B. Anexo: Álgebra lineal

B.1. Sistemas lineales y matrices.

Diremos que un sistema de n ecuaciones y m incógnitas es un sistema lineal si dicho sistema tiene la forma:

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nm}x_m = b_n \end{array}\right. \tag{B.1}$$

Una matriz de dimensiones $n \times m$ no es más que una *"caja rectangular"* con n filas y m columnas.

El sistema de ecuaciones (B.1) está completamente caracterizado por su matriz asociada que no es más que la matriz en cuya fila i ($i = 1, 2, ..., n$) están colocados los coeficientes a_{ij} ($j = 1, 2, ..., m$) y el término independiente b_i correspondientes a la i-ésima ecuación del sistema (B.1). De esta forma la matriz asociada al sistema (B.1) tendrá la forma

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} & b_n \end{pmatrix}, \tag{B.2}$$

y se denomina *matriz ampliada del sistema*. A la matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{pmatrix}, \tag{B.3}$$

se le denomina *matriz de coeficientes del sistema*.

Un sistema lineal de la forma (B.1) es *compatible* si existen los valores de $x_1, x_2, ..., x_m$ tales que se cumplan todas y cada una de las ecuaciones del sistema (B.1). Si existen dichos valores diremos que el sistema tiene solución. Si un sistema lineal de la forma (B.1) no tiene solución decimos que es un sistema *incompatible*. Un sistema compatible que tiene una única solución se llama *determinado* y si tiene infinitas soluciones *indeterminado*.

Dos sistemas de ecuaciones se denominan equivalentes si tienen el mismo conjunto de soluciones.

Un sistema de ecuaciones lineales se puede transformar en otro sistema equivalente mediante las siguientes operaciones elementales:

1. Intercambiar dos ecuaciones.

2. Multiplicar por una constante no nula una ecuación cualquiera.

3. Reemplazar una ecuación por la suma de ésta mas un múltiplo de cualquier otra ecuación.

Dada la equivalencia de los sistemas lineales con sus matrices asociadas (las matrices ampliadas) las operaciones anteriores tienen su análogo al operar con las matrices. De forma que a cada una le corresponden las siguientes operaciones por filas

1. Intercambiar dos filas.

2. Multiplicar por una constante no nula una fila cualquiera.

3. Reemplazar una fila por la suma de ésta mas un múltiplo de cualquier otra fila.

Las operaciones anteriores se denominan *operaciones elementales de filas.*

Diremos que dos matrices son equivalentes por filas si a cualquiera de ellas la podemos transformar en la otra utilizando sólo operaciones elementales de filas. Es importante recordar que las operaciones elementales de filas son *reversibles* y por tanto las operaciones elementales descritas no cambian la solución del sistema original de ecuaciones.

B.2. Formas escalonadas y el algoritmo de reducción por filas.

En adelante diremos que una fila es una fila no nula si tiene al menos un elemento diferente de cero. Definiremos elemento guía de una fila al primer elemento (por la izquierda) no nulo de la fila.

Diremos que una matriz está en *forma escalonada* si:

1. Todas las filas no nulas están por encima de las nulas.
2. El *elemento guía* de cada fila se encuentra siempre en una columna a la derecha del elemento guía de la fila anterior.
3. Debajo de cada elemento guía sólo puede haber ceros.

Las matrices escalonadas tienen típicamente la forma:

$$\begin{pmatrix} \blacksquare & * & * & * & * & * & * \\ 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \blacksquare & * & * & * & * & * \\ 0 & \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \end{pmatrix}, \tag{B.4}$$

donde $\blacksquare$ denota a los elementos guías de cada fila. Si exigimos ahora que los elementos guías sean todos iguales a 1, y que además sean los únicos elementos diferentes de cero de la columna entonces diremos que la matriz escalonada está en forma reducida. Utilizando las formas escalonadas anteriores (B.4) tenemos que sus correspondientes formas reducidas son:

$$\begin{pmatrix} 1 & 0 & * & * & 0 & 0 & * \\ 0 & 1 & * & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & * \\ 0 & 1 & 0 & 0 & 0 & * \\ 0 & 0 & 1 & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \end{pmatrix}. \tag{B.5}$$

Cualquier matriz puede ser transformada mediante transformaciones elementales de fila en una matriz con forma escalonada no teniendo siempre ésta última una forma única. Por el contrario las formas escalonadas reducidas son únicas, es decir se cumple el siguiente teorema:

Teorema B.1 *Cualquier matriz es equivalente por filas a una única matriz con forma escalonada reducida.* $\blacksquare$

Una consecuencia del teorema anterior es que los elementos guías están siempre en la misma posición en cualquiera de las formas escalonadas obtenidas de una matriz dada mediante transformaciones elementales de fila. Dichas posiciones se denominan *pivotes* y las columnas donde se encuentran se denominan *columnas pivotes.* Así, por ejemplo

$$
\begin{array}{c}
\begin{array}{ccccccc} \text{pivote} & & & & & & \\ \downarrow & & & & & & \end{array}\\
\begin{pmatrix}
\blacksquare & * & * & * & * & * & * \\
0 & \blacksquare & * & * & * & * & * \\
0 & 0 & 0 & 0 & \blacksquare & * & * \\
0 & 0 & 0 & 0 & 0 & \blacksquare & * \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}\\
\begin{array}{ccccccc} \uparrow & \uparrow & & & \uparrow & \uparrow & \end{array}\\
\text{columnas pivote}
\end{array}. \tag{B.6}
$$

Algoritmo de reducción por filas.

El algoritmo de reducción por filas consta de dos fases (partes). La primera consiste en transformar una matriz cualquiera en una equivalente por filas con forma escalonada. Esta fase se le denomina *fase hacia abajo* o *subproceso hacia abajo.* La segunda parte consiste en transformar la forma escalonada en la forma reducida. A la segunda parte se le denomina *fase hacia arriba* o *subproceso hacia arriba.*

Para obtener la matriz en forma escalonada realizamos lo siguiente:

Subproceso hacia abajo. Este proceso consta de cuatro pasos:

1. Se comienza por la primera columna de la izquierda (generalmente es una columna pivote) y se selecciona un elemento diferente de cero. Este será el primer pivote.

2. Se coloca el elemento seleccionado en el extremo superior izquierdo intercambiando, si es preciso, las filas de la matriz hasta que dicho elemento se encuentre en la posición de pivote.

3. Utilizando las operaciones elementales de filas crear ceros debajo del pivote.

4. *Tapamos* la fila (o filas si hay más de una) con la posición pivote creada en los pasos anteriores (1-3) y aplicamos los pasos 1-3 a la submatriz que nos queda (constituida por las filas restantes). Este paso se repite hasta que no queden más filas no nulas.

Para obtener la matriz en forma escalonada reducida a partir de la forma reducida obtenida mediante los pasos 1-4 debemos realizar lo siguiente:

Subproceso hacia arriba. Este proceso consta del siguiente paso:

5. Comenzando por el primer pivote de la derecha creamos, mediante operaciones elementales de filas, ceros encima del mismo. Si el pivote no es uno dividimos por su valor para que lo sea. Continuamos el proceso hacia la izquierda hasta terminar con el primer pivote de la izquierda.

B.3. Solución de un sistema de ecuaciones lineales.

La solución de un sistema de ecuaciones lineales la podemos encontrar a partir de la matriz asociada al sistema. Para ello primero transformamos la matriz en una equivalente por filas que tenga forma escalonada o escalonada reducida. Las variables correspondientes a las columnas pivote se denominan *variables principales o básicas.* El resto de las variables se denominan *variables libres.* El significado de una variable libre es que dicha variable puede tomar cualquier valor. La solución de un sistema viene dada al escribir el valor de las variables principales, las cuales pueden venir expresadas en función de las variables libres.

