UNIVERSIDAD CENTROCCIDENTAL
LISANDRO ALVARADO
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS: MENCIÓN MATEMÁTICA

Una aplicación del algoritmo de Leverrier-Faddeev: Cálculo del polinomio mínimo

AUTOR:
Janeth del Carmen Camargo Álvarez

TUTOR:
Dr. Javier Hernández Benítez

Barquisimeto, Marzo 2012.

UNIVERSIDAD CENTROCCIDENTAL
LISANDRO ALVARADO
DECANATO DE CIENCIAS Y TECNOLOGÍA
MAESTRÍA EN CIENCIAS: MENCIÓN MATEMÁTICA

Una aplicación del algoritmo de Leverrier-Faddeev: Cálculo del polinomio mínimo

AUTOR:
Janeth del Carmen Camargo Álvarez

TUTOR:
Javier Hernández Benítez

Trabajo de grado presentado ante la ilustre
Universidad Centroccidental Lisandro Alvarado para
optar al título de Magister Scientiarum, Mención: Matemática
Barquisimeto, Marzo 2012.

A Dios, en primer lugar,
porque en mi caminar,
nunca me abandonó.
A mi madre, a mi padre,
a mi hija y a Carlos,
por su paciencia y apoyo
para alcanzar este pequeño logro.

Índice general

Índice de Tablas

Agradecimientos

Quiero hacer público mi agradecimiento muy especial a mi tutor: Dr. Javier Hernández Benítez, por su valioso aporte en conocimiento, paciencia y apoyo incondicional. También agradezco a mis amigos Yackelin, Mireya, Jurancy y Miguel que siempre estuvieron allí cuando los necesité.

Resumen

El siguiente trabajo pretende detallar la evolución del Algoritmo de Leverrier-Faddeev desde su creación en 1840 hasta las últimas implementaciones usadas. Este algoritmo, permite determinar simultáneamente el polinomio característico $p_A(s)$ de una matriz cuadrada A y la Matriz Adjunta de $(sI - A)$ en términos de la base canónica en el espacio lineal de los polinomios con coeficientes complejos y grado a lo mas n. Inicialmente, Leverrier se apoyó básicamente en la Identidad de Newton para obtener los coeficientes a_k del polinomio característico. En 1949, el algoritmo sufrió una modificación debido a Faddeev, Soriau y Otros, basándose en las propiedades de la matriz adjunta y usando los coeficientes a_k calculados inicialmente, introduciendo los coeficientes B_k de la matriz adjunta de $sI - A$. En 1996 Barnett dedujo una implementación con el objetivo de expresar el polinomio característico y la matriz adjunta en función de una Base de Polinomios Ortogonales Clásicos. En el 2004 los trabajos de Hernández, Marcellán y Rodríguez generalizaron el algoritmo abarcando todos los casos estudiados para familia particulares de polinomio ortogonales clásicos que inició Barnett. La idea de nuestro trabajo es buscar un algoritmo tipo Leverrier-Faddeev que aproveche estos cálculos para obtener el polinomio mínimo de una matriz. En nuestro trabajo se presenta un teorema que determina el Polinomio Mínimo de una matriz, sin necesidad de encontrar los auto-valores de una matriz y mediante el aprovechamiento del algoritmo de Leverrier Faddeev.

Introducción

En el estudio de las matemáticas surge la inquietud de conocer las diversas aplicaciones que pudiera tener algún área o tema en específico. En concordancia con esto, en el presente trabajo se pretende detallar la evolución del Algoritmo de Leverrier-Faddeev desde su creación hasta las últimas implementaciones usadas.

El algoritmo atribuido a Leverrier, Faddeev y otros, permite, dada una matriz $A \in \mathbb{C}^{n\times n}$, determinar simultáneamente el polinomio característico de A ($p_A(s)$) y la matriz adjunta de $sI_n - A$, donde I_n denota la matriz identidad en $\mathbb{C}^{n\times n}$. Este algoritmo toma en cuenta la representación del polinomio característico y la matriz adjunta en términos de la base canónica $\{s^k\}_{k=0}^n$ en el espacio lineal de los polinomios con coeficientes complejos y grado a lo más n. A pesar de su poco valor desde el punto de vista numérico, éste es útil tanto para propósitos teóricos como cálculo simbólico, teniendo como aplicaciones la teoría de control lineal. Específicamente, $(sI - A)^{-1} = \frac{1}{p_A(s)} Adj(sI_n - A)$ es la función de transferencia de un sistema lineal de tiempo continuo con n-entradas y n-salidas.

Entre las últimas implementaciones usadas del algoritmo Leverrier-Faddeev se pueden ver en los trabajos [1], [16] y [17] donde cambian la base canónica por bases de polinomios ortogonales. En particular, la obtención de algoritmos tipo Leverrier está íntimamente ligada a la inversión de matrices polinómicas, i. e. al estudio de la función de transferencia del sistema lineal. Extensiones del algoritmo de Leverrier han sido analizadas en los trabajos pioneros de S. Barnett ([2], [3]) y J. C. Gower [10] y posteriormente ampliadas en el caso de haces lineales (véase [13], [21], [24], [29], [30] y [12]) y matrices polinómicas de grado arbitrario ([4], [6], [8]). Más recientemente, se han aplicado dichas técnicas en el estudio de inversas generalizadas de matrices polinómicas y racionales ([18], [19], [27]). La conexión con problemas de mecánica clásica (teoría de vibraciones) se ha puesto de manifiesto en [9] y [23] así como en relación con modelos 2-dimensionales de espacios de estado ([31], [32]). Finalmente, una interesante aplicación en sistemas expertos para la obtención de modelos a tiempo discreto utilizando identificación paramétrica, ha sido desarrollado en [26].

En lugar de utilizar la representación del polinomio característico en términos de la base canónica $\{s^n\}_{n\geqslant 0}$, los trabajos J. Hernández y F. Marcellán en [12] y [13] usaron bases de polinomios ortogonales clásicos (Hermite, Laguerre, Jacobi, Bessel) y obtuvieron un marco general que abarca los casos estudiados para familias particulares de polinomios ortogonales obtenidos en [3] y [30]. La clave de esa generalización es una propiedad estructural de los polinomios ortogonales clásicos.

Un pequeño aporte de este trabajo es la adaptación del algoritmo de Leverrier-Faddeev para el cálculo del polinomio mínimo de una matriz, que presentamos con detalle y mostramos algunos ejemplos en el capítulo 3.

La estructura de esta monografía es la siguiente, en el capítulo 1, estudiamos una herramienta útil para obtener en forma explícita la función de transferencia de un sistema lineal, el algoritmo de Leverrier-Faddeev. Hacemos un repaso de la evolución de este algoritmo desde su obtención al igual que presentamos sus ventajas y desventajas.

El capítulo 2 es una introducción que contiene conceptos básicos acerca de funcionales de momentos, caracterizaciones y operaciones algebraicas. También introducimos algunos aspectos de la teoría de polinomios ortogonales asociados a funcionales lineales al igual que damos ciertas caracterizaciones de los polinomios ortogonales clásicos. Al final de este capítulo estudiamos la relación entre polinomios ortogonales y funcionales lineales en el contexto de la conexión entre la matriz de momentos. En este capítulo también estudiamos la implementación del algoritmo de Leverrier-Faddeev desarrolladas en [12] y [13] en la que se puede expresar la función de transferencia de un sistema lineal en función de una base de polinomios ortogonales clásicos.

En el capítulo 3, adaptamos el algoritmo de Leverrier-Faddeev para encontrar el polinomio mínimo de una matriz $A \in \mathbb{C}^{n\times n}$. Este desarrollo es inspirado por los trabajos de J. Gower (ver [10] y [11]) en los que utiliza un algoritmo tipo Leverrier-Faddeev, pero sujeto al conocimiento a priori de los coeficientes del polinomio mínimo. También mostramos algunos ejemplos a fin de ilustrar el funcionamiento de nuestra adaptación.

Capítulo 1

Algoritmo de Leverrier-Faddeev

El algoritmo de Leverrier-Faddeev es un método que permite calcular de forma simultánea, el polinomio característico de una matriz $A \in \mathbb{C}^{n\times n}$ y la matriz adjunta de $\lambda I_n - A$, basándose en la traza de las potencias de la matriz A.

Leverrier obtuvo un algoritmo para calcular el polinomio característico de una matriz apoyado básicamente en la identidad de Newton (ver [15]), que permite establecer una relación entre los coeficientes de un polinomio y la suma de las potencias de sus raíces.

Teorema 1 (Identidad de Newton) *Sean $x_1, x_2, \cdots, x_n$, las raíces del polinomio $p(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ con $b_n = 1$.*

Denotemos el momento de orden k asociado a p, mediante $\mu_k = x_1^k + x_2^k + \cdots + x_n^k$. Entonces,

$$kb_{n-k} + \mu_1 b_{n-k+1} + \mu_2 b_{n-k+2} + \cdots + \mu_k b_n = 0, \quad k = 1, 2, \cdots, n. \tag{1.1}$$

Demostración Definamos la función simétrica elemental de $x_1, x_2, \cdots, x_n$, como:

$$s_k = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} x_{i_1}x_{i_2}\cdots x_{i_k} \quad \textit{con } k = 1, 2, \cdots, n,$$

consideremos

$$(b_1, b_2, \cdots, b_n) = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} x_{i_1}^{b_1}x_{i_2}^{b_2}\cdots x_{i_n}^{b_n}$$

donde los b_i son enteros no negativos con $b_i \geqslant b_{i+1}$ y cada sumatoria produce términos distintos. Si $b_i = 0$ para $i > t$, se escribirá $(b_1, b_2, \cdots, b_t)$ en lugar de

$(b_1, b_2, \cdots, b_n)$, esto es,

$$(b_1, b_2) = \sum x_{i_1}^{b_1} x_{i_2}^{b_2} \cdots x_{i_n}^{0} = \sum x_{i_1}^{b_1} x_{i_2}^{b_2}$$

$$(b_1) = \sum x_{i_1}^{b_1} x_{i_2}^{0} \cdots x_{i_n}^{0} = \sum x_{i_1}^{b_1}$$

Utilizando esta notación, las ecuaciones anteriores $(1), (1,1)$ y $(1,1,1)$ son exactamente las funciones simétricas elementales s_1, s_2, s_3 y las ecuaciones $(1), (2)$ y (3) son respectivamente, las funciones momentos μ_1, μ_2 y μ_3.

Para encontrar una relación entre los funcionales de momento y las funciones simétricas elementales descritas anteriormente consideremos el siguiente procedimiento, para el caso $k = 3$ y $n \geqslant 3$, podemos notar que

$$(2)(1) = (3) + (2,1),$$

respectivamente

$$\mu_2 s_1 = \mu_3 + (2,1).$$

Para eliminar $(2,1)$ usemos el hecho que

$$(1)(1,1) = (2,1) + 3(1,1,1)$$

puede ser vista equivalentemente como

$$\mu_1 s_2 = (2,1) + 3 s_3,$$

donde el lado izquierdo es un producto de una función de Momento y una función simétrica elemental. Si restamos la segunda ecuación con la primera, obtenemos exactamente la identidad de Newton:

$$\mu_3 - \mu_2 s_1 + \mu_1 s_2 - 3 s_3 = 0,$$

en efecto,

$$\begin{aligned} (2)(1) - (3) &= (1)(1,1) - 3(1,1,1) \\ \mu_2 s_1 - \mu_3 &= \mu_1 s_2 - 3 s_3 \\ \mu_3 - \mu_2 s_1 + \mu_1 s_2 - 3 s_3 &= 0. \end{aligned}$$

Generalicemos este hecho, sea $s_i = (1_i)$ una sucesión de $i-$unos y si $t \geqslant 1$, sea $(t, 1_i) = (c_1, \cdots, c_{i+1})$ donde $c_1 = t$ y $c_j = 1$ para $j > 1$. Para obtener la identidad de Newton, introducimos $\mu_1, \mu_2, \cdots, \mu_k$, escribimos t ecuaciones, donde

$$t = min(k-1, n):$$

$$\begin{array}{lcl} (k-1)(1) & = & (k) + (k-1,1) \\ (k-2)(1,1) & = & (k-1,1) + (k-2,1,1) \\ (k-3)(1,1,1) & = & (k-2,1,1) + (k-3,1,1,1) \end{array}$$

En general,

$$(k-i)(1_i) = (k-i+1, 1_{i-1}) + (k-i, 1_i) \;\; i = 1, \cdots, t.$$

Si $n \geqslant k = t+1$, la ecuación anterior nos queda

$$(1)(1_{k-1}) = (2, 1_{k-2}) + k(1_k)$$

mientras, si $k > n = t$, entonces,

$$(k-n)(1_n) = (k-n+1, 1_{n-1}).$$

Debido a que $(k-n, 1_n)$ tiene $n+1$ entradas, representa el polinomio cero. Multiplicando la $i-$ésima ecuación por $(-1)^{i-1}$ y sumando las ecuaciones, obtenemos la Identidad de Newton. En efecto,

$$\sum_{i=1}^{t}(-1)^{i-1}(k-i)(1_i) = \sum_{i=1}^{t}(-1)^{i-1}(k-i+1, 1_{i-1}) + \sum_{i=1}^{t}(-1)^{i-1}(k-i, 1_i)$$

Desarrollando cada una de las sumas por separado tenemos,

$$\sum_{i=1}^{t}(-1)^{i-1}(k-i)(1_i) = -\sum_{i=1}^{k-1}(-1)^{i}\mu_{k-i}s_i$$

$$\begin{array}{lcl} \sum_{i=1}^{t}(-1)^{i-1}(k-i+1, 1_{i-1}) & = & (k, 1_0) - (k-1, 1_1) + (k-2, 1_2) - (k-3, 1_3) + (k-4, 1_4) \\ & & + \cdots + (-1)^{t-1}(k-t+1, 1_{t-1}) \\ \sum_{i=1}^{t}(-1)^{i-1}(k-i, 1_i) & = & (k-1, 1_1) - (k-2, 1_2) + (k-3, 1_3) - (k-4, 1_4) + (k-5, 1_5) \\ & & + \cdots + (-1)^{t-2}(k-t+1, 1_{t-1}) + (-1)^{t-1}(k-t, 1_t) \end{array}$$

De esta manera,

$$\begin{aligned} -\sum(-1)^i\mu_{k-i}s_i &= ((k,1_0)-(k-1,1_1)+(k-2,1_2)-(k-3,1_3)+(k-4,1_4) \\ &\quad +\cdots+(-1)^{t-1}(k-t+1,1_{t-1}))+((k-1,1_1)-(k-2,1_2) \\ &\quad +(k-3,1_3)-(k-4,1_4)+(k-5,1_5) \\ &\quad +\cdots+(-1)^{t-2}(k-t+1,1_{t-1})+(-1)^{t-1}(k-t,1_t)), \end{aligned}$$

luego,

$$\begin{aligned} -\sum_{i=1}^{k-1}(-1)^i\mu_{k-i}s_i &= (k,1_0)+(-1)^{t-1}(k-t,1_t) \\ (k,1_0)+\sum_{i=1}^{k-1}(-1)^i\mu_{k-i}s_i+(-1)^{t-1}(k-t,1_t) &= 0 \\ (k)+\sum_{i=1}^{k-1}(-1)^i\mu_{k-i}s_i+(-1)^k k1_k &= 0 \\ \mu_k+\sum_{i=1}^{k-1}(-1)^i\mu_{k-i}s_i+(-1)^k ks_k &= 0 \end{aligned}$$

■

1.1. Deducción del Algoritmo

Sea $A \in \mathbb{C}^{n\times n}$ una matriz con polinomio característico

$$\begin{aligned} p_A(s) &= det(sI_n - A) \\ &= s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0. \end{aligned} \tag{1.2}$$

De la expresión (1.1), tenemos que,

$$ka_{n-k} + \mu_1 a_{n-k+1} + \cdots + \mu_k = 0, k = 1, 2, \cdots, n,$$

esto es, los coeficientes de p_A dados por (1.2) satisfacen

$$a_k = \frac{-1}{k}(\mu_k + a_1\mu_{k-1} + \cdots + a_{k-1}\mu_1),\ k = 1, 2, \cdots, n. \tag{1.3}$$

De esta forma, el algoritmo de Leverrier, se formula de la siguiente manera, (ver tabla 1.1).

