Métodos Numéricos II
Segundo de Grado en Matemáticas, Curso 2021/22
Facultad de Ciencias de la Universidad de Málaga
Dpto. de Análisis Matemático, Estadística e Investigación Operativa, y Matemática Aplicada
Profs. Manuel J. Castro y Francisco J. Palma (Área de Conocimiento Mat. Aplicada)

TEMA 1

Vectores y Matrices

Los **objetivos** de este tema son:

- recordar (demostrar) los resultados básicos de la teoría de matrices y de los espacios vectoriales de dimensión finita;
- operar con matrices particionadas en bloques;
- obtener resultados de reducción de una matriz a su forma triangular o diagonal;
- definir normas vectoriales y normas matriciales;
- obtener resultados de convergencia de las sucesiones de potencias de matrices;
- presentar los problemas fundamentales del *Análisis Numérico Matricial.*

1. Cuerpo de escalares

Siempre trabajamos, bien con el **cuerpo de los números reales** $\mathbb{R}$, bien con el **cuerpo de los números complejos** $\mathbb{C}$; cuando no hay lugar a confusión hablamos simplemente del **cuerpo de escalares** $\mathbb{K}$.

Suponemos que son bien conocidas la definición y las propiedades más importantes de los **números reales** $\mathbb{R}$ (cuerpo ordenado verificando el axioma del supremo).

Respecto del **cuerpo de los números complejos** recordamos que:

- se **define** $\mathbb{C} = \{z = x + y\, i \ : \ x, y \in \mathbb{R}\, ; \, i^2 = -1\}$; dado el número complejo $z = x + y\, i$, x e y reciben el nombre de **parte real** e **imaginaria** respectivamente;

- definidas la suma y producto en la forma habitual, $\mathbb{C}$ resulta tener estructura de **cuerpo**, aunque no es posible definir un orden total en el mismo;

- se verifica que $\mathbb{R} \subset \mathbb{C}$, sin más que **identificar** todo número real x con el número complejo $x + 0\, i$;

- se define el **conjugado** del número complejo $z = x + y\, i$ como $\bar{z} = x - y\, i \in \mathbb{C}$. Se verifica que $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$, $\bar{0} = 0$, $\overline{-z} = -\bar{z}$, $\overline{z_1\, z_2} = \bar{z}_1\, \bar{z}_2$, $\bar{1} = 1$, $\overline{z^{-1}} = \bar{z}^{-1}$, $\bar{i} = -i$ y $\bar{\bar{z}} = z$; por otro lado dado $z \in \mathbb{C}$, se verifica que $z \in \mathbb{R}$ si y sólo si $\bar{z} = z$;

- se define el **módulo** del número complejo $z = x + y\, i$ como $|z| = \sqrt{x^2 + y^2} \geq 0$ y su **argumento** (principal) como el ángulo φ_z ($\in [0, 2\,\pi)$) tal que $\operatorname{sen} \varphi_z = \frac{y}{|z|}$ y $\cos \varphi_z = \frac{x}{|z|}$; observamos que el par $(|z|, \varphi_z)$ **caracteriza** al número complejo z (**forma polar**). Se verifica que $|z_1 + z_2| \leq |z_1| + |z_2|$, $|z| = 0$ si y sólo si $z = 0$, $|-z| = |z|$, $|z_1\, z_2| = |z_1|\, |z_2|$, $|1| = 1$, $|z^{-1}| = |z|^{-1}$, $|i| = 1$, $|\bar{z}| = |z|$, $z\, \bar{z} = |z|^2$, $\varphi_{-z} = \varphi_z + \pi$, $\varphi_{z_1 z_2} = \varphi_{z_1} + \varphi_{z_2}$, $\varphi_1 = 0$, $\varphi_{z^{-1}} = -\varphi_z$, $\varphi_i = \frac{\pi}{2}$ y $\varphi_{\bar{z}} = -\varphi_z$; por otro lado dado $z \in \mathbb{C}$, se verifica que $z \in \mathbb{R}$ si y sólo si $\varphi_z = k\, \pi$, $k \in \mathbb{Z}$;

- los números complejos admiten una **representación gráfica** en el plano cartesiano xy, tomando el eje de las abscisas como eje real y el eje de las ordenadas como eje imaginario; se tiene entonces que el módulo de un número complejo es su distancia al origen y el argumento el ángulo formado con el eje de las abscisas;

- se define la **función exponencial compleja** mediante $e^z = e^x\, (\cos y + \operatorname{sen} y\, i)$; esta definición es una extensión de la función exponencial real, conservando las propiedades características de la misma, en particular que $e^{z_1}\, e^{z_2} = e^{z_1 + z_2}$. Se verifica además que $z = |z|\, e^{\varphi_z\, i}$ (**forma exponencial** de los números complejos);

- **teorema fundamental del Álgebra**: $\mathbb{C}$ es un cuerpo algebraicamente cerrado, por lo que todo polinomio de grado n tiene exactamente n raíces; es decir, dados $a_i \in \mathbb{C}$, $i = 0, 1, \ldots, n$, $a_0 \neq 0$, y construido el polinomio $p(z) = a_0\, z^n + a_1\, z^{n-1} + \cdots + a_{n-1}\, z + a_n$, existen n raíces $z_k \in \mathbb{C}$, $k = 1, 2, \ldots, n$, (no necesariamente diferentes dos a dos), tales que $p(z) = a_0\, (z - z_1)\, (z - z_2)\, \cdots\, (z - z_n)$; además si los coeficientes $a_i \in \mathbb{R}$ y si z_k es raíz, entonces $\bar{z}_k$ también lo es (y con la misma multiplicidad).

2. El espacio vectorial $\mathbb{K}^n$

Notamos por $\mathbb{K}^n$ el conjunto de los **vectores** X de **componentes** $x_i \in \mathbb{K}$, $i = 1, 2, \ldots, n$, que escribiremos usualmente en forma de columna, es decir,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = (x_i)_{i=1}^n = (x_i)_{i=1,2,\ldots,n} = (x_i)_i = (x_i) \in \mathbb{K}^n \,;$$

como puede observarse, cuando no hay lugar a confusión se omite expresar el recorrido de los índices utilizados.

En $\mathbb{K}^n$ se definen la **suma** y el **producto por escalares** de la forma usual, es decir, dados $X = (x_i) \in \mathbb{K}^n$, $Y = (y_i) \in \mathbb{K}^n$, $\alpha \in \mathbb{K}$, entonces

$$X + Y = (x_i + y_i)\,, \quad \alpha\, X = (\alpha\, x_i)\,.$$

El conjunto $\mathbb{K}^n$ con las operaciones así definidas resulta tener estructura de **espacio vectorial** sobre el cuerpo $\mathbb{K}$ de **dimensión** n, siendo una **base** del mismo (la **base canónica**) la formada por los **vectores elementales** $E_i \in \mathbb{K}^n$, $i = 1, 2, \ldots, n$, donde $E_i = (e_j^{(i)})_j$, siendo $e_j^{(i)} = \delta_{i,j}$ (como es habitual, $\delta_{i,j}$ representa la delta de Kronecker). Notamos por $0 = (0) \in \mathbb{K}^n$ el vector nulo de este espacio vectorial y por $-X = (-x_i) \in \mathbb{K}^n$ el opuesto del vector $X = (x_i) \in \mathbb{K}^n$.

Con la ayuda de la base canónica, dado cualquier vector $X = (x_i) \in \mathbb{K}^n$ podemos escribir

$$X = \sum_{i=1}^n x_i\, E_i = \sum_i x_i\, E_i = \sum x_i\, E_i = x_i\, E_i\,,$$

(utilizando en la última expresión la convención de sumación sobre los índices repetidos) ya que en este caso las **coordenadas** del vector respecto de dicha base coinciden con las componentes del mismo.

3. El espacio vectorial $\mathcal{M}_{m\times n}(\mathbb{K})$

Notamos por $\mathcal{M}_{m\times n}(\mathbb{K})$ el conjunto de las **matrices** A de **componentes** $a_{i,j} \in \mathbb{K}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, es decir,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = (a_{i,j})_{i=1,2,\ldots,m,\, j=1,2,\ldots,n} = (a_{i,j})_{i,j} = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})\,.$$

En $\mathcal{M}_{m\times n}(\mathbb{K})$ se definen la **suma** y el **producto por escalares** de la forma usual, es decir, dados $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B = (b_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, $\alpha \in \mathbb{K}$, entonces

$$A + B = (a_{i,j} + b_{i,j})\,, \quad \alpha\, A = (\alpha\, a_{i,j})\,.$$

El conjunto $\mathcal{M}_{m\times n}(\mathbb{K})$ con las operaciones así definidas resulta tener estructura de **espacio vectorial** sobre el cuerpo $\mathbb{K}$ de **dimensión** $m \times n$, siendo una **base** del mismo (la **base canónica**) la formada por las **matrices elementales** $E_{i,j} \in \mathcal{M}_{m\times n}(\mathbb{K})$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, donde $E_{i,j} = (e^{(i,j)}_{k,l})_{k,l}$, siendo $e^{(i,j)}_{k,l} = \delta_{i,k}\,\delta_{j,l}$. De nuevo notamos por $0 = (0) \in \mathcal{M}_{m\times n}(\mathbb{K})$ la matriz nula de este espacio vectorial y por $-A = (-a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ la opuesta de la matriz $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$.