Por ejemplo, consideremos la matriz del sistema (B.6). Dicha matriz representa un sistema de 5 ecuaciones con 6 incógnitas. Las variables principales serán x_1, x_2, x_5, x_6 y las libres x_3, x_4. De la forma de la matriz observamos que x_5, x_6 están completamente determinados por los términos independientes (recordar que la última columna está compuesta por los términos independientes de las ecuaciones) pero x_1, x_2 dependen tanto de los términos independiente como de las variables libres x_3, x_4. Es decir fijando diferentes valores de x_3, x_4 obtenemos diferentes valores de x_1, x_2. Por ejemplo, el sistema (x_1 depende de la variable libre x_3)

$$\begin{pmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ tiene la solución } \begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 \\ x_3 \text{ es libre} \end{cases} \tag{B.7}$$

De todo lo anterior deducimos el siguiente teorema de existencia y unicidad:

Teorema B.2 *Un sistema de ecuaciones lineales en compatible si y sólo si la primera columna de la derecha de la matriz ampliada del sistema no es una columna pivote, o sea, no existe ninguna fila de la forma* $(0\ 0\ 0\ \cdots\ 0\ \alpha)$ *con* α *diferente de cero. Si el sistema es compatible entonces tiene solución única (sistema compatible determinado) si no hay ninguna variable libre y tiene infinitas soluciones (sistema compatible indeterminado) si existe al menos una variable libre.* ■

Por ejemplo, los sistemas correspondientes a las dos matrices (B.4) son compatibles siendo el correspondiente a la primera matriz indeterminado (tiene dos variables libres) y el segundo determinado (no tiene variables libres). Por el contrario el sistema correspondiente a la matriz

$$\begin{pmatrix} \blacksquare & * & * & * & * & * & * \\ 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{B.8}$$

es incompatible al ser la última columna una columna pivote, o lo que es igual, al tener la fila $(0\ 0\ 0\ \cdots\ 0\ \blacksquare)$ (recordar que $\blacksquare$ denotaba al elemento guía que por definición es diferente de cero).

B.4. Vectores y ecuaciones matriciales.

B.4.1. Vectores de $\mathbb{R}^n$

Un vector de $\mathbb{R}^n$ no es más que un conjunto de n números ordenados de la forma

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \text{donde } x_1, ..., x_n \text{ son números reales.}$$

es decir un vector es una matriz de $n \times 1$. Los valores $x_1, ..., x_n$ se denominan coordenadas (o componentes) del vector x. Diremos que dos vectores son iguales si todas sus componentes son iguales, o sea

$$x = y \quad \Longleftrightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \Longleftrightarrow \quad x_1 = y_1,\ ..., x_n = y_n.$$

Llamaremos vector nulo 0 al vector cuyas componentes son todas iguales a cero.

B.4.2. Operaciones con vectores.

Sean x, y, z, w vectores de $\mathbb{R}^n$ y α, β números reales. Definiremos la suma de dos vectores x e y al vector z de $\mathbb{R}^n$ cuyas coordenadas son la suma de las coordenadas de x e y, o sea,

$$z = x + y \qquad \Longrightarrow \qquad \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

y definiremos la multiplicación de un vector x por un escalar (número) α al vector w cuyas componentes se obtienen al multiplicar las componentes de x por α, o sea,

$$w = \alpha\, x \qquad \Longrightarrow \qquad \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Dichas operaciones cumplen las propiedades

1. $x + y = y + x$
2. $(x + y) + z = x + (y + z)$
3. $x + 0 = 0 + x = x$
4. $x + (-x) = (-x) + x = 0$ donde $(-x)$ denota al vector $(-1)\,x$
5. $\alpha(x + y) = \alpha x + \alpha y$
6. $(\alpha + \beta)x = \alpha x + \beta x$
7. $\alpha(\beta x) = (\alpha\beta)x$
8. $1\,x = x$

Diremos que un vector x de $\mathbb{R}^m$ es combinación lineal de los vectores $a_1, a_2, ..., a_n$ de $\mathbb{R}^m$ con pesos $\alpha_1, ..., \alpha_n$, números reales, respectivamente si x se expresa de la forma

$$x = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n.$$

Definiremos espacio generado por los vectores $a_1, a_2, ..., a_n$ de $\mathbb{R}^m$ y lo denotaremos como $\mathrm{span}\,(a_1, ..., a_n) = L(a_1, ..., a_n)$ al subespacio de $\mathbb{R}^m$ constituido por todos las combinaciones lineales de los vectores $a_1, a_2, ..., a_n$, es decir, al conjunto de todos los vectores de la forma

$$L(a_1, ..., a_n) = \mathrm{span}\,(a_1, ..., a_n) = \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n,$$

quienes quiera sean $\alpha_1, ..., \alpha_n$ números reales.

Definiremos una ecuación vectorial a la ecuación de la forma

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b,$$

donde b es un vector dado de $\mathbb{R}^m$ y $x_1,\ x_2, ..., x_n$ son las incógnitas que se quieren encontrar. Es fácil comprobar que la ecuación vectorial anterior tiene la misma solución que el sistema de ecuaciones lineales cuya matriz ampliada tiene la forma

$$[a_1\, a_2\, \cdots a_n\, b].$$

Además b está en $\mathrm{span}\,(a_1, ..., a_n)$ si y sólo si el sistema $[a_1\, a_2\, \cdots a_n\, b]$ es compatible.

B.4.3. La ecuación $Ax = b$.

Sea A una matriz $m \times n$ con columnas $a_1, a_2, ..., a_n$ (vectores de $\mathbb{R}^m$) y x un vector de $\mathbb{R}^n$. Definiremos la multiplicación de una matriz A de $m \times n$ por un vector x de $\mathbb{R}^n$ al vector z de $\mathbb{R}^m$ definido por

$$z = Ax = [a_1\, a_2\, \cdots a_n] \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n.$$

Teorema B.3 *Sea A una matriz $m \times n$ con columnas $a_1, a_2, ..., a_n$ y b un vector de $\mathbb{R}^m$. La ecuación matricial*

$$A\,x = b$$

tiene la misma solución que la ecuación vectorial

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b,$$

la cual a su vez tiene la misma solución que el sistema cuya matriz ampliada tiene la forma

$$[a_1\, a_2\, \cdots a_n\, b].$$

■

Una consecuencia de este teorema es que la ecuación $A\,x = b$ tiene solución si y sólo si b es una combinación lineal de las columnas de A.

Teorema B.4 *Sea A una matriz $m \times n$ con columnas $a_1, a_2, ..., a_n$. Las siguientes tres afirmaciones son equivalentes:*

1) Para todo b vector de $\mathbb{R}^m$, la ecuación $A\,x = b$ tiene solución.

2) Las columnas de A generan $\mathbb{R}^m$, o sea, $\mathbb{R}^m = Span(a_1, a_2, ..., a_n)$.

3) A tiene una posición pivote en cada fila. ■

Regla para el cálculo de Ax. Para calcular la $i-$ésima componente del vector Ax (si éste está definido) tenemos que sumar los productos de la $i-$ésima fila de A por las coordenadas de x, o sea,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Teorema B.5 *Sea A una matriz $m \times n$ con columnas $a_1, a_2, ..., a_n$, x, y vectores de $\mathbb{R}^n$ y α un número real cualquiera. Las siguientes propiedades son válidas:*

1) $A\,(x + y) = A\,x + A\,y$, *2) $A(\alpha\,x) = \alpha\,(A\,x)$* ■

B.4.4. Conjunto solución de los sistemas lineales.

Todo sistema lineal de m ecuaciones con n incógnitas se puede escribir en forma matricial $Ax = b$ donde la matriz de coeficientes A es una matriz $m \times n$, x es un vector de $\mathbb{R}^n$ y b es un vector de $\mathbb{R}^m$.

1. El sistema lineal $Ax = 0$ se denomina sistema homogéneo. Dicho sistema siempre tiene la solución trivial $x = 0$. Además se cumple que el sistema homogéneo tiene soluciones no triviales si y sólo si el sistema tiene al menos una variable libre, o sea la matriz A tiene al menos una columna que no es pivote.

