Capítulo 3

Herramientas de Álgebra Lineal

3.1. Vectores y Matrices

Los elementos básicos en teoría de sistemas lineales son vectores $n \times 1$ (columna) o $1 \times n$ (fila) y matrices $n \times m$ con elementos reales, es decir, $v \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times m}$. Denotamos el elemento i del vector v como v_i, y el elemento ij de una matriz A como a_{ij} o $[A]_{ij}$.

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T, \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

En MATLAB: vectores `v = [v1;v2;...;vn]` y matrices

`A=[a11,a12,...,a1m;a21,a22,...,a2m;...]`.[1]

El producto de dos matrices $A \in \mathbb{R}^{m\times n}$ y $B \in \mathbb{R}^{r\times s}$ sólo está definido si $n = r$. En particular para vectores $v \in \mathbb{R}^n$, $w \in \mathbb{R}^m$, tenemos $vw^T \in \mathbb{R}^{n\times m}$.

Escribimos la matriz $n \times m$ con todos sus elementos cero como $0_{n\times m}$, simplemente 0 si las dimensiones (MATLAB `[n,m]=size(A)`) son claras del contexto. Para matrices cuadradas, $m = n$, la matriz nula es 0_n y la matriz identidad I_n o simplemente I (MATLAB `I=eye(n)`).

Asumimos las operaciones de suma y producto de matrices familiares. Lo interesante del producto de matrices es que en general es *no conmutativo*, es decir, AB y BA no son siempre lo mismo.

Definición 3.1 (Matriz Nilpotente). Si A es cuadrada, para cualquier entero $k \geq 0$ la potencia A^k está bien definida, con $A^0 = I$. Si existe un $k > 0$ tal que $A^k = 0$ decimos que A es *nilpotente*.

Definición 3.2 (Traza de una matriz). Para una matriz cuadrada $n \times n$, la *traza* es la suma de los elementos de su diagonal (en MATLAB `trace(A)`)

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} .$$

[1]Ojo con la transpuesta en MATLAB cuando se trabaja con vectores o matrices complejos; `'` representa la *transpuesta conjugada*, la transpuesta sin conjugar es `.'`.

Si A y B son tales que AB es cuadrada, entonces

$$\operatorname{tr}[AB] = \operatorname{tr}[BA].$$

Definición 3.3 (Determinante de una matriz). El *determinante* es otra familiar función escalar de una matriz cuadrada, $\det A$ (MATLAB `det(A)`). El determinante puede evaluarse recursivamente mediante la *expansión de Laplace*: Sea c_{ij} el *cofactor* del elemento a_{ij}, o sea el producto de $(-1)^{i+j}$ por el determinante de la submatriz $(n-1)\times(n-1)$ obtenida de eliminar en A la fila i y la columna j. Entonces para cualquier i fijo, $1 \le i \le n$,

$$\det A = \sum_{j=1}^{n} a_{ij} c_{ij}\,. \tag{3.1}$$

Esta es la expansión sobre la fila i. Una expresión análoga existe para la expansión sobre una columna j.

Vemos de (3.1) que el determinante no es más que la suma de productos de los elementos de la matriz, por lo que es una función matricial diferenciable todas las veces que uno quiera.

Si A y B son matrices $n \times n$, entonces

$$\det(AB) = \det(A)\cdot\det(B) = \det(BA).$$

Definición 3.4 (Inversa de una matriz). La matriz cuadrada A tiene una inversa, A^{-1}, $AA^{-1} = A^{-1}A = I$, sii $\det A \neq 0$.

En MATLAB, la inversa se calcula `inv(A)`. Una fórmula común para A^{-1} se basa en los cofactores de A.

Definición 3.5 (Adjunto de una matriz). El *adjunto* de A, $\operatorname{adj} A$, es la matriz cuyo elemento ij es el cofactor ji de A — o sea, la transpuesta de la matriz de cofactores.

Usando el adjunto, la fórmula de la inversa es

$$A^{-1} = \frac{\operatorname{adj} A}{\det A}$$

La inversa del producto de dos matrices cuadradas no singulares cumple $(AB)^{-1} = B^{-1}A^{-1}$.

Otra fórmula muy útil en teoría de sistemas lineales es el *lema de inversión de matrices*.

Lema 3.1 (Inversión de Matrices). Dadas matrices R, S, U, V de dimensiones compatibles, donde R y S son cuadradas y no singulares, entonces

$$(R - US^{-1}V)^{-1} = R^{-1} + R^{-1}U(S - VR^{-1}U)^{-1}VR^{-1}\,.$$

3.2. Bases y Ortonormalización

El conjunto de todos los vectores de $\mathbb{R}^n$ puede verse como un *espacio vectorial lineal* con las operaciones estándar de suma de vectores y producto por números reales. Las matrices $m \times n$ (o $n \times m$ si consideramos vectores *fila*) son *operadores lineales* definidos sobre estos espacios.

Definición 3.6 (Independencia lineal). Dos vectores x_1 e x_2 en $\mathbb{R}^n$ son *linealmente independientes (LI)* si la única combinación lineal de ellos que da el vector nulo, $\alpha_1 x_1 + \alpha_2 x_2 = 0$, es la trivial, $\alpha_1 = 0, \alpha_2 = 0$. Sino son *linealmente dependientes*.

El concepto se extiende a un conjunto de cualquier número de vectores. Si el conjunto de vectores $\{x_1, x_2, \ldots, x_m\}$ es linealmente dependiente, entonces existe un conjunto de escalares $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ donde al menos *uno* es distinto de cero, digamos $\alpha_1 \neq 0$, y

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m = 0\,.$$

Entonces podemos escribir x_1 como combinación lineal de los restantes vectores,

$$x_1 = -\frac{1}{\alpha_1}(\alpha_2 x_2 + \cdots + \alpha_m x_m).$$

Definición 3.7 (Dimensión de un espacio lineal). La *dimensión* de un espacio lineal es el máximo número de vectores LI en ese espacio; n en $\mathbb{R}^n$.

Definición 3.8 (Base de un espacio lineal). Un conjunto de vectores LI en $\mathbb{R}^n$ es una *base* de $\mathbb{R}^n$ si todo vector de $\mathbb{R}^n$ se puede escribir como una combinación lineal única de los vectores del conjunto.

Dada una base $\{q_1, q_2, \ldots, q_n\}$ de $\mathbb{R}^n$, todo vector $x \in \mathbb{R}^n$ puede expresarse en forma única como

$$\begin{aligned} x &= \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n \\ &= \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = Q\tilde{x}\,, \end{aligned}$$

donde $\tilde{x} = [\alpha_1, \alpha_2, \ldots, \alpha_n]^T$ es la *representación* del vector x con respecto a la base (o en las *coordenadas*) $\{q_1, q_2, \ldots, q_n\}$. La matriz Q es no singular por definición, y sirve para obtener la representación de cualquier vector de $\mathbb{R}^n$ en la base $\{q_1, q_2, \ldots, q_n\}$, $\tilde{x} = Q^{-1}x$. En particular, siempre existe la base formada por los *versores* $e_1 = [1, 0, \ldots, 0]^T$, $e_2 = [0, 1, \ldots, 0]^T$, etc. En ese caso $Q = I$.

Ejemplo 3.1. Consideremos los vectores $q_1 = [3, 1]^T$ y $q_2 = [2, 2]^T$. Como son LI, forman una base en $\mathbb{R}^2$. La representación del vector $x = [1, 3]^T$ en las coordenadas $\{q_1, q_2\}$ es $[-1, 2]^T$, que se obtiene de

$$\begin{aligned} \tilde{x} &= \begin{bmatrix} q_1 & q_2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1/2 & -1/2 \\ -1/4 & 3/4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \end{aligned}$$

o sea que x puede escribirse como $x = -q_1 + 2q_2$.

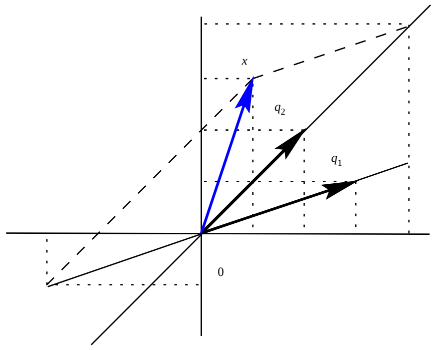

Figura 3.1: Cambio de coordenadas.

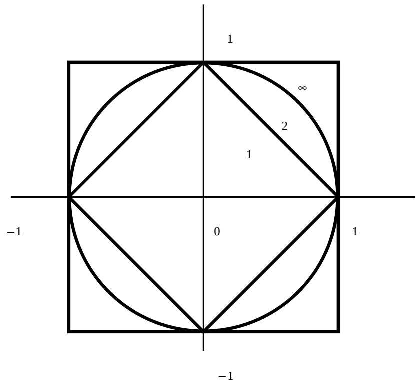

Figura 3.2: Bola unitaria en $\mathbb{R}^2$ en normas $1, 2$ e ∞.

3.2.1. Norma de vectores

El concepto de *norma* de un vector es una generalización de la idea de magnitud.

Definición 3.9 (Norma de un vector). La norma de un vector $x \in \mathbb{R}^n$, denotada $\|x\|$, es una función de $\mathbb{R}^n$ en $\mathbb{R}_0^+$ (los reales positivos más el 0) que cumple con las siguientes tres propiedades:

1. $\|x\| \geq 0$ para todo x, y $\|x\| = 0$ sii $x = 0$.
2. $\|\alpha x\| = |\alpha|\|x\|$, para todo escalar α.
3. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ para todo x_1, x_2 (*desigualdad triangular*).