Datos de entrada: $A \in \mathbb{C}^{n\times n}$.
Condición inicial $a_0 = 1$
FOR: $k = 1, 2, \cdots, n$
$\mu_k = tr(A^k)$
$a_k = \frac{-1}{k}(\mu_k + a_1\mu_{k-1} + \cdots + a_{k-1}\mu_1)$
END

Tabla 1.1: Algoritmo de Leverrier

Años después, este algoritmo fue modificado por D.K. Faddeev (ver [7]) y J.M. Soriau, entre otros. Supongamos que el polinomio característico de A es

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

donde los coeficientes $\{a_k\}_{k=1}^n$ se obtuvieron mediante el algoritmo de Leverrier. Entonces,

$$a_1 = \mu_1 = -\sum_{j=1}^{n} s_j = -trA$$

luego, por la ecuación (1.3), tenemos,

$$\begin{aligned} -2a_2 &= \mu_2 + a_1\mu_1 \\ &= tr(A^2) + a_1 tr(A) \\ &= tr(A^2 + a_1 tr(A)) \\ &= tr(A(A + a_1 I_n)). \end{aligned}$$

Si $B_k = a_k I_n + AB_{k-1},\ k = 1, 2, \cdots, n-1,\ con\ B_0 = I_n$, entonces,

$$a_2 = \frac{-1}{2} tr(AB_1).$$

Para $k = 3, ..., n$,

$$\begin{array}{rcl} -ka_k & = & (\mu_k + a_1\mu_{k-1} + \cdots + a_{k-1}\mu_1) \\ & = & tr(A^k) + a_1 tr(A^{k-1}) + \cdots + a_{k-1} tr(A) \\ & = & tr(A^k + a_1 tr(A^{k-1}) + \cdots + a_{k-1}A) \\ & = & tr(A^{k-1}(A + a_1 I_n) + a_2 A^{k-2} + \cdots + a_{k-1}A) \\ & = & tr(A^{k-1}(B_1) + a_2 A^{k-2} + \cdots + a_{k-1}A) \\ & = & tr(A^{k-2}(a_2 I_n + AB_1) + a_3 A^{k-3} + \cdots + a_{k-1}A) \\ & = & tr(A(a_{k-1}I_n + AB_{k-2})) \\ & = & tr(AB_{k-1}). \end{array}$$

De ahí,

$$a_k = -\frac{1}{k} tr(AB_{k-1}),$$

Por otra parte, puesto que como $det(sI_n - A) \neq 0$, consideremos

$$B(s) := Adj(sI_n - A) = s^{n-1}I_n + s^{n-2}A_1 + \cdots + sA_{n-2} + A_{n-1}, \tag{1.4}$$

donde $A_1, \cdots, A_{n-1}$ son matrices de orden n. De la igualdad

$$(sI_n - A)(s^{n-1}I_n + s^{n-2}A_1 + \cdots + sA_{n-2} + A_{n-1}) = p_A(s)I_n, \tag{1.5}$$

se sigue que

$$\sum_{k=0}^{n} a_k I_n s^k = I_n s^n - AA_{n-1} + \sum_{k=1}^{n-1} (A_{n-k} - AA_{n-k-1}) s^k,$$

igualando los coeficientes en cada uno de los miembros de la expresión anterior, tenemos,

$$\begin{array}{lcl} A_1 & = & a_1 I_n + A = B_1 \\ A_2 & = & a_2 I_n + AB_1 = B_2 \\ & \vdots & \\ A_{n-1} & = & a_{n-1}I_n + AB_{n-2} = B_{n-1} \\ AB_{n-1} & = & -a_n I_n. \end{array}$$

En consecuencia,

$$B(s) = Adj(sI_n - A) = s^{n-1}I_n + s^{n-2}B_1 + \cdots + sB_{n-2} + B_{n-1}.$$

DATOS DE ENTRADA: $A \in \mathbb{C}^{n\times n}$.
Condiciones iniciales: $a_0 = 1,\ \ B_0 = I_n$,
FOR: $k = 1, 2, \cdots, n-1$

$a_k = -\frac{1}{k}tr(AB_{k-1})$
$B_k = a_k I_n + AB_{k-1}$
END (For)
$a_n = -\frac{1}{n}tr(AB_{n-1})$

Tabla 1.2: Algoritmo de Leverrier-Faddeev

De esta manera, el algoritmo de Leverrier-Faddeev calcula en forma simultánea los coeficientes $\{a_k\}_{k=1}^n$ y $\{B_k\}_{k=1}^n$ correspondiente al polinomio característico de A y la matriz adjunta de $sI_n - A$.

Ejemplo 1 Consideremos la siguiente matriz,

$$A = \begin{bmatrix} 1 & -4 & -1 & -4 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 6 \end{bmatrix},$$

Si aplicamos el algoritmo de Leverrier-Faddeev tenemos

$$a_1 = -tr(AB_0) = -trA = -5,\ B_1 = a_1 I_4 + A = \begin{bmatrix} -4 & -4 & -1 & -4 \\ 2 & -2 & 5 & -4 \\ -1 & 1 & -7 & 3 \\ -1 & 4 & -1 & 1 \end{bmatrix}$$

$$a_2 = -\frac{1}{2}tr(AB_1) = 9,\ B_2 = a_2 I_4 + AB_1 = \begin{bmatrix} 2 & -1 & -10 & 5 \\ -9 & -10 & -33 & 3 \\ 5 & 9 & 26 & -3 \\ 7 & 7 & 22 & 0 \end{bmatrix}$$

$$a_3 = -tr(AB_2) = -7,\ B_3 = a_3 I_4 + AB_2 = \begin{bmatrix} -2 & 2 & 8 & -4 \\ 1 & 8 & 22 & -5 \\ 0 & -6 & -16 & 4 \\ -1 & -6 & -16 & 3 \end{bmatrix}$$

$a_4 = -tr(AB_3) = 2.$

Por lo tanto, el polinomio característico de A es

$$p_A(s) = s^4 - 5s^3 + 9s^2 - 7s + 2,$$

y la matriz adjunta de $sI_4 - A$ es

$$Adj(sI_4 - A) = s^3 I_4 + s^2 B_1 + sB_2 + B_3$$

$$= \begin{bmatrix} s^3 - 4s^2 + 2s - 2 & -4s^2 - s + 2 & -s^2 10s + 8 & -4s^2 + 5s - 4 \\ 2s^2 - 92s + 1 & s^3 - 5s^2 - s + 8 & 5s^2 - 332s + 22 & -4s^2 + 3s - 5 \\ -s^2 + 5s & s^2 + 9s - 6 & s^3 - 7s^2 + 26s - 16 & 3s^2 - 3s + 4 \\ -s^2 + 7s - 1 & s^2 + 7s - 6 & -s^2 + 22s - 16 & s^3 + s^2 + 3 \end{bmatrix}.$$

Ejemplo 2. Consideremos la matriz,

$$A = \begin{bmatrix} \frac{1}{100} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{100} & 0 & 0 & 0 \\ 0 & 0 & 0{,}99 & 0 & 0 \\ 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 10000 \end{bmatrix},$$

tenemos que el polinomio característico de A es,

$$p_A(t) = t^5 - 10101t^4 + 19901t - 99.$$

Pero, si aplicamos el algoritmo de Leverrier-Faddeev (tabla1.2), se obtiene como polinomio característico,

$$p_A(t) = t^5 - 10101t^4 + 19901t - 10{,}58459743,$$

el cual como se puede observar, muestra que este algoritmo no es adecuado desde el punto de vista computacional.

Capítulo 2

Polinomios Ortogonales.

Pretendemos dar una visión de los polinomios ortogonales, entes matemáticos de gran interés y sencillez y con un sin número de aplicaciones tanto en matemáticas como en física e ingeniería (ver[1], [4], [26]). Dentro de los polinomios ortogonales existen unas familias que merecen una mención especial, por su relevancia en el desarrollo histórico debido a la gran cantidad de aplicaciones que se encuentran en áreas como la física cuántica, ecuaciones diferenciales, etc. Aparecen cuando se estudian problemas con ecuaciones diferenciales hipergeométricas. Se trata de los Polinomios Ortogonales Clásicos: Laguerre, Hermite, Bessel y Jacobi, que constituyen las familias más importantes de polinomios ortogonales.

2.1. Funcionales de Momentos y Polinomios Ortogonales asociados a funcionales lineales.

Sea $\mathcal{U} : \mathbb{P} \mapsto \mathbb{C}$ un funcional lineal en el espacio $\mathbb{P}$ de los polinomios con coeficientes complejos y consideremos la base canónica de $\{x^n\}_{n\geqslant 0}$.

Definición 1 La sucesión de números complejos $\{m_n\}_{n\geqslant 0}$, definida por

$$\mathcal{U}(x^n) = m_n, \quad n = 0, 1, 2 \ldots$$

se denomina sucesión de momentos. El valor m_n se denomina momento de orden n y el funcional lineal $\mathcal{U}$ se llama funcional de momento asociado a la sucesión $\{m_n\}_{n\geqslant 0}$.

Si $\pi(x) = \sum_{k=0}^{n} a_k x^k$, por la linealidad de $\mathcal{U}$, tenemos,

$$\mathcal{U}(\pi(x)) = \sum_{k=0}^{n} a_k m_k.$$

La variable x siempre será considerada una variable real y $\mathcal{U}(\overline{\pi(x)}) = \sum_{k=0}^{n} \overline{a_k} m_k$, donde $\overline{z}$ denota el conjugado del número complejo z.

Definición 2 Una sucesión de polinomios $\{P_n\}_{n\geqslant 0}$ se llama sucesión de polinomios ortogonales con respecto al funcional de momento $\mathcal{U}$ y nos referiremos a éste como una SPO para $\mathcal{U}$ si y solo si

(i) P_n es de grado n ($\deg P_n(x) = n$).

(ii) $\mathcal{U}(P_n P_m) = k_n \delta_{n,m}$ *con* $k_n \neq 0$ y $\delta_{n,m}$ la delta del kronecker definida por

$$\delta_{n,m} = \begin{cases} 1 & \text{si } n = m, \\ 0 & \text{si } n \neq m. \end{cases}$$

Cuando, $k_n = 1$, la sucesión $\{P_n\}_{n\geqslant 0}$ es una sucesión de polinomios ortonormales, para $n \geqslant 0$. Cuando el coeficiente principal de cada $\{P_n\}$ es la unidad, se llama sucesión de polinomios ortogonales mónicos (SPOM) para $\mathcal{U}$.

De la definición (2) se sigue que existe una SPO para $\mathcal{U}$, entonces,

$$m_0 = \mathcal{U}(P_0(x)) \neq 0, \; P_0(x) \neq 0.$$

Por lo tanto, una SPO no puede existir si $m_0 = 0$. Otros casos menos triviales donde una SPO no existe es cuando por ejemplo, $m_0 = m_1 = m_2 = 1$. En este sentido, si $P_0(x) = a \neq 0$ *y* $P_1(x) = bx + c$, $b \neq 0$, entonces

$$\mathcal{U}(P_0(x)P_1(x)) = a(bm_1 + cm_0) = a(b + c) = 0,$$

es decir, $b = -c$ y se tiene $P_1(x)^2 = b^2(m_2 - 2m_1 + m_0) = 0$.

Definición 3 Sea $\{m_n\}_{n\geqslant 0}$ la sucesión de momentos de un funcional $\mathcal{U}$, la matriz semi-infinita H, cuyas entradas están dadas por $h_{ij} = m_{i+j}, i, j = 0, 1, 2, ...$ se denomina matriz de momentos estándar asociada al funcional $\mathcal{U}$. **H** es una matriz de Hankel, esto es, una matriz cuyos elementos de las anti-diagonales son idénticos, es decir,

$$H = \begin{bmatrix} m_0 & m_1 & m_2 & \dots \\ m_1 & m_2 & m_3 & \dots \\ m_2 & m_3 & m_4 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Para cada valor de n, denotemos por H_n a las submatrices principales de H de orden $(n+1)$,

$$H_n = \begin{bmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{bmatrix}$$

TEOREMA 2 *Sea $\mathcal{U}$ un funcional de momentos y $\{m_n\}_{n\in\mathbb{N}}$ la correspondiente sucesión de momentos. Entonces existe una SPO asociada a $\mathcal{U}$ si y solo si las submatrices H_n son no singulares, es decir, cuando $det H_n \neq 0$ para $n \in \mathbb{N}$. En este caso, $\mathcal{U}$ se llama regular o cuasi-definido.*

Demostración (ver[5], pág.11) ■

TEOREMA 3 *Sea $\{P_n(x)\}_{n\geqslant 0}$ una SPO asociada a $\mathcal{U}$. Entonces, para todo $\pi \in \mathbb{P}$ con* $\deg \pi(x) = n$,

$$\pi(x) = \sum_{k=0}^{n} \lambda_k P_k(x),$$

donde

$$\lambda_k = \frac{\mathcal{U}(P_k(x)\pi(x))}{\mathcal{U}(P_k(x)^2)}, \ \ para \ \ k = 0, 1, 2, \cdots, n.$$

COROLARIO 1 *Si $\{P_n(x)\}_{n\geqslant 0}$ y $\{Q_n(x)\}_{n\geqslant 0}$ son dos SPO asociadas al mismo funcional $\mathcal{U}$, entonces, existe una sucesión $\{\mathbf{c}_n\}_{n\geqslant 0}$ con $\mathbf{c}_n \neq 0$ tal que $Q_n(x) = \mathbf{c}_n P_n(x)$, es decir, cada $P_n(x)$ es único, salvo por factores constantes.*

Demostración (ver [5] pág 9.)