Con la ayuda de la base canónica, dada cualquier matriz $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ podemos escribir

$$A = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}\, E_{i,j} = \sum_{i,j} a_{i,j}\, E_{i,j} = \sum a_{i,j}\, E_{i,j} = a_{i,j}\, E_{i,j}\,,$$

ya que de nuevo las **coordenadas** de la matriz respecto de dicha base coinciden con sus componentes.

En el caso de que $m = n$ se habla de **matrices cuadradas** y se nota por $\mathcal{M}_n(\mathbb{K})$, en el caso $m = 1$ se habla de **matrices (o vectores) fila** y en el caso $n = 1$ se habla de **matrices (o vectores) columnas**; resulta evidente la identificación entre $\mathbb{K}^m$ y $\mathcal{M}_{m\times 1}(\mathbb{K})$.

Recordamos también algunas definiciones asociadas con una matriz A:

- **submatriz**: aquella matriz que resulta de suprimir de A cualquier número aleatorio de filas y/o columnas;
- **submatriz principal** de orden k, $k = 1, 2, \ldots, \min\{m, n\}$: la submatriz formada por las k primeras filas y columnas de A (las notamos por $A^{(k)}$);
- **diagonal principal** de una matriz cuadrada de orden n: el conjunto de elementos $a_{i,i}$, $i = 1, 2, \ldots, n$.

4. El álgebra $\mathcal{M}_n(\mathbb{K})$

Recordamos que dadas dos matrices $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ y $B = (b_{j,k}) \in \mathcal{M}_{n\times p}(\mathbb{K})$, se define su **producto** $C = (c_{i,k}) \in \mathcal{M}_{m\times p}(\mathbb{K})$ de forma que

$$c_{i,k} = \sum_{j=1}^{n} a_{i,j}\, b_{j,k} = \sum_{j} a_{i,j}\, b_{j,k} = \sum a_{i,j}\, b_{j,k} = a_{i,j}\, b_{j,k}\,,\ \ i = 1, 2, \ldots, m\,,\ \ k = 1, 2, \ldots, p\,.$$

El producto así definido resulta ser asociativo y distributivo respecto de la suma, pero no es, en general, conmutativo ni cancelativo y posee divisores de cero (en todos los casos suponiendo dimensiones conformes para poder realizar todas las operaciones necesarias).

En $\mathcal{M}_n(\mathbb{K})$ se llama **matriz identidad** I_n (o simplemente I si no hay lugar a confusión) a aquella que tiene todas sus componentes nulas, excepto las situadas sobre la diagonal principal que valen 1, es decir $I = (\delta_{i,j})$. Resulta evidente entonces que dada $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ se tiene que $I_m\, A = A\, I_n = A$, es decir, la matriz identidad es el elemento neutro para el producto.

Recopilando toda esta información, podemos concluir que el conjunto de las matrices cuadradas $\mathcal{M}_n(\mathbb{K})$,

- considerando las operaciones de suma y producto por escalares, tiene estructura de **espacio vectorial**;
- considerando las operaciones de suma y producto tiene estructura de **anillo unitario** (no conmutativo);
- considerando las operaciones de suma, producto por escalares y producto tiene estructura de **álgebra unitaria** (no conmutativa).

5. Transposición y conjugación de matrices

Dada una matriz $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{R})$, se define su **matriz transpuesta** mediante $A^t = (b_{i,j}) \in \mathcal{M}_{n\times m}(\mathbb{R})$, donde

$$b_{i,j} = a_{j,i}\,, \quad i = 1,2,\ldots,n\,, \quad j = 1,2,\ldots,m\,.$$

Análogamente, dada una matriz $A = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{C})$, se define su **matriz conjugada** (adjunta) mediante $A^* = (b_{i,j}) \in \mathcal{M}_{n\times m}(\mathbb{C})$, donde

$$b_{i,j} = \bar{a}_{j,i}\,, \qquad i = 1,2,\ldots,n\,, \quad j = 1,2,\ldots,m\,.$$

Indicamos que se puede definir también la matriz transpuesta de una matriz compleja, pero es una noción que carece de mayor interés práctico (como veremos más adelante los conceptos de transposición y de conjugación de una matriz van asociados al de producto escalar, y, en el caso complejo, la sola transposición de la matriz no verifica las propiedades adecuadas).

Se puede comprobar fácilmente que dados $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$, $C \in \mathcal{M}_{n\times p}(\mathbb{K})$, $\alpha \in \mathbb{K}$, se verifican los siguientes **resultados**:

$$\begin{aligned}
(A^t)^t = A\,; &\quad (\text{resp.} \quad (A^*)^* = A\,;) \\
(A+B)^t = A^t + B^t\,; &\quad (\text{resp.} \quad (A+B)^* = A^* + B^*\,;) \\
0^t = 0\,; &\quad (\text{resp.} \quad 0^* = 0\,;) \\
(-A)^t = -A^t\,; &\quad (\text{resp.} \quad (-A)^* = -A^*\,;) \\
(\alpha\, A)^t = \alpha\, A^t\,; &\quad (\text{resp.} \quad (\alpha\, A)^* = \bar{\alpha}\, A^*\,;) \\
(A\,C)^t = C^t\, A^t\,; &\quad (\text{resp.} \quad (A\,C)^* = C^*\, A^*\,;) \\
I^t = I\,. &\quad (\text{resp.} \quad I^* = I\,.)
\end{aligned}$$

6. Partición en bloques de matrices

Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$; supongamos que $m = m_1 + m_2 + \cdots + m_M$ y $n = n_1 + n_2 + \cdots + n_N$, donde $m_I \geq 1$, $I = 1, 2, \ldots, M$ y $n_J \geq 1$, $J = 1, 2, \ldots, N$. Asociada a esta descomposición de los índices de las filas y las columnas de la matriz, podemos también considerar la **partición en bloques de submatrices** de la matriz A en la forma

$$A = \left(\begin{array}{c|c|c|c} A_{1,1} & A_{1,2} & \cdots & A_{1,N} \\ \hline A_{2,1} & A_{2,2} & \cdots & A_{2,N} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{M,1} & A_{M,2} & \cdots & A_{M,N} \end{array}\right) = (A_{I,J})_{I,J} = (A_{I,J}) \in \mathcal{M}_{m\times n}(\mathbb{K}),$$

donde $A_{I,J} \in \mathcal{M}_{m_I \times n_J}(\mathbb{K})$, $I = 1, 2, \ldots, M$, $J = 1, 2, \ldots, N$. Cuando una **matriz cuadrada sea particionada en bloques**, supondremos que las submatrices diagonales $A_{I,I}$ son cuadradas.

El interés de estas particiones en bloques es que ciertas **operaciones con matrices** permanecen "formalmente" válidas cuando los coeficientes $a_{i,j}$ de la matriz se cambian por submatrices $A_{I,J}$ de la partición (supuesto siempre que los órdenes de las submatrices permiten realizar las operaciones deseadas). Así, dos matrices del mismo orden se dice que están **particionadas conformemente para la suma**, cuando tienen el mismo número de filas y columnas de submatrices y, dos a dos, las submatrices correspondientes tienen el mismo orden de manera que pueden sumarse. Por otro lado, dos matrices de órdenes tales que puedan multiplicarse se dice que están **particionadas conformemente para el producto**, cuando el número de columnas de submatrices de la primera coincide con el número de filas de submatrices de la segunda y, dos a dos, los órdenes de las submatrices que deben multiplicarse y posteriormente sumarse permiten dichas operaciones.

De forma análoga, **transponer** (resp. **conjugar**) una matriz particionada en bloques equivale a transponer (resp. conjugar) los bloques de submatrices que conforman la partición, y transponer (resp. conjugar) cada uno de estos bloques individualmente.