Dado un vector $x = [x_1 x_2 \dots x_n]^T$, tres normas típicas en $\mathbb{R}^n$ son

$$\begin{aligned} \|x\|_1 &\triangleq \textstyle\sum_{i=1}^n |x_i| && \text{norma-1} \\ \|x\|_2 &\triangleq \sqrt{x^T x} = \sqrt{\textstyle\sum_{i=1}^n |x_i|^2} && \text{norma-2 o } \textit{euclídea} \\ \|x\|_\infty &\triangleq \max_i |x_i| && \text{norma-}\infty \end{aligned}$$

En MATLAB calculamos estas normas con `norm(x,1)`, para la norma 1, `norm(x,2) = norm(x)`, para la norma 2, y `norm(x,inf)` para la norma ∞.

El conjunto de vectores de norma no mayor que la unidad, $B_1 = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$, se llama *bola unitaria*, que no necesariamente es una bola *esférica*; la forma de la bola unitaria depende de la norma (ver Figura 3.2). La norma 2, también llamada *euclídea*, es efectivamente la longitud del vector medida desde el origen (que no es el caso de las otras normas). A menos que aclaremos lo contrario, vamos a trabajar siempre con la norma 2.

3.2.2. Ortonormalización

Un vector x de $\mathbb{R}^n$ está *normalizado* si $\|x\| = \sqrt{x^T x} = 1$. Dos vectores x_1 y x_2 son *ortogonales* si $x_1^T x_2 = x_2^T x_1 = 0$. Un conjunto de vectores $x_i, i = 1, 2, \dots, m$, es *ortonormal* si

$$x_i^T x_j = \begin{cases} 0 & \text{si } i \neq j\,, \\ 1 & \text{si } i = j\,. \end{cases}$$

Dado un conjunto de vectores LI $\{p_1, p_2, \dots, p_m\}$, podemos obtener un conjunto ortonormal de vectores $\{q_1, 1_2, \dots, 1_m\}$ usando el *procedimiento de ortonormalización de Schmidt*,

$$\begin{aligned}
u_1 &\triangleq p_1 & q_1 &\triangleq u_1/\|u_1\|\,, \\
u_2 &\triangleq p_2 - (q_1^T p_2)q_1 & q_2 &\triangleq u_2/\|u_2\|\,, \\
&\cdots & &\cdots \\
u_m &\triangleq p_m - \textstyle\sum_{k=1}^{m-1}(q_k^T p_m)q_k & q_m &\triangleq u_m/\|u_m\|\,.
\end{aligned}$$

Si $Q = [q_1 q_2 \dots q_m]$ es una matriz $n \times m$, $m \leq n$, con todas sus columnas ortonormales, entonces $Q^T Q = I$. ¿Qué puede decirse entonces de QQ^T?

3.3. Ecuaciones Lineales Algebraicas

Consideremos el conjunto de ecuaciones lineales algebraicas

$$Ax = y \tag{3.2}$$

donde $A \in \mathbb{R}^{m\times n}$ e $y \in \mathbb{R}^m$ son una matriz y un vector dados como datos, y $x \in \mathbb{R}^n$ es el vector incógnita a resolver. Este es un problema de m ecuaciones y n incógnitas escalares, donde puede que m sea menor, igual o mayor que n. La existencia de solución dependerá de la *imagen* de A.

Definición 3.10 (Imagen y rango de una matriz). El *espacio imagen*, o simplemente *imagen*, de A es el espacio generado haciendo todas las posibles combinaciones lineales de las columnas de A. La dimensión de la imagen de A es el *rango* de A.

Definición 3.11 (Vector nulo y kernel de una matriz). Un vector x se llama *vector nulo* de A si $Ax = 0$. El espacio de todos los vectores nulos de A se llama el *kernel* o *espacio nulo* de A.

En MATLAB `orth(A)` da una base ortonormal de la imagen de A, `rank(A)` da el rango, y `null(A)` da una base ortonormal del kernel. Una relación importante que cumplen siempre la dimensión del kernel y el rango de una matriz A es la siguiente:

$$\text{dimensión de } \ker(A) = \text{número de columnas de } A - \text{rango}(A) \tag{3.3}$$

Ejemplo 3.2. Sea la matriz

$$A = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} \triangleq [a_1\, a_2\, a_3\, a_4]\,.$$

El rango de A es 2, ya que a_1 y a_2 son LI pero a_3 y a_4 son combinación lineal de a_1 y a_2. Podemos tomar entonces el conjunto $\{a_1, a_2\}$ como una base de la imagen de A. La ecuación (3.3) implica que la dimensión del kernel de A es 2, y puede verificarse que los vectores

$$n_1 = \begin{bmatrix} 1 & 1 & -1 & 0 \end{bmatrix}^T \qquad \text{y} \qquad n_2 = \begin{bmatrix} 0 & 2 & 0 & -1 \end{bmatrix}^T$$

son vectores nulos de A ($An_1 = 0 = An_2$). Como n_1 y n_2 son LI, $\{n_1, n_2\}$ es una base del kernel de A.

Hemos definido el rango de A como el número de columnas LI de A, pero como también es igual al número de *filas* LI de A, necesariamente tenemos la relación

$$\operatorname{rango}(A) \leq \operatorname{mín}(m, n)\,.$$

Con las definiciones de rango y kernel de una matriz, volvemos al problema de obtener la solución de la ecuación algebraica (3.2). Los siguientes dos teoremas resumen todos los casos posibles.

Teorema 3.1 (Existencia de solución). Dada una matriz $A \in \mathbb{R}^{m\times n}$ y un vector $y \in \mathbb{R}^m$,

1. Existe una solución $x \in \mathbb{R}^n$ de la ecuación $Ax = y$ sii y pertenece a la imagen de A, o sea,
$$\operatorname{rango}(A) = \operatorname{rango}([A\ y])$$
donde $[A\ y] \in \mathbb{R}^{m\times(n+1)}$.
2. Existe una solución x de $Ax = y$ para todo $y \in \mathbb{R}^m$ sii $\operatorname{rango}(A) = m$ (A es de *rango fila pleno*).

Teorema 3.2 (Parametrización de todas la soluciones). Dada una matriz $A \in \mathbb{R}^{m\times n}$ y un vector $y \in \mathbb{R}^m$, sea x_p una solución de $Ax = y$ y sea $k \triangleq n - \operatorname{rango}(A)$ la dimensión del kernel de A. Entonces

1. Si $k = 0$ (A tiene rango columna pleno), entonces la solución x_p es *única*.
2. Si $k > 0$ sea $\{n_1, n_2, \ldots, n_k\}$ una base del kernel de A. Entonces para cualquier conjunto de k números reales $\{\alpha_i, i = 1, 2, \ldots, k\}$ el vector
$$x = x_p + \alpha_1 n_1 + \alpha_2 n_2 + \cdots + \alpha_k n_k$$
es también solución de $Ax = y$.

Ejemplo 3.3. Consideremos la ecuación $Ax = y$

$$\begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -4 \\ -8 \\ 0 \end{bmatrix}. \tag{3.4}$$

Puede verse fácilmente que y está en la imagen de A y que $x_p = \begin{bmatrix} 0 & -4 & 0 & 0 \end{bmatrix}^T$ es una solución. Con la base $\{n_1, n_2\}$ del kernel de A obtenida en el Ejemplo 3.2 podemos expresar la solución general de (3.4) como

$$\begin{aligned} x &= x_p + \alpha_1 n_1 + \alpha_2 n_2 \\ &= \begin{bmatrix} 0 \\ -4 \\ 0 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \end{aligned}$$

donde α_1 y α_2 pueden tomar cualquier valor en $\in \mathbb{R}$.

A veces podemos encontrar el problema formulado en términos de vectores fila, $xA = y$, donde x, y son $1 \times n$ y $1 \times m$. Este problema puede tratarse de igual forma al caso columna considerando el problema transpuesto equivalente $A^T x^T = y^T$.

En MATLAB la solución de $Ax = y$ puede obtenerse como `x = A\y`. El símbolo $\backslash$ denota división matricial por izquierda. Para el caso $xA = y$ usamos la división matricial por derecha `x = y/A`.

Para el caso de sistemas con matriz A cuadrada (que vamos a encontrar con frecuencia) la existencia de soluciones del sistema de ecuaciones algebraicas (3.2) se reduce a las condiciones dadas en el siguiente corolario.

Corolario 3.1 (Sistemas cuadrados). Sea $Ax = y$ donde A es cuadrada. Entonces

1. Si A es no singular existe una solución *única* para cada y, dada por $x = A^{-1}y$. En particular la única solución de $Ax = 0$ es $x = 0$.

2. La ecuación homogénea $Ax = 0$ tiene soluciones no nulas sii A es singular. El número de soluciones LI es igual a la dimensión del kernel de A.