Una de las características mas importantes de los polinomios ortogonales es que existe una conección entre tres polinomios consecutivos a través de una simple relación, llamada Fórmula Fundamental de Recurrencia.

Teorema 4 *(Fórmula Fundamental de Recurrencia). Sea $\mathcal{U}$ un funcional de momentos cuasi-definido y $\{P_n(x)\}_{n\in\mathbb{N}}$ la correspondiente sucesión de polinomios ortogonales mónicos. Entonces existe dos sucesiones de números complejos $\{\beta_n\}_{n=0}^{\infty}$ y $\{\gamma_n\}_{n=1}^{\infty}$ con $\gamma_n \neq 0$ para todo $n \in \mathbb{N}$ tales que*

$$xP_n(x) = P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x),\ n \geqslant 0 \tag{2.1}$$

$$P_0(x) = 1,\ P_1(x) = x - \beta_0.$$

Demostración (ver [5], pág 18.)

La representación matricial del operador asociado con la multiplicación por x en $\mathbb{P}$ respecto a la base de polinomios ortogonales mónicos es una matriz tridiagonal:

$$J = \begin{bmatrix} \beta_0 & 1 & 0 & 0 & \cdots \\ \gamma_1 & \beta_1 & 1 & 0 & \cdots \\ 0 & \gamma_2 & \beta_2 & 1 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

además, el polinomio característico de la submatriz principal J_{n+1} de dimensión $(n+1)\times(n+1)$ es el polinomio P_{n+1}, con $n \geqslant 0$.

Definición 4 Sea $E \subset (-\infty, \infty)$. Un funcional de momentos $\mathcal{U}$ se dice definido positivo en E, si y solo si $\mathcal{U}(p(x)) > 0$ para todo polinomio $p(x) \in \mathbb{P}$, no nulo, no negativo y todo $x \in \mathbb{R}$. En un caso definido positivo, los momentos del funcional son reales (ver [5]) y existe una medida de Borel positiva μ soportada en un subconjunto $\mathbb{E}$ de la recta real tal que el funcional lineal $\mathcal{U}$ tiene una representación integral

$$\mathcal{U}(p(x)) = \int_{\mathbb{E}} p(x)d\mu(x).$$

Definición 5 Sea $\phi(x) \in \mathbb{P}$, un polinomio con $deg\ \phi = n$, el funcional lineal $\phi\mathcal{U}$, es llamado multiplicación a la izquierda de $\mathcal{U}$ por ϕ y se define como

$$(\phi(x)\mathcal{U})(p(x)) = \mathcal{U}(\phi(x)p(x))$$

Asimismo, llamaremos multiplicación por la izquierda de $\mathcal{U}$ por el polinomio inverso de ϕ, al funcional $\phi^{-1}\mathcal{U}$, definido por

$$(\phi^{-1}(x)\mathcal{U})(p(x)) = \mathcal{U}\left(\frac{p(x) - L_{p,\phi}(x)}{\phi(x)}\right)$$

donde $L_{p,\phi}(x)$ es el polinomio interpolador de grado $deg(p) - 1$ que interpola a p en las raíces de ϕ. Particularmente, si $\alpha \in \mathbb{C}$, llamaremos multiplicación por la izquierda de $\mathcal{U}$ por la función racional $(x-\alpha)^{-1}$ al funcional lineal $\widetilde{\mathcal{U}}$ definido por

$$\widetilde{\mathcal{U}} := (x-\alpha)^{-1}\mathcal{U}(p(x)) = \mathcal{U}(\frac{p(x)-p(\alpha)}{x-\alpha}),\ p \in \mathbb{P}.$$

Definición 6 La derivada distribucional usual para $\mathcal{U}$ está dada por

$$(D\mathcal{U})(p(x)) = -\mathcal{U}(p'(x)),$$

para todo $p(x) \in \mathbb{P}$.

Sea $\mathcal{U}$ un funcional de momento regular, definido sobre el espacio vectorial de los polinomios con coeficientes complejos $\mathbb{P}$ y sea $\{P_n\}_{n\geqslant 0}$ la correspondiente sucesión de polinomios ortogonales mónicos, el conjunto de todos los funcionales de momento $\mathbb{P}^*$, sobre $\mathbb{P}$ es el dual algebraico de $\mathbb{P}$ y debido a que $\{P_n, n \in \mathbb{N}\}$ es una base algebraica en $\mathbb{P}$, podemos asociar la base dual $\{\alpha_n\}_{n\in\mathbb{N}}$ de siguiente forma

$$\langle \alpha_n, P_m \rangle = \delta_{m,n}.$$

Si $\mathbf{v}$ es un elemento de $\mathbb{P}^*$, podemos expresar éste como

$$\mathbf{v} = \sum_{i=0}^{\infty} \lambda_i \alpha_i,\ donde\ \lambda_i = \langle \mathbf{v}, P_i \rangle\ para\, i = 0, 1, 2, ...$$

Como una consecuencia inmediata, si $\mathbf{v} \in \mathbb{P}^*$ satisface $\langle \mathbf{v}, P_i \rangle = 0$, para $i \geqslant j$, entonces,

$$\mathbf{v} = \sum_{i=0}^{j-1} \lambda_i \alpha_i.$$

Lema 1 *Sea $\{P_n\}_{n\geqslant 0}$ una sucesión de polinomios ortogonales mónicos, con respecto a $\mathbf{u}$ y $\{\alpha_n\}_{n\geqslant 0}$ la correspondiente base dual, entonces,*

$$\alpha_i = \frac{P_i(x)}{\langle \mathbf{u}, P_i^2(x) \rangle}\mathbf{u},\ para\ i \in \mathbb{N}.$$

Demostración Como $\{\alpha_n\}_{n\geqslant 0}$ es una base dual en $\mathbb{P}^*$ por y $\{P_n\}_{n\geqslant 0}$ es una $SPOM$ con respecto a $\mathbf{u}$ entonces,

$$\langle \alpha_n, P_m(x) \rangle = \delta_{n,m}$$

luego,

$$\begin{aligned}
\langle \frac{P_i(x)}{\langle \mathbf{u}, P_i^2(x)\rangle}\mathbf{u}, P_n(x)\rangle &= \frac{1}{\langle \mathbf{u}, P_i^2(x)\rangle}\langle P_i(x)\mathbf{u}, P_n(x)\rangle \\
&= \frac{1}{\langle \mathbf{u}, P_i^2(x)\rangle}\langle \mathbf{u}, P_i(x)P_n(x)\rangle \\
&= \delta_{i,n}
\end{aligned}$$

Por lo tanto, $\alpha_i = \dfrac{P_i(x)}{\langle \mathbf{u}, P_i^2\rangle}\mathbf{u}$, *para* $i \in \mathbb{N}$. ■

Lema 2 *Si* $\{P_n\}_{n\geqslant 0}$ *y* $\{Q_n\}_{n\geqslant 0}$ *son sucesiones de polinomios mónicos y* α_n, β_n *las correspondientes bases duales en* $\mathbb{P}^*$ *donde* $Q_n = \dfrac{P'_{n+1}}{(n+1)}$, *entonces,*

$$D\beta_n = -(n+1)\alpha_{n+1}.$$

Demostración Como $\{\alpha_n\}_{n\geqslant 0}$ y $\{\beta_n\}_{n\geqslant 0}$ son las correspondientes bases duales en $\mathbb{P}^*$ de las sucesiones $\{P_n\}_{n\geqslant 0}$, $\{Q_n\}_{n\geqslant 0}$, respectivamente, entonces

$$\langle \alpha_n, P_m\rangle = \delta_{n,m} = \langle \beta_n, Q_m\rangle,$$

luego,

$$\begin{aligned}
\langle D\beta_n, P_{m+1}\rangle &= -\langle \beta_n, P'_{m+1}\rangle, \\
&= -(m+1)\langle \beta_n, Q_m\rangle, \\
&= -(m+1)\delta_{n,m}\,,
\end{aligned}$$

en consecuencia, $D\beta_n = -(n+1)\alpha_{n+1}$. ■

Definición 7 Sea $\mathbf{u}$ un funcional regular de momentos y $\{m_n\}_{n\geqslant 0}$ la correspondiente sucesión de momentos. La función Stieljes asociada a $\mathbf{u}$ está dada por

$$S(\mathbf{u})(z) = S(z) = -\sum_{n\geqslant 0}\frac{m_n}{z^{n+1}}.$$

2.2. Polinomios Ortogonales Clásicos, algunas caracterizaciones.

Definición 8 Sea $\mathcal{L} : \mathbb{P} \mapsto \mathbb{C}$ un funcional lineal cuasi-definido y $\{P_n\}_{n\geqslant 0}$ la sucesión de polinomios ortogonales mónicos respecto a $\mathcal{L}$. Diremos que $\mathcal{L}$ es un funcional lineal clásico, si existen polinomios ϕ y ψ con $deg\,\phi \leqslant 2$ y $deg\,\psi = 1$ tales que

$$D(\phi\mathcal{L}) = \psi\mathcal{L}. \tag{2.2}$$

En este caso, se dice que $\{P_n\}_{n\geqslant 0}$ es una sucesión de polinomios ortogonales clásicos respecto a $\mathcal{L}$. La ecuación (2.2) se denomina ecuación distribucional de Pearson.

Teorema 5 ([22]) *Sea $\mathcal{L}$ un funcional lineal cuasidefinido, entonces los siguientes enunciados son equivalentes*

(i) *$\mathcal{L}$ es un funcional lineal clásico.*

(ii) *Los polinomios mónicos $\{P_n\}_{n\geqslant 0}$ ortogonales respecto al funcional $\mathcal{L}$ son autofunciones del operador diferencial de segundo orden $\phi D^2 + \psi D$, esto significa que existe $\lambda_n \neq 0$ tal que*

$$\phi P_n'' + \psi P_n' + \lambda_n P_n = 0 \tag{2.3}$$

Los polinomios $\{P_n\}_n \geqslant 0$ que satisfacen la ecuación (2.3) son denominados polinomios de tipos hipergeométricos, como son los polinomios de Hermite, Laguerre, Jacobi y Bessel.

- Si $deg\,\phi = 0$, tenemos un polinomio mónico de Hermite, $H_n(x)$.

- Si $deg\,\phi = 1$, se obtiene un polinomio de Laguerre de parámetro $\alpha, L_n^{\alpha}(x)$.

- Si $deg\,\phi = 2$, obtenemos dos familias:

  - Si ϕ tiene raíces distintas α, β, tenemos un polinomio de Jacobi de parámetros α *y* β, $P_n^{(\alpha,\beta)}(x)$.

  - Si ϕ tiene raíces iguales a α, tenemos un polinomio mónico de Bessel, de parámetro α, $B_n^{(\alpha)}(x)$.

Teorema 6 *Si $\{P_n\}_{n\geqslant 0}$ es una sucesión de polinomios ortogonales mónicos clásicos, asociada al funcional de momento $\mathbf{u}$, entonces, la correspondiente función de Stieltjes satisface la ecuación diferencial*

$$\phi(x)S'(x) + [\phi'(x) - \psi(x)]S(x) + (a-p)m_0 = 0,$$

donde ϕ y ψ son polinomios dados por (2.2) y a, p, son los coeficientes principales de ϕ y ψ respectivamente.

Demostración Sean $\phi(z) = az^2 + bz + c$ y $\psi(z) = pz + q$,
entonces,

$$\phi(z)S'(z) + [\phi' z) - \psi(z)]S(z) + (a-p)m_0$$

$$= (az^2 + bz + c)\sum_{n\geqslant 0}(n+1)\frac{m_n}{z^{n+2}} - [(2az+b) - (pz+q)]\sum_{n\geqslant 0}\frac{m_n}{z^{n+1}} + (a-p)m_0$$

$$= \sum_{n\geqslant 0}(n+1)(az^2+bz+c)\frac{m_n}{z^{n+2}} - \sum_{n\geqslant 0}[(2a-p)z + b - q]\frac{m_n}{z^{n+1}} + (a-p)m_0$$

$$= \sum_{n\geqslant 0}((n-1)a+p)\frac{m_n}{z^n} + \sum_{n\geqslant 0}(nb+q)\frac{m_n}{z^{n+1}} + \sum_{n\geqslant 0} nc\frac{m_n}{z^{n+1}}$$

$$= \sum_{n\geqslant 0}[(na+p)m_{n+1} + (nb+q)m_n + ncm_{n-1}]\frac{1}{z^{n+1}} = 0$$

■

Proposición 7 *Sea $\{P_n\}_{n\geqslant 0}$ una $SPOM$ asociado al funcional de momentos $\mathbf{u}$. Entonces $\{P_n\}_{n\geqslant 0}$ es clásico si y solo si para cada $n \geqslant 0$ existen parámetro $\lambda_n \in \mathbb{C}$ tal que P_n satisface la ecuación diferencial*

$$A(x)y'' + B(x)y' + C(x)y = \lambda_n y$$

.