Dada $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ si notamos por $A_i^f \in \mathcal{M}_{1\times n}(\mathbb{K})$ $i = 1, 2, \ldots, m$, cada una de sus filas y por $A_j^c \in \mathcal{M}_{m\times 1}(\mathbb{K})$, $j = 1, 2, \ldots, n$, cada una de sus columnas, podemos escribir

$$A = \left(\begin{array}{c} A_1^f \\ \hline A_2^f \\ \hline \vdots \\ \hline A_m^f \end{array}\right) = \left(\begin{array}{c|c|c|c} A_1^c & A_2^c & \cdots & A_n^c \end{array}\right) \in \mathcal{M}_{m\times n}(\mathbb{K}),$$

Queda claro entonces que considerando la matriz identidad $I \in \mathcal{M}_n(\mathbb{K})$ y con ayuda de los vectores elementales, podemos escribir

$$I = \left(\begin{array}{c} E_1^t \\ \hline E_2^t \\ \hline \vdots \\ \hline E_n^t \end{array}\right) = \left(\begin{array}{c|c|c|c} E_1 & E_2 & \cdots & E_n \end{array}\right) \in \mathcal{M}_n(\mathbb{K}).$$

7. Producto escalar ordinario en $\mathbb{K}^n$

Dados dos vectores X e Y de $\mathbb{R}^n$ (resp. $\mathbb{C}^n$), se define su **producto escalar ordinario** mediante

$$\langle X, Y\rangle = Y^t\,X = \sum_i y_i\,x_i = y_i\,x_i \in \mathbb{R}\,. \quad \text{(resp.} \quad \langle X, Y\rangle = Y^*\,X = \sum_i \bar{y}_i\,x_i = \bar{y}_i\,x_i \in \mathbb{C}\,.)$$

Observamos que la definición anterior verifica las propiedades necesarias de todo **producto escalar** como **forma bilineal** (resp. **sesquilineal**) **simétrica** (resp. **hermítica**) y **definida positiva**, es decir,

$$\begin{aligned}
\langle X_1 + X_2, Y\rangle &= \langle X_1, Y\rangle + \langle X_2, Y\rangle\,, && \text{para todo} \quad X_1, X_2, Y \in \mathbb{K}^n\,,\\
\langle X, Y_1 + Y_2\rangle &= \langle X, Y_1\rangle + \langle X, Y_2\rangle\,, && \text{para todo} \quad X, Y_1, Y_2 \in \mathbb{K}^n\,,\\
\alpha\,\langle X, Y\rangle &= \langle \alpha\,X, Y\rangle = \langle X, \alpha\,Y\rangle\,, && \text{para todo} \quad \alpha \in \mathbb{R}\,, \quad X, Y \in \mathbb{R}^n\,,\\
\text{(resp.} \quad \alpha\,\langle X, Y\rangle &= \langle \alpha\,X, Y\rangle = \langle X, \bar{\alpha}\,Y\rangle\,, && \text{para todo} \quad \alpha \in \mathbb{C}\,, \quad X, Y \in \mathbb{C}^n\,,)\\
\langle X, Y\rangle &= \langle Y, X\rangle\,, && \text{para todo} \quad X, Y \in \mathbb{R}^n\,.\\
\text{(resp.} \quad \langle X, Y\rangle &= \overline{\langle Y, X\rangle}\,, && \text{para todo} \quad X, Y \in \mathbb{C}^n\,,)\\
\langle X, X\rangle &> 0\,, && \text{para todo} \quad X \in \mathbb{K}^n - \{0\}\,.
\end{aligned}$$

Observamos que en el caso complejo, la propiedad de hermítica implica que $\langle X, X\rangle \in \mathbb{R}$ para todo $X \in \mathbb{C}^n$, por lo que tiene perfecto sentido la última propiedad señalada antes.

Una vez definido un producto escalar en $\mathbb{K}^n$, recordamos que un vector X se dice **unitario** si $\langle X, X\rangle = 1$, que dos vectores X e Y se dicen **ortogonales** si $\langle X, Y\rangle = 0$ y se dicen **ortonormales** si son ortogonales y, además, ambos son unitarios; estas definiciones se ortogonalidad y ortonormalidad se extienden de forma natural a un conjunto cualquiera de vectores $\{X_i\}_{i\in I}$.

Asociado también con el concepto de producto escalar, recordamos los siguientes **resultados**:

- dado cualquier vector no nulo $X \in \mathbb{K}^n$, el vector $\dfrac{1}{\sqrt{\langle X, X\rangle}}\,X$ es unitario y proporcional a X (mantiene la misma dirección y sentido);
- todo sistema ortogonal es linealmente independiente;
- a partir de cualquier sistema de vectores linealmente independiente, el **proceso de ortonormalización de Gram-Schmidt** permite construir un nuevo sistema ortonormal, de manera que en todo momento el subespacio engendrado por los k primeros vectores del nuevo sistema es el mismo que el subespacio engendrado por los k primeros vectores del sistema original.

8. Estructura de una matriz

En cuanto a la **distribución** de los elementos nulos y no nulos, se dice que una matriz $A \in \mathcal{M}_n(\mathbb{K})$ tiene **estructura**:

- **diagonal** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i \neq j$;
- **tridiagonal** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $|i - j| > 1$;
- **pentadiagonal** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $|i - j| > 2$;
- **banda** si existe $p \in \mathbb{N}$ (llamado semi-anchura de la banda, "pequeño" en comparación con n) tal que $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $|i - j| > p$;
- **triangular superior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i > j$;
- **estrictamente triangular superior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i \geq j$;
- **triangular inferior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i < j$;
- **estrictamente triangular inferior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i \leq j$;
- **Hessenberg superior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $i > j + 1$;
- **Hessenberg inferior** si $a_{i,j} = 0$, para todo $i, j = 1, 2, \ldots, n$, con $j > i + 1$.

Indicamos también que si la matriz A está **particionada en bloques**, todos los conceptos anteriores se generalizan de forma natural cambiando $a_{i,j}$ por $A_{I,J}$; de esta manera se habla de matriz diagonal por bloques, tridiagonal por bloques, etc.

Para las matrices diagonales y triangulares (tanto superior como inferior) se verifica que la **suma**, el **opuesto**, el **producto por un escalar** y el **producto de matrices** de cada uno de estos tipos es del mismo tipo, siendo además los **elementos diagonales** de la matriz resultante la suma, el opuesto, el producto por el escalar o el producto de los correspondientes elementos diagonales de las matrices en cuestión.

Evidentemente la **transpuesta** o **conjugada** de cualquier matriz diagonal, tridiagonal, pentadiagonal o banda es del mismo tipo, y la de cualquier matriz triangular o Hessenberg superior es inferior (y viceversa).

9. Matrices inversibles

Recordamos que $\mathcal{M}_n(\mathbb{K})$ tiene estructura de anillo unitario pero no de cuerpo, ya que el producto de matrices, además de no ser conmutativo, no asegura la existencia de inverso de cualquier matriz no nula. No obstante, una matriz $A \in \mathcal{M}_n(\mathbb{K})$ se dice **inversible**, si existe otra matriz $A^{-1} \in \mathcal{M}_n(\mathbb{K})$, llamada su **matriz inversa**, tal que $A\,A^{-1} = A^{-1}\,A = I$.

Referente a la inversibilidad de una matriz $A \in \mathcal{M}_n(\mathbb{K})$, se comprueban fácilmente los siguientes **resultados**:

- la matriz inversa, de existir, es única; además, dicha matriz inversa es a su vez inversible, siendo $(A^{-1})^{-1} = A$;
- la matriz nula 0 no es inversible;
- si A es inversible, entonces $-A$ también lo es, siendo $(-A)^{-1} = -A^{-1}$;
- si A es inversible y $\alpha \neq 0$, entonces $\alpha\,A$ es inversible, siendo $(\alpha\,A)^{-1} = \frac{1}{\alpha}\,A^{-1}$;
- si A y B son inversibles, también lo es $A\,B$, siendo $(A\,B)^{-1} = B^{-1}\,A^{-1}$;
- la matriz identidad I es inversible, siendo $I^{-1} = I$;
- si $\mathbb{K} = \mathbb{R}$ y A es inversible, también lo es A^t, siendo $(A^t)^{-1} = (A^{-1})^t$;
- si $\mathbb{K} = \mathbb{C}$ y A es inversible, también lo es A^*, siendo $(A^*)^{-1} = (A^{-1})^*$;
- si A es diagonal o triangular, entonces es inversible si y sólo si sus elementos diagonales son no nulos, en cuyo caso su inversa tiene la misma estructura y los elementos diagonales de la inversa son los inversos de los elementos diagonales de la matriz inicial;
- las columnas de una matriz inversible forman un sistema linealmente independiente (y por tanto una base) de vectores de $\mathbb{K}^n$ (y viceversa).