   El conjunto solución de un sistema homogéneo se escribe como el conjunto de todas las combinaciones lineales de ciertos vectores $s_1, s_2, ..., s_p$ cuyo número coincide con el de las variables libres del sistema. Para obtener los vectores $s_1, s_2, ..., s_p$ se puede hacer lo siguiente:
   1) Escribir el vector solución x expresando las variables principales en función de las libres.
   2) El vector s_1 se obtiene del vector x anterior al sustituir la primera variable libre por 1 y las demás por cero. El segundo vector s_2 se obtiene del vector x anterior al sustituir la segunda variable libre por 1 y las demás por cero. Y así sucesivamente.
   Por ejemplo si la solución x de $Ax = 0$ viene dada por (dos variables libres x_2, x_3)

$$\begin{pmatrix} \frac{3}{2}x_2 - \frac{1}{2}x_3 \\ x_2 \\ x_3 \end{pmatrix}, \text{ entonces } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \underbrace{\begin{pmatrix} \frac{3}{2} \\ 1 \\ 0 \end{pmatrix}}_{s_1} + x_3 \underbrace{\begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}}_{s_2}$$

2. El sistema lineal $Ax = b$ con $b \neq 0$ se denomina sistema no homogéneo.

   Teorema B.6 *Supongamos que $Ax = b$ tiene solución para cierto b de $\mathbb{R}^m$. Sea p una solución de $Ax = b$ (solución particular). Entonces el conjunto solución del sistema no homogéneo $Ax = b$ es el conjunto de todos los vectores de $\mathbb{R}^m$ de la forma*

$$x = p + x_h, \tag{B.9}$$

   donde x_h es la solución general del sistema homogéneo correspondiente $Ax_h = 0$.

B.4.5. Dependencia e independencia lineal.

Un conjunto de vectores $a_1, a_2, ..., a_n$ de $\mathbb{R}^n$ se denomina linealmente independiente si la ecuación vectorial

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0,$$

tiene como única solución la trivial $x_1 = \cdots = x_n = 0$.

Un conjunto de vectores $a_1, a_2, ..., a_n$ se denomina linealmente dependiente si existen los valores $x_1, x_2, \cdots, x_n$ no todos iguales a cero tales que se verifique la ecuación vectorial

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0.$$

Teorema B.7 *Un conjunto $S = \{a_1, a_2, ..., a_p\}$ de dos o más vectores es linealmente dependiente si y sólo si al menos uno de los vectores del conjunto es combinación lineal de los demás.*

Como consecuencia del teorema anterior se deduce que:

1) Dos vectores a_1 y a_2 son linealmente dependientes si y sólo si son proporcionales, es decir, si existe un número real α tal que $a_1 = \alpha a_2$ o $a_2 = \alpha a_1$.

2) Si $a_1, a_2, ..., a_n$ son vectores linealmente dependientes y $a_1 \neq 0$ entonces para cierto $j > 1$ el vector a_j es combinación lineal de los anteriores, es decir,

$$a_j = x_1 a_1 + x_2 + a_2 + \cdots + x_{j-1} a_{j-1}.$$

NOTA: Esto no implica que si tenemos un conjunto de vectores linealmente dependientes $a_1, a_2, ..., a_n$ necesariamente cada vector es combinación de los anteriores.

Teorema B.8 *Un conjunto $S = \{a_1, a_2, ..., a_n\}$ de dos o más vectores de $\mathbb{R}^m$ con $n > m$ es necesariamente un conjunto linealmente dependiente.*

Teorema B.9 *Un conjunto $S = \{a_1, a_2, ..., a_n\}$ de dos o más vectores de $\mathbb{R}^m$ con alguno de los vectores $a_i = 0$ $(1 \leq i \leq n)$ es necesariamente un conjunto de vectores linealmente dependientes, o sea si alguno de los vectores de S es el vector nulo entonces S es un conjunto de vectores linealmente dependientes.*

B.5. Álgebra de matrices.

Vamos a estudiar las operaciones elementales de las matrices. El elemento a_{ij} de una matriz A es el que se encuentra situado en la intercepción de la fila i con la columna j:

$$\begin{array}{cc} & \begin{array}{cccccc} & & & \text{columna } j & & \\ & & & \downarrow & & \end{array} \\ \begin{array}{c} \\ \\ \\ \text{fila } i \;\rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & \boxed{a_{ij}} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \end{array}. \tag{B.10}$$

B.5.1. Suma de matrices y multiplicación por un escalar.

Sean $A = [a_1\, a_2\, \cdots\, a_n]$ y $B = [b_1\, b_2\, \cdots\, b_n]$ matrices de dimensiones $m \times n$, donde $a_1, a_2, ..., a_n$ y $b_1, b_2, ..., b_n$ son vectores de $\mathbb{R}^m$. Definiremos la suma de dos matrices A y B a la matriz C cuyos vectores columnas coinciden con la suma de los correspondientes vectores columnas de A y B, es decir:

$$C = A + B = [a_1\, a_2\, \cdots\, a_n] + [b_1\, b_2\, \cdots\, b_n] = [a_1 + b_1\; a_2 + b_2\, \cdots\, a_n + b_n],$$

o en términos de los elementos matriciales: $C \equiv ||c_{ij}|| = ||a_{ij} + b_{ij}||$. Sea α un número real cualquiera y $A = [a_1\, a_2\, \cdots\, a_n]$ una matriz de $m \times n$. Definiremos el producto de una matriz por un escalar a la matriz cuyas columnas coinciden con las de A multiplicadas por α, es decir:

$$\alpha A = \alpha[a_1\, a_2\, \cdots\, a_n] = [\alpha a_1\; \alpha a_2\, \cdots\, \alpha a_n],$$

o bien $\alpha\, A \equiv \alpha ||a_{ij}|| = ||\alpha\, a_{ij}||$.

Teorema B.10 *Sean A, B y C matrices de $m \times n$ y α, β números reales. Entonces:*

1. $A + B = B + A$.
2. $(A + B) + C = A + (B + C)$.
3. $A + 0 = 0 + A = A$, *donde* $0 \equiv ||0||$ *es la matriz nula.*
4. $\alpha(A + B) = \alpha A + \alpha B$.
5. $(\alpha + \beta)A = \alpha\, A + \beta\, A$.
6. $\alpha(\beta\, A) = (\alpha\, \beta)\, A$.

B.5.2. Multiplicación de matrices.

Sea $A = [a_1\, a_2\, \cdots\, a_n]$ matriz de $m \times n$ y $B = [b_1\, b_2\, \cdots\, b_p]$ de $n \times p$. Se define la matriz producto $C = A \cdot B$ a la matriz de $m \times p$ definida por:

$$A \cdot B = A[b_1\, b_2\, \cdots\, b_p] = [Ab_1\, Ab_2\, \cdots\, Ab_p], \tag{B.11}$$

o elemento a elemento:

$$||(A \cdot B)_{ij}|| = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Nótese que cada columna de $A \cdot B$ es una combinación lineal de las columnas de A:

$$A \cdot B = [Ab_1\, Ab_2\, \cdots\, Ab_p] = [\sum_{k=1}^{n} b_{k1} a_k \;\; \sum_{k=1}^{n} b_{k2} a_k \;\; \cdots \;\; \sum_{k=1}^{n} b_{kp} a_k].$$

Una interpretación de la multiplicación de matrices se puede dar a partir de la composición de las aplicaciones lineales: Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una aplicación lineal con matriz A de $m \times n$ y $S : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ una aplicación lineal con matriz B de $n \times p$, entonces a la aplicación $Q : \mathbb{R}^p \longrightarrow \mathbb{R}^m$ definida por $Q(x) = T(S(x))$ (composición de T y S) le corresponde la matriz $A \cdot B$ de $m \times p$.

Teorema B.11 *Sean A de $m \times n$, B y C matrices de dimensiones apropiadas para que estén definidas las operaciones y α, β números reales. Entonces:*

1. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.
2. $A \cdot 0 = 0 \cdot A = 0$, *donde* $0 \equiv ||0||$ *es la matriz nula.*
3. $A \cdot (B + C) = A \cdot B + A \cdot C$.
4. $(B + C) \cdot A = B \cdot A + C \cdot A$.
5. $\alpha(A \cdot B) = (\alpha \cdot A) \cdot B$.
6. $I_m \cdot A = A \cdot I_n = A$, *donde* I_k *es la matriz identidad de* $k \times k$.