Demostración Como $\{P_n\}_{n\geqslant 0}$ es clásico, por corolario (1) tenemos que $\{Q_n\}_{n\geqslant 0}$ es una $SPOM$ asociado a $\mathbf{v} = \phi\mathbf{u}$, por lo tanto,

$$\langle \phi\mathbf{u}, P_n'(x^m)'\rangle = \begin{cases} 0, & \text{si} \quad m < n\ ; \\ l_n \neq 0, & \text{si} \quad m = n. \end{cases} \tag{2.4}$$

Así, tenemos

$$\begin{aligned} \langle \phi\mathbf{u}, P_n'(x^m)'\rangle &= \langle \phi\mathbf{u}, (x^m P_n')' - x^m P_n''\rangle \\ &= -(\langle \psi\mathbf{u}, x^m P_n'\rangle + \langle \mathbf{u}, x^m \phi P_n''\rangle) \\ &= -\langle \mathbf{u}, x^m(\phi P_n'' + \psi P_n')\rangle \end{aligned}$$

por lo tanto, haciendo

$$P_n^* = \phi P_n'' + \psi P_n' \ \ para \ \ n \geqslant 1 \ donde \ \ P_0^* = 1,$$

y comparando con (2.4), concluimos que $\{P_n^*\}_{n\geqslant 0}$ es también una $SPOM$ con respecto a $\mathbf{u}$. Como consecuencia del corolario (2), tenemos que existe $c_n \neq 0$ tal que

$$P_n^* = c_n P_n, \quad n = 0, 1, 2, ...$$

Así, podemos escribir, $\phi P_n'' + \psi P_n' = c_n P_n, \ \ n \geqslant 1,$ es decir, P_n satisface la ecuación diferencial de segundo orden

$$A(x)y'' + B(x)y' + C(x)y = \lambda_n y$$

con $A(x) = \phi(x)$, $B(x) = \psi(x)$, $C(x) = 0$, $\lambda_n = c_n \, (n \geqslant 1)$, $\lambda_0 = 0$. Recíprocamente, supongamos que P_n satisface

$$A(x)y'' + B(x)y' + C(x)y = \lambda_n \ \ para \ \ n \geqslant 0.$$

Sea $\mathbf{v} = D(A\mathbf{u})$, consideremos a $Q_n = \dfrac{P_{n+1}'}{n+1}$ como una base de $\{P_n\}_{n\geqslant 0}$, entonces,

$$\begin{aligned}
\langle \mathbf{v}, Q_n \rangle &= \frac{1}{n+1}\langle \phi(A\mathbf{u}), P_{n+1}' \rangle \\
&= -\frac{1}{n+1}\langle \mathbf{u}, AP_{n+1}'' \rangle \\
&= -\frac{1}{n+1}\langle \mathbf{u}, -BP_{n+1}' + (\lambda_{n+1} - C)P_{n+1} \rangle \\
&= \langle B\mathbf{u}, \frac{P_{n+1}'}{n+1} \rangle - \frac{\lambda_{n+1} - C}{n+1}\langle \mathbf{u}, P_{n+1} \rangle \\
&= \langle B\mathbf{u}, Q_n \rangle, \ \ n \geqslant 0
\end{aligned}$$

Por lo tanto, los funcionales de momentos $\mathbf{v}$ y $B\mathbf{u}$ son iguales para los elementos de la base $\{Q_n\}_{n\geqslant 0}$ de $\mathbb{P}$, por linealidad, $D(A\mathbf{u}) = B\mathbf{u}$, y $deg\, A \leqslant 2$, $deg\, B = 1$, en consecuencia, $\{P_n\}_{n\geqslant 0}$ es clásico. ■

Proposición 8 *Sea $\{P_n\}_{n\geqslant 0}$ es una $SPOM$ con respecto a $\mathbf{u}$. Entonces, $\{P_n\}_{n\geqslant 0}$ es clásico si y solo si existe un polinomio ϕ con deg $\phi \leqslant 2$ y una sucesión de polinomios $\{k_n\} \neq 0$ tal que*

$$k_n^{-1} P_n(x)\mathbf{u} = D^n(\phi^n(x)\mathbf{u}).$$

Demostración Supongamos que se cumple

$$k_n^{-1}P_n(x)\mathbf{u} = D^n(\phi^n(x)\mathbf{u}),$$

entonces, para $n = 1$ se tiene que

$$k_1^{-1}P_1(x)\mathbf{u} = D'(\phi(x)\mathbf{u}),$$

de allí que $\{P_n\}_{n\geqslant 0}$ es una $SPOM$ con respecto a $\mathbf{u}$. Recíprocamente, si $\{P_n\}_{n\geqslant 0}$ es clásico, entonces $D\phi\mathbf{u} = \psi\mathbf{u}$ donde $deg\phi \leqslant 2$ y $deg(\psi) = 1$. Fijando n, probaremos que existe una constante $k_n \neq 0$ tal que el funcional lineal $k_n^{-1}P_n(x)\mathbf{u}$ y $D^n(\phi^n)(x)\mathbf{u}$ son iguales para los elementos de la base $\{P_i\}$ de $\mathbb{P}$. En efecto, si $i \leqslant n$ tenemos

$$\langle D^n(\phi^n\mathbf{u}), P_i\rangle = (-1)^n\langle \phi^n\mathbf{u}, P_i^{(n)}\rangle = (-1)^n n!\langle \mathbf{u}, \phi^n\rangle\delta_{in}.$$

Para $i > n$, podemos escribir $P_i = P_{m+n}, \quad m \geqslant 1$, luego,

$$\begin{aligned}
\langle D^n(\phi^n\mathbf{u}), P_i\rangle &= (-1)^n\langle \phi^n\mathbf{u}, P_i^n\rangle \\
&= (m+1)(m+2)...(m+n)(-1)^n\langle \phi^n\mathbf{u}, \frac{P_{n+m}^{(n)}}{(m+1)(m+2)...(m+n)}\rangle
\end{aligned},$$

del corolario (2), tenemos

$$\begin{aligned}
\langle D^n(\phi^n\mathbf{u}), P_i\rangle &= (m+1)(m+2)...(m+n)(-1)^n\langle \mathbf{v}_n, Q_{m,n}\rangle = 0 \\
&= (-1)^n n!\langle \mathbf{u}, \phi^n\rangle\delta_{i,n}.
\end{aligned}$$

Haciendo

$$k_n^{-1} = \frac{-1^n n!\langle \mathbf{u}, \phi^n\rangle}{\langle \mathbf{u}, P_n^2\rangle}$$

tenemos,

$$\begin{aligned}
\langle k_n^{-1}P_n\mathbf{u}, P_i\rangle &= k_n^{-1}\langle \mathbf{u}, P_nP_i\rangle \\
&= (-1)^n n!\langle \mathbf{u}, \phi^n\rangle\delta_{in}
\end{aligned}$$

Por lo tanto,

$$k_n^{-1}P_n(x)\mathbf{u} = D^n(\phi^n(x)\mathbf{u}).$$

■

2.3. Polinomios de Hermite.

La familia de los polinomios mónicos de Hermite $\{H_n\}_{n\geqslant 0}$ es ortogonal respecto al funcional definido positivo

$$\mathcal{L}_H(p(x)) = \int_{-\infty}^{+\infty} p(x)e^{-x^2}dx.$$

La familia $\{H_n\}_{n\geqslant 0}$ satisface la relación de recurrencia

$$xH_n(x) = H_{n+1}(x) + \frac{n}{2}H_{n-1}(x),\ n \geqslant 0, \tag{2.5}$$

con el convenio $H_{-1}(x) = 0$. El funcional lineal $\mathcal{L}_H$ satisface la ecuación distribucional de Pearson (2.2) con $\phi(x) = 1$ y $\psi(x) = -2x$

2.4. Polinomios de Laguerre.

Los polinomios mónicos de Laguerre $\{L_n^{(\alpha)}\}_{n\geqslant 0}$ con $\alpha > -1$ constituyen una familia uniparamétrica ortogonal respecto al funcional definido positivo

$$\mathcal{L}_L(p(x)) = \int_{0}^{+\infty} p(x)x^\alpha e^{-x}dx.$$

Los polinomios $\phi(x) = x$, $\psi(x) = -x + \alpha + 1$ son los asociados a $\mathcal{L}_L$ en la ecuación distribucional de Pearson (2.2). La sucesión $\{L_n^\alpha\}_{n\geqslant 0}$ satisface la relación de recurrencia

$$xL_n^{(\alpha)}(x) = L_{n+1}^{(\alpha)}(x) + (2n + \alpha + 1)L_n^{(\alpha)}(x) + (n(n + \alpha))L_{n-1}^{(\alpha)}(x),\ n \geqslant 0, \tag{2.6}$$

con el convenio $L_{-1}^{(\alpha)}(x) = 0$.

2.5. Polinomios de Jacobi.

Los polinomios mónicos de Jacobi $\{P_n^{(\alpha,\beta)}\}_{n\geqslant 0}$ constituyen una familia dependiente de dos parámetros $\alpha, \beta > -1$ y son ortogonales respecto al funcional definido positivo

$$\mathcal{L}_P(p(x)) = \int_{-1}^{1} p(x)(1 - x)^\alpha(1 + x)\beta dx.$$

La relación de recurrencia que satisface $\{P_n^{(\alpha,\beta)}\}_{n\geqslant 0}$ es

$$xP_n^{(\alpha,\beta)}(x) = P_{n+1}^{(\alpha,\beta)}(x) + \beta_n P_n^{(\alpha,\beta)}(x) + \gamma_n P_{n-1}^{\alpha,\beta}(x),\ n \geqslant 0, \tag{2.7}$$

con el convenio $P_{-1}^{(\alpha,\beta)}(x) = 0$, donde

$$\begin{aligned} \beta_n &= \frac{\beta^2 - \alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta)+2} \\ \gamma_n &= \frac{4n(n+\alpha)(n+\beta)(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)} \end{aligned}$$

En la ecuación (2.2) los polinomios asociados a $\mathcal{L}_P$ son

$$\phi(x) = 1 - x^2 \ y \ \psi(x) = -(\alpha+\beta+2)x + \beta - \alpha.$$

2.6. Polinomios de Bessel.

Los polinomios mónicos de Bessel $\{B_n^{\alpha}\}_{n\geqslant 0}$ constituyen una familia uniparamétrica con parámetro $\alpha \neq -2, -3, -4, ...$ y son ortogonales respecto al funcional lineal cuasi-definido

$$\mathcal{L}_B(p(x)) = \int_T p(x) x^{\alpha} e^{-\frac{2}{x}} dx$$

donde $T = z \in \mathbb{C} : |z| = 1$.

Los polinomios $\{B_n^{\alpha}\}_{n\geqslant 0}$ satisfacen la relación de recurrencia

$$xB_n^{(\alpha)}(x) = B_{n+1}^{(\alpha)}(x) + \beta_n B_n^{(\alpha)}(x) + \gamma_n B_{n-1}^{\alpha}(x),\ n \geqslant 0, \tag{2.8}$$

con el convenio que $B_{-1}^{\alpha}(x) = 0$, donde

$$\begin{aligned} \beta_n &= -\frac{2\alpha}{(2n+\alpha)(2n+\alpha+2)}, \\ \gamma_n &= -\frac{4n(n+\alpha)}{(2n+\alpha-1)(2n+\alpha)^2(2n+\alpha+1)}. \end{aligned}$$

El funcional lineal $\mathcal{L}_B$ satisface la ecuación distribucional de Pearson (2.2) con $\phi(x) = x^2$ y $\psi(x) = (\alpha+2)x + 2$. La siguiente tabla resume los valores de los coeficientes de la relación de recurrencia (2.1)

	β_n	γ_n
Hermite	0	$\frac{n}{2}$
Laguerre	$2n+\alpha+1$	$n(n+\alpha)$
Jacobi	$\frac{\beta^2-\alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}$	$\frac{4n(n+\alpha)(n+\beta)(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)}$
Bessel	$-\frac{2\alpha}{(2n+\alpha)(2n+\alpha+2)}$	$-\frac{4n(n+\alpha)}{(2n+\alpha-1)(2n+\alpha)^2(2n+\alpha+1)}$

Tabla 2.1: Coeficientes en la relación (2.1)

Otra caracterización de los polinomios ortogonales clásicos, viene dada por

Proposición 9 (ver [22]) *Los siguientes enunciados son equivalentes.*

(i) $\{P_n\}_{n\geqslant 0}$ *es una sucesión de polinomios ortogonales clásicos.*

(ii) La sucesión de polinomios mónicos $\{Q_n\}_{n\geqslant 0}$ *dada por* $Q_n = \dfrac{P'_{n+1}}{n+1}$, *es también una sucesión de polinomios ortogonales.*

(iii) En el desarrollo de P_n *en términos de la base polinómica* $\{Q_n\}_{n\geqslant 0}$, *los coeficientes para* $0 \leqslant k \leqslant n-3$ *se anulan, es decir:*

$$P_n(x) = Q_n(x) + r_n Q_{n-1}(x) + s_n Q_{n-2}(x). \tag{2.9}$$

La familia $\{Q_n\}_{n\geqslant 0}$ *es ortogonal respecto al funcional* $\phi\mathcal{L}$, *donde* ϕ *es el polinomio que aparece en la ecuación de Pearson. Los coeficientes* r_n, s_n *de la relación* (2.9) *para las distintas familias de polinomios ortogonales clásicos están dadas en la siguiente tabla*

	r_n	s_n
Hermite	0	0
Laguerre	n	0
Jacobi	$\frac{2n(\alpha-\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}$	$-\frac{4n(n-1)(n+\alpha)(n+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)^2(2n+\alpha+\beta+1)}$
Bessel	$\frac{4n}{(2n+\alpha)(2n+\alpha+2)}$	$\frac{4n(n-1)}{(2n+\alpha-1)(2n+\alpha)^2(2n+\alpha+1)}$

Tabla 2.2: Coeficientes en la relación (2.9)

2.7. Algoritmo de Leverrier-Faddeev y los polinomios ortogonales clásicos

Consideremos una matriz $A \in \mathbb{C}^{n\times n}$ *y una familia de polinomios ortogonales clásicos mónicos* $\{P_n\}_{n\geqslant 0}$. *Si expresamos el polinomio característico de* A *y la matriz* $\tilde{A}(s) := \mathrm{Adj}\ (sI_n - A)$ *en función de la base de* $\mathbb{P}$ *formada por dichos polinomios se tiene*

$$p_A(s) = P_n(s) + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k(s), \tag{2.10}$$

$$\tilde{A}(s) = P_{n-1}(s) I_n + \sum_{k=0}^{n-2} P_k(s) \widehat{B}_{n-k-1}. \tag{2.11}$$

Deduciremos nuestro algoritmo partiendo de la identidad que establece el siguiente

Lema 3 *Dada* $A \in \mathbb{C}^{n\times n}$, *sea* $p_A(s)$ *su polinomio característico. Entonces*

$$\frac{d}{ds} p_A(s) \;=\; \mathrm{tr}\,\mathrm{Adj}\,\big(sI_n - A\big). \tag{2.12}$$

Demostración:
De la fórmula para la derivada de un determinante (ver **[25]***), tenemos*

$$\begin{aligned} \frac{d}{ds} p_A(s) &= \frac{d}{ds} \det (sI_n - A) \\ &= \sum_{k=1}^{n} \det \mathbf{D}_k, \end{aligned}$$

donde $\mathbf{D}_k,\ k = 1, \dots, n$, *son matrices de dimensión* $n \times n$ *cuyas entradas coinciden con las de la matriz* $sI_n - A$ *excepto en la fila* k, *que se sustituye por la fila* k *de la matriz* I_n. *Entonces*

$$\frac{d}{ds} p_A(s) = \sum_{k=1}^{n} \det(sI_n - A)_{(k|k)}. \tag{2.13}$$

La matriz $(sI_n - A)_{(k|k)}$ *es de dimensión* $(n-1) \times (n-1)$, *y se obtiene eliminando las* k*-ésimas fila y columna de la matriz* $sI_n - A$. *En consecuencia, los términos* $\det(sI_n - A)_{(k|k)},\ k = 1, 2, \dots, n$, *son las entradas de la diagonal de la matriz* $\mathrm{Adj}\,(sI_n - A)$ *y por lo tanto, el segundo miembro de la expresión (2.13) representa la traza de la matriz* $\mathrm{Adj}\,(sI_n - A)$, *de donde se sigue el enunciado.* ■