3.4. Transformaciones de Semejanza

Una matriz cuadrada A mapea vectores de $\mathbb{R}^n$ en $\mathbb{R}^n$,

$$Ax = y\,. \tag{3.5}$$

Si representamos los vectores de $\mathbb{R}$ con respecto a una nueva base $\{q_1, q_2, \ldots, q_n\} \triangleq Q$, sabemos de § 3.2 que tenemos las relaciones $x = Q\bar{x}$ e $y = Q\bar{y}$. Así, la ecuación (3.5) puede alternativamente escribirse como

$$Ax = y \quad \Leftrightarrow \quad AQ\bar{x} = Q\bar{y} \quad \Leftrightarrow \quad (Q^{-1}AQ)\bar{x} = \bar{y}$$

La matriz $\bar{A} \triangleq (Q^{-1}AQ)$ es la representación de A con respecto a la base $\{q_1, q_2, \ldots, q_n\}$. La transformación

$$\bar{A} = Q^{-1}AQ \qquad A = Q\bar{A}Q^{-1}$$

se llama *transformación de semejanza*, y las matrices A y $\bar{A}$ se dicen *semejantes*.

Recordemos que dada una base $\{q_1, q_2, \ldots, q_n\}$, todo vector x de $\mathbb{R}^n$ puede escribirse en forma única como $x = Q\bar{x}$, donde $Q \triangleq [q_1, q_2, \ldots, q_n]$, y $\bar{x}$ es la *representación de x en esa base*. En particular, si consideramos la base canónica $\{e_1, e_2, \ldots, e_n\}$ (formada por los versores unitarios), entonces la columna i de una matriz A no es más que la representación del vector Ae_i con respecto a la base canónica $\{e_1, e_2, \ldots, e_n\}$. Escribiendo la relación $\bar{A} = Q^{-1}AQ$ de la forma

$$\begin{aligned} & AQ = Q\bar{A} \\ \Leftrightarrow \quad & A\begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix} = Q\begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \ldots & \bar{a}_n \end{bmatrix} \\ \Leftrightarrow \quad & \begin{bmatrix} Aq_1 & Aq_2 & \ldots & Aq_n \end{bmatrix} = \begin{bmatrix} Q\bar{a}_1 & Q\bar{a}_2 & \ldots & Q\bar{a}_n \end{bmatrix} \end{aligned}$$

concluimos que la columna i de $\bar{A}$ es la representación de Aq_i en la base $\{q_1, q_2, \ldots, q_n\}$.

Ejemplo 3.4. Consideremos la matriz

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \quad \text{y la base} \quad Q = \begin{bmatrix} 0 & -1 & -4 \\ 0 & 0 & 2 \\ 1 & 1 & -3 \end{bmatrix}$$

Calculamos

$$\bar{A} = Q^{-1}AQ = \begin{bmatrix} 0 & 0 & 17 \\ 1 & 0 & -15 \\ 0 & 1 & 5 \end{bmatrix}$$

Podemos comprobar, por ejemplo, que el vector, digamos $z = Aq_1 = A \left[\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} -1 \\ 0 \\ 1 \end{smallmatrix}\right]$ verifica la ecuación $z = Q\bar{a}_1 = Q \left[\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right]$, es decir que $\bar{a}_1$ es la representación del vector z en la base $\{q_1, q_2, \ldots, q_n\}$.

3.5. Forma Diagonal y Forma de Jordan

3.5.1. Autovalores y autovectores

Los autovalores y autovectores de una matriz juegan un rol clave en el estudio de sistemas lineales estacionarios en ecuaciones de estado. En esta sección repasamos sus definiciones e introducimos las formas canónicas asociadas y otras que utilizaremos más adelante en el análisis de estabilidad y el cálculo de controladores por realimentación de estados para estos sistemas.

Camille Jordan (1838-1922)

Definición 3.12 (Autovalores, autovectores y espectro de una matriz cuadrada). Un número $\lambda \in \mathbb{C}$ es un *autovalor* de la matriz $A \in \mathbb{R}^{n\times n}$ si existe un vector no nulo $v \in \mathbb{R}^n$ tal que $Av = \lambda v$. Este vector v es un *autovector* (por derecha) de A asociado al autovalor λ. El conjunto de *todos* los autovalores de una matriz A se llama el *espectro* de A.

Los autovalores de A se encuentran resolviendo la ecuación algebraica

$$(\lambda I - A)v = 0\,. \tag{3.6}$$

Esta es un sistema de ecuaciones lineales algebraicas, que como vimos en § 3.3, admite soluciones no nulas si la matriz $(\lambda I - A)$ es singular (es decir, tiene determinante nulo).

Definición 3.13 (Polinomio característico). El *polinomio característico* de la matriz cuadrada A es

$$\Delta(\lambda) = \det(\lambda I - A)\,.$$

Notar que el polinomio $\Delta(\lambda)$ es mónico, es decir que el coeficiente del término de mayor orden es 1, de grado n, y de coeficientes reales. Para cada raíz de $\Delta(\lambda)$ la matriz $(\lambda I - A)$ es singular y, por lo tanto, la ecuación algebraica (3.6) admite al menos una solución no nula.

Concluimos que toda raíz de $\Delta(\lambda)$ es un autovalor de A; y como $\Delta(\lambda)$ tiene grado n, la matriz A necesariamente tiene n autovalores (aunque no necesariamente todos distintos; puede haber repetidos).

En MATLAB los autovalores de la matriz A se calculan con la función `r = eig(A)`, donde $r = [\lambda_1, \dots, \lambda_n]^T$ es un vector con todos los autovalores de A; la función `poly(r)` da el polinomio característico de A.

3.5.2. Forma Companion

En general, el cálculo del polinomio característico de una matriz requiere la expansión de $\det(\lambda I - A)$. Sin embargo, para ciertas matrices el polinomio característico es evidente. Estas son las matrices en la *forma companion*, (MATLAB `A = compan(p)`, donde `p` es un polinomio)

$$\begin{bmatrix} 0 & 0 & 0 & -\alpha_4 \\ 1 & 0 & 0 & -\alpha_3 \\ 0 & 1 & 0 & -\alpha_2 \\ 0 & 0 & 1 & -\alpha_1 \end{bmatrix} \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} \begin{bmatrix} -\alpha_1 & 1 & 0 & 0 \\ -\alpha_2 & 0 & 1 & 0 \\ -\alpha_3 & 0 & 0 & 1 \\ -\alpha_4 & 0 & 0 & 0 \end{bmatrix}$$

con el mismo polinomio característico $\Delta(\lambda) = \lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4$. Otro caso en que los autovalores (y el polinomio característico) son particularmente obvios es cuando la matriz A es diagonal,

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Aquí los autovalores de A son directamente los elementos de su diagonal. Pero, *¿podemos siempre representar A en forma diagonal?* No siempre, depende de si A tiene:

Caso 1. autovalores todos reales y distintos

Caso 2. autovalores complejos y distintos

Caso 3. autovalores repetidos

Analizamos cada caso por separado en la siguiente sección.

3.5.3. Forma de Jordan

La forma de Jordan de una matriz A es la forma *diagonal en bloques* más simple a la que puede llevarse la matriz A mediante transformaciones de semejanza. Una matriz diagonal en bloques es una matriz que tiene matrices cuadradas no nulas en la diagonal, y ceros fuera de ella, por ejemplo

$$A = \begin{bmatrix} \boxed{\begin{matrix} \alpha & \beta \\ \gamma & \delta \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} \boxed{\epsilon} & 0 \\ 0 & \boxed{\phi} \end{matrix} \end{bmatrix}.$$

Vemos la forma de Jordan en los tres casos enumerados al final de la sección anterior.

Caso 1: autovalores de A todos reales y distintos

En este caso los autovectores correspondientes son LI. Usamos el conjunto de autovectores, $\{v_1, v_2, \dots, v_n\}$, como base. Sea $\bar{A}$ la representación de A en esta base. La columna 1 de $\bar{A}$ es la representación de $Av_1 = \lambda_1 v_1$ en la base nueva,

$$Av_1 = \lambda_1 v_1 = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

y concluimos que la columna 1 de $\bar{A}$ es $[\lambda_1\, 0\, \dots\, 0]^T$. La columna 2 es la representación de $Av_2 = \lambda_2 v_2$ con respecto a la base nueva, o sea $[0\, \lambda_2\, 0\, \dots\, 0]^T$. Siguiendo este procedimiento, llegamos a que

$$\bar{A} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Concluimos que *toda matriz con autovalores distintos tiene una representación en forma diagonal* (es diagonalizable).

Caso 2: autovalores de A complejos y distintos

Cuando todos los autovalores son distintos, pero algunos son complejos, no podemos llevar A a una forma diagonal real (aunque si compleja). Siendo A real, los autovalores complejos aparecen siempre en pares conjugados $\lambda = \sigma \pm j\omega$, y dan lugar a pares de autovectores también complejos conjugados $v = u \pm jw$.

En este caso no se puede llevar A a una forma diagonal real, pero sí a una forma *diagonal en bloques* real. Para esto, la nueva base debe armarse con los autovectores reales, las partes reales de los autovectores complejos, y las partes imaginarias de los autovectores complejos.

Ejemplo 3.5 (Forma block-diagonal real de una matriz con autovalores complejos). Supongamos que la matriz A tiene dos autovalores reales y dos pares de autovalores complejos,

$$\{\lambda_1, \lambda_2, \sigma_1 + j\omega_1, \sigma_1 - j\omega_1, \sigma_2 + j\omega_2, \sigma_2 - j\omega_2\}$$

con autovectores

$$\{v_1, v_2, u_1 + jw_1, u_1 - jw_1, u_2 + jw_2, u_2 - jw_2\}.$$

Formamos entonces la base

$$\{v_1, v_2, u_1, w_1, u_2, w_2\}\,.$$

En esta base la matriz A tiene la representación

$$\bar{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1 & \omega_1 & 0 & 0 \\ 0 & 0 & -\omega_1 & \sigma_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_2 & \omega_2 \\ 0 & 0 & 0 & 0 & -\omega_2 & \sigma_2 \end{bmatrix}.$$

Vemos que hay un bloque de 2×2 por cada par de autovalores complejos conjugados.