10. Tipos importantes de matrices

Una matriz $A \in \mathcal{M}_n(\mathbb{K})$ se dice que es:

- **simétrica**, si $\mathbb{K} = \mathbb{R}$ y $A^t = A$ (**antisimétrica** si $A^t = -A$);
- **hermítica**, si $\mathbb{K} = \mathbb{C}$ y $A^* = A$ (**antihermítica** si $A^* = -A$);
- **ortogonal**, si $\mathbb{K} = \mathbb{R}$ y $A\,A^t = A^t\,A = I$ (en este caso $A^{-1} = A^t$);
- **unitaria**, si $\mathbb{K} = \mathbb{C}$ y $A\,A^* = A^*\,A = I$ (en este caso $A^{-1} = A^*$);
- **normal**, si $A\,A^* = A^*\,A$ (observamos que todos los tipos de matrices anteriores son normales);
- **definida positiva**, si es simétrica (resp. hermítica) y

$$X^t\,A\,X > 0\,, \quad \text{para todo} \quad X \in \mathbb{R}^n - \{0\}\,.$$
$$(\text{resp.} \quad X^*\,A\,X > 0\,, \quad \text{para todo} \quad X \in \mathbb{C}^n - \{0\}\,.)$$

Respecto a la definición de matriz definida positiva realizamos las siguientes **puntualizaciones**. En primer lugar indicamos que cambiando en la misma el signo $>$ por $\geq$, $<$ ó $\leq$, se obtienen, respectivamente, definiciones de matriz **definida no negativa** (semi-definida positiva), **definida negativa** o **definida no positiva** (semi-definida negativa). Por otro lado observamos que en el caso complejo, al obligar que la matriz sea hermítica se implica que $X^*\,A\,X \in \mathbb{R}$ para todo $X \in \mathbb{C}^n$, por lo que la definición tiene perfecto sentido. Por último, el concepto de matriz definida positiva no debe confundirse con el de **matriz positiva** (concepto menos utilizado); una matriz se dice positiva si todos sus elementos son positivos.

Referente a las **matrices ortogonales-unitarias**, destacamos el hecho de que sus columnas, no sólo forman un sistema linealmente independiente (y por tanto una base) de $\mathbb{K}^n$ al ser dichas matrices inversibles, sino que además constituyen un **sistema ortonormal de vectores** para el producto escalar ordinario (y viceversa).

11. Traza

Dada una matriz $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ se define su **traza** mediante

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{i,i} = a_{i,i}\,,$$

es decir la suma de los elementos de la diagonal principal.

Dados $A, B \in \mathcal{M}_n(\mathbb{K})$, $C = (c_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})$, $D \in \mathcal{M}_{n\times m}(\mathbb{K})$, $\alpha \in \mathbb{K}$, se comprueban fácilmente los siguientes **resultados**:

$$\begin{aligned}
\operatorname{tr}(A+B) &= \operatorname{tr}(A) + \operatorname{tr}(B)\,;\\
\operatorname{tr}(0) &= 0\,;\\
\operatorname{tr}(-A) &= -\operatorname{tr}(A)\,;\\
\operatorname{tr}(\alpha\, A) &= \alpha\ \operatorname{tr}(A)\,;\\
\operatorname{tr}(C\, D) &= \operatorname{tr}(D\, C)\,;\\
\operatorname{tr}(I_n) &= n\,;\\
\operatorname{tr}(A^t) &= \operatorname{tr}(A)\,;\\
\operatorname{tr}(A^*) &= \overline{\operatorname{tr}(A)}\,;\\
\operatorname{tr}(C^t\, C) &= \sum_{i=1}^{m}\sum_{j=1}^{n} {c_{i,j}}^2\,;\\
\operatorname{tr}(C^*\, C) &= \sum_{i=1}^{m}\sum_{j=1}^{n} |c_{i,j}|^2\,.
\end{aligned}$$

12. Determinante

Notamos por Σ_n el grupo de todas las permutaciones σ del conjunto $\{1, 2, \ldots, n\}$; dicho conjunto posee $n!$ elementos. Dada $\sigma \in \Sigma_n$, notamos por $\varepsilon_\sigma \in \{1, -1\}$ la signatura de dicha permutación (depende de la paridad del número de permutaciones elementales necesarias para llevar σ al orden natural).

Utilizando los conceptos anteriores, dada una matriz $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ se define su **determinante** como

$$\det(A) = \sum_{\sigma \in \Sigma_n} \varepsilon_\sigma \, a_{1,\sigma(1)} \, a_{2,\sigma(2)} \, \cdots \, a_{n,\sigma(n)} \, .$$

Recordamos también que una matriz A se dice **regular** si su determinante es no nulo y **singular** en caso contrario.

Dados $A, B \in \mathcal{M}_n(\mathbb{K})$, $\alpha \in \mathbb{K}$, se verifican las siguientes **resultados**:

$$\begin{array}{rrcl}
 & \det(0) & = & 0\,; \\
 & \det(-A) & = & (-1)^n \det(A)\,; \\
 & \det(\alpha\, A) & = & \alpha^n \det(A)\,; \\
 & \det(A\,B) & = & \det(A)\,\det(B)\,; \\
 & \det(I) & = & 1\,; \\
 & \det(A^t) & = & \det(A)\,; \\
 & \det(A^*) & = & \overline{\det(A)}\,; \\
\text{A es inversible si y sólo si es regular, y en ese caso} & \det(A^{-1}) & = & (\det(A))^{-1}\,; \\
\text{si A es diagonal o triangular} & \det(A) & = & \displaystyle\prod_{i=1}^{n} a_{i,i}\,.
\end{array}$$

Observamos que una matriz diagonal o triangular es regular y, por tanto, inversible si y sólo si todos sus elementos diagonales son no nulos, en cuyo caso los elementos diagonales de la inversa son los respectivos inversos de los elementos diagonales de la matriz inicial.

13. Cofactores y rango

Para cada elemento $a_{i,j}$, $i, j = 1, 2, \ldots, n$, de una matriz $A \in \mathcal{M}_n(\mathbb{K})$, se define su **menor complementario** $M_{i,j}$ como el determinante de la submatriz que resulta de suprimir de A la fila i-ésima y la columna j-ésima a la que pertenece dicho elemento; de forma análoga se define su **cofactor (adjunto)** $A_{i,j}$ como $A_{i,j} = (-1)^{i+j} M_{i,j}$. Finalmente se define la **matriz de cofactores** de la matriz A, $\operatorname{cof}(A) \in \mathcal{M}_n(\mathbb{K})$, como la matriz formada por los cofactores de cada uno de los elementos, es decir, $\operatorname{cof}(A) = (A_{i,j})$, y la **matriz adjunta** de la matriz A, $\operatorname{adj}(A) \in \mathcal{M}_n(\mathbb{K})$, como la transpuesta de la anterior, es decir $\operatorname{adj}(A) = \operatorname{cof}(A)^t = (A_{j,i})$ (aún en el caso complejo, sólo se transpone la matriz). Esta matriz adjunta verifica la **propiedad fundamental**

$$A \operatorname{adj}(A) = \operatorname{adj}(A)\, A = \det(A)\, I\,.$$

Observamos que esta relación equivale a que el producto de cualquier línea (fila o columna) de la matriz A por sus respectivos cofactores vale $\det(A)$, mientras que por los cofactores de una línea paralela, diferente de la que se está considerando, es nulo, es decir,

$$\begin{aligned} \sum_{k=1}^{n} a_{i,k}\, A_{j,k} &= \det(A)\, \delta_{i,j}\,, \quad i, j = 1, 2, \ldots, n\,, \\ \sum_{k=1}^{n} a_{k,i}\, A_{k,j} &= \det(A)\, \delta_{i,j}\,, \quad i, j = 1, 2, \ldots, n\,, \end{aligned}$$

fórmulas conocidas como **desarrollo de un determinante por los cofactores de una línea**. Otra consecuencia importante de la relación anterior es que en caso de que la matriz A sea regular, entonces su **inversa** viene dada por

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)\,.$$

Por otro lado, dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, se define su **rango**, notado por $\mathrm{r}(A)$, como el orden de la mayor submatriz cuadrada de determinante no nulo.

