbf{A} = \begin{array}{cl} \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1m} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{im} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nm} \end{bmatrix} & \longrightarrow \text{Fila i} \\ \downarrow & \\ \text{Columna j} & \end{array} \tag{1}$$

en el que los a_{ij} son elementos de F.

El arreglo de números se encierra, como se muestra en (1), entre corchetes (se pudo usar también paréntesis) para indicar que se quiere considerar como un solo objeto.

Los números en una línea horizontal en la matriz conforman una **fila** de la matriz, y los números en una línea vertical una **columna**.

Nos referiremos a la fila conformada por los elementos

$$a_{i1}, a_{i2}, \ldots, a_{im}$$

como la **i-ésima fila** o **fila i**, y a la columna conformada por

$$a_{1j}, a_{2j}, \ldots, a_{nj}$$

como la **j-ésima columna** o **columna j**.

Los números que conforman la matriz serán llamados **elementos** o **componentes** de la matriz, y el elemento a_{ij} (el cual está simultáneamente en la fila i y en la columna j) se dirá el **elemento** o **componente** (i, j) de la matriz.

Si una matriz tiene n filas y m columnas (como en (1)), diremos que es una **matriz de orden $n \times m$** (léase "n por m"), o que es una **matriz $n \times m$**.

Notaremos $\mathcal{M}_{n\times m}(F)$ el conjunto de todas las matrices $n \times m$ sobre F.

En forma abreviada, una matriz $\mathbf{A}$ como la dada en (1) será denotada $\mathbf{A} = [a_{ij}]$. (En lugar de las letras i, j puede usarse otras). Observe que en esta notación no se indica el orden de la matriz.

Una matriz $n \times m$ en la que $n = m$ se dirá **cuadrada**. En este caso en lugar de decir que es de orden $n \times n$, diremos que es de **orden** n.

Si $\mathbf{A} = [a_{ij}]$ es una matriz cuadrada de orden n, la diagonal conformada por las componentes a_{ii}, será llamada la **diagonal principal** de $\mathbf{A}$, y en ocasiones **la diagonal** de $\mathbf{A}$.

Una matriz cuadrada cuyos elementos fuera de la diagonal principal son todos nulos se dirá una **matriz diagonal**.

Cualquier matriz cuyos elementos son todos nulos se dirá una **matriz nula** o una **matriz cero**. La matriz nula de orden $n \times m$ será denotada $\mathbf{0}_{n\times m}$, o simplemente $\mathbf{0}$ cuando no haya lugar a confusión.

Una matriz $n \times 1$ (de una sola columna) la llamaremos una **matriz columna** o un **vector columna**. Notaremos F^n el conjunto de vectores columna con n componentes en F. Cada columna de una matriz puede verse como un vector columna, y con frecuencia notaremos una matriz $n \times m$ $\mathbf{A} = [a_{ij}]$ en la forma

$$\mathbf{A} = [\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_m] \tag{2}$$

donde $\mathbf{a}_j \in F^n$ es la j-ésima columna de $\mathbf{A}$, esto es,

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

Matrices de una sola fila serán llamadas **matrices fila** o **vectores fila**. Las filas de una matriz pueden verse como vectores filá, y también a menudo notaremos una matriz $n \times m$ $\mathbf{A} = [a_{ij}]$ en la forma

$$\mathbf{A} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{3}$$

siendo α_i la i-ésima fila de $\mathbf{A}$, es decir,

$$\alpha_i = [\, a_{i1} \quad a_{i2} \quad \dots \quad a_{im} \,].$$

Llamaremos **submatriz** de una matriz $\mathbf{A}$, cualquier matriz que resulte al suprimir algunas filas y/o algunas columnas de $\mathbf{A}$. Por ejemplo, para

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 0 & 5 \\ 0 & 3 & 8 \end{bmatrix} \tag{4}$$

las matrices siguientes son submatrices:

$$\mathbf{A}_1 = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 0 & 5 \end{bmatrix}, \qquad \mathbf{A}_2 = [\, 0 \quad 3 \quad 8 \,] \tag{5}$$

$$\mathbf{B}_1 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{B}_2 = [\, 0 \quad 3 \,], \qquad \mathbf{B}_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \qquad \mathbf{B}_4 = [\, 8 \,]. \tag{6}$$

Observe que tanto las filas como las columnas de $\mathbf{A}$ son submatrices de $\mathbf{A}$.

En la escritura (2) de una matriz $\mathbf{A}$, ésta aparece expresada en términos de sus columnas. Diremos que está **separada** o **particionada** en columnas; de la misma forma cuando $\mathbf{A}$ esté expresada como en (3) diremos que está **separada** o **particionada** en filas. Tales maneras de expresar una matriz resultarán de mucha utilidad, como se apreciará a lo largo de éste y los siguientes capítulos, pero no son las únicas formas útiles de **particionar** una matriz, es decir, de expresarla separada en submatrices. Por ejemplo, la matriz $\mathbf{A}$ en (4) podría convenir expresarla, para cierto propósito, particionada como

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 0 & 5 \\ \dots & \dots & \dots \\ 0 & 3 & 8 \end{bmatrix}$$

esto es, como

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}$$

siendo $\mathbf{A}_1$ y $\mathbf{A}_2$ las submatrices que aparecen en (5).

esto es, como

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}$$

siendo $\mathbf{A}_1$ y $\mathbf{A}_2$ las submatrices que aparecen en (5).

Quizá esa misma matriz, para otro fin, convenga particionarla como

$$\mathbf{A} = \left[\begin{array}{cc:c} 2 & 0 & 4 \\ 1 & 0 & 5 \\ \hdashline \dots & \dots & \dots \\ \hdashline 0 & 3 & 8 \end{array}\right]$$

es decir, como

$$\mathbf{A} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_3 \\ \mathbf{B}_2 & \mathbf{B}_4 \end{bmatrix}$$

siendo $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4$ las submatrices que aparecen en (6).