Es importante recordar que la multiplicación de matrices no es conmutativa, es decir para cualquiera sean las matrices A y B tenemos que $A \cdot B \neq B \cdot A$.

Si A es una matriz cuadrada $n \times n$ y k es un número entero no negativo ($k = 0, 1, 2, ...$), definiremos la potencia $k-$ ésima de A por la fórmula

$$A^k = \underbrace{A \cdot A \cdots A}_{k \text{ veces}} \cdot I_n, \qquad \text{donde } A^0 \equiv I_n.$$

B.5.3. Transposición de matrices.

Dada una matriz A de $m \times n$, la matriz transpuesta de A, que denotaremos por A^T, es la matriz de $n \times m$ cuyas filas coinciden con las columnas de A. Si $A = ||a_{ij}||$, entonces $A^T = ||a_{ji}||$.

Teorema B.12 *Sean A de $m \times n$, B una matriz de dimensiones apropiadas para que estén definidas las operaciones y α un número real. Entonces:*

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $(\alpha\, A)^T = \alpha\, A^T$.
4. $(A \cdot B)^T = B^T \cdot A^T$.

Como consecuencia de este teorema tenemos que

$$(A_1 \cdot A_2 \cdots A_n)^T = A_n^T \cdot A_{n-1}^T \cdots A_1^T.$$

B.5.4. La inversa de una matriz cuadrada.

Sea A una matriz de $n \times n$. Si existe una matriz C tal que

$$A \cdot C = I_n, \qquad C \cdot A = I_n,$$

entonces a C se le denomina matriz inversa de A y se le denota A^{-1}. Las matrices A para las cuales existe la inversa A^{-1} se denominan invertibles. Si A es invertible entonces A^{-1} es única.

Teorema B.13 *Sea A una matriz cuadrada de 2×2 invertible*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \textit{entonces} \qquad A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

El número $ad - bc$ se llama determinante de la matriz A de 2×2 y se le denota por $det\, A$. Nótese que una matriz A de 2×2 tiene inversa si y sólo si $det\, A \neq 0$.

Teorema B.14 *Sea A una matriz cuadrada de $n \times n$ invertible. Entonces el sistema $A\,x = b$ es compatible determinado para cualquiera sea b vector de $\mathbb{R}^n$, y su solución se expresa por la fórmula $x = A^{-1}\, b$.*

Teorema B.15 *Sean A y B matrices cuadradas de $n \times n$ invertibles. Entonces:*

1. $(A^{-1})^{-1} = A$.
2. $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.
3. *A^T es invertible y $(A^T)^{-1} = (A^{-1})^T$.*

Como consecuencia de este teorema tenemos que

$$(A_1 \cdot A_2 \cdots A_n)^{-1} = A_n^{-1} \cdot A_{n-1}^{-1} \cdots A_1^{-1}.$$

B.5.5. Matrices elementales.

Una matriz elemental de $n \times n$ es aquella que se obtiene al realizar una operación elemental de filas sobre la matriz identidad de $n \times n$, es decir son las que se obtienen al intercambiar dos filas, al multiplicar por una constante no nula una fila cualquiera y al reemplazar una fila por la suma de ésta mas una combinación lineal de cualquier otra. Por ejemplo, la matriz elemental correspondiente al intercambio de la 1 y 2 filas de una matriz de 3×3 será

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

la que corresponde a cambiar la 3 fila por la 3 más α veces la primera será

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix},$$

y la que multiplica la 2 fila por el número α será

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Se puede comprobar que si multiplicamos la matriz E_1 por una matriz cualquiera A de $3 \times p$ la matriz $E_1 \cdot A$ coincide con la matriz que se obtiene a partir de A al intercambiar la 1 y 2 filas.

Análogamente, $E_2 \cdot A$ coincide con la matriz que se obtiene de A al cambiar la 3 fila por la 3 más α veces la 1 fila y $E_3 \cdot A$ coincide con la matriz que se obtiene al multiplicar por α la 2 fila. Es decir al multiplicar una matriz elemental E de $m \times m$ por A de $m \times n$ se obtiene una nueva matriz $E \cdot A$ que coincide con la matriz que se se obtiene al realizar la operación elemental sobre A y dicha matriz E se obtiene realizando sobre la matriz identidad I_m de $m \times m$ la misma operación elemental que se quiere realizar sobre A. Puesto que todas las operaciones elementales son invertibles ello implica que las matrices elementales son invertibles y la inversa de E se obtiene realizando sobre la matriz identidad I_m la operación inversa. Por ejemplo, la inversa de E_1 es E_1. La inversa de E_2 es

$$E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha & 0 & 1 \end{pmatrix},$$

y la de E_3 es

$$E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{\alpha} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Teorema B.16 *Una matriz cuadrada A de $n \times n$ es invertible si y sólo si A es equivalente por filas a la matriz identidad I_n de $n \times n$ en cuyo caso toda secuencia de operaciones elementales que reducen A a I_n, reducen I_n a A^{-1}.*

Ello implica que $A \sim E_k \cdot E_{k-1} \cdots E_1 \cdot A = I_n$. Luego $A = (E_k \cdot E_{k-1} \cdots E_1)^{-1}$ lo cual implica que $A^{-1} = E_k \cdot E_{k-1} \cdots E_1 \cdot I_n$.

Como consecuencia de este teorema se tiene el siguiente algoritmo para calcular A^{-1}: Dada la matriz A de $n \times n$ invertible:

1. Construimos la matriz $[A\ I_n]$.
2. Reducimos por filas la matriz A a I_n, entonces $[A\ I_n] \sim [I_n\ A^{-1}]$.

Nótese que encontrar la inversa de la matriz A equivale a encontrar una matriz B tal que $A \cdot B = I$, o sea, $A[b_1 \cdots b_n] = [A\, b_1 \cdots A\, b_n] = [e_1 \cdots e_n] = I_n$. Es decir es equivalente a resolver las n ecuaciones

$$A\, b_1 = e_1, \quad A\, b_2 = e_2, \quad \cdots \quad A\, b_n = e_n,$$

donde $e_1, ..., e_n$ son las columnas de la matriz identidad.

Teorema B.17 *(*Teorema de inversión de matrices*) Sea A una matriz cuadrada A de $n \times n$. Las siguientes expresiones son equivalentes:*

1. *A es invertible.*
2. *A es equivalente por filas a la matriz identidad I_n de $n \times n$. (A tiene n posiciones pivote)*
3. *A es un producto de matrices elementales.*
4. *La ecuación $A\,x = 0$ solo tiene la solución trivial. (Las columnas de A son linealmente independientes.)*
5. *La ecuación $A\,x = b$ tiene al menos una solución para cada b vector de $\mathbb{R}^n$, o sea, b es un elemento del subespacio generado por las columnas de A (la solución de $A\,x = b$ es única).*
6. *Existe un C tal que $A \cdot C = I$.*
7. *Existe un D tal que $D \cdot A = I$.*

8. *A^T es invertible.*

Una consecuencia de este teorema es que si $A \cdot B = I$ o $B \cdot A = I$, entonces A y B son invertibles y $A = B^{-1}$, $B = A^{-1}$.

B.6. Determinantes.

B.6.1. Definición de determinante de una matriz cuadrada.

Sea A una matriz cuadrada

$$\begin{array}{cc} & \begin{array}{c}\text{columna } j\\ \downarrow\end{array} \\ \text{fila } i \;\rightarrow & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & a_{2j} & a_{2j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j-1} & a_{i-1j} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i1} & a_{i2} & \cdots & a_{ij-1} & \boxed{a_{ij}} & a_{ij+1} & \cdots & a_{in} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j-1} & a_{i+1j} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix} . \end{array} \tag{B.12}$$

Denotaremos por A_{ij} a la submatriz que se obtiene a partir de A si quitamos la fila i y la columna j, o sea

$$A_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}.$$

Sea A una matriz $n \times n$ ($n \geq 2$). El determinante $\det A$ de la matriz A es la suma de los términos $\pm a_{1j} \det A_{1j}$ donde a_{1j} son los elementos de la primera fila de A y A_{1j} son las correspondientes submatrices obtenidas eliminando la fila 1 y la columna j. Más exactamente:

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + a_{1n}(-1)^{n+1} \det A_{1n}, \tag{B.13}$$

donde suponemos que $\det a = a$ si a es un número real cualquiera.