Notemos que $\tilde{A}(s)$ *satisface*

$$(sI_n - A)\tilde{A}(s) = p_A(s) I_n. \tag{2.14}$$

En consecuencia, deducimos

$$(sI_n - A) \left(P_{n-1}(s) I_n + \sum_{k=0}^{n-2} P_k(s) \widehat{B}_{n-k-1} \right) = P_n(s) I_n + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k(s) I_n. \tag{2.15}$$

Si usamos la relación de recurrencia a tres términos (2.1) correspondiente a la familia ortogonal $\{P_n\}_{n \geqslant 0}$, *en la expresión (2.15) se obtiene*

$$\begin{aligned} &\left[P_n(s) + \beta_{n-1} P_{n-1}(s) + \gamma_{n-1} P_{n-2}(s) \right] I_n - \\ &- P_{n-1}(s) A + \sum_{k=0}^{n-2} \left(P_{k+1}(s) + \beta_k P_k(s) + \gamma_k P_{k-1}(s) \right) \widehat{B}_{n-k-1} - \\ &- \sum_{k=0}^{n-2} P_k(s) A \widehat{B}_{n-k-1} = P_n(s) I + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k(s) I_n. \end{aligned}$$

Si reordenamos la expresión anterior, obtenemos

$$P_n(s)I_n + P_{n-1}(s)\Big(\beta_{n-1}I_n + \widehat{B}_1 - A\Big) + P_{n-2}(s)\Big(\gamma_{n-2}I_n + \widehat{B}_2 + \beta_{n-2}\widehat{B}_1 - A\widehat{B}_1\Big)+$$
$$+\sum_{k=1}^{n-3} P_k(s)\Big(\widehat{B}_{n-k} + \beta_k\widehat{B}_{n-k-1} + \gamma_{k+1}\widehat{B}_{n-k-2} - A\widehat{B}_{n-k-1}\Big)+$$
$$+P_0(s)\Big(\beta_0\widehat{B}_{n-1} + \gamma_1\widehat{B}_{n-2} - A\widehat{B}_{n-1}\Big) \;=\; P_n(s)I + \sum_{k=0}^{n-1}\widehat{a}_{n-k}P_k(s)I_n.$$

Igualando los coeficientes de P_k, $k = 0, 1, \dots, n-1$, *en los dos miembros de la anterior expresión se tiene*

$$\begin{array}{rcl} A\widehat{B}_0 & = & -\widehat{a}_1 I_n + \beta_{n-1}\widehat{B}_0 + \widehat{B}_1, \\ A\widehat{B}_1 & = & -\widehat{a}_2 I_n + \gamma_{n-1}\widehat{B}_0 + \beta_{n-2}\widehat{B}_1 + \widehat{B}_2, \\ \vdots & & \\ A\widehat{B}_{n-k-1} & = & -\widehat{a}_{n-k} I + \gamma_{k+1}\widehat{B}_{n-k-2} + \beta_k\widehat{B}_{n-k-1} + \widehat{B}_{n-k}, \\ & & k = 1, 2, \dots, n-3, \\ & & \\ A\widehat{B}_{n-1} & = & -\widehat{a}_n I_n + \gamma_1\widehat{B}_{n-2} + \beta_0\widehat{B}_{n-1}, \end{array} \tag{2.16}$$

con $\widehat{B}_0 = I_n$.

En forma matricial

$$A\begin{bmatrix} \widehat{B}_{n-1} \\ \vdots \\ \widehat{B}_0 \end{bmatrix} = M\begin{bmatrix} \widehat{B}_{n-1} \\ \vdots \\ \widehat{B}_0 \end{bmatrix} \tag{2.17}$$

donde $M = J_n - [0|\widehat{a}]$. J_n *es la matriz tridiagonal de dimensión* $n \times n$, $n \geqslant 1$, *asociada a la familia* $\{P_n\}_{n\geqslant 0}$, *i. e.*

$$J_n = \begin{bmatrix} \beta_0 & \gamma_1 & 0 & \cdots & 0 \\ 1 & \beta_1 & \gamma_2 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & & & \gamma_{n-1} \\ 0 & \cdots & 0 & 1 & \beta_{n-1} \end{bmatrix} \quad y \quad \widehat{a} = \begin{bmatrix} \widehat{a}_n \\ \widehat{a}_{n-1} \\ \vdots \\ \widehat{a}_1 \end{bmatrix}. \tag{2.18}$$

*La matriz M se denomina matriz compañera (comrade matrix) de A respecto a la base ortogonal $\{P_n\}_{n\geqslant 0}$ (véase [**1**] pp. 372). Además*

$$\operatorname{tr} A = -\sum_{j=0}^{n-1} \beta_j - \widehat{a}_1.$$

Por otra parte, de la relación (2.12) tenemos

$$P_n'(s) + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k'(s) = nP_{n-1}(s) + \sum_{k=0}^{n-2} P_k(s) \operatorname{tr} \widehat{B}_{n-k-1}, \quad n = 2, 3, \dots . \tag{2.19}$$

Debido a que $\{P_n\}_{n\geqslant 0}$ es una familia de polinomios ortogonales clásicos, entonces de la Proposición 9, obtenemos

$$P_k(s) = \frac{P_{k+1}'(s)}{k+1} + r_k \frac{P_k'(s)}{k} + s_k \frac{P_{k-1}'(s)}{k-1}, \qquad k = 2, 3, \dots$$

De esta forma, sustituyendo la expresión anterior en (2.19) se obtiene

$$\begin{aligned} P_n'(s) + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k'(s) \quad &= \quad P_n'(s) + r_{n-1} \frac{n}{n-1} P_{n-1}'(s) + s_{n-1} \frac{n}{n-2} P_{n-2}'(s) + \\ &\quad + \sum_{k=2}^{n-2} \left(\frac{P_{k+1}'(s)}{k+1} + r_k \frac{P_k'(s)}{k} + s_k \frac{P_{k-1}'(s)}{k-1} \right) \operatorname{tr} \widehat{B}_{n-k-1} + \\ &\quad + \operatorname{tr} \widehat{B}_{n-1} P_1'(s) + \operatorname{tr} \widehat{B}_{n-2} \left(\frac{P_2'(s)}{2} + r_1 P_1'(s) \right), \end{aligned} \tag{2.20}$$

o, equivalentemente,

$$\begin{aligned} P_n'(s) + \sum_{k=0}^{n-1} \widehat{a}_{n-k} P_k'(s) \quad &= \quad P_n'(s) + \left(\tfrac{n}{n-1} r_{n-1} + \tfrac{1}{n-1} \operatorname{tr} \widehat{B}_1 \right) P_{n-1}'(s) + \\ &\quad + \sum_{k=1}^{n-2} \frac{1}{k} \left(\operatorname{tr} \widehat{B}_{n-k} + r_k \operatorname{tr} \widehat{B}_{n-k-1} + s_{k+1} \operatorname{tr} \widehat{B}_{n-k-2} \right) P_k'(s). \end{aligned} \tag{2.21}$$

Si identificamos los coeficientes de P_k', $k = 1, 2, \dots, n-1$, en ambos miembros de la expresión anterior obtenemos

$$\begin{aligned} (n-1)\widehat{a}_1 \quad &= \quad n r_{n-1} + \operatorname{tr} \widehat{B}_1, \\ (n-k)\widehat{a}_k \quad &= \quad \operatorname{tr} \widehat{B}_k + r_{n-k} \operatorname{tr} \widehat{B}_{k-1} + s_{n-k+1} \operatorname{tr} \widehat{B}_{k-2}, \\ &\quad\quad k = 2, 3, \dots, n-1. \end{aligned} \tag{2.22}$$

Finalmente, combinando (2.16) y (2.22), se obtiene

$$\widehat{a}_k = \frac{1}{k}\left[(\beta_{n-k} - r_{n-k})\mathrm{tr}\,\widehat{B}_{k-1} + (\gamma_{n-k+1} - s_{n-k+1})\mathrm{tr}\,\widehat{B}_{k-2} - \mathrm{tr}\,\left(A\widehat{B}_{k-1}\right)\right], \quad (2.23)$$

para $k = 1, 2, \ldots, n$. *Como conclusión tenemos el siguiente*

Teorema 10 *Dada* $A \in \mathbb{C}^{n\times n}$*, sean* $p_A(s)$ *su polinomio característico dado por (2.10) y* $\tilde{A}(s)$ *la matriz adjunta de* $sI_n - A$ *dada por (2.11). Entonces*

(i) *Para* $k = 1, \ldots, n$,

$$\widehat{a}_k = \frac{1}{k}\left[(\beta_{n-k} - r_{n-k})\mathrm{tr}\,\widehat{B}_{k-1} + (\gamma_{n-k+1} - s_{n-k+1})\mathrm{tr}\,\widehat{B}_{k-2} - \mathrm{tr}\,\left(A\widehat{B}_{k-1}\right)\right], \quad (2.24)$$

con $\widehat{B}_{-1} = 0,\ r_0 = 0,\ s_1 = 0.$

(ii) *Para* $k = 1, 2, \ldots, n-1$

$$\widehat{B}_k = \widehat{a}_k I - \gamma_{n-k+1}\widehat{B}_{k-2} - \beta_{n-k}\widehat{B}_{k-1} + A\widehat{B}_{k-1}. \quad (2.25)$$

Así, el algoritmo queda de la forma mostrada en la tabla (2.3) (para los detalles, ver [14])

DATOS DE ENTRADA: $A \in \mathbb{C}^{n\times n}$, $\{\beta_k\}_{k=0}^{n-1},\ \{\gamma_k\}_{k=1}^{n},\ \{r_k\}_{k=0}^{n-1},\ \{s_k\}_{k=1}^{n}$.
Condiciones iniciales: $\widehat{B}_{-1} = 0,\ \widehat{B}_0 = I_n$.
FOR $k = 1, 2, \ldots, n-1$

$$\widehat{a}_k = \tfrac{1}{k}\left[(\beta_{n-k} - r_{n-k})\mathrm{tr}\,\widehat{B}_{k-1} + (\gamma_{n-k+1} - s_{n-k+1})\mathrm{tr}\,\widehat{B}_{k-2} - \mathrm{tr}\,\left(A\widehat{B}_{k-1}\right)\right],$$
$$\widehat{B}_k = \widehat{a}_k I - \beta_{n-k}\widehat{B}_{k-1} - \gamma_{n-k+1}\widehat{B}_{k-2} + A\widehat{B}_{k-1}.$$

END (FOR)

$$\widehat{a}_n = \frac{1}{n}\left[\beta_0\mathrm{tr}\,\widehat{B}_{n-1} + \gamma_1\mathrm{tr}\,\widehat{B}_{n-2} - \mathrm{tr}\,\left(A\widehat{B}_{n-1}\right)\right].$$

Tabla 2.3: Algoritmo de Leverrier-Faddeev usando $SPOM$ clásicos

Las expresiones (2.24) y (2.25) pueden simplificarse si las aplicamos a ciertas familias de polinomios clásicos.

Polinomios de Hermite. *Por el Teorema 10 tenemos*

$$\widehat{a}_k = \frac{n-k+1}{2k}\mathrm{tr}\,\widehat{B}_k - 2 - \frac{1}{k}\mathrm{tr}\,\left(AB_{k-1}\right) \tag{2.26}$$

y

$$\widehat{B}_k = a_k I_n - \frac{n-k+1}{2}\widehat{B}_{k-2} + AB_{k-1}. \tag{2.27}$$

Si consideramos la traza en (2.27) y la usamos en (2.26) obtenemos

$$\mathrm{tr}\,\widehat{B}_k = (n-k)a_k.$$

Por otra parte, si sustituimos la expresión anterior en (2.26), entonces

$$\widehat{a}_k = \frac{(n-k+1)(n-k+2)}{2k}\widehat{a}_{k-2} - \frac{1}{k}\mathrm{tr}\,\left(AB_{k-1}\right). \tag{2.28}$$

Polinomios de Laguerre. *De acuerdo con el Teorema 10*

$$k\widehat{a}_k = (n-k+\alpha+1)\mathrm{tr}\,\widehat{B}_{k-1} + (n-k+1)(n-k+\alpha+1)\widehat{B}_{k-2} - \mathrm{tr}\,\left(A\widehat{B}_{k-1}\right) \tag{2.29}$$

y

$$\widehat{B}_k = \widehat{a}_k I_n - (n-k+1)(n-k+\alpha+1)\widehat{B}_{k-2} - (2(n-k)+\alpha+1)\widehat{B}_{k-1} + AB_{k-1}. \tag{2.30}$$

Aplicando trazas en (2.30) y usando (2.29) se tiene

$$\mathrm{tr}\,\widehat{B}_k = (n-k)\widehat{a}_k - n - k\mathrm{tr}\,\widehat{B}_{k-1}.$$

En consecuencia, se deduce

$$\widehat{a}_k = \frac{(n-k+1)(n-k+\alpha+1)}{k}\widehat{a}_{k-1} - \frac{1}{k}\mathrm{tr}\,\left(A\widehat{B}_{k-1}\right). \tag{2.31}$$

Polinomios de Jacobi. *Del Teorema 10 se deduce*

$$k\widehat{a}_k = \frac{\beta-\alpha}{2(n-k)+\alpha+\beta+2}\mathrm{tr}\,\widehat{B}_{k-1} + \frac{4(n-k+1)(n-k+1+\alpha)(n-k+1+\beta)}{(2(n-k)+\alpha+\beta+2)^2(2(n-k)+\alpha+\beta+3)}\mathrm{tr}\,\widehat{B}_{k-2} - \mathrm{tr}\,\left(A\widehat{B}_{k-1}\right). \tag{2.32}$$

Si $\alpha=\beta$, entonces aparece la familia de los polinomios de Gegenbauer. En este caso el funcional lineal es simétrico y, en consecuencia, la expresión anterior se reduce a

$$\widehat{a}_k = \frac{n-k+1}{k(2(n-k)+2\alpha+3)}\mathrm{tr}\,\widehat{B}_{k-2} - \frac{1}{k}\mathrm{tr}\,\left(AB_{k-1}\right). \tag{2.33}$$