Caso 3: autovalores de A repetidos

Este es un caso especial. Cuando A tiene autovalores repetidos puede que no podamos llevarla a una forma diagonal. Sin embargo, siempre se puede llevar a una forma triangular o diagonal en bloques. Ilustramos el procedimiento para un caso concreto.

Supongamos que $A \in \mathbb{R}^{n\times n}$ tiene un autovalor λ con multiplicidad n, es decir, un solo autovalor repetido n veces. Para simplificar consideremos $n = 4$. Sabemos que la matriz $(A - \lambda I)$ es singular. Ahora, caben las posibilidades de que el kernel (espacio nulo) de $(A - \lambda I)$ tenga dimensión $1, 2, 3$ o 4.

Si el kernel de $(A - \lambda I)$ tiene dimensión 4, entonces sabemos, del Teorema 3.2, que hay cuatro soluciones independientes (no nulas, pues la nula es la solución particular en este caso), correspondientes a una base del kernel $\{n_1, n_2, n_3, n_4\}$.

$$(A - \lambda I)x = 0 \quad \Rightarrow \quad x = \begin{bmatrix} n_1 & n_2 & n_3 & n_4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}.$$

En este caso hay 4 autovectores independientes, y A es diagonalizable usando la base $\{n_1, n_2, n_3, n_4\}$.

Veamos ahora el otro caso extremo, el kernel de A tiene dimensión 1. En este caso sólo podemos encontrar *una* solución independiente. Necesitamos tres vectores independientes más para armar una base. Una forma de lograrlo es usando el concepto de *autovalores generalizados*.

Definición 3.14 (Autovector generalizado). Sea $A \in \mathbb{R}^{n\times n}$ y λ un autovalor de A. Un vector v es un autovector generalizado de A de grado n asociado al autovalor λ si

$$(A - \lambda I)^n v = 0 \quad \text{y} \quad (A - \lambda I)^{n-1} v \neq 0$$

Si $n = 1$ el problema se reduce al caso recién visto. Para $n = 4$ definimos los vectores

$$\begin{aligned} v_4 &\triangleq v \\ v_3 &\triangleq (A - \lambda I)v_4 = (A - \lambda I)v \\ v_2 &\triangleq (A - \lambda I)v_3 = (A - \lambda I)^2 v \\ v_1 &\triangleq (A - \lambda I)v_2 = (A - \lambda I)^3 v\,. \end{aligned}$$

Los vectores v_1, v_2, v_3, v_4 forman una *cadena de autovectores generalizados de longitud 4*, y cumplen las propiedades $(A-\lambda I)v_1 = 0$, $(A-\lambda I)^2 v_2 = 0$, $(A-\lambda I)^3 v_3 = 0$ y $(\lambda I - A)^4 v_4 = 0$. Estos vectores son LI por definición, y además cumplen con las relaciones

$$\begin{aligned} Av_1 &= \lambda v_1 \\ Av_2 &= v_1 + \lambda v_2 \\ Av_3 &= v_2 + \lambda v_3 \\ Av_4 &= v_3 + \lambda v_4. \end{aligned}$$

A partir de estas relaciones, la representación de A con respecto a la base $\{v_1, v_2, v_3, v_4\}$ resulta

$$J = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}. \tag{3.7}$$

Es fácil chequear que las columnas de J son la representación de $Av_i, i = 1,2,3,4$, en la base $\{v_1, v_2, v_3, v_4\}$. Esta matriz triangular es un *bloque de Jordan de orden 4*.

Si el kernel de $(A - \lambda I)$ tiene dimensión 2, entonces hay *dos* cadenas de autovectores generalizados cuya longitudes sumadas dan 4. Pueden ser $\{v_1, v_2\}$ y $\{u_1, u_2\}$, $\{v_1\}$ y $\{u_1, u_2, u_3\}$, $\{v_1, v_2, v_3\}$ y $\{u_1\}$. En este caso la base se forma con los vectores de las dos cadenas.

Igualmente, si el kernel de $(A - \lambda I)$ tiene dimensión 3, entonces hay *tres* cadenas de autovectores generalizados, etc. En el primer caso visto, en que el kernel de $(A - \lambda I)$ tiene dimensión 4, tenemos cuatro cadenas triviales de longitud 1.

En resumen, si $A \in \mathbb{R}^{4\times 4}$ tiene un autovalor λ con multiplicidad 4, existe una matriz no singular Q tal que la matriz semejante $\bar{A} = Q^{-1}AQ$ tiene una de las siguientes formas de Jordan

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

La forma general a la que podemos llevar una matriz va a ser diagonal en bloques, donde los bloques pueden ser de 1×1 para autovalores reales distintos, 2×2 para pares de autovalores complejos conjugados, y bloques de Jordan para autovalores múltiples.

La forma de Jordan puede usarse para establecer propiedades generales de matrices. Por ejemplo, $\det A = \det(Q^{-1}\bar{A}Q) = \det Q \det Q^{-1} \det \bar{A}$. Como $\bar{A}$ es triangular, $\det \bar{A}$ es el producto de los autovalores,

$$\det A = \prod_{i=1}^{n} \lambda_i(A),$$

de lo que concluimos que *A es no singular sii no tiene autovalores nulos*.

Terminamos esta sección con una propiedad útil de los bloques de Jordan, que es la de *nilpotencia* de la matriz $(J - \lambda I)$.

Definición 3.15 (Matriz Nilpotente). Una matriz A es *nilpotente* si existe algún $k \in \mathbb{N}$ tal que $A^k = 0$.

Por definición del bloque de Jordan J en (3.7), la matriz $(J - \lambda I)$ es nilpotente con $k = 4$, ya que

$$(J-\lambda I) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (J-\lambda I)^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (J-\lambda I)^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$(J-\lambda I)^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

En MATLAB podemos usar `[q,d] = eig(A)` para calcular los autovalores y la base de autovectores si A es diagonalizable. Si A *no* es diagonalizable, Chen (1999) menciona la función `[q,d] = jordan(A)`, que puede usarse para matrices de dimensión moderada, pero disponible solamente en las versiones de MATLAB con el `symbolic toolbox`).

3.6. Funciones de Matrices Cuadradas

3.6.1. Polinomios

Vimos que para cualquier entero $k \neq 0$ la potencia A^k está bien definida. Dado un polinomio $f(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_m$ definimos $f(A)$ como

$$f(A) = A^m + a_1 A^{m-1} + \cdots + a_m I\,.$$

Si A es diagonal en bloques, digamos $A = \left[\begin{smallmatrix} A_1 & 0 \\ 0 & A_2 \end{smallmatrix}\right]$ es fácil verificar que

$$A^k = \begin{bmatrix} A_1^k & 0 \\ 0 & A_2^k \end{bmatrix} \quad \text{y también} \quad f(A) = \begin{bmatrix} f(A_1) & 0 \\ 0 & f(A_2) \end{bmatrix}.$$

Dada la transformación de semejanza $\bar{A} = Q^{-1}AQ$ o $A = Q^{-1}\bar{A}Q$, como

$$A^k = (Q\bar{A}Q^{-1})^k = Q\bar{A}Q^{-1}QAQ^{-1}\cdots = Q\bar{A}^kQ^{-1}\,,$$

tenemos que

$$f(A) = Qf(\bar{A})Q^{-1} \quad \text{o} \quad f(\bar{A}) = Q^{-1}f(A)Q\,.$$

Un resultado importante en el cálculo matricial es el Teorema de Cayley-Hamilton, que establece que *toda matriz satisface su propio polinomio característico.*

Teorema 3.3 (Cayley-Hamilton). Sea $\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$ el polinomio característico de A. Entonces

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1}A + \alpha_n I = 0\,. \tag{3.8}$$

○

El teorema de Cayley-Hamilton implica que A^n se puede escribir como una combinación lineal de las potencias de A de 0 a $n-1$. Si multiplicamos (3.8) por A obtenemos que A^{n+1} se puede escribir como combinación lineal de $\{A, A^2, \ldots, A^n\}$, que a su vez se puede escribir en términos de $\{I, A, \ldots, A^{n-1}\}$.

En conclusión, *todo polinomio* $f(A)$ puede expresarse, para apropiados β_i, como una combinación lineal de las potencias de A de 0 a $n-1$.

$$f(A) = \beta_0 I + \beta_1 A + \cdots + \beta_{n-1}A^{n-1}\,.$$

3.6.2. Polinomio mínimo

El teorema de Cayley-Hamilton establece que toda matriz satisface su polinomio característico, es decir, si $\Delta(\lambda) = \det(\lambda I - A)$, entonces $\Delta(A) = 0$. Pero, *¿podrá A satisfacer un polinomio de grado* menor *al de $\Delta(\lambda)$?* La respuesta depende de la multiplicidad de los autovalores de A. Cuando los autovalores son todos de multiplicidad 1 (todos distintos), entonces *el polinomio característico es el polinomio de menor grado que A satisface* (***polinomio mínimo***), o sea n en este caso.