Utilizando la definición anterior obtenemos para una matriz 2×2 el resultado:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

que coincide con el determinante definido en el capítulo anterior (teorema B.13).

Usualmente a la cantidad $(-1)^{i+j} \det A_{ij}$ se le conoce como cofactor (o menor) asociado al elemento a_{ij}. Utilizando esta notación tenemos que la definición anterior se puede reescribir de la forma:

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n},$$

que se conoce como *desarrollo del determinante de A por su primera fila.*

Teorema B.18 *Si A es de $n \times n$ entonces:*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

Las dos fórmulas anteriores se conocen como desarrollo del determinante de A por su $i-$ésima fila *y* desarrollo del determinante de A por su $j-$ésima columna*, respectivamente.*

Teorema B.19 *Si A es una matriz triangular (inferior o superior) entonces* $\det A$ *es el producto de los elementos de la diagonal, o sea,* $\det A = a_{11} \cdot a_{22} \cdots a_{nn}$.

B.6.2. Propiedades de los determinantes.

Los dos teoremas que enunciaremos a continuación son los más importantes de este apartado y nos permitirán calcular los determinantes de matrices de orden alto ($n \geq 4$) de manera eficiente.

Teorema B.20 *Sea A una matriz $n \times n$. Sea B una matriz que se obtiene de A al realizar sobre A alguna de las operaciones elementales de filas y sea E la matriz elemental que realiza dicha operación. Por ejemplo,*

1. *Si E_1 es la matriz que intercambia dos filas, la matriz $B = E_1 \cdot A$ se obtiene al intercambiar dos filas de A.*

2. *Si E_2 es la matriz que multiplica por una constante r no nula una fila cualquiera, la matriz $B = E_2 \cdot A$ se obtiene al multiplicar por r la correspondiente fila de A.*

3. *Si E_3 es la matriz que reemplaza una fila por la suma de ésta mas un múltiplo de cualquier otra fila, la matriz $B = E_3 \cdot A$ se obtiene al reemplazar una fila de A por la suma de ésta mas un múltiplo de cualquier otra fila.*

Entonces, $\det B = \det(E \cdot A) = \alpha \det A$ *donde* $\alpha = \begin{cases} -1 & \text{si } E \text{ es del tipo } E_1 \\ r & \text{si } E \text{ es del tipo } E_2 \\ 1 & \text{si } E \text{ es del tipo } E_3 \end{cases}$ *Además, si $A = I$ es la matriz identidad y por tanto* $\det I = 1$*, entonces* $\det E = \alpha$*, es decir el determinante de cualquier matriz elemental vale o -1 ó r ó 1 en dependencia si dicha matriz es del tipo E_1, E_2 o E_3, respectivamente.*

Como consecuencia del teorema anterior tenemos que si una fila es múltiplo de un número r entonces tenemos:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ r\,a_{i1} & r\,a_{i2} & \cdots & r\,a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

En particular si una fila de A es nula, $\det A = 0$ y si una fila de A es proporcional a otra $\det A = 0$.

Teorema B.21 *Sea A una matriz $n \times n$ y sea A^T su traspuesta, entonces* $\det A = \det A^T$.

Este segundo teorema nos permite enunciar todos los apartados del teorema anterior cambiando la palabra *filas* por *columnas.*

Es conocido de apartados anteriores que cualquier matriz A se puede, mediante transformaciones elementales, transformar en una matriz triangular superior U (forma escalonada). Además si resulta que la matriz A tiene n filas pivote entonces el determinante de A será proporcional al $\det U$. Luego, si tenemos n filas pivote todos los pivotes son distintos de cero y $\det A \neq 0$. Pero si A tiene n filas pivote entonces A es invertible. Este razonamiento nos conduce al siguiente teorema.

Teorema B.22 *Una* A $n\times n$ *es invertible si y sólo si* $\det A$ *es diferente de cero.*

Teorema B.23 *Sean* A *y* B *matrices* $n\times n$. *Entonces* $\det(A\cdot B)=(\det A)\cdot(\det B)$.

B.6.3. Regla de Kramer.

Sea A una matriz $n\times n$ y b un vector cualquiera de $\mathbb{R}^n$. Denotaremos por $A_j(b)$ a la matriz en la cual la $j-$ ésima columna de A está reemplazada por el vector b, es decir si A y b son

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij-1} & a_{ij} & a_{ij+1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}, \quad b=\begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix},$$

entonces $A_j(b)$ es la matriz

$$A_j(b)=\begin{pmatrix} a_{11} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij-1} & b_i & a_{ij+1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}.$$

Teorema B.24 *Sea* A *matriz* $n\times n$ *invertible. Entonces para todo* b *de* $\mathbb{R}^n$ *el sistema*

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij-1} & a_{ij} & a_{ij+1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}}_{A}\underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}}_{x}=\underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}}_{b},$$

es compatible y tiene una única solución que viene dada por

$$x_1=\frac{\det A_1(b)}{\det A}, \quad x_2=\frac{\det A_2(b)}{\det A}, \quad \cdots, \quad x_n=\frac{\det A_n(b)}{\det A}.$$

Es importante tener en cuenta que este método es ineficiente para calcular la solución de sistemas de más de 3 ecuaciones pues en el están involucrados determinantes. Es recomendable resolver los sistemas por el método de reducción de filas descrito en la primera parte del curso (ver *Algoritmo de reducción por filas*).

Finalmente, recordemos que "*invertir*" una matriz es equivalente a resolver n sistemas de ecuaciones de la forma:

$$A\,b_1=e_1, \quad A\,b_2=e_2, \quad \cdots \quad A\,b_n=e_n,$$

donde $e_1,...,e_n$ son las columnas de la matriz identidad. Por tanto para calcular la $j-$ésima columna de A^{-1} tenemos que resolver el sistema $A\,x=e_j$ cuya solución es, según la regla de Kramer, $x_1=\dfrac{\det A_1(e_j)}{\det A}, \cdots, x_n=\dfrac{\det A_n(e_j)}{\det A}$, pero

$$\det A_1(e_j)=(-1)^{1+j}\det A_{j1}=C_{j1},\cdots,\det A_n(e_j)=(-1)^{n+j}\det A_{jn}=C_{jn},$$

luego la columna j de A^{-1} es

$$\frac{1}{\det A}\begin{pmatrix} C_{j1} \\ C_{j2} \\ \vdots \\ C_{jn} \end{pmatrix}.$$

Todo esto nos conduce al teorema

Teorema B.25 *Sea A una matriz invertible $n \times n$. Entonces*

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} C_{11} & C_{21} & \cdots & C_{j1} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{j2} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{1i} & C_{2i} & \cdots & C_{ji} & \ddots & C_{ni} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{jn} & \cdots & C_{nn} \end{pmatrix},$$

donde $C_{lm} = (-1)^{l+m} \det A_{lm}$ son los cofactores (menores) del elemento a_{lm} de la matriz A y A_{lm} es la submatriz que se obtiene al eliminar la fila l y la columna m de A. La matriz $||C_{lm}||$ se denomina matriz adjunta de A.

Es importante notar que este método es ineficiente para calcular inversas (recurre a la regla de Kramer que es ineficiente es si misma) por lo que es recomendable para encontrar A^{-1} aplicar el algoritmo descrito en el teorema B.16 del capítulo anterior.