La expresión (2.25) queda simplificada de la siguiente manera

$$\widehat{B}_k = \widehat{a}_k I_n - \frac{(n-k+1)(n-k+2\alpha+1)}{(2(n-k)+2\alpha+1)(2(n-k)+\alpha+3)}\widehat{B}_{k-1}. \tag{2.34}$$

2.8. Ejemplo

Consideremos la matriz

$$A = \begin{bmatrix} 1 & -4 & -1 & -4 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 6 \end{bmatrix}.$$

Su polinomio característico es

$$p_A(s) = s^4 - 5s^3 + 9s^2 - 7s + 2.$$

Tomemos como base la familia de los polinomios de Hermite $\{H_n\}_{n\geqslant 0}$. *Los parámetros que definen esta familia son:* $\beta_n = r_n = s_n = 0$, $n \geqslant 0$, *y* $\gamma_n = \dfrac{n}{2}$, $n \geqslant 1$. *Teniendo en cuenta las expresiones (2.28) y (2.27)*

$$a_1 = -\mathrm{tr}\, A = -5,$$

así como

$$\widehat{B}_1 = a_1 I_4 + A = \begin{bmatrix} -4 & -4 & -1 & -4 \\ 2 & -5 & 5 & -4 \\ -1 & 1 & -7 & 3 \\ -1 & 4 & -1 & 1 \end{bmatrix}.$$

Por otra parte,

$$a_2 = 3 - \frac{1}{2}\mathrm{tr}\ (AB_1) = 12,$$

junto a

$$\widehat{B}_2 = a_2 I_4 - \frac{3}{2}B_0 + AB_1 = \begin{bmatrix} \frac{7}{2} & -1 & -10 & 5 \\ -9 & -\frac{17}{2} & -33 & 3 \\ 5 & 9 & \frac{55}{2} & -3 \\ 7 & 7 & 22 & \frac{3}{2} \end{bmatrix}.$$

Continuando el proceso

$$a_3 = a_1 - \frac{1}{3}\mathrm{tr}\ (AB_2) = -\frac{29}{2},$$

y

$$\widehat{B}_3 = a_3 I_4 - B_1 + AB_2 = \frac{1}{2}\begin{bmatrix} -8 & 0 & 15 & -12 \\ 4 & 11 & 49 & -14 \\ -1 & -11 & -39 & 11 \\ -3 & -8 & -33 & 7 \end{bmatrix}.$$

Finalmente,

$$a_4 = \frac{1}{4}a_2 - \frac{1}{4}\mathrm{tr}\ (AB_3) = \frac{29}{4}.$$

En consecuencia, el polinomio característico de la matriz A, expresado como combinación lineal de la base $\{H_n\}_{n\geqslant 0}$, resulta ser

$$a(s) = H_4(s) - 5H_3(s) + 12H_2(s) - \frac{29}{2}H_1(s) + \frac{29}{4}H_0(s),$$

y la matriz adjunta de $sI_4 - A$,

$$\mathrm{Adj}\ (sI_4 - A) = H_3(s)I_4 + H_2(s)\widehat{B}_1 + H_1(s)\widehat{B}_2 + H_0(s)\widehat{B}_3.$$

Si consideramos la familia $\left\{L_n^{(\alpha)}\right\}_{n\geqslant 0}$ de los polinomios de Laguerre con parámetro α se tiene $\beta_n = 2n + \alpha + 1$, $r_n = n$, $n \geqslant 0$, $\gamma_n = n(n+\alpha)$, $n \geqslant 1$, y $s_n = 0$, $n \geqslant 1$. De (2.31) y (2.30) deducimos $a_1 = 4(4+\alpha) - \mathrm{tr}\, A = 4\alpha + 11$, y

$$B_1 = \begin{bmatrix} 5+3\alpha & -4 & -1 & -4 \\ 2 & 4+3\alpha & 5 & -4 \\ -1 & 1 & 2+3\alpha & 3 \\ -1 & 4 & -1 & 10+3\alpha \end{bmatrix}.$$

Por otra parte,

$$a_2 = 6\alpha^2 + 27\alpha + 36,$$

junto con

$$B_2 = \begin{bmatrix} 3\alpha^2+11\alpha+4 & -8\alpha-17 & -2\alpha-14 & -8\alpha-11 \\ 4\alpha-1 & 3\alpha^2+5\alpha-12 & 10\alpha-13 & -8\alpha-13 \\ -2\alpha+1 & 2\alpha+13 & 3\alpha^2+\alpha+16 & 6\alpha+9 \\ -2\alpha+3 & 8\alpha+23 & 2\alpha+18 & 3\alpha^2+17\alpha+22 \end{bmatrix}.$$

Continuando el proceso

$$a_3 = 4\alpha^3 + 21\alpha^2 + 47\alpha + 35,$$

$$B_3 = \begin{bmatrix} 4\alpha^3+21\alpha^2+47\alpha+35 & -4\alpha^2-13\alpha-7 & -\alpha^2-13\alpha-4 & -4\alpha^2-7\alpha-7 \\ 2\alpha^2-3\alpha-4 & 4\alpha^3+21\alpha^2+47\alpha+35 & 5\alpha^2-18\alpha-1 & -4\alpha^2-9\alpha-10 \\ -\alpha^2+2\alpha+3 & \alpha^2+12\alpha+5 & 4\alpha^3+21\alpha^2+47\alpha+35 & 3\alpha^2+6\alpha+7 \\ -\alpha^2+4\alpha+4 & 4\alpha^2+19\alpha+9 & -\alpha^2+19\alpha+4 & 4\alpha^3+21\alpha^2+47\alpha+35 \end{bmatrix}.$$

Finalmente,

$$a_4 = \alpha^4 + 5\alpha^3 + 14\alpha^2 + 15\alpha + 7$$

Por tanto, la representación del polinomio característico de A en la base $\left\{L_n^{(\alpha)}\right\}_{n\geqslant 0}$ es

$$\begin{aligned} p_A(s) = \; & L_4^{(\alpha)}(s) - (4\alpha+11)\, L_3^{(\alpha)}(s) + (6\alpha^2+27\alpha+36)\, L_2^{(\alpha)}(s) \\ & - (4\alpha^3+21\alpha^2+47\alpha+35)\, L_1^{(\alpha)}(s) + (\alpha^4+5\alpha^3+14\alpha^2+15\alpha+7)\, L_0^{(\alpha)}(s), \end{aligned}$$

$$\mathrm{Adj}\,(sI_4 - A) = L_3^{(\alpha)}(s)I_4 + L_2^{(\alpha)}(s)B_1 + L_1^{(\alpha)}(s)B_2 + L_0^{(\alpha)}(s)B_3.$$

Capítulo 3

Polinomio Mínimo

Sea $A \in \mathbb{C}^{n\times n}$, dado un polinomio $p(t) \in \mathbb{P}[t]$, digamos

$$p(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0,$$

podemos definir

$$p(A) \equiv A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I_n.$$

Existe una relación muy importante entre polinomios y las matrices, una de ellas es el polinomio característico. Dicha relación viene dada por el Teorema de Cayley-Hamilton.

Teorema 11 (Teorema de Cayley-Hamilton [15]) *Sea $p_A(t)$ el polinomio característico de una matriz $A \in \mathbb{C}^{n\times n}$. Entonces, $p_A(A) = 0$*

Una idea natural es ver si existe otro polinomio de grado menor o igual al polinomio característico que también aniquile a la matriz A, o en otras palabras; dada una matriz $A \in M_n$, existirá un polinomio de grado mínimo que aniquila a A?. Por el Teorema de Cayley-Hamilton, el grado de ese polinomio debe ser menor o igual a n.

Teorema 12 *(ver[15]) Existe un único polinomio mónico $q_A(t)$ de grado mínimo que aniquila a A. El grado de este polinomio, es a lo mas n. Si $p(t)$ es otro polinomio tal que $p(A) = 0$, entonces, $q_A(t)$ divide a $p(t)$.*

Demostración El polinomio característico es un ejemplo de un polinomio de grado $n \neq 0$ que aniquila A, así, existe un entero positivo $m \leqslant n$ y un polinomio mónico $q(t)$ de grado m tal que $q(A) = 0$. Si $p(t)$ aniquila a A, y si $q(t)$ es un polinomio mónico de grado mínimo que aniquila a A, entonces, el grado de $q(t)$ debe ser menor o igual que el grado de $p(t)$. Por el algoritmo Euclidiano, existe un polinomio $h(t)$ y un polinomio $r(t)$ de grado menor que el de $q(t)$ tal que $p(t) = q(t)h(t) + r(t)$. Pero

$$p(A) = q(A)h(A) + r(A) = 0.h(A) + r(A) = 0,$$

de allí, $r(A) = 0$. Si $r(t) \neq 0$, podemos normalizar a éste y obtener un polinomio mónico de grado menor que el grado de $q(t)$ que aniquila a A. Ya que esto contradice la propiedad minimal de $q(t)$, concluimos que $r(t) \equiv 0$ y por tanto $q(t)$ divide a $p(t)$ con cociente h(t). Si existen dos polinomios mónicos de mínimo grado que aniquilan a A, cada uno divide al otro, por lo tanto, sus grados son los mismos y uno debe ser múltiplo escalar del otro, pero como son mónicos, la única posibilidad del escalar es 1, por lo tanto, ambos polinomios son los mismos. ∎

Definición 9 Sea $A \in \mathbb{C}^{n\times n}$ la matriz dada, el único polinomio mónico $q_A(t)$ de menor grado que aniquila a A se llama polinomio mínimo de A.

Corolario 2 *Matrices similares tienen el mismo polinomio mínimo.*

Demostración Si $A, B, S \in \mathbb{C}^{n\times n}$ y si $A = SBS^{-1}$, entonces

$$q_B(A) = q_B(SBS^{-1}) = Sq_B(B)S^{-1} = 0$$

así, el grado de $q_B(t)$ no es menor que el grado de $q_A(t)$. Pero $B = S^{-1}AS$, entonces, el mismo argumento muestra que el grado de q_A no es menor que el grado de $q_B(t)$. En consecuencia, los dos polinomios mónicos tienen el mismo grado minimal y ambos aniquilan a A y por el Teorema (12), los polinomios son los mismos. ∎

Corolario 3 *El polinomio mínimo $q_A(t)$ divide al polinomio característico $p_A(t)$, además, $q_A(\lambda) = 0$ si y solo si λ es un autovalor de A. En consecuencia, cualquier raíz de $p_A(t)$ es una raíz de $q_A(t)$.*

Demostración Debido a que $p_A(A) = 0$ y del hecho de que existe un polinomio $h(t)$ tal que $p_A(t) = h(t)q_A(t)$, por el Teorema 12, se tiene que $q_A(t)$ divide a $p_A(t)$, además, la factorización obtenida muestra que cualquier raíz de $q_A(t) = 0$ es una raíz de $p_A(t) = 0$, es decir, cualquier raíz de $q_A(t) = 0$ es un autovalor de A. Si λ es un autovalor de A y si x es su autovector asociado, entonces $Ax = \lambda x$ y $0 = q_A(A)x = q_A(\lambda)x$, por lo tanto $q_A(\lambda) = 0$. ∎

Este último corolario muestra que

- Si $P_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{s_i}$, $1 \leqslant s_i \leqslant n$, $s_1+s_1+\cdots+s_m = n$ con $\lambda_1, \lambda_2, \cdots, \lambda_m$ distintos, entonces

$$q_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{r_i}, \quad 1 \leqslant r_i \leqslant s_i.$$

- Cuando $P_A(t)$ tiene raices simples, $P_A(t) = q_A(t)$.

En principio esto da un algoritmo para hallar el polinomio mínimo de una matriz A.

(1) Calcular los m autovalores de A (λ_i) y sus multiplicidades $(1 \leqslant s_i \leqslant n)$.

(2) Hacer $P_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{s_i}$.

(3) Comenzar con $r_i = 1$ y evaluar $P_A(A)$ hasta que $P_A(A) = 0$.

Numéricamente, esto no es un buen algoritmo, si se trata de factorizaciones para polinomios característicos de una matriz muy grande, aunque si puede ser muy efectivo para cálculos manuales con matrices pequeñas. Otro enfoque para calcular el polinomio mínimo se presenta a través de la relación íntima entre la forma canónica de Jordan de una matriz y su polinomio característico.

3.1. Forma canónica de Jordan y Polinomio mínimo

Definición 10 Un bloque de Jordan $J_k(\lambda)$ es una matriz triangular superior de orden k de la forma

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}. \tag{3.1}$$

Esta matriz tiene $(k-1)$ términos $+1$ en la diagonal superior, el escalar λ aparece k veces en la diagonal superior, el resto de las entradas son 0 y $J_1(\lambda) = [\lambda]$. Una matriz de Jordan, $J \in M_n$ es una suma directa de bloques de Jordan,

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_k}(\lambda_k) \end{bmatrix}, \quad n_1 + n_2 + \cdots + n_k = n, \tag{3.2}$$

para el cual los n_i y los valores λ_i no necesariamente son distintos. Nótese, que si cada bloque de Jordan $J_{n_i}(\lambda_i)$ en (3.2) es unidimensional, esto es, todo $n_i = 1$ y $k = n$, entonces, la matriz de Jordan J es diagonal. Si algún bloque de Jordan $J_m(\lambda)$ en (3.2) tiene $m > 1$, entonces J no solo no es diagonal, sino que no es diagonalizable. Si $J_m(\lambda) = S\Lambda S^{-1}$ con Λ diagonal, entonces, necesariamente

$$\Lambda = diag(\lambda, \lambda, \cdots, \lambda) = \lambda I.$$

Así,

$$J_m(\lambda) - \lambda I = S\Lambda S^{-1} - \lambda I = \lambda I - \lambda I = 0,$$

el cual no se cumple en el caso que $m > 1$. Por otro lado, existe un autovector de J asociado con cada bloque separado de Jordan; es decir, existe una base vectorial estándar asociada con la primera entrada diagonal de cada $J_m(\lambda)$ en J. La matriz de Jordan (3.2)tiene una definida estructura que hace aparente ciertas propiedades básicas de la matriz y algunas matrices similares a ella, como por ejemplo las siguientes:

- El número k de bloques de Jordan es el número de autovectores linealmente independiente de j.
- La matriz J es diagonalizable si y solo si $k = n$, es decir,si $n_i = 1$ y $k = n$, entonces la matriz de Jordan J es diagonal y no necesariamente lo es para $m > 1$.
- El número de bloques de Jordan correspondientes a los autovalores es la multiplicidad geométrica de dichos autovalores, el cual corresponde a la dimensión del autoespacio asociado.
- El tamaño de todos los bloques de Jordan en la matriz general de Jordan (3.2) son determinados por el conocimiento del rango de ciertas potencias. Si λ_1

Teorema 13 *(ver [15]) Sea $A \in \mathbb{C}^{n\times n}$ una matriz cuyos autovalores distintos son $\lambda_1, \lambda_2, \cdots, \lambda_m$. El polinomio mínimo de A es*

$$q_A(t) = \prod_{i=1}^{m} (t - \lambda_i)^{r_i}, \tag{3.3}$$

donde r_i es el orden del bloque de Jordan mas grande de A correspondientes al autovalor λ_i.