Cuando hay autovalores con multiplicidad mayor que 1 (repetidos), el polinomio mínimo podrá ser de grado menor que n. El polinomio mínimo se puede expresar como

$$\Psi(\lambda) = \prod_i (\lambda - \lambda_i)^{\tilde{n}_i},$$

donde $\tilde{n}_i$ es la dimensión del bloque de Jordan más grande asociado al autovalor λ_i. Como el autovalor λ_i puede tener más de un bloque de Jordan asociado, concluimos que $\tilde{n}_i \leq n$.

Ejemplo 3.6. En la matriz

$$A = \begin{bmatrix} \boxed{\begin{matrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} \boxed{\lambda_1} & 0 \\ 0 & \boxed{\lambda_2} \end{matrix} \end{bmatrix}$$

El autovalor λ_1 tiene multiplicidad 3, y dos bloques de Jordan asociados: órdenes 2 y 1. El autovalor λ_2 tiene multiplicidad 1.

$$\begin{aligned} \Delta(\lambda) &= \prod_i (\lambda - \lambda_i)^{n_i} = (\lambda - \lambda_1)^3 (\lambda - \lambda_2), \quad \text{polinomio característico} \\ \Psi(\lambda) &= \prod_i (\lambda - \lambda_i)^{\tilde{n}_i} = (\lambda - \lambda_1)^2 (\lambda - \lambda_2), \quad \text{polinomio mínimo}. \end{aligned}$$

El polinomio mínimo es siempre factor del polinomio característico.

Como consecuencia del teorema de Cayley-Hamilton vimos que para todo polinomio $f(\lambda)$, el polinomio matricial $f(A)$ se puede expresar como una combinación lineal de las potencias de A, $\{I, A, \ldots, A^{n-1}\}$. Si el polinomio mínimo se conoce (y por lo tanto los $\tilde{n}_i$), en realidad $f(A)$ se puede expresar como una combinación lineal de un conjunto menor, $\{I, A, \ldots, A^{\tilde{n}_i - 1}\}$, que es aún mejor.

3.6.3. Evaluación de funciones matriciales

Una forma de calcular $f(A)$ (con $\Psi(\lambda)$ si se conoce, y sino con $\Delta(\lambda)$) es mediante la fórmula de división de polinomios:

$$f(\lambda) = q(\lambda)\Delta(\lambda) + h(\lambda),$$

donde $q(\lambda)$ es el *polinomio cociente* y $h(\lambda)$ el *polinomio resto*. Tenemos

$$f(A) = q(A)\Delta(A) + h(A) = q(A)0 + h(A) = h(A),$$

o sea que todo se reduce a determinar los coeficientes del polinomio $h(\lambda) = \eta_{n-1}\lambda^{n-1} + \cdots + \eta_1\lambda + \eta_0$.

La división de polinomios es útil si el grado de $f(\lambda)$ no es mucho mayor que el de $\Delta(\lambda)$. De lo contrario es mejor determinar los coeficientes de $h(\lambda)$ evaluando $f(\lambda)$ en los autovalores de A y planteando un sistema de ecuaciones algebraicas.

Si los autovalores λ_i de A son todos distintos, los η_i de $h(\lambda)$ se pueden resolver del sistema de n ecuaciones con n incógnitas

$$f(\lambda_i) = q(\lambda_i)\Delta(\lambda_i) + h(\lambda_i) = h(\lambda_i)\,, \qquad \text{para } i = 1, 2, \ldots, n\,.$$

Si A tiene autovalores repetidos, hay que derivar $f(\lambda) = q(\lambda)\Delta(\lambda) + h(\lambda)$ hasta obtener las ecuaciones faltantes.

Ejemplo 3.7. Queremos calcular A^{100} con

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

Planteamos el problema como el cálculo de $f(\lambda) = \lambda^{100}$ evaluado en A. El polinomio característico de A es $\Delta(\lambda) = (\lambda + 1)^2$. Sea $h(\lambda) = \eta_1\lambda + \eta_0$.

En el espectro de A (dos elementos iguales en este caso, $\lambda = -1$) tenemos

$$\begin{aligned} f(-1) = h(-1) \quad &\Rightarrow \qquad (-1)^{100} = -\eta_1 + \eta_0 \\ f'(-1) = h'(-1) \quad &\Rightarrow \quad 100 \cdot (-1)^{99} = \eta_1\,, \end{aligned}$$

y así concluimos que $\eta_1 = -100$, y $\eta_0 = -99$, o sea, $h(\lambda) = -100\lambda - 99$. Finalmente,

$$\begin{aligned} A^{100} &= \eta_1 A + \eta_0 I \\ &= -100 \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} - 99 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -99 & -100 \\ 100 & 101 \end{bmatrix}. \end{aligned}$$

Resumimos el resultado que usamos en el ejemplo en un teorema.

Teorema 3.4 (Evaluación de una función matricial). Sean dados $f(\lambda)$ y una matriz A $n \times n$ con polinomio característico $\Delta(\lambda) = \prod_{i=1}^{m}(\lambda - \lambda_i)^{n_i}$, donde $\sum_{i=1}^{m} n_i = n$. Sea $h(\lambda) = \eta_{n-1}\lambda^{n-1} + \cdots + \eta_1\lambda + \eta_0$ el polinomio de orden $n-1$, con coeficientes η_i a determinar, tal que $f(A) = h(A)$. Entonces los coeficientes η_i pueden calcularse resolviendo el sistema de n ecuaciones algebraicas

$$f^{(k)}(\lambda_i) = h^{(k)}(\lambda_i)\,, \quad \text{para } k = 0, 1, \ldots, n_i - 1, \text{ y } i = 1, 2, \ldots, m,$$

donde

$$f^{(k)}(\lambda_i) \triangleq \left.\frac{d^k f(\lambda)}{d\lambda^k}\right|_{\lambda=\lambda_i} \quad \text{y} \quad h^{(k)}(\lambda_i) \triangleq \left.\frac{d^k h(\lambda)}{d\lambda^k}\right|_{\lambda=\lambda_i}.$$

3.6.4. Funciones matriciales no necesariamente polinomiales

Dada una función $f(\lambda)$ más general, una forma de definir $f(A)$ es usando el Teorema 3.4. Es decir: calculamos el polinomio $h(\lambda)$ de orden $n-1$ igualando $f(\lambda) = h(\lambda)$ en el espectro de A, y así *definimos* $f(A) = h(A)$.

Ejemplo 3.8. Queremos calcular e^{At} para $A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$. El polinomio característico de A es $\Delta(\lambda) = (\lambda-1)^2(\lambda-2)$. Sea $h(\lambda) = \eta_2\lambda^2 + \eta_1\lambda + \eta_0$. Aplicando el Teorema 3.4 tenemos que [2]

$$\begin{aligned} f(1) = h(1) \quad &\Rightarrow\ e^t = \eta_2 + \eta_1 + \eta_0 \\ f'(1) = h'(1) \quad &\Rightarrow te^t = 2\eta_2 + \eta_1 \\ f(2) = h(2) \quad &\Rightarrow e^{2t} = 4\eta_2 + 2\eta_1 + \eta_0 \,. \end{aligned}$$

Haciendo las cuentas obtenemos $\eta_0 = -2te^t + e^{2t}$, $\eta_1 = 3te^t + 2e^t - 2e^{2t}$, y $\eta_2 = e^{2t} - e^t - te^t$. Finalmente

$$\begin{aligned} e^{At} = h(A) &= (e^{2t} - e^t - te^t)A^2 + (3te^t + 2e^t - 2e^{2t})A + (-2te^t + e^{2t})I \\ &= \begin{bmatrix} 2e^t - e^{2t} & 0 & 2e^t - 2e^{2t} \\ 0 & e^t & 0 \\ e^{2t} - e^t & 0 & 2e^{2t} - e^t \end{bmatrix} . \end{aligned}$$

Ejemplo 3.9. Consideremos el bloque de Jordan de orden 4

$$A = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} .$$

Su polinomio característico es simplemente $\Delta(\lambda) = (\lambda - \lambda_1)^4$. Si en vez de seleccionar $h(\lambda) = \eta_3\lambda^3 + \eta_2\lambda^2 + \eta_1\lambda + \eta_0$, elegimos

$$h(\lambda) = \eta_3(\lambda-\lambda_1)^3 + \eta_2(\lambda-\lambda_1)^2 + \eta_1(\lambda-\lambda_1) + \eta_0 \,,$$

La condición $f(\lambda) = h(\lambda)$ en el espectro de A da las expresiones

$$\eta_0 = f(\lambda_1)\,, \quad \eta_1 = f'(\lambda_1)\,, \quad \eta_2 = \frac{f^{(2)}(\lambda_1)}{2!}\,, \quad \eta_3 = \frac{f^{(3)}(\lambda_1)}{3!} \,.$$

La propiedad de nilpotencia de $(A - \lambda I)$ cuando A es un bloque de Jordan,

$$(\lambda_1 I - A) = \begin{bmatrix} 0&1&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0 \end{bmatrix}, (\lambda_1 I - A)^2 = \begin{bmatrix} 0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0 \end{bmatrix}, (\lambda_1 I - A)^3 = \begin{bmatrix} 0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix}, (\lambda_1 I - A)^4 = \begin{bmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{bmatrix},$$

en este ejemplo permite obtener la expresión general para $f(A)$

$$f(A) = \begin{bmatrix} f(\lambda_1) & f'(\lambda_1)/1! & f^{(2)}(\lambda_1)/2! & f^{(3)}(\lambda_1)/3! \\ 0 & f(\lambda_1) & f'(\lambda_1)/1! & f^{(2)}(\lambda_1)/2! \\ 0 & 0 & f(\lambda_1) & f'(\lambda_1)/1! \\ 0 & 0 & 0 & f(\lambda_1) \end{bmatrix} .$$

Así, por ejemplo, para la función exponencial matricial $f(\lambda) = e^{\lambda t}$, tenemos

$$e^{At} = \begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! & t^3 e^{\lambda_1 t}/3! \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! \\ 0 & 0 & e^{\lambda_1 t} & te^{\lambda_1 t} \\ 0 & 0 & 0 & e^{\lambda_1 t} \end{bmatrix} .$$

[2] Ojo que la diferenciación en la segunda línea es con respecto a λ, no t.