Para concluir esta parte dedicada a determinantes veamos una propiedad interesante que satisfacen los determinantes y que es consecuencia de los teoremas B.18 y B.20:

Sea $A = [a_1\, a_2 \cdots a_n]$ una matriz de $n \times n$ con columnas $a_1, ..., a_n$. Definamos la aplicación lineal:

$$T(x) = \det([a_1 \cdots x \cdots a_n]) = \det A_k(x), \qquad x \text{ de } \mathbb{R}^n, \qquad k = 1, 2, ..., n.$$

La aplicación anterior es lineal pues para todo α de $\mathbb{R}$ y x, y de $\mathbb{R}^n$:

$$\begin{aligned} T(\alpha\, x) &= \det([a_1 \cdots \alpha\, x \cdots a_n]) = \alpha \det([a_1 \cdots x \cdots a_n]) = \alpha\, T(x), \\ T(x+y) &= \det([a_1 \cdots x+y \cdots a_n]) = \det([a_1 \cdots x \cdots a_n]) + \det([a_1 \cdots y \cdots a_n]) = \\ &= T(x) + T(y). \end{aligned}$$

Esta propiedad se conoce como la multilinealidad del determinante

B.7. Espacios Vectoriales.

En esta sección vamos a definir los espacios vectoriales que son la generalización de los espacios $\mathbb{R}^n$ ya estudiados.

B.7.1. Definición y ejemplos.

Sea V un conjunto de elementos sobre el cual están definidas las operaciones *suma "+"* de dos elementos x, y de V y *multiplicación "·"* de un escalar (número real) α por un elemento de V. Diremos que V es un espacio vectorial si se cumplen las siguientes propiedades (axiomas): Sean x, y, z elementos arbitrarios de V y α, β números reales. Entonces V es un espacio vectorial si

1. Para todos x e y, *vectores* de V, el vector suma, $w = x + y$, también es un vector de V y se cumple que:

a) $x + y = y + x$

b) $(x + y) + z = x + (y + z)$

c) Existe un elemento "nulo" de V, tal que $x + 0 = 0 + x = x$

d) Existe el elemento $(-x)$ "opuesto" a x, tal que $x + (-x) = (-x) + x = 0$ donde $(-x)$ es vector opuesto de x

2. Para todo x *vector* de V, el vector que se obtiene al multiplicar por un escalar, $w = \alpha \cdot x$, también es un vector de V y se cumple que:

a) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$

b) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$

c) $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$

d) $1 \cdot x = x$

Ejemplos.
1.) El conjunto de los vectores de $\mathbb{R}^n$ cuando la suma de dos vectores y la multiplicación por un escalar es la definida en la sección 2.
2.) El conjunto de las matrices $m \times n$ cuando la suma de dos matrices y la multiplicación por un escalar es la definida en la sección 3. Dicho espacio lo denotaremos por $\mathbb{R}^{m\times n}$
3.) El conjunto de los polinomios de grado a lo sumo n, que denotaremos por $\mathbb{P}_n$, o sea,

$$\mathbb{P}_n = \{p_n(t) = a_0 + a_1\, t + \cdots + a_n\, t^n, \quad a_0, ..., a_n \text{ números reales.}\},$$

donde definiremos la suma de dos polinomios y la multiplicación por un escalar de la siguiente forma:

$$p(t) = a_0 + a_1\, t + \cdots + a_n\, t^n, \quad q(x) = b_0 + b_1\, t + \cdots + b_n\, t^n,$$

$$(p + q)(t) \equiv p(t) + q(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n,$$

$$(\alpha \cdot p)(t) \equiv \alpha\, p(t) = (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_n)t^n.$$

Además, $p_n = 0$, si y sólo si $a_0 = a_1 = \cdots = a_n = 0$. 4.) El conjunto de las funciones continuas en el intervalo $[a, b]$, que denotaremos por $C_{[a,b]}$, cuando la suma de dos funciones f y g y la multiplicación por un escalar α están dadas por

$$(f + g)(t) \equiv f(t) + g(t), \quad (\alpha \cdot f)(t) \equiv \alpha \cdot f(t).$$

B.7.2. Subespacios vectoriales.

Sea V un espacio vectorial. Diremos que un subconjunto H de elementos de V es un subespacio vectorial de V si H es a su vez un espacio vectorial respecto a las mismas operaciones *suma "+"* y *multiplicación "·"* que V.

Ejemplos.
1.) Dado un espacio vectorial V, son subespacios vectoriales "triviales" los subespacios $H = \{0\}$ (conjunto que tiene como único elemento, el nulo) y $H = V$ (el mismo espacio vectorial).
2.) Para $V = C_{[a,b]}$, $H = \mathbb{P}_n$ es un subespacio vectorial, para cualquier $n = 0, 1, 2, ...$ entero.
3.) Para $V = \mathbb{P}_n$, $H = \mathbb{P}_k$ es un subespacio vectorial para todo $k < n$.

Teorema B.26 *Un subconjunto H de elementos de V es un subespacio vectorial de V si y sólo si se cumplen las siguientes tres condiciones.*

1. *El elemento nulo de V pertenece a H.*

2. *Para todos x e y,* vectores *de H, el vector $w = x + y$ también es un vector de H.*

3. Para todos x vector de H, el vector $w = \alpha \cdot x$ también es un vector de H.

Teorema B.27 *Dado un conjunto de vectores $\{v_1, v_2, ..., v_p\}$ de un espacio vectorial V, el conjunto* span $(v_1, v_2, ..., v_p)$ *es un subespacio vectorial de V. Dicho subespacio vectorial comúnmente se denomina subespacio generado por los vectores $v_1, v_2, ..., v_p$.*

B.7.3. Conjuntos linealmente independientes. Bases.

Un conjunto de vectores $v_1, v_2, ..., v_p$ de un espacio vectorial V se denomina linealmente independiente si la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0,$$

tiene como única solución la trivial $x_1 = \cdots = x_p = 0$.

Un conjunto de vectores $v_1, v_2, ..., v_p$ se denomina linealmente dependiente si existen los valores $x_1, x_2, \cdots, x_p$ no todos iguales a cero tales que se verifique la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0.$$

Para estudiar la dependencia lineal de los vectores de V podemos utilizar los mismos teoremas estudiados en la sección 2 cambiando los espacios $\mathbb{R}^n$ por los correspondientes espacios vectoriales V. Por tanto las siguientes afirmaciones son ciertas:

1. *Un conjunto $S = \{v_1, v_2, ..., v_p\}$ de dos o más vectores es linealmente dependiente si y sólo si al menos uno de los vectores del conjunto es combinación lineal de los demás.*

2. *Un conjunto $S = \{v_1, v_2, ..., v_p\}$ de dos o más vectores de V con alguno de los vectores $v_i = 0$ $(1 \leq i \leq p)$ es necesariamente un conjunto de vectores linealmente dependientes, o sea si alguno de los vectores de S es el vector nulo entonces S es un conjunto de vectores linealmente dependientes.*

3. *Dos vectores v_1 y v_2 de V son linealmente dependientes si y sólo si son proporcionales, es decir, si existe un número real α tal que $v_1 = \alpha v_2$ o $v_2 = \alpha v_1$*

Teorema B.28 *El conjunto de dos o más vectores $v_1, v_2, ..., v_p$ de V con $v_1 \neq 0$ es un conjunto linealmente dependiente si y sólo si para cierto j, $1 < j < p$, el vector v_j es combinación lineal de los anteriores, es decir,*

$$v_j = x_1 v_1 + x_2 v_2 + \cdots + x_{j-1} v_{j-1}.$$

NOTA: Esto no implica que si tenemos un conjunto de vectores linealmente dependientes $v_1, v_2, ..., v_p$ necesariamente cada vector es combinación de los anteriores.

Los vectores linealmente independientes de un espacio vectorial juegan un papel fundamental en el estudio de los sistemas lineales gracias a la siguiente definición

Dado un subespacio vectorial H del espacio vectorial V diremos que el conjunto de vectores $B = \{b_1, b_2, ..., b_p\}$ de V es una base de H si

i) B es un conjunto de vectores linealmente independientes

ii) $H = Span(b_1, b_2, ..., b_p)$, o sea, B genera a todo H.

En particular si H coincide con V, entonces B es una base de todo el espacio vectorial V.