Si se desea consultar acerca del concepto de matrices **particionadas** o **separadas** se puede recurrir, por ejemplo, a los textos [*Noble-Daniel*], [*Searle*] y [*Maltsev*].

Por último, recordamos que dos matrices $\mathbf{A} = [a_{ij}]$, $\mathbf{B} = [b_{ij}]$ en $\mathcal{M}_{n\times m}(F)$ se dicen **iguales** (notamos $\mathbf{A} = \mathbf{B}$) si $a_{ij} = b_{ij} \quad \forall i = 1, \dots, n \quad \forall j = 1, \dots, m$.

Suma de matrices

Sean $\mathbf{A} = [a_{ij}]$, $\mathbf{B} = [b_{ij}]$ en $\mathcal{M}_{n\times m}(F)$. La matriz $n \times m$ $\mathbf{C} = [c_{ij}]$ en la que $c_{ij} = a_{ij} + b_{ij}$ para cada $i \in \{1, \dots, n\}$ y cada $j \in \{1, \dots, m\}$, se llama la matriz **suma** de $\mathbf{A}$ y $\mathbf{B}$. Notamos $\mathbf{C} = \mathbf{A} + \mathbf{B}$.

Propiedades básicas de la suma

$\forall \mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{M}_{n\times m}(F)$

1. $\mathbf{A} + \mathbf{B} \in \mathcal{M}_{n\times m}(F)$
2. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
3. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
4. $\mathbf{A} + 0_{n\times m} = \mathbf{A}$
5. Si $\mathbf{A} = [a_{ij}]$, la matriz $-\mathbf{A} = [-a_{ij}]$, en $\mathcal{M}_{n\times m}(F)$, es tal que $\mathbf{A} + (-\mathbf{A}) = 0_{n\times m}$.

La **diferencia** entre matrices se define como entre números: Si $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times m}(F)$, entonces

$$\mathbf{B} - \mathbf{A} = \mathbf{B} + (-\mathbf{A}).$$

Multiplicación por escalar

Sean $\mathbf{A} = [a_{ij}] \in \mathcal{M}_{n\times m}(F)$ y $k \in F$. El producto de k por $\mathbf{A}$ es la matriz $n \times m$, notada $k\mathbf{A}$, cuyo elemento en la posición (i, j) es ka_{ij}, esto es,

$$k\mathbf{A} = [ka_{ij}] .$$

Propiedades básicas de la multiplicación por escalar

- $\forall \mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{M}_{n\times m}(F)$, $\forall k, s \in F$

1. $k\mathbf{A} \in \mathcal{M}_{n\times m}(F)$
2. $(k + s)\mathbf{A} = k\mathbf{A} + s\mathbf{A}$
3. $k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$
4. $k(s\mathbf{A}) = (ks)\mathbf{A}$
5. $1\mathbf{A} = \mathbf{A}$.

Las propiedades enunciadas para las operaciones anteriores en $\mathcal{M}_{n\times m}(F)$ hacen de este conjunto un espacio vectorial sobre el cuerpo F. (Más adelante recordaremos el concepto general de espacio vectorial).

En particular F^n (conjunto de vectores columna con n componentes en F), con esas operaciones, es un espacio vectorial sobre F.

Producto de matrices

a) **Producto básico de una matriz fila $1 \times m$ por una matriz columna $m \times 1$:** El producto $\boldsymbol{\alpha}\mathbf{b}$ de la matriz fila $\boldsymbol{\alpha} = [a_1 \quad a_2 \quad \dots \quad a_m]$ por la matriz columna $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ se define como:

$$\boldsymbol{\alpha}\mathbf{b} = [a_1 \quad a_2 \quad \dots \quad a_m] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = [a_1b_1 + a_2b_2 + \cdots + a_mb_m] \tag{7}$$

producto que identificaremos con el escalar

$$a_1b_1 + a_2b_2 + \cdots + a_mb_m .$$

b) **Producto de una matriz $n \times m$ por una columna $m \times 1$:**
Sea $\mathbf{A} = [a_{ij}]$ una matriz $n \times m$ con filas $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ y sea $\mathbf{b}$ un vector columna como en a). Definimos el producto $\mathbf{Ab}$ como se indica a continuación:

$$\mathbf{Ab} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_i \\ \vdots \\ \boldsymbol{\alpha}_n \end{bmatrix} \mathbf{b} = \begin{bmatrix} \boldsymbol{\alpha}_1\mathbf{b} \\ \vdots \\ \boldsymbol{\alpha}_i\mathbf{b} \\ \vdots \\ \boldsymbol{\alpha}_n\mathbf{b} \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1m}b_m \\ \vdots \\ a_{i1}b_1 + a_{i2}b_2 + \cdots + a_{im}b_m \\ \vdots \\ a_{n1}b_1 + a_{n2}b_2 + \cdots + a_{nm}b_m \end{bmatrix}. \tag{8}$$

Observaciones.

i) Si $\mathbf{A} \in \mathcal{M}_{n\times m}(F)$ y $\mathbf{b} \in F^m$ entonces $\mathbf{Ab} \in F^n$.

ii) La i-ésima componente de $\mathbf{Ab}$ es el escalar

$$\boldsymbol{\alpha}_i\mathbf{b} = a_{i1}b_1 + a_{i2}b_2 + \cdots + a_{im}b_m = \sum_{k=1}^{m} a_{ik}b_k\,. \tag{9}$$

iii) Note que, también,

$$\mathbf{Ab} = \mathbf{A}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = b_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + b_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + b_m\begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}.$$

Por tanto, si $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ (en F^n) son las columnas de $\mathbf{A}$, podemos escribir

$$\mathbf{Ab} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_m]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_m\mathbf{a}_m\,. \tag{10}$$

Una expresión como $b_1\mathbf{a}_1 + b_2\mathbf{a}_2 + \cdots + b_m\mathbf{a}_m$ $(b_1, b_2, \ldots, b_m \in F)$ se dirá una **combinación lineal** (*c.l.*) de los vectores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Los escalares $b_1, \ldots, b_m$ se dirán los coeficientes de la *c.l.* (El concepto de *c.l.* en un espacio vectorial cualquiera será introducido en la sección siguiente).