3.6.5. Series de Potencias

Otra forma de definir una función matricial $f(A)$ es a través de la serie de potencias de $f(\lambda)$. Supongamos que $f(\lambda)$ se puede representar por

$$f(\lambda) = \sum_{i=0}^{\infty} \beta_i \lambda^i$$

con radio de convergencia ρ. Si todos los autovalores de A tienen magnitud menor que ρ, entonces podemos definir $f(A)$ como

$$f(A) = \sum_{i=0}^{\infty} \beta_i A^i \,. \tag{3.9}$$

Esta definición es en realidad equivalente a la anteriormente vista en el Teorema 3.4. No lo probamos, pero mostramos un caso particular.

Ejemplo 3.10. Consideremos otra vez el bloque de Jordan del Ejemplo 3.9,

$$A = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} .$$

Supongamos que $f(\lambda)$ tiene el siguiente desarrollo en serie de potencias alrededor de λ_1

$$f(\lambda) = f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1) + \frac{f''(\lambda_1)}{2!}(\lambda - \lambda_1)^2 + \cdots$$

Entonces

$$f(A) = f(\lambda_1)I + f'(\lambda_1)(A - \lambda_1 I) + \frac{f''(\lambda_1)}{2!}(A - \lambda_1 I)^2 + \cdots$$

Por la propiedad de nilpotencia del bloque de Jordan, $(\lambda I - A)^k = 0$ para $k \geq n = 4$, la serie de potencias se reduce inmediatamente a la expresión que sacamos en el Ejemplo 3.9.

3.6.6. Diferenciación e Integración Matricial

Se define elemento a elemento,

$$\int_0^t A(\tau)d\tau, \quad \frac{d}{dt}A(t) \qquad \text{son respectivamente} \qquad \int_0^t a_{ij}(\tau)d\tau, \quad \frac{d}{dt}a_{ij}(t) \,.$$

Con estas definiciones no es difícil probar que vale la propiedad

$$\frac{d}{dt}\left[A(t)B(t)\right] = \dot{A}(t)B(t) + A(t)\dot{B}(t) \,,$$

el *teorema fundamental del cálculo*,

$$\frac{d}{dt}\int_0^t A(\tau)d\tau = A(t) \,,$$

y la *regla de Leibniz*

$$\frac{d}{dt}\int_{f(t)}^{g(t)} A(t,\tau)d\tau = A(t,g(t))\dot{g}(t) - A(t,f(t))\dot{f}(t) + \int_{f(t)}^{g(t)} \frac{\partial}{\partial t}A(t,\tau)d\tau \,.$$

3.6.7. La Exponencial Matricial

Una función de particular interés en este curso es e^{At}. Como la serie de Taylor $e^{\lambda t} = 1 + \lambda t + \frac{\lambda^2 t^2}{2!} + \cdots + \frac{\lambda^n t^n}{n!} + \cdots$ converge para todo λ y t finitos, tenemos que

$$e^{At} = I + tA + \frac{t^2}{2!}A + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \,. \tag{3.10}$$

En MATLAB e^A se calcula con la función `expm(A)`,[3] que implementa la *expansión de Padé*.

Usando (3.10) es fácil probar las siguientes dos propiedades de e^{At}

$$\begin{aligned} e^0 &= I\,, \\ e^{A(t_1+t_2)} &= e^{At_1} e^{At_2}\,, \\ \left(e^{At}\right)^{-1} &= e^{-At}\,. \end{aligned}$$

(¿Cómo probamos la tercera?) Notar que en general $e^{(A+B)t} \neq e^{At}e^{Bt}$ (¿por qué?).

Diferenciando término a término (3.10) obtenemos

$$\begin{aligned} \frac{d}{dt} e^{At} &= \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k \\ &= A \left(\sum_{k=1}^{\infty} \frac{t^k}{(k)!} A^k \right) \\ &= \left(\sum_{k=1}^{\infty} \frac{t^k}{(k)!} A^k \right) A\,, \end{aligned}$$

así tenemos que $\frac{d}{dt} e^{At} = Ae^{At} = e^{At}A$.

3.7. Ecuación de Lyapunov

Es la siguiente ecuación matricial

$$AM + MB = C\,, \tag{3.11}$$

donde A y B son matrices constantes, respectivamente de dimensiones $n \times n$ y $m \times m$. Las matrices C y la *incógnita* M son $n \times m$.

La ecuación (3.11) es lineal en M y puede escribirse como sistema de ecuaciones algebraicas en la forma estándar $Ax = y$. Veámoslo para $n = 3$ y $m = 2$:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} + \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

Haciendo las cuentas y expresando M y C como vectores apilando sus columnas, llegamos a

$$\begin{bmatrix} a_{11}+b_{11} & a_{12} & a_{13} & b_{12} & 0 & 0 \\ a_{21} & a_{22}+b_{11} & a_{23} & 0 & b_{12} & 0 \\ a_{31} & a_{32} & a_{33}+b_{11} & 0 & 0 & b_{12} \\ b_{21} & 0 & 0 & a_{11}+b_{22} & a_{12} & a_{13} \\ 0 & b_{21} & 0 & a_{21} & a_{22}+b_{22} & a_{23} \\ 0 & 0 & b_{21} & a_{31} & a_{32} & a_{33}+b_{22} \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{12} \\ m_{22} \\ m_{32} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \\ c_{12} \\ c_{22} \\ c_{32} \end{bmatrix}$$

[3] Ojo no confundir con `exp(A)`, que da la matriz de las exponenciales de los elementos de A.

es decir $\mathcal{A}\vec{M} = \vec{C}$, donde $\mathcal{A}$ es la matriz $(n \cdot m) \times (n \cdot m)$ de la ecuación anterior, y $\vec{M}$ y $\vec{C}$ son las matrices M y C convertidas en vectores $(n \cdot m) \times 1$.[4]

Como $\mathcal{A}$ es cuadrada, por el Corolario 3.1, esta ecuación tiene solución única si la matriz $\mathcal{A}$ es invertible, es decir si no tiene ningún autovalor nulo. Puede probarse que si α_i y β_i son respectivamente los autovalores de A y B, los autovalores de $\mathcal{A}$ son $\alpha_i + \beta_j$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$. En consecuencia, la ecuación de Lyapunov tendrá solución única si no existen i, j tales que $\alpha_i + \beta_j = 0$.

En MATLAB la ecuación (3.11) se resuelve con `lyap(A,B,-C)`.

3.8. Formas Cuadráticas

Dada una matriz $M \in \mathbb{R}^{n\times n}$, la función escalar $x^T M x$, donde $x \in \mathbb{R}^n$, es una *forma cuadrática*. Sin pérdida de generalidad se puede tomar M como *simétrica*, $M = M^T$, ya que

$$x^T(Q + Q^T)x = 2x^T Q x, \quad \text{para todo } x \in \mathbb{R}^n.$$

Los autovalores de una matriz simétrica *son todos reales*, ya que para todo autovalor λ con autovector v de $M = M^T$,

1. el escalar $v^* M v$ (donde v^* denota la transpuesta conjugada de v) es real (sea v real o no),
$$(v^* M v)^* = v^* M^* v = v^* M v, \quad \text{y así}$$
2. λ debe ser real, dado que
$$v^* M v = v^* \lambda v = \lambda (v^* v)\,.$$

Toda matriz real simétrica M es diagonalizable, es decir, el orden de su mayor bloque de Jordan es 1. Si no lo fuera, es decir, si existiera un autovector generalizado v de orden mayor que 1 asociado a algún autovalor repetido λ, tendría que verificarse que

$$(\lambda I - M)^k v = 0, \quad \text{para algún } k > 1, \text{ y} \tag{3.12}$$
$$(\lambda I - M)^{k-1} v \neq 0. \tag{3.13}$$

Pero entonces

$$\begin{aligned} \left((\lambda I - M)^{k-1} v\right)^* (\lambda I - M)^{k-1} v &= v^* (\lambda I - M^*)^{k-1} (\lambda I - M)^{k-1} v \\ &= v^* (\lambda I - M)^{2k-2} v \\ &= v^* (\lambda I - M)^{k-2} (\lambda I - M)^k v \end{aligned} \tag{3.14}$$

debería ser nulo por (3.12) y no nulo por (3.13). Una contradicción que prueba que $M = M^T$ no puede tener bloques de Jordan de orden mayor que 1, y por lo tanto debe ser semejante a una matriz diagonal: existe una matriz $Q \in \mathbb{R}^{n\times n}$ no singular tal que $M = QDQ^{-1}$, donde $D \in \mathbb{R}^{n\times n}$ es diagonal.