Por ejemplo, si tomamos una matriz $n \times n$ invertible, entonces sus columnas $a_1, ..., a_n$ son linealmente independientes y además $\mathbb{R}^n = Span(a_1, ..., a_n)$. Por tanto $B = a_1, ..., a_n$ es una base de

$\mathbb{R}^n$. En particular, si $A = I_n$, la matriz identidad $n \times n$, las columnas $e_1, e_2, ..., e_n$ de la misma son una base de $\mathbb{R}^n$ la cual se conoce como base canónica de $\mathbb{R}^n$.

Otro ejemplo lo constituye el conjunto de vectores $S = \{1, t, t^2, ..., t^n\}$ del espacio vectorial $\mathbb{P}_n$. Es fácil comprobar que dichos vectores son linealmente independientes y que $\operatorname{span}(1, t, t^2, ..., t^n) = \mathbb{P}_n$. S se conoce como la base canónica de $\mathbb{P}_n$.

Teorema B.29 *Sea $S = \{v_1, v_2, ..., v_p\}$ un conjunto de vectores de un espacio vectorial V y sea $H = Span(v_1, v_2, ..., v_p)$ un subespacio de V.*

i) Supongamos que un vector de S, digamos el v_j, $1 \leq j \leq p$, es combinación lineal de los restantes. Entonces si suprimimos a dicho vector v_k del conjunto S, H no cambia, o sea,

$$H = Span(v_1, ..., v_{k-1}, v_k, v_{k+1}, ..., v_p) = \operatorname{span}(v_1, ..., v_{k-1}, v_{k+1}, ..., v_p).$$

ii) Si H no es el espacio $\{0\}$, entonces algún subconjunto de S es una base de H.

Este teorema nos dice que a partir de un conjunto de vectores que generan un espacio vectorial siempre es posible "extraer" una base de dicho espacio vectorial. Por ejemplo, consideremos el espacio generado por los vectores

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad H = Span\left[v_1, v_2, v_3\right].$$

Es evidente que $H = Span[v_1, v_2, v_3] = \operatorname{span}[v_1, v_2]$ pues el tercer vector es combinación lineal de los dos anteriores. Además, $S = \{v_1, v_2, v_3\}$ no es una base de H pues v_1, v_2, v_3 no son linealmente independientes. Sin embargo, el conjunto $B = \{v_1, v_2\}$ que se obtiene al suprimir el tercer elemento si que es una base de H.

B.7.4. Sistemas de coordenadas.

Cuando estudiamos el espacio $\mathbb{R}^n$ (sección 2) definimos a los vectores de $\mathbb{R}^n$ mediante sus coordenadas. Ello era evidente pues

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 e_1 + \cdots x_n e_n.$$

Además, es evidente que dado un vector x de $\mathbb{R}^n$ sus coordenadas, es decir los valores $x_1, ..., x_n$, son únicos. El siguiente teorema generaliza este resultado para cualquier espacio vectorial V.

Teorema B.30 *Sea $B = \{b_1, b_2, ..., b_n\}$ una base del espacio vectorial V. Entonces para todo x de V existe un único conjunto de valores $\{\alpha_1, ..., \alpha_n\}$ tal que*

$$x = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n.$$

Este teorema nos permite definir las coordenadas de un vector cualquiera de un espacio vectorial V.

Supongamos que $B = \{b_1, b_2, ..., b_n\}$ es una base del espacio vectorial V. Denominaremos coordenadas del vector x de V en la base B al conjunto de los valores $\{\alpha_1, ..., \alpha_n\}$ tales que

$$x = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_n b_n,$$

y lo denotaremos por $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}_B$ o $[x]_B$.

Consideremos nuevamente las coordenadas en $\mathbb{R}^n$. Como vimos al principio de este apartado los elementos de cualquier vector X de $\mathbb{R}^n$ coinciden con las coordenadas de dicho vector en la base canónica $C = \{e_1, e_2, ..., e_n\}$. Supongamos ahora que pasamos a una nueva base $B = \{b_1, b_2, ..., b_n\}$. Entonces

$$x = x_1 e_1 + \cdots + x_n e_n = \alpha_1 b_1 + \cdots + \alpha_n b_n = [b_1\, b_2\, \cdots\, b_n] \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}_B = P_{CB} \cdot [x]_B,$$

y por tanto, si las coordenadas de x en la base canónica C son $x_1, ..., x_n$, la siguiente relación será cierta:

$$[x]_C = P_{CB} \cdot [x]_B,$$

donde P_{CB} es la matriz de cambio de base B en C que pasa las coordenadas de un vector en la base B a las escritas en C. Nótese que las columnas de la matriz P_{CB} coinciden con los vectores de la nueva base escritos en la base canónica. Es evidente que la matriz P_{CB} es invertible (todas sus columnas son independientes y por tanto el $\det P_{CB} \neq 0$) por tanto la matriz inversa P_{CB}^{-1} será la matriz de cambio de base de base canónica a la base B, o sea

$$[x]_B = P_{CB}^{-1} \cdot [x]_C = P_{BC} \cdot [x]_C.$$

B.7.5. La dimensión de un espacio vectorial.

Hemos visto que cualquier espacio V con una base de n elementos es isomorfo a $\mathbb{R}^n$. Resulta que dicho número n es una propiedad intrínseca de V, su dimensión.

Teorema B.31 *Si un espacio vectorial V tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier conjunto con más de n vectores de V es linealmente dependiente.*

Este teorema es una generalización del teorema B.8.

Teorema B.32 *Si un espacio vectorial V tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier otra base de V tendrá que tener n vectores de V.*

Un espacio vectorial es de dimensión finita n si V está generado por una base de n elementos, es decir si $V = Span(b_1, ..., b_n)$, donde $B = \{b_1, ..., b_b\}$ es una base de V y lo escribiremos de la forma $\dim V = n$. En el caso que $V = \{0\}$ sea el espacio vectorial nulo, $\dim\{0\} = 0$. Si V no puede ser generado por una base finita de vectores, entonces diremos que V es de dimensión infinita y lo denotaremos por $\dim V = \infty$.

Por ejemplo, $\dim \mathbb{R}^n = n$, $\dim \mathbb{P}_n = n + 1$, $\dim C_{[a,b]} = \infty$.

Teorema B.33 *Sea H un subespacio vectorial de un espacio vectorial de dimensión finita V, ($\dim V = n < \infty$). Entonces todo conjunto de vectores linealmente independientes de H se puede ampliar hasta convertirlo en una base de H. Además $\dim H \leq \dim V$.*

Teorema B.34 *Sea V un espacio vectorial n-dimensional ($\dim V = n < \infty$) y sea $S = \{b_1, b_2, ..., b_n\}$ un conjunto de exactamente n vectores. Entonces:*

a) Si S es un conjunto de vectores linealmente independientes, S es una base de V.

b) Si S genera a todo V, es decir si $V = Span(b_1, b_2, ..., b_n)$, S es una base de V.

B.7.6. Cambio de base.