Así que (10) nos dice que $\mathbf{Ab}$ es una *c.l.* de las columnas de $\mathbf{A}$ en la que las componentes del vector $\mathbf{b}$ son los coeficientes. Este hecho será usado frecuentemente.

c) **Producto de una matriz $\mathbf{n \times m}$ por una matriz $\mathbf{m \times p}$:**
Si $\mathbf{A}$ es una matriz $n \times m$ y $\mathbf{B}$ es una matriz $m \times p$, el producto $\mathbf{AB}$ se define de tal

forma que cada columna de $\mathbf{AB}$ se obtenga multiplicando $\mathbf{A}$ por la correspondiente columna de $\mathbf{B}$ (como en (8)), es decir, si $\mathbf{B} = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_p\,]$ ($\mathbf{b}_j$ es la j-ésima columna de $\mathbf{B}$) entonces

$$\mathbf{AB} = [\,\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \dots \quad \mathbf{Ab}_p\,]. \tag{11}$$

Note que $\mathbf{AB}$ es una matriz $n \times p$.

Si el número de columnas de la matriz $\mathbf{A}$ no coincide con el número de filas de la matriz $\mathbf{B}$, el producto $\mathbf{AB}$ no está definido.

Observaciones.

Digamos que $\mathbf{A} = [a_{ij}]$ y $\mathbf{B} = [b_{ij}]$. Consideremos $\mathbf{A}$ y $\mathbf{B}$ particionadas en filas y en columnas como se indica a continuación:

$$\mathbf{A} = [\,\mathbf{a}_1 \quad \dots \quad \mathbf{a}_m\,] = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_n \end{bmatrix}, \qquad \mathbf{B} = [\,\mathbf{b}_1 \quad \dots \quad \mathbf{b}_p\,] = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_m \end{bmatrix}$$

i)

$$\mathbf{AB} = \mathbf{A}\,[\,\mathbf{b}_1 \quad \dots \quad \mathbf{b}_p\,] = [\,\mathbf{Ab}_1 \quad \dots \quad \mathbf{Ab}_p\,] \qquad \text{(según (11))}$$

$$\begin{aligned} \text{Columna } j \text{ de } \mathbf{AB} &= \mathbf{A}(\text{Columna } j \text{ de } \mathbf{B}) \\ &= \mathbf{Ab}_j \\ &= [\,\mathbf{a}_1 \quad \dots \quad \mathbf{a}_m\,] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{bmatrix} \\ &= \underbrace{b_{1j}\mathbf{a}_1 + \cdots + b_{mj}\mathbf{a}_m}_{\text{c.l. de las columnas de } \mathbf{A}}. \qquad \text{(según (10))} \end{aligned}$$

ii)

$$\begin{aligned} \text{Componente } (i,j) \text{ de } \mathbf{AB} &= i\text{-ésima componente de } \mathbf{Ab}_j \\ &= \boldsymbol{\alpha}_i \mathbf{b}_j \qquad \text{(ver (9))} \\ &= [\,a_{i1} \quad \dots \quad a_{im}\,] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{bmatrix} \\ &= \sum_{k=1}^{m} a_{ik} b_{kj} \end{aligned}$$

UNIVERSIDAD DE ... BIBLIOTECA

$$\mathbf{AB} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_i \\ \vdots \\ \boldsymbol{\alpha}_n \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_j & \dots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1\mathbf{b}_1 & \dots & \boldsymbol{\alpha}_1\mathbf{b}_j & \dots & \boldsymbol{\alpha}_1\mathbf{b}_p \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{\alpha}_i\mathbf{b}_1 & \dots & \boldsymbol{\alpha}_i\mathbf{b}_j & \dots & \boldsymbol{\alpha}_i\mathbf{b}_p \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{\alpha}_n\mathbf{b}_1 & \dots & \boldsymbol{\alpha}_n\mathbf{b}_j & \dots & \boldsymbol{\alpha}_n\mathbf{b}_p \end{bmatrix}$$

iii) Se puede probar que

$$\mathbf{AB} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_n \end{bmatrix} \mathbf{B} = \begin{bmatrix} \boldsymbol{\alpha}_1\mathbf{B} \\ \vdots \\ \boldsymbol{\alpha}_n\mathbf{B} \end{bmatrix}. \tag{12}$$

y que

$$\boldsymbol{\alpha}_i\mathbf{B} = [a_{i1} \quad \dots \quad a_{im}] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_m \end{bmatrix} = \underbrace{a_{i1}\boldsymbol{\beta}_1 + \dots a_{im}\boldsymbol{\beta}_m}_{\text{c.l. de las filas de } \mathbf{B}}. \tag{13}$$

Además que,

$$\mathbf{AB} = [\mathbf{a}_1 \quad \dots \quad \mathbf{a}_m] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_m \end{bmatrix} = \mathbf{a}_1\boldsymbol{\beta}_1 + \dots + \mathbf{a}_m\boldsymbol{\beta}_m. \tag{14}$$

En el ejercicio 1.1.1 se pide probar las igualdades (12), (13) y (14).