Demostración (ver[15]) ■

En la práctica, este método de calcular el polinomio mínimo, tampoco es de mucha ayuda, ya que por lo general, es mas difícil determinar la forma canónica de Jordan de una matriz que su polinomio mínimo. Después de todo , si unicamente se conocen los autovalores de una matriz, su polinomio mínimo puede ser determinado por simples cálculos de ensayo y error. Existen, sin embargo consecuencias teóricas importantes.

El siguiente lema muestra una fórmula para calcular la inversa de una matriz triangular superior, la cual necesitaremos para utilizar un resultado posterior.

Lema 4 *Dada una matriz A triangular inferior, con unos en su diagonal principal y de orden k*

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ a_{10} & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \\ a_{k-1,0} & \dots & a_{k-1,k-2} & 1 \end{bmatrix},$$

la matriz inversa de A es

$$C = \begin{bmatrix} 1 & 0 & \dots & 0 \\ c_{10} & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \\ c_{k-1,0} & \dots & c_{k-1,k-2} & 1 \end{bmatrix},$$

donde c_{ij} se obtienen de forma recursiva mediante la fórmula

$$\begin{aligned} c_{i,i-1} &= -a_{i,i-1}, \quad {\scriptstyle i=1,2,\ldots,k} \\ c_{i,j} &= -a_{i,j} - \sum_{\ell=j+1}^{i-1} a_{i,\ell} c_{\ell,j}; \quad i = 2, \ldots, k, \;\; j = 0, 1, \ldots, i-2 \end{aligned} \tag{3.4}$$

y $c_{i,i} = 1$; $i = 1, \ldots, k$.

Demostración Para demostrar este lema, probaremos que $AC = I_k$, donde I_k es la matriz identidad de orden K. Sea $B = AC$, entonces, B es también una matriz triangular inferior con unos en su diagonal principal, es decir, B es de la forma

$$B = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ b_{10} & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ b_{k-1,1} & \ldots & b_{k-1,k-2} & 1 \end{bmatrix}.$$

Observemos que $b_{i,i} = a_{i,i}c_{i,i} = 1$ para todo $i = 1, \ldots, k$. entonces bastará con probar que $b_{i,j} = 0$ para $i > j$ donde $i = 1, \ldots, k-1$ y $j = i-1, \ldots, k-2$. En efecto, calcularemos en primer lugar los términos $c_{i,j}$ donde $(i > j)$ que resultan de aplicar la fórmula (3.4),

$$\begin{aligned} c_{10} &= -a_{10} \\ c_{20} &= -a_{20} - a_{21}c_{10} \\ &\vdots \\ c_{(i,0)} &= -a_{i,0} - a_{i,1}c_{10} - a_{i,2}c_{20} - \ldots - a_{i,i-1}c_{i-1,0} \end{aligned}$$

$$\begin{aligned} c_{21} &= -a_{21} \\ c_{31} &= -a_{31} - a_{32}c_{21} \\ &\vdots \\ c_{i,1} &= -a_{i,1} - a_{i,2}c_{21} - \cdots - a_{i,i-1}c_{i-1,1} \end{aligned}$$

para $i = 2, \ldots, k-1$ así,

$$c_{i,j} = -a_{i,j} - a_{i,j+1}c_{j+1,j} - a_{i,j+2}c_{j+2,j} - \cdots - a_{i,i-1}c_{i-1,j}$$

para $i = 2, \ldots, k$ y $j = 0, \ldots, i-2$. Multiplicando las matrices A y C, obtenemos para la primera columna $b_{i,0}$, donde $i = 1, \ldots, k-1$,

$$\begin{aligned} b_{10} &= a_{10} + -a_{10} = 0 \\ b_{20} &= a_{20} + a_{21}(-a_{10}) + (-a_{20} - a_{21}c_{10}) \\ &= a_{20} - a_{21}a_{10} - a_{20} + a_{21}a_{10} = 0 \\ &\vdots \\ b_{i,0} &= (\ a_{i,0} + a_{i,1}c_{10} + \ldots + a_{i,i-1}c_{i-1,0}) + c_{i,0} \\ &= (a_{i,0} + a_{i,1}c_{10} + \ldots + a_{i,i-1}c_{i-1,0}) + (-a_{i,0} - a_{i,1}c_{10} - a_{i,2}c_{20} - \ldots \\ &= -a_{i,i-1}c_{i-1,0}) \\ &= 0 \end{aligned}$$

Para la segunda columna $b_{(i,1)}$ con , $i = 2, \ldots, k-1$ tenemos

$$\begin{aligned}
b_{21} &= a_{21} + (-a_{21}) = 0 \\
b_{31} &= a_{31} + a_{32}(-a_{21}) + (-a_{31} - a_{32}(-a_2 1)) \\
&= a_{31} - a_{32}a_{21} - a_{31} + a_{32}a_{21} = 0 \\
&\vdots \\
b_{i,1} &= a_{i,1} + a_{i,2}c_{21} + \cdots + a_{i,i-1}c_{i-1,1} + c_{i,1} \\
&= a_{i,1} + a_{i,2}c_{21} + \cdots + a_{i,i-1}c_{i-1,1} + (-a_{i,1} - a_{i,2}c_{21} - \cdots - a_{i,i-1}c_{i-1,1}) \\
&= 0
\end{aligned}$$

y para cualquier término $b_{i,j}$, donde $i = j+1, \ldots, k-1; (i > j)$ obtenemos

$$\begin{aligned}
b_{i,j} &= a_{i,j} + a_{i,j+1}c_{j+1,j} + \cdots + a_{i,i-1}c_{i-1,j} + c_{i,j} \\
&= (a_{i,j} + a_{i,j+1}c_{j+1,j} + \cdots + a_{i,i-1}c_{i-1,j}) + (-a_{i,j} \\
&\quad -a_{i,j+1}c_{j+1,j} - \cdots - a_{i,i-1}c_{i-1,j}) \\
&= 0
\end{aligned}$$

Así, $B = AC = I_k$, por lo tanto, C es la matriz inversa de A, donde los términos $c_{i,j}$, se obtienen aplicando la fórmula (3.4). ■

A continuación, el siguiente teorema nos dará la idea de calcular el polinomio mínimo de una matriz, pero sin necesidad de conocer sus autovalores y mediante el aprovechamiento del algoritmo de Leverrier-Faddeev.

Teorema 14 *Dada una matriz $A \in \mathbb{M}_n$ cuyas columnas están dadas por*

$$A_1 = [a_{11}, a_{21}, ..., \cdots, a_{n1}]^t, \cdots, A_n = [a_{1n}, a_{2n}..., \cdots, a_{nn}]^t,$$

la transformación $T : \mathbb{C}^{n\times n} \mapsto \mathbb{C}^{n^2}$, definida por:

$$T(A) = (a_{11}, a_{21}, \cdots, a_{n1}, a_{12}, a_{22}, \cdots, a_{n2}, a_{1n}, a_{2n}, \cdots, a_{nn})$$

satisface lo siguiente.

(i) *T es un isomorfismo.*

(ii) *Los vectores definidos por,*

$$v_0 = T(I), v_1 = T(A), v_2 = T(A^2), ..., v_n = T(A^n),$$

forman un conjunto $\{v_0, v_1, .., v_n\}$ linealmente dependiente.

(iii) *Sean $w_0, w_1, ..., w_k$ los vectores ortonormales construidos en el proceso de Gram-Schmidt, con los v_i y tal que w_k es el primer vector para el cual*

$$w_k = \zeta_0 v_0 + \zeta_1 v_1 + \cdots + \zeta_k v_k = 0$$

entonces ,

$$T^{-1}(w_k) = \zeta_0 I + \zeta_1 A + \cdots + \zeta_k A^k = 0,$$

y el polinomio mínimo de A es

$$q_A(t) = \frac{\zeta_k t^k + \cdots + \zeta_1 t + \zeta_0}{\zeta_k}$$

donde $\zeta_k \neq 0$.

Demostración (i) Sean $a,\, b \in \mathbb{C}^{n^2}$ y $X, Y \in \mathbb{C}^{n\times n}$, entonces,

$$\begin{aligned} T(aX+bY) &= (ax_{11}+by_{11}, \cdots, ax_{n1}+by_{n1}, \cdots, ax_{1n}+by_{1n}, \cdots, ax_{nn}+by_{nn}) \\ &= (ax_{11}, \cdots, ax_{n1}, \cdots, ax_{1n}, \cdots, ax_{nn}) + (by_{11}, \cdots, by_{n1}, \cdots, by_{1n}, \cdots, by_{nn}) \\ &= a(x_{11}, \cdots, x_{n1}, \cdots, x_{1n}, \cdots, x_{nn}) + b(y_{11}, \cdots, y_{n1}, \cdots, y_{1n}, \cdots, y_{nn}) \\ &= aT(X) + bT(Y), \end{aligned}$$

luego,T es una transformación lineal. Por otra parte tenemos,

$$\begin{aligned} X \in Ker(T) &\Leftrightarrow T(X) = \Theta_{(\mathbb{C}^{n^2})} \\ &\Leftrightarrow (x_{11}, \cdots, x_{n1}, \cdots, x_{1n}, \cdots, x_{nn}) = \Theta_{(\mathbb{C}^{n^2})} \\ &\Leftrightarrow X = (0, 0, \cdots, 0)_{\mathbb{C}^{n^2}} \end{aligned}$$

por lo tanto, $Ker(T) = \Theta_{(\mathbb{C}^{n^2})}$, luego, T es una transformación inyectiva. Ahora, como

$$Im(T) = \{T(X) \in \mathbb{C}^{n^2} : X \in \mathbb{C}^{n\times n}\} = \mathbb{C}^{n^2},$$

se tiene que T es sobre y por lo tanto, una transformación biyectiva, lo que prueba que es un isomorfismo.

(ii) Consideremos el polinomio característico de A

$$p_A(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 I$$

y observemos que los coeficientes de éste no son todos nulos (p_A es un polinomio mónico). Por el teorema de Cayley-Hamilton, tenemos

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0.$$

Aplicando la transformación lineal T a la igualdad anterior, tenemos

$$T(A^n) + a_{n-1}T(A^{n-1}) + \cdots + a_1 T(A) + a_0 T(I) = T(0),$$

es decir,

$$v_n + a_{n-1}v_{n-1} + \cdots + a_1 v + a_0 v_0 = 0,$$

donde los coeficientes a_i no todos son nulos. Por lo tanto, $\{v_0, v_1, .., v_n\}$ es linealmente dependiente.

(iii) Apliquemos el proceso Gram-Schmidt a los $n+1$ vectores $v_0, v_1, ..., v_n$. Entonces,

$$\begin{aligned} w_0 &= v_0 \\ w_1 &= v_1 - \alpha_{10} v_0 \\ w_2 &= v_2 - \alpha_{21} w_1 - \alpha_{2,0} w_0 \\ &\ \vdots \\ w_n &= v_n - \sum_{j=0}^{n-1} \alpha_{n,j} w_j = 0 \end{aligned}$$

donde $\alpha_{i,j} = \dfrac{\langle v_i, w_j \rangle}{\|w_j\|^2}$, $i = 1, \ldots, n$; $j = 0, \ldots, n-1$, $(i > j)$. Despejemos los vectores v_i en términos de los w_i, entonces,

$$\begin{aligned} v_0 &= w_0 \\ v_1 &= w_1 + \alpha_{10} w_0 \\ v_2 &= w_2 + \alpha_{2,1} w_1 - \alpha_{20} w_0 \\ &\ \vdots \\ v_n &= w_n + \sum_{j=0}^{n-1} \alpha_{n,j} w_j \end{aligned} \tag{3.5}$$

Por la dependencia lineal de los vectores $v_0, v_1, \ldots, v_n$, existe algún $k \neq 0$ tal que $w_k = 0$. Supongamos que w_k es el primer vector que se anula y escribamos

$$V = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_k \end{bmatrix}, \quad W = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{k-1} \\ w_k \end{bmatrix} \quad \text{y} \quad \Lambda = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha_{10} & 1 & 0 & \cdots & 0 \\ \alpha_{20} & \alpha_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_{k-1,0} & \alpha_{k-1,1} & \cdots & \alpha_{k-1,k-2} & 1 \end{bmatrix},$$

entonces, (3.5) se expresa en forma matricial como

$$V = \Lambda W,$$

donde

$$\alpha_{(i,j)} = \frac{trA_iW_j^t}{trW_jW_j^t}.$$

Claramente la matriz Λ es invertible, ya que es una matriz triangular inferior con unos en la diagonal. Si denotamos los coeficientes de Λ^{-1} por

$$\Lambda^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \zeta_{10} & 1 & 0 & \cdots & 0 \\ \zeta_{20} & \zeta_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \zeta_{(k-1,0)} & \zeta_{(k-1,1)} & \cdots & \zeta_{(k-1,k-2)} & 1 \end{bmatrix},$$

tenemos,

$$W = \Lambda^{-1}V,$$

en consecuencia,

$$w_k = \zeta_{(k-1,0)}v_0 + \zeta_{(k-1,1)}v_1 + \cdots + \zeta_{(k-1,k-2)}v_k = 0,$$

tomando, $\beta_i = \zeta_{(k-1,i)}$ (para $i = 0, \ldots, k-2$), tenemos

$$\beta_0T(I) + \beta_1T(A) + \cdots + \beta_kT(A^k) = 0,$$

lo cual implica que

$$\beta_0I + \beta_1A + \cdots + \beta_kA^k = 0,$$

luego, el polinomio $q(t) = \beta_0 + \beta_1t + \cdots + \beta_kt^k$, es un aniquilador de A. Supongamos que existe otro polinomio $r(s) = \beta v_0 + \beta v_1 + \ldots + \beta_n v_s$ que es aniquilador de A con $s \leqslant (k-1)$, entonces, como $\{v_0, \ldots, v_s\}$ es linealmente independiente para todos $s \leqslant k-1$, luego, $\beta_0 = \beta_1, \ldots, \beta_s = 0$ así,

$$\beta_0T(I) + \beta_1T(A) + \cdots + \beta_sT(A^s) = T(\beta_0I + \beta_1A + \cdots + \beta_sA^s) = 0.$$