Sea V un espacio vectorial de dimensión finita n y sea $B = \{b_1, b_2, ..., b_n\}$ y $D = \{d_1, d_2, ..., d_n\}$ dos bases de V. Como D es una base de V y los elementos de B son elementos de V, éstos se podrán escribir en la base D, o sea,

$$\begin{array}{c} b_1 = a_{11}d_1 + a_{12}d_2 + \cdots + a_{1n}d_n, \\ \vdots \\ b_k = a_{k1}d_1 + a_{k2}d_2 + \cdots + a_{kn}d_n, \\ \vdots \\ b_1 = a_{n1}d_1 + a_{n2}d_2 + \cdots + a_{nn}d_n. \end{array} \quad \Longrightarrow \quad b_1 = \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}_D, \cdots, b_n = \begin{pmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nn} \end{pmatrix}_D.$$

Supongamos que las coordenadas de un vector x de V en la base B son $x_1, x_2, ..., x_n$ y en la base D son $y_1, y_2, ..., y_n$, es decir,

$$x = x_1b_1 + \cdots + x_nb_n \longrightarrow [x]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_B, \text{ y } x = y_1d_1 + \cdots + y_nd_n \longrightarrow [x]_D = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_D,$$

Entonces tenemos que

$$[x]_D = [x_1b_1 + \cdots + x_nb_n]_D = x_1[b_1]_D + \cdots x_n[b_n]_D = ([b_1]_D\,[b_2]_D \cdots [b_n]_D)[x]_B.$$

Es decir si conocemos como se expresan los vectores de la base B en la base D podemos encontrar la matriz de cambio de base B en D, cuyas columnas $a_1, a_2, ..., a_n$ no son más que las coordenadas de los vectores de la base B escritos en la base D. Por tanto tenemos que $[x]_D = P_{DB}[x]_B$, o en forma matricial

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_D = \begin{pmatrix} a_{11} & \cdots & a_{k1} & \cdots & a_{n1} \\ a_{12} & \cdots & a_{k2} & \cdots & a_{n2} \\ \vdots & & & & \\ a_{1n} & \cdots & a_{kn} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_B.$$

Además, P_{DB} es una matriz invertible (sus columnas son linealmente independientes), luego existe la inversa P_{DB}^{-1} la cual es la matriz de cambio de base D en B, o sea,

$$[x]_B = P_{DB}^{-1}[x]_D.$$

B.8. El problema de autovalores de una matriz.

Sea A una matriz de $n \times n$. Denominaremos al vector x de $\mathbb{R}^n$, autovector de A asociado al autovalor λ, al vector no nulo $x \neq 0$, tal que

$$A\,x = \lambda\,x, \qquad x \neq 0. \tag{B.14}$$

Es fácil comprobar que si x es un autovector asociado al autovalor λ, entonces el sistema

$$(A - \lambda\,I)x = 0, \qquad \text{donde } I \qquad \text{es la matriz identidad } n \times n, \tag{B.15}$$

tiene soluciones no triviales. Por tanto, dado un autovalor λ de A, el conjunto de los autovectores asociados a λ, que coincide con una base del núcleo de A, es un subespacio vectorial de $\mathbb{R}^n$. Dicho espacio se denomina *autoespacio* de A correspondiente al autovalor λ. Es importante destacar que conocido el autovalor λ, encontrar el autoespacio consiste en encontrar las soluciones de un sistema homogéneo de ecuaciones, o sea, resolver $(A - \lambda\,I)x = 0$.

Teorema B.35 *Sea A una matriz triangular de $n \times n$. Entonces los autovalores de A coinciden con los elementos de la diagonal principal $\lambda = a_{ii}$, $i = 1, 2, .., n$.*

Es importante destacar que las operaciones elementales de filas cambian los autovalores de una matriz. Es decir si U es la forma escalonada de A los autovalores de A y U son, en general, distintos.

Teorema B.17 (*Teorema de inversión de matrices*, conclusión.) Sea A una matriz cuadrada A de $n \times n$. Las siguientes expresiones son equivalentes:

1. A es invertible.
2. $\lambda = 0$ no es un autovalor de A

El teorema anterior es evidente pues si $\lambda = 0$ es autovalor de A, entonces $A\,x = 0$ tiene que tener soluciones no triviales y por tanto A no puede ser invertible.

Teorema B.36 *Sea A una matriz de $n \times n$ que tiene p autovalores distintos $\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_p$. Entonces los autovectores $v_1, v_2, ..., v_p$ asociados a los autovalores $\lambda_1, \lambda_2, ..., \lambda_p$ son linealmente independientes.*

B.8.1. La ecuación característica.

¿Cómo calcular los autovalores? La respuesta a esta pregunta no es difícil pues los autovalores λ son tales que el sistema (B.15), $(A - \lambda\,I)x = 0$, tiene soluciones no triviales. Pero un sistema homogéneo tiene soluciones no triviales si y sólo si

$$\det(A - \lambda\,I) = 0. \tag{B.16}$$

La ecuación anterior se denomina *ecuación característica* de A. Lo anterior lo podemos resumir como:

Un número λ es un autovalor de la matriz A de $n \times n$ si y sólo si λ satisface la ecuación característica de A (B.16), $\det(A - \lambda\,I) = 0$.

Es fácil comprobar que la expresión $\det(A-\lambda\,I) = 0$ es un polinomio de grado n en λ. Por tanto el polinomio $p_n(\lambda) = \det(A - \lambda\,I)$ se denomina polinomio característico de A y los autovalores de A serán las raíces de dicho polinomio. O sea, λ es un autovalor de A si y sólo si $p_n(\lambda) = 0$.

Matrices similares. Sean A y B dos matrices $n \times n$. Diremos que A es similar a B si existe una matriz invertible P, tal que

$$A = P \cdot B \cdot P^{-1}, \qquad \text{o} \qquad B = P^{-1} \cdot A \cdot P. \tag{B.17}$$

Es evidente que si A es similar a B, B es similar a A, para demostrarlo es suficiente tomar $Q = P^{-1}$. Por tanto diremos simplemente que las matrices A y B son *similares* si se cumple (B.17). La transformación que a cada matriz le hace corresponder otra similar, o sea, $A \longrightarrow P^{-1} \cdot A \cdot P$, se denomina transformación de similitud.

Teorema B.37 *Si las matrices A y B de $n \times n$ son similares, entonces tienen el mismo polinomio característico y, por tanto, los mismos autovalores.*

B.8.2. Diagonalización de matrices.

Una matriz A de $n \times n$ es *diagonalizable* si A es similar a una matriz diagonal D, o sea, si existe una matriz P invertible y otra D diagonal, tales que $A = P \cdot D \cdot P^{-1}$.

Teorema B.38 *Una matriz A de $n \times n$ es* diagonalizable *si y sólo si A tiene n autovectores linealmente independientes. Si $A = P \cdot D \cdot P^{-1}$, donde D es diagonal y P invertible, entonces los elementos de la diagonal de D son los autovalores de A y las columnas de P son los correspondientes autovectores.*

Este teorema se puede reformular diciendo que A es *diagonalizable* si y sólo si sus autovectores forman una base de $\mathbb{R}^n$.

Algoritmo para diagonalizar una matriz A.

1. Encontrar los autovalores de A resolviendo la ecuación (B.16).

2. Encontrar los autovectores linealmente independientes correspondientes a cada uno de los autovalores calculados en el paso 1. Si tenemos en total n autovectores linealmente independientes entonces el teorema B.38 nos asegurará que nuestra matriz es diagonalizable, si no hay n autovectores independientes entonces A no es diagonalizable y detenemos el proceso. Si A es diagonalizable continuamos como sigue.

3. Construimos la matriz P cuyas n columnas son los n autovectores linealmente independientes.

4. Construimos la matriz diagonal D cuyos elementos de la diagonal coinciden con los autovalores de A. Es importante colocar en cada columna de D el autovalor asociado al autovector de la correspondiente columna de P.

Es recomendable comprobar que hemos calculado correctamente P y D. Para ello comprobamos que $A \cdot P = P \cdot D$.

Teorema B.39 *(Condición suficiente para que una matriz sea diagonalizable I)*
Si una matriz A de $n \times n$ tiene n autovalores distintos, entonces es diagonalizable.

Teorema B.40 *(Condición suficiente para que una matriz sea diagonalizable II)*
Sea A una matriz de $n \times n$ con p autovalores distintos $\lambda_1, \lambda_2, ..., \lambda_p$. Sea B_k, $k = 1, 2, ..., p$ una base del correspondiente autoespacio asociado al autovalor λ_k y sea n_k la dimensión de dicho autoespacio. Construyamos el conjunto de vectores que contiene a todos los conjuntos B_k, o sea

$$B = B_1 \cup B_2 \cup \cdots \cup B_p, \qquad \dim(B) = n_1 + n_2 + \cdots + n_p.$$

Entonces B es un conjunto de vectores de $\mathbb{R}^n$ linealmente independientes. Además, A es diagonalizable si y sólo si $\dim(B) = n$, o sea, si B contiene exactamente n vectores.