14. Sistemas de ecuaciones lineales

Consideremos un **sistema de m ecuaciones lineales con n incógnitas**,

$$\left\{ \begin{array}{ccccccccc} a_{1,1}\,x_1 & + & a_{1,2}\,x_2 & + & \cdots & + & a_{1,n}\,x_n & = & b_1 \\ a_{2,1}\,x_1 & + & a_{2,2}\,x_2 & + & \cdots & + & a_{2,n}\,x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}\,x_1 & + & a_{m,2}\,x_2 & + & \cdots & + & a_{m,n}\,x_n & = & b_m \end{array} \right\},$$

que podemos escribir en forma matricial como $A\,X = B$, donde

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = (a_{i,j}) \in \mathcal{M}_{m\times n}(\mathbb{K})\,.$$

es la **matriz de coeficientes**, y

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = (x_j) \in \mathbb{K}^n \quad \text{y} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix} = (b_i) \in \mathbb{K}^m$$

son, respectivamente, el **vector de incógnitas** y el **vector de segundos miembros**; introducimos también la **matriz ampliada**

$$(A|B) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right) = (a_{i,j}|b_i) \in \mathcal{M}_{m\times(n+1)}(\mathbb{K})\,.$$

Recordamos el **teorema de Rouché-Fröbenius** que dice que la condición necesaria y suficiente para que dicho sistema tenga solución (sea **compatible**), es que el rango de la matriz de coeficientes coincida con el rango de la matriz ampliada; en este caso, si dicho rango r coincide con n, la solución es única (**determinado**) y si $r < n$ hay infinitas soluciones (**indeterminado**) que dependen de $n - r$ parámetros. A **efectos prácticos**, todas aquellas ecuaciones que no intervienen a la hora de determinar el rango pueden eliminarse del sistema, pues se tratan de combinaciones lineales de las otras ecuaciones, y las $n - r$ incógnitas que no intervienen en la determinación del rango pueden pasarse el segundo miembro como incógnitas secundarias y resolver el sistema en función de las mismas.

Recordamos también la **regla de Cramer** para resolver un sistema $A\,X = B$ compatible y determinado de n ecuaciones con n incógnitas, que dice que

$$x_j = \frac{\det(A_1^c|\cdots|B|\cdots|A_n^c)}{\det(A)}\,, \quad j = 1, 2, \ldots, n\,,$$

donde en el numerador aparece el determinante de la matriz que resulta de sustituir en A la j-ésima columna por B.

15. Matrices de transformación

Sea $\sigma_{k,l}$ una permutación elemental de $\{1, 2, \ldots, n\}$ (permutación en la que sólo cambian entre sí los elementos k y l) y construyamos la matriz, llamada **matriz de permutación**, $P_{\sigma_{k,l}} = (\delta_{i,\sigma_{k,l}(j)}) \in \mathcal{M}_n(\mathbb{K})$. Se comprueba fácilmente que dicha matriz tiene determinante -1 y que es simétrica-hermítica y ortogonal-unitaria, por lo que coincide con su inversa; además dada $A \in \mathcal{M}_n(\mathbb{K})$ las matrices $P_{\sigma_{k,l}}\, A$ y $A\, P_{\sigma_{k,l}}$ quedan caracterizadas por permutar las filas (en el primer caso) y las columnas (en el segundo) de A siguiendo la permutación $\sigma_{k,l}$. Resulta entonces inmediato concluir que el determinante de una matriz, cuando se permutan dos líneas paralelas entre sí, cambia de signo.

Consideremos ahora la matriz $C_{k,l}(\alpha) = I + \alpha\, E_{k,l} \in \mathcal{M}_n(\mathbb{K})$, $k, l \in \{1, 2, \ldots, n\}$, $k \neq l$, $\alpha \in \mathbb{K}$. De forma similar al caso anterior, se comprueba fácilmente que dicha matriz tiene determinante 1 y que su inversa es $C_{k,l}(-\alpha)$; dada ahora una matriz $A \in \mathcal{M}_n(\mathbb{K})$, las matrices $C_{k,l}(\alpha)\, A$ y $A\, C_{k,l}(\alpha)$ quedan caracterizadas por ser análogas a A excepto la fila k-ésima a la que se le suma una combinación lineal de α veces la fila l-ésima en el primer caso, y la columna l-ésima a la que se le suma una combinación lineal de α veces la columna k-ésima en el segundo caso. Se concluye entonces que el determinante de una matriz, cuando a una línea se le suma un múltiplo escalar de otra línea paralela, no varía.

$$P_{\sigma_{k,l}} = \begin{array}{c c} & \begin{array}{ccccccc} & & (k) & & (l) & & \end{array} \\ \begin{array}{c} \\ \\ (k) \\ \\ (l) \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array} \qquad C_{\sigma_{k,l}}(\alpha) = \begin{array}{c c} & \begin{array}{ccccccc} & & (k) & & (l) & & \end{array} \\ \begin{array}{c} \\ \\ (k) \\ \\ (l) \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & \alpha & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}$$

Finalmente consideramos la matriz $P_k(\alpha) = I + (\alpha - 1)\, E_{k,k} \in \mathcal{M}_n(\mathbb{K})$, $k \in \{1, 2, \ldots, n\}$, $\alpha \in \mathbb{K}$. Se comprueba fácilmente que dicha matriz tiene determinante α, que, supuesta inversible, su inversa es $P_k(\alpha^{-1})$ y que dada $A \in \mathcal{M}_n(\mathbb{K})$, las matrices $P_k(\alpha)\, A$ y $A\, P_k(\alpha)$ quedan caracterizadas por coincidir con A, excepto la fila (en el primer caso) y columna (en el segundo) k-ésima que queda multiplicada por α. Se concluye entonces que el determinante de una matriz, cuando una línea se multiplica por un escalar, se multiplica también por dicho escalar.

$$P_k(\alpha) = \begin{array}{c c} & \begin{array}{ccccc} & & (k) & & \end{array} \\ \begin{array}{c} \\ \\ (k) \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{array}$$

16. Autovalores y autovectores de matrices

Dada una matriz $A \in \mathcal{M}_n(\mathbb{K})$, se llaman **autovalores** (o **valores propios**) λ de A a las n raíces reales o complejas, simples o múltiples del **polinomio característico** de A, que se define como

$$p(\lambda) = \det(A - \lambda\, I) = (-1)^n\, (\lambda^n + a_1\, \lambda^{n-1} + \cdots + a_{n-1}\, \lambda + a_n) \in \mathcal{P}_n[\lambda]\,.$$

Se dice que un vector $X \in \mathbb{C}^n - \{0\}$ es un **autovector** (o **vector propio**) asociado al autovalor λ si

$$A\, X = \lambda\, X \iff (A - \lambda\, I)\, X = 0\,.$$

Notamos que la relación anterior puede verse como un sistema de n ecuaciones lineales con n incógnitas homogéneo, que tendrá solución no trivial al ser la matriz de coeficientes de rango estrictamente inferior a n por la propia definición de autovalor. Observamos también que una matriz real puede tener autovalores (y por tanto autovectores) complejos; en ese caso, si $\lambda \in \mathbb{C}$ es autovalor, $\bar{\lambda}$ también lo es.

Asociados con los autovalores de una matriz se tienen también las siguientes importantes definiciones; se llama **espectro** de la matriz $A \in \mathcal{M}_n(\mathbb{K})$ a

$$\mathrm{sp}(A) = \sigma(A) = \{\lambda \,:\, \lambda \text{ es autovalor de } A\} \subset \mathbb{C}\,;$$

se llama **radio espectral** de la matriz A a

$$\rho(A) = \text{máx}\{|\lambda| \,:\, \lambda \text{ es autovalor de } A\} \geq 0\,.$$

Respecto de los autovalores de una matriz, se verifican los siguientes **resultados**:

- los autovectores asociados a un mismo autovalor λ, vector nulo incluido, tienen estructura de subespacio vectorial, notado E_λ;
- autovectores asociados a autovalores diferentes son linealmente independientes;
- los autovalores de los múltiplos escalares, potencias e inversa de una matriz, son los correspondientes múltiplos escalares, potencias e inversos de los autovalores de la matriz, además asociados a los mismos autovectores;
- los autovalores de las matrices diagonales y triangulares son sus elementos diagonales;
- una matriz y su transpuesta tienen los mismos autovalores;
- una matriz y su conjugada tienen autovalores conjugados;
- una matriz es singular si y sólo si tiene a 0 como autovalor;
- los autovalores de una matriz simétrica-hermítica son reales y los autovectores asociados a autovalores diferentes son ortogonales respecto del producto escalar ordinario;
- los autovalores de una matriz antisimétrica-antihermítica tienen parte real nula;
- los autovalores de una matriz ortogonal-unitaria son de módulo uno.