Propiedades básicas del producto

En las siguientes propiedades $\mathbf{A}, \mathbf{B}, \mathbf{C}$ denotan matrices con componentes en F y $k \in F$. Se supone que las operaciones que aparecen en cada propiedad están definidas.

1. Si $\mathbf{A}$ es $n \times m$ y $\mathbf{B}$ es $m \times p$ el producto $\mathbf{AB}$ es una matriz $n \times p$ sobre F
2. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
3. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
4. $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$
5. Si $\mathbf{A}$ es $n \times m$, $\mathbf{A}0_{m\times p} = 0_{n\times p}$ y $0_{k\times n}\mathbf{A} = 0_{k\times m}$
6. $k(\mathbf{AB}) = (k\mathbf{A})\mathbf{B} = \mathbf{A}(k\mathbf{B})$
7. La matriz diagonal $n \times n$ siguiente

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

es tal que si $\mathbf{A}$ es $n \times m$ y $\mathbf{B}$ es $k \times n$ entonces

$$\mathbf{I}_n\mathbf{A} = \mathbf{A} \quad \text{y} \quad \mathbf{B}\mathbf{I}_n = \mathbf{B}$$

$\mathbf{I}_n$ será llamada matriz **identidad** o matriz **unidad** de orden n.

Advertencias.

i) El producto de matrices no es una operación conmutativa. A continuación precisamos esta afirmación.

Digamos que $\mathbf{A}$ y $\mathbf{B}$ son matrices $n \times m$ y $m \times p$ respectivamente, de modo que el producto $\mathbf{AB}$ está definido.

Si $n \neq p$ el producto $\mathbf{BA}$ no está definido.

Si $n = p$ pero $p \neq m$ entonces $\mathbf{AB}$ y $\mathbf{BA}$ están definidos pero $\mathbf{AB}$ es $n \times n$ mientras que $\mathbf{BA}$ es $m \times m$.

El único caso en que $\mathbf{AB}$ y $\mathbf{BA}$ están definidos y son del mismo orden es cuando $n = p = m$. Resulta que aún en esta situación puede ocurrir $\mathbf{AB} \neq \mathbf{BA}$, como se muestra en el siguiente ejemplo:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

ii) Es posible que el producto de dos matrices no nulas sea la matriz nula. Por ejemplo

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Inversas

Sea $\mathbf{A}$ una matriz $n \times m$. Una matriz $\mathbf{G}$, $m \times n$, tal que $\mathbf{GA} = \mathbf{I}_m$ es llamada una **inversa a izquierda** de $\mathbf{A}$. Similarmente, una matriz $\mathbf{H}$, $m \times n$, tal que $\mathbf{AH} = \mathbf{I}_n$ es llamada una **inversa a derecha** de $\mathbf{A}$.

Teorema 1.1.1. *Sea $\mathbf{A}$ una matriz $n \times m$. Si $\mathbf{A}$ posee una inversa a izquierda $\mathbf{G}$ $(m \times n)$ y también una inversa a derecha $\mathbf{H}$ $(m \times n)$ entonces $\mathbf{G} = \mathbf{H}$.*

Prueba.

$$\mathbf{H} = \mathbf{I}_m\mathbf{H} = (\mathbf{GA})\mathbf{H} = \mathbf{G}(\mathbf{AH}) = \mathbf{G}\mathbf{I}_n = \mathbf{G}\,. \quad \blacksquare$$

Comentario.
Se puede probar (ver ejercicio 1.5.2) que si la matriz $\mathbf{A}$ posee inversa a izquierda y también inversa a derecha entonces $\mathbf{A}$ es necesariamente cuadrada.

Combinando este hecho con el resultado del teorema anterior obtenemos, entre otras cosas, que una matriz cuadrada tiene a lo más una inversa bilateral (a ambos lados). En el caso de poseer una, esa única inversa bilateral se llamará la **inversa** de $\mathbf{A}$, la notaremos $\mathbf{A}^{-1}$ y diremos que $\mathbf{A}$ es **invertible** o **no singular**.

De manera que, $\mathbf{A}^{-1}$ (la inversa de $\mathbf{A}$), cuando existe, es la única matriz tal que

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n = \mathbf{A}^{-1}\mathbf{A}$$

si $\mathbf{A}$ es $n \times n$.

Es posible que una matriz cuadrada posea inversas sólo a un lado? El siguiente resultado nos dice que ésto no es posible.

Teorema 1.1.2. *Si $\mathbf{A}$ es una matriz cuadrada, entonces $\mathbf{A}$ posee una inversa a ambos lados $\mathbf{A}^{-1}$ o no posee ni inversas a izquierda, ni inversas a derecha. Equivalentemente si la matriz cuadrada $\mathbf{A}$ posee una inversa $\mathbf{B}$ a algún lado entonces $\mathbf{A}$ es invertible y $\mathbf{B} = \mathbf{A}^{-1}$.*

(Más adelante se puede probar este resultado. Ver ejercicio 1.5.2).

Teorema 1.1.3. *Si $\mathbf{A}$, $\mathbf{B}$ son matrices de orden n invertibles y k un escalar, $k \neq 0$, entonces*

1. *$\mathbf{A}^{-1}$ es invertible y $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$*
2. *$\mathbf{AB}$ es invertible y $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$*
3. *$k\mathbf{A}$ es invertible y $(k\mathbf{A})^{-1} = k^{-1}\mathbf{A}^{-1}$.*

Transpuesta, conjugada y trauspuesta hermitiana

Sea $\mathbf{A} = [a_{ij}]$ una matriz $n \times m$.