Notar que como M es simétrica y D diagonal,

$$M = QDQ^{-1} = (QDQ^{-1})^T = (Q^{-1})^T D Q^T\,,$$

que implica que $Q^T = Q^{-1}$ y $QQ^T = I$. Toda matriz no singular Q con esta propiedad se llama *ortogonal*, y sus columnas son vectores *ortonormales*.

[4] $\mathcal{A}$ puede escribirse en forma compacta como $\mathcal{A} = I_{m\times m} \otimes A + B \otimes I_{n\times n}$ ($\otimes$: producto de Kronecker).

Teorema 3.5. Para toda matriz real simétrica M existe una matriz ortogonal Q tal que

$$M = QDQ^T \quad \text{o} \quad D = Q^T MQ\,,$$

donde D es una matriz diagonal con los autovalores de M, que son todos reales, en la diagonal.

Una desigualdad importante para una matriz simétrica M es la *desigualdad de Rayleigh-Ritz*, que dice que para cualquier $x \in \mathbb{R}^n$

$$\lambda_{\text{mín}} x^T x \leq x^T M x \leq \lambda_{\text{máx}} x^T x\,,$$

donde $\lambda_{\text{mín}}$ y $\lambda_{\text{máx}}$ son respectivamente el menor y mayor autovalor de M.

3.8.1. Matrices definidas y semi-definidas positivas

Definición 3.16 (Matriz definida y semi-definida positiva). Una matriz simétrica M se dice *definida positiva*, que denotamos $M > 0$, si $x^T M x > 0$ para todo vector $x \in \mathbb{R}^n$ no nulo. Es *semi-definida positiva*, que denotamos $M \geq 0$, si $x^T M x \geq 0$ para todo vector $x \in \mathbb{R}^n$ no nulo.

Si M es definida positiva, entonces $x^T M x = 0$ sii $x = 0$. Si M es semi-definida positiva, entonces existe algún x no nulo tal que $x^T M x = 0$.

Existen tests para determinar si una matriz es definida o semi-definida positiva basados en las propiedades del signo de los determinantes de diversas submatrices. Uno de estos tests se basa en los *menores principales*.

Definición 3.17 (Menores principales). Sea una matriz simétrica $M \in \mathbb{R}^{n\times n}$ con entradas $\{m_{ij}\}$. Entonces dados los enteros $p = 1, \ldots, n$ y $\{i_1, i_2, \ldots, i_p\}$, con $1 \leq i_1 \leq i_2 \leq \cdots \leq i_p \leq n$, definimos los *menores principales* de la matriz M como

$$M(i_1, i_2, \ldots, i_p) = \det \begin{bmatrix} m_{i_1,i_1} & m_{i_1 i_2} & \cdots & m_{i_1 i_p} \\ m_{i_2,i_1} & m_{i_2 i_2} & \cdots & m_{i_2 i_p} \\ \cdots & \cdots & \cdots & \cdots \\ m_{i_p,i_1} & m_{i_p i_2} & \cdots & m_{i_p i_p} \end{bmatrix}.$$

Los escalares $M(1, 2, \ldots, p)$, $p = 1, 2, \ldots, n$, que son simplemente los determinantes de las submatrices de la esquina superior izquierda de M,

$$M(1) = m_{11}, \quad M(1,2) = \det \left[\begin{smallmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{smallmatrix}\right],$$
$$M(1,2,3) = \det \left[\begin{smallmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{smallmatrix}\right], \quad \cdots$$

son los *primeros menores principales* de M.

Teorema 3.6 (Matriz definida (semi-definida) positiva). Una matriz simétrica M es definida positiva (semi-definida positiva) sii cualquiera de las siguientes condiciones se satisface:

1. Cada uno de sus autovalores es positivo (no negativo).

2. Todos sus *primeros menores principales* son positivos (todos su *menores principales* son no negativos).

3. Existe una matriz no singular $N \in \mathbb{R}^{n\times n}$ (una matriz singular $N \in \mathbb{R}^{n\times n}$, o una matriz $N \in \mathbb{R}^{m\times n}$, con $m < n$) tal que $M = N^T N$.

Una matriz simétrica M es *definida negativa* (*semi-definida negativa*) si $-M$ es definida positiva (*semi-definida positiva*).

Ejemplo 3.11. La matriz simétrica

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix}$$

es definida positiva si y sólo si $m_{11} > 0$ y $m_{11}m_{22} - m_{12}^2 > 0$. Es semi-definida positiva si y sólo si $m_{11} \geq 0, m_{22} \geq 0$, y $m_{11}m_{22} - m_{12}^2 \geq 0$.

3.9. La Descomposición en Valores Singulares (SVD)

Consideremos una matriz $A \in \mathbb{R}^{m\times n}$ y definamos $M = A^T A$. Claramente M es $n \times n$, simétrica, y semi-definida positiva. Sus autovalores λ_i son no negativos. Como

$$\det(\lambda I_m - AA^T) = \lambda^{m-n} \det(\lambda I_n - A^T A)$$

las matrices cuadradas $A^T A$ y AA^T comparten los mismos autovalores positivos y sólo difieren en el número de autovalores nulos.

Supongamos que hay r autovalores positivos de M, y sea $p = \text{mín}(m, n)$. Entonces podemos ordenar

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_p\,,$$

Los *valores singulares* de la matriz A son

$$\sigma_i \triangleq \sqrt{\lambda_i}, \quad i = 1, \ldots, p\,.$$

Por el Teorema 3.6, para $M = A^T A$ existe una matriz ortogonal Q tal que

$$Q^T A^T A Q = D = S^T S,$$

donde D es una matriz diagonal con los σ_i^2 en la diagonal. La matriz S es $m \times n$ con los σ_i en la diagonal.

Teorema 3.7 (Descomposición en Valores Singulares). Para toda matriz $A \in \mathbb{R}^{m\times n}$, existen matrices ortonormales $U \in \mathbb{R}^{m\times m}$ y $V \in \mathbb{R}^{n\times n}$ tales que

$$U^T A V = S \triangleq \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_p & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n}$$

donde $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$.

Los vectores en las columnas de $U = [\, u_1, \dots, u_m \,]$ y $V = [\, v_1, \dots, v_n \,]$ son los *vectores singulares izquierdos* y *derechos* respectivamente de A.

Es fácil verificar comparando las columnas de las ecuaciones $AV = US$ y $A^T U = VS^T$ que

$$\left.\begin{array}{rcl} Av_i & = & \sigma_i u_i \\ A^T u_i & = & \sigma_i v_i \end{array}\right\} i = 1, \dots, p = \text{mín}(m, n)\,.$$

La descomposición en valores singulares (SVD)[5] revela muchas propiedades de la matriz A. Por ejemplo, si r es el índice del valor singular positivo más pequeño, entonces

1. el rango de A es r,
2. los vectores $\{v_{r+1}, \dots, v_n\}$ son una base ortonormal del kernel de A,
3. los vectores $\{u_1, \dots, u_r\}$ son una base ortonormal de la imagen de A,
4. tenemos la representación
$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T\,.$$

Los valores singulares representan precisamente las longitudes de los semiejes del hiper-elipsoide $E = \{Ax : \|x\| = 1\}$. La Figura 3.3 muestra el conjunto E para una matriz A con $\sigma_1 = 0{,}8, \sigma_2 = 0{,}6, \sigma_3 = 0{,}4$.

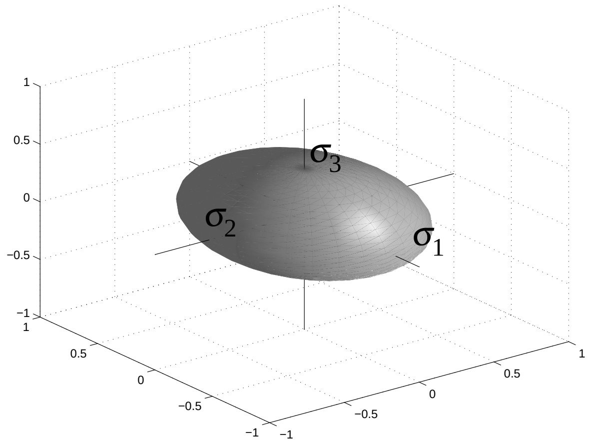

Figura 3.3: Hiper-elipsoide E en $\mathbb{R}^3$.

En MATLAB la función `[U,S,V]=svd(A)` calcula los valores y vectores singulares.

3.10. Normas de Matrices

El concepto de norma de vectores se extiende a matrices. Sea A una matriz $m \times n$. La norma de A puede definirse como

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| \tag{3.15}$$

[5]Singular Value Decomposition.

Esta norma definida desde la norma de los vectores se llama *norma inducida*. Usando diferentes normas vectoriales (1, 2, ∞, etc.) obtenemos diferentes normas inducidas. En este curso vamos a utilizar la norma inducida por la norma vectorial euclídea, $\|x\|_2$, que es la que se conoce como *norma espectral* de una matriz.