17. Matrices semejantes

Dos matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ se dicen **semejantes**, si existe una matriz inversible, llamada **matriz de paso**, $P \in \mathcal{M}_n(\mathbb{K})$ tal que

$$P^{-1}\,A\,P = B \iff A\,P = P\,B \iff P^{-1}\,A = B\,P^{-1} \iff A = P\,B\,P^{-1}\,.$$

Una matriz se dice **diagonalizable** (resp. **triangularizable**) si es semejante a una matriz diagonal (resp. triangular).

Respecto de las matrices semejantes, se verifican los siguientes **resultados**:

- dos matrices semejantes tienen la misma traza;
- dos matrices semejantes tienen el mismo determinante;
- dos matrices semejantes tienen el mismo polinomio característico y, por tanto, los mismos autovalores;
- si A y B son semejantes con matriz de paso P y $X \in \mathbb{C}^n - \{0\}$ es un autovector de A asociado al autovalor λ, entonces $Y = P^{-1}\,X \in \mathbb{C}^n - \{0\}$ es autovector de B asociado al mismo autovalor.

Sea $A \in \mathcal{M}_n(\mathbb{K})$ una matriz **diagonalizable**; existen por tanto $D \in \mathcal{M}_n(\mathbb{K})$ matriz diagonal y $P \in \mathcal{M}_n(\mathbb{K})$ matriz inversible, tales que $P^{-1}\,A\,P = D$. Ponemos $D = \mathrm{diag}\{\lambda_i\}$, donde $\lambda_i \in \mathrm{sp}(D) = \mathrm{sp}(A)$ y denotamos por X_i las columnas de la matriz P, es decir, $P = (X_1|X_2|\cdots|X_n)$. Observamos entonces que

$$\begin{aligned} P^{-1}\,A\,P = D &\iff A\,P = P\,D \\ &\iff A\,(X_1|X_2|\cdots|X_n) = (X_1|X_2|\cdots|X_n)\;\mathrm{diag}\{\lambda_i\} \\ &\iff A\,(X_1|X_2|\cdots|X_n) = (\lambda_1\,X_1|\lambda_2\,X_2|\cdots|\lambda_n\,X_n) \\ &\iff A\,X_i = \lambda_i\,X_i\,, \quad i = 1, 2, \ldots, n\,, \end{aligned}$$

es decir, cada columna X_i de la matriz P es un autovector de A asociado al respectivo autovalor λ_i que aparece en la diagonal de la matriz D. El hecho de que P sea inversible significa que sus columnas forman una base de $\mathbb{K}^n$. Por tanto, decir que A es diagonalizable es equivalente a decir que existe una base de $\mathbb{K}^n$ formada por autovectores; además debido a que puede haber autovalores múltiples deberá ocurrir que el subespacio de los autovectores asociados a un mismo autovalor debe tener la misma dimensión (**dimensión geométrica**) que la multiplicidad (**dimensión algebraica**) del autovalor como raíz del polinomio característico.

18. Reducción de matrices

Teorema 1 (Teorema de Schur) *Se verifica que:*

- *dada* $A \in \mathcal{M}_n(\mathbb{C})$*, existe* $U \in \mathcal{M}_n(\mathbb{C})$ *matriz de paso unitaria, tal que* $U^{-1} A U = T$*, con* $T \in \mathcal{M}_n(\mathbb{C})$ *triangular (*A *es triangularizable con matriz de paso unitaria);*
- *dada* $A \in \mathcal{M}_n(\mathbb{C})$*, es normal si y sólo si existe* $U \in \mathcal{M}_n(\mathbb{C})$ *matriz de paso unitaria, tal que* $U^{-1} A U = D$*, con* $D \in \mathcal{M}_n(\mathbb{C})$ *diagonal (*A *es diagonalizable con matriz de paso unitaria);*
- *dada* $A \in \mathcal{M}_n(\mathbb{R})$ *simétrica, existe* $O \in \mathcal{M}_n(\mathbb{R})$ *matriz de paso ortogonal, tal que* $O^{-1} A O = D$*, con* $D \in \mathcal{M}_n(\mathbb{R})$ *diagonal (*A *es diagonalizable con matriz de paso ortogonal).*

Comentarios sobre el Teorema anterior:

- las matrices de paso no son únicas (basta considerar, por ejemplo, que $A = I$);
- los elementos diagonales de las matrices triangulares o diagonales que aparecen en el teorema son los autovalores de la misma que, al ser semejante a la matriz inicial, son también sus autovalores;
- se concluye entonces que

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i\,, \quad \det(A) = \prod_{i=1}^{n} \lambda_i\,;$$

- si A es normal, entonces es diagonalizable, y al ser la matriz de paso (cuyas columnas son autovectores de A) unitaria, significa que es posible encontrar una base ortonormal de autovectores de A. Además concluimos que autovectores asociados a autovalores distintos son ortogonales respecto del producto escalar ordinario de $\mathbb{C}^n$;
- toda matriz real que conmute con su transpuesta (considerada como compleja es normal) es diagonalizable, pero la matriz diagonal y la matriz de paso unitaria pueden ser complejas; solamente en el caso de que sea simétrica podemos asegurar que la matriz diagonal es real (formada por sus autovalores) y la matriz de paso (formada por sus autovectores que son ortonormales) es ortogonal.

19. Valores singulares

Con la ayuda de los autovalores se puede dar una **caracterización** de las **matrices definidas positiva**; así, se puede probar que una matriz simétrica-hermítica (recordamos que sus autovalores son siempre reales) es definida positiva (resp. no negativa) si y sólo si sus autovalores son positivos (resp. no negativos).

Por otro lado, dada una matriz $A \in \mathcal{M}_n(\mathbb{R})$ (resp. $A \in \mathcal{M}_n(\mathbb{C})$), la matriz $A^t\,A$ (resp. $A^*\,A$) siempre es definida no negativa (es definida positiva si y sólo si A es regular); esto implica que los autovalores de $A^t\,A$ (resp. $A^*\,A$) (que denotaremos por η_i, $i = 1, 2, \ldots, n$) son siempre reales y mayores o iguales que cero (estrictamente mayores que cero si y sólo si A es regular).

Dada $A \in \mathcal{M}_n(\mathbb{R})$ (resp. $A \in \mathcal{M}_n(\mathbb{C})$), se llaman **valores singulares** de A, notados por μ_i, $i = 1, 2, \ldots, n$, a las raíces cuadradas positivas de los autovalores de $A^t\,A$ (resp. $A^*\,A$), es decir, $\mu_i = +\sqrt{\eta_i}$. A partir de las consideraciones anteriores queda claro que esta definición tiene perfecto sentido.

Teorema 2 (Diagonalización mediante valores singulares) *Dada $A \in \mathcal{M}_n(\mathbb{K})$, existen $U, V \in \mathcal{M}_n(\mathbb{K})$ ortogonales-unitarias tales que $U^{-1}\,A\,V = D$, con $D \in \mathcal{M}_n(\mathbb{R})$ diagonal y siendo dicha matriz diagonal la compuesta por los valores singulares de A, es decir, $D = \mathrm{diag}\{\mu_i\}$ (A es pseudo-diagonalizable con matrices de paso unitarias).*

20. El espacio normado $\mathbb{K}^n$

Se llama **norma** (vectorial) sobre el espacio vectorial a toda aplicación $\|\cdot\|\colon \mathbb{K}^n \to \mathbb{R}^+$ tal que

$$\begin{aligned} \|X\| = 0 \quad &\text{si y sólo si} \quad X = 0\,, \\ \|X+Y\| \le \|X\| + \|Y\|\,, \quad &\text{para todo} \quad X, Y \in \mathbb{K}^n \quad \text{(desigualdad triangular)}, \\ \|\alpha\, X\| = |\alpha|\, \|X\|\,, \quad &\text{para todo} \quad \alpha \in \mathbb{K}\,, \quad X \in \mathbb{K}^n\,; \end{aligned}$$

el par formado por $(\mathbb{K}^n, \|\cdot\|)$ recibe el nombre de **espacio (vectorial) normado**.

Algunos **ejemplos** de norma en $\mathbb{K}^n$:

- $\|X\|_1 = \displaystyle\sum_{i=1}^{n} |x_i|$;
- $\|X\|_2 = \left(\displaystyle\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$ (**norma euclídea**);
- $\|X\|_\infty = \displaystyle\max_{i=1,2,\ldots,n} |x_i|$;
- $\|X\|_p = \left(\displaystyle\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$ $(p \ge 1)$;

observamos que, en particular, $(\mathbb{K}, |\cdot|)$ (valor absoluto o módulo) es un espacio normado.