La **transpuesta** de $\mathbf{A}$, notada $\mathbf{A}^T$, es la matriz $m \times n$ cuya componente (i, j) es a_{ji}; es decir, la transpuesta de $\mathbf{A}$ es la matriz $m \times n$ $\mathbf{A}^T = [b_{ij}]$ en la que $b_{ij} = a_{ji}$ para cada $i \in \{1, \ldots, m\}$ y cada $j \in \{1, \ldots, n\}$. (Las columnas de $\mathbf{A}^T$ son las filas de $\mathbf{A}$, y las filas de $\mathbf{A}^T$ son las columnas de $\mathbf{A}$).

La **conjugada** de $\mathbf{A}$, notada $\overline{\mathbf{A}}$, es la matriz $n \times m$ cuya componente (i, j) es $\overline{a_{ij}}$ (el **conjugado** del número a_{ij}): $\overline{\mathbf{A}} = (\overline{a_{ij}})$.

La matriz $\overline{\mathbf{A}^T}$ (la conjugada de la transpuesta de $\mathbf{A}$) la llamaremos **transpuesta hermitiana de A**; la notaremos $\mathbf{A}^H$. La componente (i,j) de $\mathbf{A}^H$ es $\overline{a_{ji}}$.

Cuando $\mathbf{A}$ es real, es decir, todas sus componentes son reales, $\mathbf{A}^T$ es real, $\overline{\mathbf{A}} = \mathbf{A}$ y $\mathbf{A}^H = \mathbf{A}^T$.

Si $\mathbf{A} = [\mathbf{a}_1 \quad \dots \quad \mathbf{a}_m]$ ($\mathbf{a}_j \in F^n$ es la j-ésima columna de $\mathbf{A}$), entonces

$$\mathbf{A}^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}, \qquad \overline{\mathbf{A}} = [\overline{\mathbf{a}_1} \quad \dots \quad \overline{\mathbf{a}_m}], \qquad \mathbf{A}^H = \begin{bmatrix} \mathbf{a}_1^H \\ \vdots \\ \mathbf{a}_m^H \end{bmatrix}.$$

Propiedades

Para $\mathbf{A}, \mathbf{B}$ matrices con componentes complejas se tiene:

1. $(\mathbf{A}^T)^T = \mathbf{A}$. $\overline{\overline{\mathbf{A}}} = \mathbf{A}$ y $(\mathbf{A}^H)^H = \mathbf{A}$
2. $(\mathbf{A} \pm \mathbf{B})^T = \mathbf{A}^T \pm \mathbf{B}^T$, $\overline{(\mathbf{A} \pm \mathbf{B})} = \overline{\mathbf{A}} \pm \overline{\mathbf{B}}$ y $(\mathbf{A} \pm \mathbf{B})^H = \mathbf{A}^H \pm \mathbf{B}^H$
3. $(k\mathbf{A})^T = k\mathbf{A}^T$, $\overline{k\mathbf{A}} = \overline{k}\,\overline{\mathbf{A}}$ y $(k\mathbf{A})^H = \overline{k}\mathbf{A}^H$.
4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$, $\overline{\mathbf{AB}} = \overline{\mathbf{A}}\,\overline{\mathbf{B}}$ y $(\mathbf{AB})^H = \mathbf{B}^H\mathbf{A}^H$
5. Si $\mathbf{A}$ es invertible entonces $\mathbf{A}^T$, $\overline{\mathbf{A}}$ y $\mathbf{A}^H$ también lo son. Además

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T, \quad (\overline{\mathbf{A}})^{-1} = \overline{(\mathbf{A}^{-1})} \quad \text{y} \quad (\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H.$$

Matrices simétricas y hermitianas

Sea $\mathbf{A} = [a_{ij}]$ una matriz cuadrada, $n \times n$.

Si $\mathbf{A} = \mathbf{A}^T$, diremos que $\mathbf{A}$ es **simétrica** (en este caso $a_{ij} = a_{ji}$, $\forall i, j \in \{1, \dots, n\}$).

Si $\mathbf{A} = \mathbf{A}^H$, diremos que $\mathbf{A}$ es **hermitiana** (esta vez $a_{ij} = \overline{a_{ji}}$, $\forall i, j \in \{1, \dots, n\}$).

Cuando $\mathbf{A}$ es real, $\mathbf{A}$ es hermitiana si y sólo si $\mathbf{A}$ es simétrica.

Ejemplo 1.1.1

$\mathbf{A} = \begin{bmatrix} i & 2 \\ 2 & 1 \end{bmatrix}$ es simétrica, porque $\mathbf{A}^T = \mathbf{A}$; pero $\mathbf{A}$ no es hermitiana.

Cómo deben ser los elementos en la diagonal principal de una matriz hermitiana? ▲

Ejemplo 1.1.2

Sea $\mathbf{A}$ una matriz $n \times m$. Como $\mathbf{A}^T$ es $m \times n$ entonces $\mathbf{A}^T\mathbf{A}$ es $m \times m$ y $\mathbf{A}\mathbf{A}^T$ es $n \times n$:

además, $\mathbf{A}^T\mathbf{A}$ y $\mathbf{A}\mathbf{A}^T$ son ambas simétricas:

$$(\mathbf{A}\mathbf{A}^T)^T = (\mathbf{A}^T)^T\mathbf{A}^T = \mathbf{A}\mathbf{A}^T \quad \text{y} \quad (\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^T)^T = \mathbf{A}^T\mathbf{A}\,.$$

Análogamente, las matrices $\mathbf{A}^H\mathbf{A}$ y $\mathbf{A}\mathbf{A}^H$ son ambas hermitianas. ▲

Matrices y sistemas de ecuaciones lineales. Método de eliminación de Gauss.