Por lo tanto, $r(t) = \beta_0 + \beta_1t + \ldots + \beta_st^s$ de grado $s(\leqslant k-1)$ no aniquila a A, en consecuencia, $q(t) = \beta_0 + \beta_1t + \cdots + \beta_kt^k$ es el polinomio de menor grado que aniquila a A. Dividiendo por el coeficiente $\beta_0 \neq k$, concluimos que el polinomio mínimo de A es

$$q_A(t) = \frac{\beta_kt^k + \cdots + \beta_1t + \beta_0}{\beta_k}$$

con $\beta_k \neq 0$. ■

Así, el método para calcular el polinomio mínimo de una matriz cuadrada queda determinado por el siguiente algoritmo:

Dato de entrada: A.
Condiciones iniciales $B_0 = I_n = A_0 = W_0,\ sw = 0,\ k = 1$
WHILE $sw = 0$ y $k \leqslant n - 1$
$a_k = -\frac{1}{k} tr(AB_{k-1})$
$B_k = a_k I_n + AB_k - 1$
$A_k = B_k - \sum_{j=1}^{k} a_j A_{k-j};$
$W_k = A_k$
FOR $j = 0 \ldots k - 1$
    $\alpha_{k,j} = \frac{tr(A_k W_j^t)}{tr(W_j W_j^t)}$
    $W_k = W_k - \alpha_{(k,j)} W_j$
END
IF $W_k = 0$
    $p = k,\ sw = 1$
ELSE
    $k = k + 1.$
END (IF)
END (WHILE)
IF $sw = 0$
$a_n = -\frac{1}{n} tr(AB_{n-1})$
SALIDA: $q_A(t) = t^n + a_1 t^{n-1} + \cdots + a_{n-1} t + a_n$
ELSE
FOR $i = 1 \ldots p$
    $\zeta_{(i,i-1)} = -\alpha_{(i,i-1)}$
    FOR $j = 0 \ldots i - 2$
        $\zeta_{(i,j)} = -\alpha_{(i,j)} - \sum_{\ell=j+1}^{i-1} \alpha_{(i,\ell)} \zeta_{(\ell,j)}$
    END
END
SALIDA: $q_A(t) = t^p + \zeta_{(p,p-1)} t^{p-1} + ... + \zeta_{(p,1)} t + \zeta_{(p,0)}$
END (IF)

Tabla 3.1: Algoritmo para determinar el polinomio mínimo de una matriz.

3.2. Ejemplos

Ejemplo 1. Consideremos la matriz dada

$$A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}.$$

Si aplicamos el algoritmo anterior, obtenemos los siguientes resultados:

$$a_1 = -tr(AB_0) = -trA = -8,$$

$$B_1 = a_1 I_3 + AB_0 = \begin{bmatrix} -5 & -3 & 2 \\ -1 & -3 & -2 \\ -1 & 3 & -8 \end{bmatrix},$$

$$A_1 = B_1 - a_1 A_0 = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix},$$

$$\alpha_{1,0} = \frac{tr(A_1 W_0^t)}{tr(W_0 W_0^t)} = \frac{8}{3};$$

$$W_1 = W_1 - \alpha_{1,0} W_0 = \begin{bmatrix} \frac{1}{3} & -3 & 2 \\ -1 & \frac{7}{3} & -2 \\ -1 & 3 & \frac{-8}{3} \end{bmatrix},$$

$$a_2 = \frac{-1}{2} tr(AB_1) = 20;$$

$$B_2 = a_2 B_0 + AB_1 = \begin{bmatrix} 6 & 6 & -4 \\ 2 & 2 & 4 \\ 2 & -6 & 12 \end{bmatrix},$$

$$A_2 = B_2 - a_2 B_0 - a_1 A_1 = \begin{bmatrix} 10 & -18 & 12 \\ -6 & 22 & -12 \\ -6 & 18 & -8 \end{bmatrix},$$

$$\alpha_{2,0} = \frac{tr(A_2 W_0^t)}{tr(W_0 W_0^t)} = 8;$$

$$\alpha_{2,1} = \frac{tr(A_2 W_1^t)}{tr(W_1 W_1^t)} = 6;$$

$$W_2 = A_2 - \alpha_{2,0} W_0 - \alpha_{2,1} W_1 = 0.$$

salida: $p = 2$; $\zeta_{2,1} = -6$; $\zeta_{2,0} = 8$.
Por lo tanto, el polinomio mínimo de A es: $q_A(t) = t^2 - 6t + 8$

Ejemplo 2. Dada la matriz

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

Los resultados del algoritmo son los siguientes:

$$a_1 = -tr(AB_0) = -trA = 0$$

$$B_1 = a_1 B_0 + AB_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

$$A_1 = B_1 - a_1 A_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

$$\alpha_{1,0} = \frac{tr(A_1 W_0^t)}{tr(W_0 W_0^t)} = 0;$$

$$W_1 = W_1 - \alpha_{1,0} W_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

$$a_2 = \frac{-1}{2} tr(AB_1) = -3;$$

$$B_2 = a_2 B_0 + AB_1 = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix},$$

$$A_2 = B_2 - a_2 B_0 - a_1 A_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

$$\alpha_{2,0} = \frac{tr(A_2 W_0^t)}{tr(W_0 W_0^t)} = 2;$$

$$\alpha_{2,1} = \frac{tr(A_2 W_1^t)}{tr(W_1 W_1^t)} = 1;$$

$$W_2 = A_2 - \alpha_{2,0} W_0 - \alpha_{2,1} W_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

salida: $p = 2$; $\zeta_{21} = -2$; $\zeta_{21} = -1$.
Por lo tanto, el polinomio mínimo de A es $q_A(t) = t^2 - 2t - 1$.

Ejemplo 3. Dada la siguiente matriz,

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

obtenemos los siguientes resultados:

$$a_1 = -tr(AB_0) = -trA = -1$$

$$B_1 = a_1 B_0 + AB_0 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & -3 \end{bmatrix},$$

$$A_1 = B_1 - a_1 A_0 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 1 & 0 & -2 \end{bmatrix},$$

$$\alpha_{1,0} = \frac{tr(A_1 W_0^t)}{tr(W_0 W_0^t)} = \frac{1}{3};$$

$$W_1 = W_1 - \alpha_{1,0} W_0 = \begin{bmatrix} \frac{2}{3} & 0 & -1 \\ 0 & \frac{5}{3} & 1 \\ 1 & 0 & \frac{-7}{3} \end{bmatrix},$$

$$a_2 = \frac{-1}{2} tr(AB_1) = -3;$$

$$B_2 = a_2B_0 + AB_1 = \begin{bmatrix} -4 & 0 & 2 \\ 1 & -1 & -1 \\ -2 & 0 & 2 \end{bmatrix},$$

$$A_2 = B_2 - a_2B_0 - a_1A_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 4 & 0 \\ -1 & 0 & 3 \end{bmatrix},$$

$$\alpha_{2,0} = \frac{tr(A_2W_0^t)}{tr(W_0W_0^t)} = \frac{7}{3};$$

$$\alpha_{2,1} = \frac{tr(A_2W_1^t)}{tr(W_1W_1^t)} = \frac{-1}{5};$$

$$W_2 = A_2 - \alpha_{2,0}W_0 - \alpha_{2,1}W_1 = \begin{bmatrix} \frac{-11}{5} & 0 & \frac{4}{5} \\ 1 & 2 & \frac{1}{5} \\ \frac{-4}{5} & 0 & \frac{1}{5} \end{bmatrix},$$

$W_2 \neq 0$ $a_3 = (\frac{-1}{3})tr(AB_2) = 2$
salida: $a_1 = -1$, $a_2 = -3$, $a_3 = 2$.
Por lo tanto, el polinomio mínimo de A es: $q_A(t) = t^2 - t - 3t + 2$

Conclusiones

En el siguiente trabajo realizamos un análisis de la evolución del algoritmo de Leverrier-Faddeev desde su creación hasta las últimas implementaciones usadas.

Un pequeño aporte de este trabajo es la adaptación del algoritmo de Leverrier-Faddeev para el cálculo del polinomio mínimo de una matriz sin necesidad de encontrar auto-valores y mediante el aprovechamiento del algoritmo de Leverrier-Faddeev.

A lo largo de esta monografía y del análisis de los casos estudiados, surgen de manera natural una serie de problemas que pretendemos abordar en un futuro.

1. Implementación del algoritmo de Leverrier-Faddeev en el caso de matrices por bloques. Dichos tipos de matrices aparecen en el tratamiento de sistemas lineales, en particular en el estudio de filtros de predicción para sistemas de múltiples entradas y salidas (véase [16] y [17]).

2. Encontrar un algoritmo tipo Leverrier-Faddeev para el cálculo del polinomio mínimo de una matriz polinómica de grado arbitrario. Los resultados expuestos en [20], dan un algoritmo pero usando la transformada rápida de Fourier, sería interesante comparar las dos estrategias.

3. El polinomio mínimo guarda una relación con las inversa generalizada de una matriz, en [28] calculan la inversa de Moore-Penrose de una matriz usando un algoritmo tipo Leverrier-Faddeev. Se estudiará la viabilidad de implantar el algoritmo presentado aquí para tratar de obtener resultados similares al trabajo antes mencionado.

Bibliografía

[1] S. Barnett. *Polynomials and Linear Control Systems.* Marcel Dekker, New York, 1983.

[2] S. Barnett. Leverrier's algorithm: a new proof and extension. *SIAM J. Matrix Anal. Appl.*, 10:551–556, 1989.

[3] S. Barnett. Leverrier's algorithm for orthogonal polynomial bases. *Linear Alg. and Appl.*, 236:245–263, 1996.

[4] J. C. Basilio. Inversion of polynomial matrices via state-space. *Linear Alg. and Appl.*, 357:259–271, 2002.

[5] T. S. Chihara. *An Introduction to Orthogonal Polynomials.* Gordon and Breach, New York, 1978.

[6] E. Emre and O. Hüseying. Generalization of Leverrier's algorithm to polynomial matrices of arbitrary degree. *IEEE Trans. Automat. Control*, 20:136, Feb. 1975.

[7] P. K. Faddeev and V. N. Faddeeva. *Computational Methods of Linear Algebra.* W. H. Freeman and Co., San Francisco, CA., 1963.

[8] J. S. Frame. Matrix functions and applications IV: Matrix functions and constituent matrices. *IEEE Spectrum*, 1:123–131, 1964.

[9] G. M. L. Gladwell. *Inverse Problems in Vibrations.* Kluwer Academic Publishers, Dordrecht, 1986.

[10] J. C. Gower. A modified Leverrier-Faddeev algorithm for matrices with multiple eigenvalues. *Linear Alg. and Appl.*, 31:61–70, 1980.

[11] J. C. Gower. An application of the modified Leverrier-Faddeev algorithm to the spectral decomposition of symmetric block-circulant matrices. *Comp. Statist. and Data Anal.*, 50:89–106, 2006.

[12] J. Hernández and F. Marcellán. An extension of Leverrier-Faddeev algorithm using basis of classical orthogonal polynomials. *Facta Universitatis*, 19:73–92, 2004.

[13] J. Hernández and F. Marcellán. Classical orthogonal polynomials and Leverrier-Faddeev algorithm for the matrix pencil $sE - A$. *Inter. J. Math. and Math. Sciences*, 2006:Article ID 74507, 13 pages, doi:10.1155/IJMMS/2006/74507, 2006.

[14] J. Hernández, F. Marcellán, and C. Rodríguez. Leverrier-Faddeev algorithm and classical orthogonal polynomials. *Rev. Acad. Colomb. Ciencias Exactas, Físicas y Naturales*, 28(106):39–47, 2004.

[15] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990.

[16] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, N. J., 1980.

[17] T. Kailath and B. Porat. State-space generators for orthogonal polynomials. In V. Mandrekar and H. Salehi, editors, *Prediction Theory and Harmonic Analysis*, pages 131–163. The Pesi Massani Volume, North-Holland, 1983.

[18] N. P. Karampetakis and P. S. Stanimirović. Symbolic implementation of Leverrier-Faddeev algorithm and applications. In 8^{th} *IEEE Mediterranean Conference on Control and Automation*, Patras, Greece, 2000.

[19] N. P. Karampetakis and P. S. Stanimirović. On the computation of the Drazin inverse of a polynomial matrix. In 1^{st} *IFAC Symposium on System Structure and Control*, Prague, Czech Republic, 2001.

[20] N. P. Karampetakis and P. Tzekis. On the computation of the minimal polynomial of a polynomial matrix. *Int. J. Appl. Math. Comp. Sci.*, 15(3):339–349, 2005.

[21] F. L. Lewis. Further remark on the Cayley-Hamilton theorem and Leverrier's method for the matrix pencil $s\mathrm{E} - \mathrm{A}$. *IEEE Trans. Automat. Control*, AC-31:869–870, 1986.

[22] F. Marcellán, A. Branquinho, and J. C. Petronilho. Classical orthogonal polynomials: A functional approach. *Acta Appl. Math.*, 34:283–303, 1994.

[23] W. Marszalek and H. Unbehauen. Second order generalized linear systems arising in analysis of the flexible beams. In *31st Conference on the Decision and Control*, pages 3514–3518, Tucson, Arizona, 1992.

[24] B. G. Mertzios. Leverrier's algorithm for singular systems. *IEEE Trans. Automat. Control*, AC-29:652–653, 1984.

[25] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, PA, 2000.

[26] M. de la Sen, A. J. Garrido, O. Barambones, and F. J. Maseda. An expert network for the obtaining of approximate discrete-time models for LTI systems under real sampling using parameter identification. In *ICSE Sixteenth International Conference on Systems Engineering, The Institution of Electrical Engineers IEE*, volume 1, pages 132–139, Coventry, UK, 2003.

[27] P. S. Stanimirović. A finite algorithm for generalized inverses of polynomial and rational matrices. *Appl. Math. and Comp.*, 144:199–214, 2003.

[28] P. S. Stanimirović and M. B. Tasić. On the Leverrier-Faddeev algorithm for computing the Moore-Penrose inverse. *Appl. Math. and Comp.*, 35:135–141, 2011.

[29] G. Wang and Y. Lin. A new extension of Leverrier's algorithm. *Linear Alg. and Appl.*, 180:227–238, 1993.

[30] G. Wang and L. Qiu. Leverrier-Chebyshev algorithm for the singular pencils. *Linear Alg. and Appl.*, 345:1–8, 2002.

[31] Z. Yun and Y. Chengwu. Algorithms for the computation of the transfer function matrix for two-dimensional regular and singular general state-space models. *Automatica*, 31:1311–1315, 1995.

[32] B. Zheng and G. Wang. Leverrier's algorithm and Cayley-Hamilton theorem for 2-D system. *Appl. Math. and Comp.*, 160:725–738, 2005.