Todo **producto escalar** $\langle \cdot, \cdot \rangle$ en $\mathbb{K}^n$ induce una norma sin más que poner

$$\|X\| = \sqrt{\langle X, X \rangle}\,, \quad \text{para todo} \quad X \in \mathbb{K}^n\,;$$

además, esta norma verifica la **desigualdad de Cauchy-Schwarz**

$$|\langle X, Y \rangle| \le \|X\|\, \|Y\|\,, \quad \text{para todo} \quad X, Y \in \mathbb{K}^n\,;$$

Lo recíproco no es cierto; hay normas que no son inducidas por ningún producto escalar; se comprueba que la norma inducida por un producto escalar verifica la llamada **regla del paralelogramo**, es decir,

$$\|X+Y\|^2 + \|X-Y\|^2 = 2\,\|X\|^2 + 2\,\|Y\|^2\,, \quad \text{para todo} \quad X, Y \in \mathbb{K}^n\,;$$

recíprocamente, si una norma verifica la regla del paralelogramo, ésta es inducida por un producto escalar.

Es inmediato comprobar que la norma euclídea es la norma inducida por el producto escalar ordinario de $\mathbb{K}^n$.

21. Normas, espacios métricos y espacios topológicos

Toda norma $\|\cdot\|$ induce en $\mathbb{K}^n$ una estructura de **espacio métrico** (y por tanto de **espacio topológico**), sin más que poner

$$\mathrm{d}(X,Y) = \|X - Y\|\,, \quad \text{para todo} \quad X, Y \in \mathbb{K}^n\,.$$

Esta estructura de espacio métrico y topológico nos permite establecer los conceptos de **límite** de una sucesión y de **continuidad** de una función:

- una sucesión $\{X_k\}_{k\in\mathbb{N}}$ de vectores de $\mathbb{K}^n$ tiene límite $X \in \mathbb{K}^n$, si para todo $\varepsilon > 0$ existe $k_\varepsilon \in \mathbb{N}$ tal que si $k \geq k_\varepsilon$, entonces $\|X_k - X\| < \varepsilon$;
- una función $f\colon D \subset \mathbb{K}^n \to \mathbb{K}^m$ es continua en $X \in \mathbb{K}^n$, si para todo $\varepsilon > 0$ existe $\delta_\varepsilon > 0$ tal que si $Y \in D$ con $\|Y - X\| < \delta_\varepsilon$, entonces $\|f(Y) - f(X)\| < \varepsilon$.

Dos normas $\|\cdot\|$ y $\|\cdot\|'$ se dicen **equivalentes**, si existen $C, C' > 0$ tales que

$$C\,\|X\|' \leq \|X\| \leq C'\,\|X\|'\,, \quad \text{para todo} \quad X \in \mathbb{K}^n\,;$$

el concepto de normas equivalentes implica que las métricas asociadas son también equivalentes, es decir, definen la misma topología, por lo que los conceptos de límite y continuidad antes mencionados son independientes de las normas equivalentes utilizadas en $\mathbb{K}^n$.

Algunos **resultados** relacionados con los conceptos métricos y topológicos mencionados:

- la propia aplicación norma es una función continua;
- en $\mathbb{K}^n$ todas las normas son equivalentes;
- toda aplicación lineal es continua;
- en cuanto $\mathcal{M}_{m\times n}(\mathbb{K})$ tiene estructura de espacio vectorial, le son de aplicación todas las definiciones anteriores.

22. El álgebra normada $\mathcal{M}_n(\mathbb{K})$

En el álgebra $\mathcal{M}_n(\mathbb{K})$, se llama **norma matricial** a toda aplicación $\|\cdot\|\colon \mathcal{M}_n(\mathbb{K}) \to \mathbb{R}^+$ tal que

$$\begin{aligned} \|A\| = 0 \quad &\text{si y sólo si} \quad A = 0\,, \\ \|A+B\| \leq \|A\| + \|B\|\,, \quad &\text{para todo} \quad A, B \in \mathcal{M}_n(\mathbb{K})\,, \\ \|\alpha\, A\| = |\alpha|\, \|A\|\,, \quad &\text{para todo} \quad \alpha \in \mathbb{K}\,, \quad A \in \mathcal{M}_n(\mathbb{K})\,, \\ \|A\, B\| \leq \|A\|\, \|B\|\,, \quad &\text{para todo} \quad A, B \in \mathcal{M}_n(\mathbb{K})\,. \end{aligned}$$

Toda norma matricial es una norma vectorial, cuando $\mathcal{M}_n(\mathbb{K})$ se considera solamente como espacio vectorial de dimensión n^2.

Teorema 3 (Normas matriciales subordinadas) *Sea $\|\cdot\|_V$ una norma vectorial sobre $\mathbb{K}^n$; para cada $A \in \mathcal{M}_n(\mathbb{K})$ se verifica que*

$$\sup_{X \in \mathbb{K}^n - \{0\}} \frac{\|A\, X\|_V}{\|X\|_V} = \sup_{X \in \mathbb{K}^n\,;\, \|X\|_V \leq 1} \|A\, X\|_V = \sup_{X \in \mathbb{K}^n\,;\, \|X\|_V = 1} \|A\, X\|_V\,;$$

además, dichos supremos son finitos y se alcanzan, por lo que podemos poner máximos. Definiendo entonces

$$\|A\|_M = \max_{X \in \mathbb{K}^n - \{0\}} \frac{\|A\, X\|_V}{\|X\|_V} = \max_{X \in \mathbb{K}^n\,;\, \|X\|_V \leq 1} \|A\, X\|_V = \max_{X \in \mathbb{K}^n\,;\, \|X\|_V = 1} \|A\, X\|_V\,,$$

se obtiene una norma matricial en $\mathcal{M}_n(\mathbb{K})$, llamada **norma matricial subordinada a la norma vectorial** *dada (normalmente se omiten los subíndices V ó M ya que no hay lugar a confusión).*

En toda **norma matricial subordinada** se verifica que

$$\|A\, X\| \leq \|A\|\, \|X\|\,, \quad \text{para todo} \quad A \in \mathcal{M}_n(\mathbb{K})\,, \quad X \in \mathbb{K}^n\,,$$

siendo además $\|A\|$ la menor constante $\alpha \geq 0$ tal que

$$\|A\, X\| \leq \alpha\, \|X\|\,, \quad \text{para todo} \quad X \in \mathbb{K}^n\,;$$

por último, en toda norma matricial subordinada la norma de la matriz identidad es siempre uno.

Teorema 4 (Normas matriciales subordinadas más usuales) *Dada* $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$*, las normas matriciales subordinadas a las normas vectoriales más usuales quedan caracterizadas por:*

$$\begin{aligned}
\|A\|_1 &= \max_{j=1,2,\ldots,n} \sum_{i=1}^{n} |a_{i,j}| \,; \\
\textit{si } \mathbb{K}=\mathbb{R}, \quad \|A\|_2 &= \sqrt{\rho(A^t\,A)} = \sqrt{\rho(A\,A^t)} = \|A^t\|_2 \,; \\
\textit{si } \mathbb{K}=\mathbb{C}, \quad \|A\|_2 &= \sqrt{\rho(A^*\,A)} = \sqrt{\rho(A\,A^*)} = \|A^*\|_2 \,; \\
\|A\|_\infty &= \max_{i=1,2,\ldots,n} \sum_{j=1}^{n} |a_{i,j}| \,.
\end{aligned}$$

Además, $\|\cdot\|_2$ *es invariante por transformación ortogonal-unitaria, es decir,*

$$\|A\|_2 = \|U\,A\|_2 = \|A\,U\|_2 = \|U^{-1}\,A\,U\|_2 \,,$$

para toda $U \in \mathcal{M}_n(\mathbb{K})$ *ortogonal-unitaria. Por último, si la matriz* A *es normal*

$$\|A\|_2 = \rho(A) \quad \textit{(norma espectral).}$$

Algunas **consecuencias** del Teorema anterior:

- se verifica que $\|A\|_2$ es el mayor valor singular de A;
- si A es simétrica-hermítica-antisimétrica-antihermítica entonces $\|A\|_2 = \rho(A)$ y si es ortogonal-unitaria entonces $\|A\|_2 = \rho(A) = 1$;
- las normas $\|\cdot\|_1$ y $\|\cdot\|_\infty$ son de fácil cálculo (no ocurre lo mismo con $\|\cdot\|_2$);
- existen normas y matrices ($\|\cdot\|_2$ y las normales) para las que $\|A\| = \rho(A)$; pero existen matrices para las que esto no es cierto cualquiera que sea la norma considerada (basta tomar una matriz no nula con autovalores nulos);
- dada $A \in \mathcal{M}_n(\mathbb{R})$, se verifica que

$$\max_{X\in\mathbb{R}^n-\{0\}} \frac{\|A\,X\|_p}{\|X\|_p} = \max_{X\in\mathbb{C}^n-\{0\}} \frac{\|A\,X\|_p}{\|X\|_p} \,, \quad p = 1, 2, \infty \,.$$