Consideremos el siguiente sistema de n ecuaciones lineales, en las m incógnitas $x_1, \ldots, x_m$

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + \ldots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + \ldots + a_{nm}x_m &= b_n \end{aligned} \tag{15}$$

en el que los coeficientes a_{ij} y los términos independientes b_i están en F. Este sistema puede escribirse en la forma matricial

$$\begin{bmatrix} a_{11} & \ldots & a_{1m} \\ a_{21} & \ldots & a_{2m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{16}$$

esto es, en la forma

$$\mathbf{A}\mathbf{x} = \mathbf{b}\,. \tag{17}$$

donde $\mathbf{A} = [a_{ij}]$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

Nos referiremos a $\mathbf{A}$ como la **matriz de coeficientes** del sistema (15), y a la matriz

$$[\mathbf{A} \vdots \mathbf{b}] = \begin{bmatrix} a_{11} & \ldots & a_{1m} & \vdots & b_1 \\ a_{21} & \ldots & a_{2m} & \vdots & b_2 \\ \vdots & \ldots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & a_{nm} & \vdots & b_n \end{bmatrix}$$

como la **matriz aumentada** o **ampliada** de ese sistema. Esta última matriz determina completamente el sistema (15).

Una **solución** del sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ (en F^m) es cualquier vector $\mathbf{x}^* \in F^m$ tal que $\mathbf{A}\mathbf{x}^* = \mathbf{b}$. Cuando el sistema tenga soluciones diremos que es **soluble** o **consistente**, en caso contrario diremos que es **no soluble** o **inconsistente**.

Si $\mathbf{b} = \mathbf{0}$ el sistema se dice **homogéneo**. En este caso $\mathbf{x}^* = \mathbf{0}$ es una solución que llamaremos **solución trivial**.

Resolver un sistema consiste en encontrar todas sus soluciones. Algunos sistemas pueden resolverse fácilmente, como se ilustra en los ejemplos siguientes.

Ejemplo 1.1.3
Consideremos el sistema

$$\begin{aligned} x_1 \qquad\qquad &= 4 \\ -2x_1 + x_2 \qquad &= 2 \\ 7x_1 - 2x_2 + x_3 &= 8 \end{aligned}$$

De la primera ecuación se observa que debe ser $x_1 = 4$; sustituyendo este valor de x_1 en la segunda ecuación se obtiene $x_2 = 4$; finalmente sustituyendo en la tercera ecuación x_1 por 4 y x_2 por 10 y despejando x_3, obtenemos $x_3 = 0$. Diremos que el sistema ha sido resuelto por **sustitución hacia adelante** o **progresiva**. La matriz aumentada del sistema es

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ -2 & 1 & 0 & 2 \\ 7 & -2 & 1 & 8 \end{array}\right] \qquad \blacktriangle$$

Ejemplo 1.1.4
Consideremos el sistema

$$\begin{aligned} x_1 + x_2 - x_3 &= 3 \\ 2x_2 - x_3 &= 7 \\ 2x_3 &= 6 \end{aligned}$$

Despejando x_3 en la tercera ecuación obtenemos $x_3 = 3$; sustituyendo en la segunda ecuación x_3 por su valor y despejando x_2 obtenemos $x_2 = 5$; finalmente sustituyendo en la primera ecuación x_3 por 3 y x_2 por 5 obtenemos $x_1 = -5$. Por tanto la única solución del sistema es $x_1 = -5$, $x_2 = 5$, $x_3 = 3$.

Observe que la incógnita x_1 solamente aparece en la primera ecuación y que x_2 aparece en las dos primeras ecuaciones pero no en la tercera. La matriz aumentada del sistema es

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 2 & -1 & 7 \\ 0 & 0 & 2 & 6 \end{array}\right] \qquad \blacktriangle$$

Ejemplo 1.1.5
Consideremos ahora el sistema

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ x_2 + x_3 &= 1 \end{aligned}$$

En éste sistema se observa que podemos dar a x_3 cualquier valor y que por cada valor que le demos obtenemos correspondientes soluciones del sistema: si $x_3 = t$ (t cierto número real), despejando x_2 en la segunda ecuación se obtiene $x_2 = 1 - t$; finalmente llevando los valores de x_2 y x_3 a la primera ecuación se obtiene $x_1 = 2 - t$.

Así que las triplas

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2-t \\ 1-t \\ t \end{bmatrix}, \quad t \in \mathbb{R}$$

son soluciones (en $\mathbb{R}^3$) del sistema. Es fácil verificar que toda solución (en $\mathbb{R}^3$) es de la forma anterior, por lo que esa expresión describe el conjunto de todas las soluciones. Tal expresión puede escribirse también así:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Observe que en éste sistema de dos ecuaciones, la variable x_1 aparece solamente en la primera ecuación. La matriz aumentada del sistema es

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \end{array}\right] \qquad \blacktriangle$$

Nos referiremos al método empleado en los ejemplos 1.1.4 y 1.1.5, para resolver el sistema dado, como **sustitución hacia atrás** o **regresiva**.

Es claro que no todos los sistemas pueden ser resueltos utilizando procedimientos como los usados en los ejemplos discutidos. En general, para resolver un sistema que no tiene forma simple, el procedimiento básico es transformarlo, mediante ciertas manipulaciones de las ecuaciones, en un sistema **equivalente** (es decir, con las mismas soluciones) más simple (esto es, en uno cuyas soluciones puedan encontrarse más fácilmente).