Otra propiedad importante de la descomposición en valores singulares de A es que la norma espectral de A está dada por su mayor valor singular,

$$\|A\| = \sigma_1 .$$

Ejemplo 3.12. Si λ_1 y λ_2 son números reales, la norma espectral de

$$A = \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_2 \end{bmatrix}$$

está dada por (3.15) como

$$\|A\| = \max_{\sqrt{x_1^2+x_2^2}=1} \sqrt{(\lambda_1 x_1 + x_2)^2 + \lambda_2^2 x_2^2} .$$

Para no tener que resolver un problema de optimización con restricciones, calculamos $\|A\|$ a través de σ_1 computando los autovalores de $A^T A$. El polinomio característico de $A^T A$ es

$$\begin{aligned} \det(\lambda I - A^T A) &= \det \begin{bmatrix} \lambda - \lambda_1^2 & -\lambda_1 \\ -\lambda_1 & \lambda - \lambda_2^2 - 1 \end{bmatrix} \\ &= \lambda^2 - (1 + \lambda_1^2 + \lambda_2^2)\lambda + \lambda_1^2 \lambda_2^2 . \end{aligned}$$

Las raíces de esta cuadrática están dadas por

$$\lambda = \frac{1 + \lambda_1^2 + \lambda_2^2 \pm \sqrt{(1 + \lambda_1^2 + \lambda_2^2)^2 - 4\lambda_1^2 \lambda_2^2}}{2}$$

El radical puede escribirse de forma que su positividad es obvia. Eligiendo el signo positivo, y después de un poco de álgebra llegamos a

$$\|A\| = \frac{\sqrt{(\lambda_1 + \lambda_2)^2 + 1} + \sqrt{(\lambda_1 - \lambda_2)^2 + 1}}{2} . \tag{3.16}$$

De (3.16) se ve que

$$\|A\| = \sigma_1 \geq \frac{|\lambda_1 + \lambda_2| + |\lambda_1 - \lambda_2|}{2} = \max(|\lambda_1|, |\lambda_2|) .$$

La propiedad de que σ_1 es mayor que la magnitud mayor de los autovalores de A es general, y más aún, para cualquier autovalor no nulo λ_i de $A \in \mathbb{R}^{n\times n}$

$$\sigma_p \leq |\lambda_i| \leq \sigma_1$$

La norma espectral de una matriz A en MATLAB se calcula con `norm(A)`.

Algunas propiedades útiles de la norma espectral de matrices:

$$\begin{aligned} \|Ax\| &\leq \|A\| \|x\| \\ \|A + B\| &\leq \|A\| + \|B\| \\ \|AB\| &\leq \|A\| \|B\| . \end{aligned}$$

3.11. Resumen

En este capítulo hemos repasado

- Las definiciones básicas relativas a vectores y matrices, traza, determinante, e inversa.
- Bases y conjuntos ortonormales, independencia lineal, el concepto de norma, y el método de ortonormalización de Schmidt.
- Los resultados principales para la solución de ecuaciones lineales algebraicas, y los conceptos de imagen, rango y kernel de una matriz. En particular, la ecuación $Ax = y$
  - tiene solución sii y está en la imagen de A.
  - tiene solución única si $\ker(A) = \{0\}$ ($\dim \ker(A) = 0$).
  - tiene infinitas soluciones si $\dim \ker(A) > 0$.
- Transformaciones de semejanza de matrices, que permiten representar una matriz en diferentes coordenadas.
- Formas diagonales y de Jordan, que muestran la estructura básica de una matriz dada. No toda matriz puede diagonalizarse, pero siempre se puede obtener la forma de Jordan (diagonal en bloques).
- Funciones polinomiales de matrices y el teorema de Cayley-Hamilton, que dice que toda matriz satisface su polinomio característico.
- El *polinomio mínimo*, que es el polinomio de menor orden $\Psi(\lambda)$ tal que $\Psi(A) = 0$. Es siempre un factor del polinomio característico de la matriz (coincide con éste si A no tiene autovalores repetidos).
- Que una función de una matriz cuadrada, $f(A)$ puede definirse evaluando $f(\lambda)$ en el *espectro* de A, y por lo tanto, como una combinación lineal de $\{I, A, \ldots, A^{n-1}\}$ (o de $\{I, A, \ldots, A^{\tilde{n}-1}\}$ donde $\tilde{n} \leq n$ es el orden del polinomio mínimo de A).
- Diferenciación e integración de matrices, definida elemento a elemento.
- La ecuación de Lyapunov $AM + MB = C$, que es una ecuación lineal algebraica matricial que tiene solución única $M \in \mathbb{R}^{n\times m}$ sii los autovalores α_i de $A \in \mathbb{R}^{n\times n}$, y β_j de $B \in \mathbb{R}^{m\times m}$ son tales que $\alpha_i + \beta_j \neq 0, \forall i = 1, \ldots, n; j = 1, \ldots, m$.
- Las formas cuadráticas $x^T Mx$, que son funciones escalares cuadráticas en los elementos del vector x, donde M puede tomarse como simétrica (o ser reemplada por $(M + M^T)/2$).
- Que las matrices simétricas tienen siempre autovalores reales, y que son siempre diagonalizables.
- Matrices simétricas definidas, y semi-definidas positivas, que son aquellas M para las que $x^T Mx$ es positiva, y no negativa, respectivamente.
- La descomposición en valores singulares de una matriz A. Los valores singulares son la raíz cuadrada de los autovalores de $A^T A$.
- La norma espectral de una matriz A, que es la inducida por la norma euclídea de vectores, y es el máximo valor singular, σ_1 de A.

3.12. Ejercicios

Ejercicio 3.1. Dados los vectores $x_1 = [2\ -3\ 1]^T$ y $x_2 = [1\ 1\ 1]^T$

1. Calcular sus normas 1, 2 e ∞.
2. Encontrar vectores ortonormales que generen el mismo subespacio.

Ejercicio 3.2. Determinar rango y bases de la imagen y el kernel para cada una de las siguientes matrices:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 4 & 1 & -1 \\ 3 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ejercicio 3.3. Dada la ecuación

$$\begin{bmatrix} 2 & -1 \\ -3 & 3 \\ -1 & 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

¿Cuantas soluciones tiene? ¿Qué pasa si $y = [1\ 1\ 1]^T$?

Ejercicio 3.4. Dada la ecuación

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

1. Encontrar la forma de la solución general.
2. Encontrar la solución de mínima norma euclídea.

Ejercicio 3.5. Dadas

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad b_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad b_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

encontrar las representaciones de la matriz A con respecto a las bases $\{b_1, Ab_1, A^2b_1, A^3b_1\}$ y $\{b_2, Ab_2, A^2b_2, A^3b_2\}$.

Ejercicio 3.6. Representar en forma de Jordan las siguientes matrices

$$A_1 = \begin{bmatrix} 1 & 4 & 10 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -4 & -3 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 4 & 3 \\ 0 & 20 & 16 \\ 0 & -25 & -20 \end{bmatrix}.$$

Ejercicio 3.7. Mostrar que si λ es un autovalor de A con autovector v, entonces $f(\lambda)$ es un autovalor de $f(A)$ con autovector v.

Ejercicio 3.8. Calcular e^{At} para las matrices A_1 y A_3 del Ejercicio 3.6.

Ejercicio 3.9. Mostrar que funciones de la misma matriz conmutan: $f(A)g(A) = g(A)f(A)$.

Ejercicio 3.10. Probar que la transformada de Laplace de e^{At} es

$$\mathcal{L}\left[e^{At}\right] = (sI - A)^{-1}.$$

Ejercicio 3.11. Sean todos los autovalores de A distintos, y sea q_i un autovector por derecha de A asociado a λ_i, $Aq_i = \lambda_i q_i$. Sea $Q = \left[\begin{smallmatrix} q_1 & q_2 & \cdots & q_n \end{smallmatrix}\right]$ y $P \triangleq Q^{-1} = \left[\begin{smallmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{smallmatrix}\right]$, donde p_i es la fila i de P. Mostrar que

1. p_i es un autovector por izquierda de A asociado a λ_i: $p_i A = \lambda_i p_i$.
2. $(sI - A)^{-1} = \sum \dfrac{1}{s - \lambda_i} q_i p_i$.

Ejercicio 3.12. Encontrar M solución de la ecuación de Lyapunov con $A = \left[\begin{smallmatrix} 0 & 1 \\ -2 & -2 \end{smallmatrix}\right]$, $B = 3$, $C = \left[\begin{smallmatrix} 3 \\ 3 \end{smallmatrix}\right]$. ¿Existe solución única?

Ejercicio 3.13. ¿Son las siguientes matrices definidas o semi-definidas positivas?

$$A_1 = \begin{bmatrix} 2 & 3 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_1^2 & a_1 a_2 & a_1 a_3 \\ a_2 a_1 & a_2^2 & a_2 a_3 \\ a_3 a_1 & a_3 a_2 & a_3^2 \end{bmatrix}$$

Ejercicio 3.14. Calcular los valores singulares de las siguientes matrices

$$A_1 = \begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 2 \\ 2 & 4 \end{bmatrix}$$

Ejercicio 3.15. Calcular la norma espectral de las siguientes matrices

$$A_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 - j & 0 \\ 0 & 1 + j \end{bmatrix}$$

Ejercicio 3.16. Dada una constante $\alpha > 1$, mostrar cómo definir una matriz 2×2 A tal que los autovalores de A sean los dos $1/\alpha$ y $\|A\| \geq \alpha$.

Ejercicio 3.17. Si A es invertible probar que $\|A^{-1}\| \geq \|A\|^{-1}$.