Teorema 5 (Normas matriciales y radio espectral) *Dada* $A \in \mathcal{M}_n(\mathbb{K})$ *y dada cualquier norma matricial* $\|\cdot\|$*, se verifica que*

$$\rho(A) \leq \|A\| \,.$$

No obstante, dado $\varepsilon > 0$*, existe una norma matricial subordinada* $\|\cdot\|$ *tal que*

$$\|A\| \leq \rho(A) + \varepsilon \,.$$

Teorema 6 (Norma euclídea o de Schur) *La aplicación*

$$\left.\begin{array}{rcl} \|\cdot\|_E\colon \mathcal{M}_n(\mathbb{K}) & \to & \mathbb{R}^+\,, \\ A & \mapsto & \|A\|_E = \left(\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{i,j}|^2\right)^{1/2}\,, \end{array}\right\}$$

define una norma matricial no subordinada (si $n > 1$), llamada **norma euclídea** *o de* **Schur**, *que es invariante por transformación ortogonal-unitaria, es decir,*

$$\|A\|_E = \|U\,A\|_E = \|A\,U\|_E = \|U^{-1}\,A\,U\|_E\,,$$

para toda $U \in \mathcal{M}_n(\mathbb{K})$ ortogonal-unitaria, y que verifica

$$\|A\|_2 \leq \|A\|_E \leq \sqrt{n}\,\|A\|_2\,.$$

La norma matricial euclídea recibe su nombre del hecho de ser la norma vectorial euclídea de $\mathcal{M}_n(\mathbb{K})$, considerado solamente como espacio vectorial de dimensión n^2; dicha norma es de fácil cálculo y permite dar acotaciones para la norma espectral.

Teorema 7 (Perturbaciones de la identidad) *Sea $\|\cdot\|$ una norma matricial subordinada y sea $B \in \mathcal{M}_n(\mathbb{K})$ tal que*

$$\|B\| < 1\,;$$

entonces $I + B$ (perturbación de la identidad) es inversible y

$$\|(I+B)^{-1}\| \leq \frac{1}{1-\|B\|}\,.$$

Por otro lado, si una matriz de la forma $I + B$ es singular, entonces

$$\|B\| \geq 1$$

para cualquier norma matricial $\|\cdot\|$.

23. Sucesiones de vectores y de matrices

Siempre que tenemos una norma tenemos una estructura topológica sobre dicho espacio, lo que nos permite tratar con conceptos como el de convergencia de una sucesión; por tanto, tanto en $\mathbb{K}^n$ como en $\mathcal{M}_{m\times n}(\mathbb{K})$ tiene sentido hablar de la **convergencia de una sucesión de vectores** $\{X_k\}_{k\in\mathbb{N}}$ o de **matrices** $\{A_k\}_{k\in\mathbb{N}}$.

Algunas **observaciones** relativas a la convergencia:

- el límite de una sucesión, de existir, es único;
- la convergencia de una sucesión es independiente de normas equivalentes;
- como en un espacio normado de dimensión finita todas las normas son equivalentes, el concepto de convergencia es independiente de la norma elegida; tomando, en particular, la norma $\|\cdot\|_\infty$ se concluye que la convergencia de una sucesión de vectores o de matrices equivale a la convergencia de las sucesiones de escalares formadas por las componentes de los vectores o las matrices.

Teorema 8 (Sucesión de potencias de una matriz) *Dada $A \in \mathcal{M}_n(\mathbb{K})$, las condiciones que siguen son equivalentes:*

1. *la sucesión $\{A^k\}_{k\in\mathbb{N}}$ converge hacia 0;*
2. *para todo $X \in \mathbb{K}^n$, la sucesión $\{A^k\, X\}_{k\in\mathbb{N}}$ converge hacia 0;*
3. $$\rho(A) < 1\,;$$
4. *para al menos una norma matricial subordinada $\|\cdot\|$, se tiene que*
$$\|A\| < 1\,.$$

Teorema 9 *Dada $A \in \mathcal{M}_n(\mathbb{K})$ y dada cualquier una norma matricial $\|\cdot\|$, se tiene que*

$$\lim_{k\to\infty} \|A^k\|^{1/k} = \rho(A)\,.$$

24. Problemas fundamentales del Análisis Numérico Matricial

Los problemas fundamentales del **Análisis Numérico Matricial** son:

- **resolución de un sistema lineal**, es decir, dada $A \in \mathcal{M}_n(\mathbb{K})$ inversible y dado $B \in \mathbb{K}^n$, encontrar $X \in \mathbb{K}^n$ tal que

$$A\,X = B\,;$$

- **cálculo de valores y vectores propios de una matriz**, es decir, dada $A \in \mathcal{M}_n(\mathbb{K})$ encontrar (todos o algunos) $\lambda \in \mathbb{C}$ y $X \in \mathbb{C}^n - \{0\}$ tales que

$$A\,X = \lambda\,X\,.$$

Algunos **comentarios** respecto de estos problemas:

- en los problemas surgidos de modelar una situación real (ingeniería, física, mecánica, química, medicina, etc.) la talla de la matriz A, suele ser "grande" o "extremadamente grande"; sin embargo, los datos suelen ser reales (recordamos que aún así los resultados de un problema de autovalores pueden ser complejos);
- estos problemas son resueltos, utilizando alguno de los algoritmos que serán estudiados, con la ayuda de un ordenador; esto implica inevitablemente errores de redondeo en el almacenamiento de los datos, en la realización de los cálculos, problemas de cancelación, underflow, overflow, etc.;
- como consecuencia de lo anterior, es necesario que los algoritmos que se utilicen para resolver los problemas precisen del mínimo número posible de operaciones, lo cual también es importante desde el punto de vista del tiempo necesario para su ejecución; esto implica que una etapa en el estudio de los diferentes métodos sea contar el número de operaciones elementales necesarias;
- las matrices que aparecen en la práctica suelen tener "buenas" propiedades que deben ser explotadas al máximo: matrices "vacías", banda, simétricas, definidas positivas, etc. El carácter vacío de una matriz hace que normalmente no sean almacenadas en un ordenador en su totalidad, sino que solamente se guarden aquellos elementos distintos de cero.

Los **métodos de resolución de sistemas lineales** son de dos tipos:

- **métodos directos**: son aquellos que conducen a la solución exacta del problema (si no hubiera errores de redondeo) en un número finito de operaciones elementales. Por ejemplo, las fórmulas de Cramer son un método directo, si bien no es en absoluto práctico debido al muy elevado número de operaciones que hay que efectuar;

- **métodos iterativos**: la solución X es el límite de una sucesión $\{X_k\}_{k\in\mathbb{N}}$ de "soluciones aproximadas" calculadas; como evidentemente se está obligado a parar el cálculo de los términos de esta sucesión en un cierto k_0, se introduce un error de truncatura medido por $\|X_{k_0} - X\|$. En consecuencia, los métodos iterativos llevan implícito el estudio de estimaciones del error de truncatura (recordamos que X no es conocido), elección del punto inicial X_0, elección adecuada del test de parada, etc.

Los **métodos de resolución de los problemas de valores y vectores propios** son necesariamente **iterativos**; nótese que en caso contrario esto implicaría que es posible calcular las raíces de un polinomio con un número finito de operaciones elementales y esto contradice, aún aceptando que la extracción de raíces n-ésimas fuera una operación elemental, a la teoría de Galois en cuanto $n \geq 5$. En este sentido se puede probar que cualquier polinomio de grado n

$$p(\lambda) = \lambda^n + a_1\,\lambda^{n-1} + \cdots + a_{n-1}\,\lambda + a_n\,,$$

es, módulo un factor $(-1)^n$, el polinomio característico de la matriz

$$A = \begin{pmatrix} -a_1 & -a_2 & -a_3 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

llamada **matriz de compañía del polinomio**.

Para terminar indicamos que la aproximación mediante diferentes métodos numéricos de las ecuaciones que modelan diversas situaciones reales (problemas de mecánica de sólidos y de fluidos, equilibrio térmico, vibraciones, interpolación y aproximación, optimización, etc.), conducen a la resolución de problemas del tipo mencionado.