Las siguientes manipulaciones o transformaciones, llamadas **operaciones elementales**, realizadas sobre las ecuaciones de un sistema de ecuaciones lineales, no alteran el conjunto de soluciones de tal sistema:

i) Intercambiar dos ecuaciones

ii) Multiplicar una ecuación por un escalar no nulo

iii) Sumar a una ecuación un múltiplo de otra

En lugar de trabajar con las ecuaciones del sistema es más cómodo hacerlo con las filas de su matriz ampliada. A las operaciones anteriores entre ecuaciones corresponden las siguientes entre filas de una matriz:

i) Intercambiar dos filas

ii) Multiplicar una fila por un escalar no nulo

iii) Sumar a una fila un múltiplo de otra

Nos referiremos a estas últimas como **operaciones elementales fila**. En el siguiente ejemplo se ilustra el uso de estas operaciones para transformar un sistema dado en uno equivalente más simple.

Ejemplo 1.1.6
Resolver el sistema

$$\begin{array}{rrrrl} x_1 & +2x_2 & +3x_3 & = & 4 \\ x_1 & +3x_2 & +4x_3 & = & 5 \\ -x_1 & & -x_3 & = & -2 \end{array} \tag{18}$$

Solución. Trabajaremos con la matriz aumentada del sistema

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 5 \\ -1 & 0 & -1 & -2 \end{array}\right] \tag{19}$$

Nuestro primer propósito es obtener ceros en la primera columna debajo de la primera fila (esto es, **eliminar** x_1 en las ecuaciones segunda y tercera) lo que se consigue realizando las siguientes operaciones:

"sustraer de la segunda fila en (19) la primera", y

"sumar a la tercera fila en (19) la primera"

obteniendose la matriz

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \end{array}\right] \tag{20}$$

Queremos ahora obtener cero en la segunda columna, debajo de la segunda fila (esto es, **eliminar** x_2 en la tercera ecuación del sistema representado por (20)) para lo cual basta

"sustraer de la tercera fila en (20), dos veces la segunda"

obteniéndose la matriz

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{21}$$

El sistema (18) es equivalente al sistema representado por (21):

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ x_2 + x_3 &= 1 \\ 0x_1 + 0x_2 + 0x_3 &= 0 \end{aligned} \tag{22}$$

Como la última ecuación en (22) se satisface cualesquiera sean los valores que demos a x_1, x_2, x_3, este sistema es equivalente al sistema

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ x_2 + x_3 &= 1 \end{aligned} \tag{23}$$

cuyo conjunto de soluciones puede encontrarse por sustitución regresiva, como se mostró en el ejemplo 1.1.5. ▲

El procedimiento usado para transformar el sistema (18) en el sistema (22), se conoce como **Método de Eliminación de Gauss** o **Método de Eliminación Gaussiana**. En términos de matrices el propósito fundamental de este método es transformar la matriz aumentada $[\mathbf{A} \vdots \mathbf{b}]$ de un sistema dado $\mathbf{Ax} = \mathbf{b}$, en una matriz con **forma escalonada** $[\mathbf{U} \vdots \mathbf{c}]$ (como la matriz en (21)). El sistema $\mathbf{Ux} = \mathbf{c}$, equivalente al sistema $\mathbf{Ax} = \mathbf{b}$, se resuelve por sustitución regresiva.

A continuación precisamos lo que entenderemos por "estar en forma escalonada".

Una matriz $\mathbf{U}$, $n \times m$, se dice que está en **forma escalonada** si satisface las siguientes condiciones, en las que llamamos **pivote** a la primera componente distinta de cero en cada fila no nula de $\mathbf{U}$:

- Primero aparecen las filas no nulas
- Debajo de cada pivote, en la columna correspondiente, todas las componentes son nulas
- El pivote de cada fila no nula está a la derecha del pivote de la fila anterior, es decir, el número de ceros al comienzo de una fila aumenta a medida que descendemos.

En el ejemplo 1.1.6 no necesitamos intercambios de filas para lograr el propósito buscado, pero en ciertos casos puede necesitarse de tales intercambios. Operaciones del tipo "multiplicar una fila por un escalar no nulo" no son necesarias; pueden o nó usarse.

En un sistema $\mathbf{Ux} = \mathbf{c}$ con $[\mathbf{U} \vdots \mathbf{c}]$ en forma escalonada, las variables correspondientes a las columnas con pivotes las llamaremos **básicas** y a las otras **libres**. (En el ejemplo 1.1.5 x_3 es libre y x_1, x_2 son básicas). Al hacer sustitución regresiva, podemos dar a las variables libres cualquier valor y despejar luego las básicas (así se procedió en el ejemplo 1.1.5).

Un refinamiento del Método de Eliminación de Gauss es el **Método de Eliminación de Gauss-Jordan** en el cual el proceso de eliminación se lleva hasta el punto en que la matriz aumentada del sistema, además de estar en forma escalonada, tiene todos sus pivotes iguales a uno y ceros encima de cada pivote en las columnas correspondientes a éstos. Una matriz con estas características de dice que está en **forma escalonada reducida**.

Un tratamiento detallado del Método de Eliminación de Gauss se encuentra, por ejemplo, en [*Noble-Daniel*], [*Strang*], [*Res-Fran-Mu*], [*Florey*], [*Grossman*].

Operaciones elementales y matrices elementales

Ya nos hemos referido antes a las denominadas operaciones elementales sobre las filas de una matriz. Si en ellas se sustituye "fila" por "columna" obtenemos las correspondientes operaciones elementales sobre las columnas:

i) Intercambiar dos columnas.

ii) Multiplicar una columna por un escalar no nulo

iii) Sumar a una columna un múltiplo escalar de otra.

Las matrices obtenidas de una matriz unidad $\mathbf{I}_n$ realizando una operación elemental sobre sus filas (columnas) se llaman **matrices elementales fila (columna)**. Todas ellas son invertibles y sus inversas son también matrices elementales.

Resulta que realizar una operación elemental sobre las filas (columnas) de una matriz $\mathbf{A}$ equivale a multiplicar a izquierda (derecha) de $\mathbf{A}$ por la correspondiente matriz elemental.

Ejemplo 1.1.7

Consideremos la matriz aumentada $[\mathbf{A} \vdots \mathbf{b}]$ en (19), ejemplo 1.1.6.

Las matrices elementales correspondientes a las operaciones elementales fila realizadas sobre tal matriz (en el orden en que allí se hicieron) son

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Si denotamos $[\mathbf{U} \vdots \mathbf{c}]$ la matriz aumentada del sistema "reducido" (21), entonces

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1[\mathbf{A} \vdots \mathbf{b}] = [\mathbf{U} \vdots \mathbf{c}].$$

Si $\mathbf{P} = \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$, entonces

$$\mathbf{P}[\mathbf{A} \vdots \mathbf{b}] = [\mathbf{U} \vdots \mathbf{c}].$$

En particular se tiene que

$$\mathbf{PA} = \mathbf{U} \quad \text{y} \quad \mathbf{Pb} = \mathbf{c}. \quad \blacktriangle$$

Ejercicios 1.1

1.1.1 Probar las igualdades (12), (13) y (14).

1.1.2 Sea $\mathbf{A} \in \mathcal{M}_{n\times m}(F)$. Probar que si $\mathbf{Ax} = 0$, para cada $\mathbf{x} \in F^m$, entonces $\mathbf{A} = \mathbf{0}$.

1.1.3 Sean $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times m}(F)$. Probar que si $\mathbf{Ax} = \mathbf{Bx}$, para cada $\mathbf{x} \in F^m$, entonces $\mathbf{A} = \mathbf{B}$.

1.1.4 a) Suponga que cada columna de una matriz $\mathbf{C}$ es *c.l.* de las columnas de una matriz $\mathbf{A}$. Probar que $\mathbf{C} = \mathbf{AB}$ para alguna matriz $\mathbf{B}$.

b) Suponga que cada fila de una matriz $\mathbf{C}$ es *c.l.* de las filas de una matriz $\mathbf{B}$. Probar que $\mathbf{C} = \mathbf{AB}$ para alguna matriz $\mathbf{A}$.

1.1.5 Sea $\mathbf{A} = [a_{ij}] \in \mathcal{M}_{n\times n}(F)$. La **traza** de $\mathbf{A}$, $Tr(\mathbf{A})$, es la suma de los elementos sobre la diagonal principal de $\mathbf{A}$:

$$Tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Probar que $\forall \mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times n}(F)$, $\forall k \in F$

a) $Tr(\mathbf{A} + \mathbf{B}) = Tr(\mathbf{A}) + Tr(\mathbf{B})$

b) $Tr(k\mathbf{A}) = kTr(\mathbf{A})$

c) $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$

d) $Tr(\mathbf{P}^{-1}\mathbf{AP}) = Tr(\mathbf{A})$ si $\mathbf{P} \in \mathcal{M}_{n\times n}(F)$ es invertible

e) $Tr(\mathbf{A}^T) = Tr(\mathbf{A})$.

1.1.6 a) Sea $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{R})$. Probar que

i) $\mathbf{A}^T\mathbf{A} = 0 \Longrightarrow \mathbf{A} = 0$

ii) $\mathbf{AA}^T = 0 \Longrightarrow \mathbf{A} = 0$

iii) $T_r(\mathbf{A}^T\mathbf{A}) = 0 \Longrightarrow \mathbf{A} = 0$.

b) Exhibir una matriz no nula $\mathbf{A}$ y no real para la cual $\mathbf{A}^T\mathbf{A} = \mathbf{0}$.

c) Sea $\mathbf{A} \in \mathcal{M}_{n\times m}(\mathbb{C})$.

Probar que

i) $\mathbf{A}^H\mathbf{A} = 0 \Longrightarrow \mathbf{A} = 0$

ii) $\mathbf{A}\mathbf{A}^H = 0 \Longrightarrow \mathbf{A} = 0$

iii) $T_r(\mathbf{A}^H\mathbf{A}) = 0 \Longrightarrow \mathbf{A} = 0$.

1.1.7 (Producto directo de matrices)

Sean $\mathbf{A}$ y $\mathbf{B}$ matrices $n \times m$ y $p \times q$ respectivamente. Si $\mathbf{A} = [a_{ij}]$, el **producto directo** de $\mathbf{A}$ por $\mathbf{B}$ es la matriz

$$\mathbf{C} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ \vdots & & \vdots \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{bmatrix}.$$

Notamos $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$. Observe que $\mathbf{C}$ está particionada en tantas submatrices como elementos tiene $\mathbf{A}$, siendo cada submatriz de la forma $a_{ij}\mathbf{B}$. En forma abreviada podemos escribir

$$\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}].$$

Este producto es también llamado **producto de Kronecker** o **producto de Zehfuss**. Está definido para cualquier par de matrices $\mathbf{A}$, $\mathbf{B}$ y no es una operación conmutativa.

Probar que

a) $\mathbf{I}_m \otimes \mathbf{A} = \begin{bmatrix} \mathbf{A} & 0 & \dots & 0 \\ 0 & \mathbf{A} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{A} \end{bmatrix}$

b) $\mathbf{I}_m \otimes \mathbf{I}_n = \mathbf{I}_{mn}$

c) Si k es un escalar,

$$(k\mathbf{A}) \otimes \mathbf{B} = k(\mathbf{A} \otimes \mathbf{B}) = \mathbf{A} \otimes (k\mathbf{B})$$

d) $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$

d) $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$

f) Si $\mathbf{A}$ y $\mathbf{B}$ son invertibles, $\mathbf{A} \otimes \mathbf{B}$ es invertible y

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}.$$