CUADERNOS DE TRABAJO

FACULTAD DE ESTUDIOS ESTADÍSTICOS

Matrices y polinomios ortogonales

Venancio Tomeo Perucha y Emilio Torrano Giménez

Cuaderno de Trabajo número 01/2019

UNIVERSIDAD
COMPLUTENSE
MADRID

Los Cuadernos de Trabajo de la Facultad de Estudios Estadísticos constituyen una apuesta por la publicación de los trabajos en curso y de los informes técnicos desarrollados desde la Facultad para servir de apoyo tanto a la docencia como a la investigación.

Los Cuadernos de Trabajo se pueden descargar de la página de la Biblioteca de la Facultad www.ucm.es/BUCM/est/, en la página del Repositorio Institucional UCM E-Prints Complutense y en la sección de investigación de la página del centro www.ucm.es/centros/webs/eest/

CONTACTO:

Biblioteca de la Facultad de Estudios Estadísticos

Universidad Complutense de Madrid

Av. Puerta de Hierro, S/N

28040 Madrid

Tlf. 913944035

buc_est@buc.ucm.es

ISSN: 2341-2550

Matrices y polinomios ortogonales

Venancio Tomeo Perucha
Dpto. Álgebra, Geometría y Topología
Facultad Estudios Estadísticos, UCM
tomeo@ucm.es

Emilio Torrano Giménez
Dpto. Matemática Aplicada
Escuela Superior de Ingenieros Informáticos, UPM
emilio@fi.upm.es

28 de mayo de 2019

Resumen:

El presente trabajo está enfocado a la construcción de la matriz D, Hessenberg superior con subdiagonal positiva, correspondiente al operador multiplicación por z en la base de los polinomios ortogonales. Para ello se comienza con el estudio de la matriz de Gram relativa a un producto escalar y de las propiedades de mínimo de los polinomios mónicos y de los núcleos.

La matriz D se construye a partir de una matriz M, infinita, hermitiana definida positiva (HDP), asociada a cualquier producto interno. Esta matriz D generaliza la tridiagonal de Jacobi.

Se muestran diferentes ejemplos matriciales y otros resultados interesantes sobre la transformación de semejanza, las matrices Toeplitz y el teorema de Szegö.

Palabras clave:

Polinomios ortogonales, matrices de Hessenberg, matriz de Gram, matrices hermitianas definidas positivas, matrices de momentos, matrices tridiagonales, matrices de Toeplitz, medidas, transformación de semejanza.

Contenido

1 Determinantes y matrices de Gram **3**

1.1 Ecuaciones normales y matriz de Gram 3

1.2 Aproximación en espacios con producto interno 7

1.3 Otras propiedades de los determinantes de Gram 10

2 Aplicación a los P.O. por el camino algebraico **13**

2.1 Obtención de los miembros de la SPON 14

2.2 Propiedad de mínimo de los mónicos 15

2.3 Propiedad de mínimo de los núcleos y su expresión mediante determinantes 15

3 Construcción de la matriz D **20**

3.1 Propiedades. Generalización de la matriz de Jacobi 20

3.2 La descomposición LR y la matriz $\widetilde{D}$ con unos en la subdiagonal 27

4 Ejemplos **29**

4.1 Medida de Lebesgue en $\mathbb{T}$ 29

4.2 Un producto escalar sin medida asociada 30

4.3 Una D no acotada. Medida de Laguerre 31

4.4 Un ejemplo con matriz M que desconocemos si es o no de momentos 32

4.5 Un ejemplo construido a partir de un operador hiponormal que no es subnormal 33

4.6 Medida planar de Lebesgue en el disco unidad 34

4.7 Otro ejemplo sobre la circunferencia unidad 35

5 Cuatro resultados sencillos **36**

5.1 Autovalores y autovectores de D_n^t 36

5.2 Obtención de M a partir de D 39

5.3 Transformaciones de semejanza del soporte y/o espectro 41

5.4 Expresión paramétrica de D en el caso en que M sea de Toeplitz 45

6 El teorema de Szegö **48**

1 Determinantes y matrices de Gram

1.1 Ecuaciones normales y matriz de Gram

Teorema 1.1. *Sean $\{x_k\}_{k=1}^n$ elementos independientes de un espacio vectorial*[1] *X dotado de producto interno $\langle \cdot\ ,\ \cdot \rangle$, y sea $\{z_k\}_{k=1}^n$ dicha base ortonormalizada, entonces cualquier elemento y de dicho espacio cumple*

$$\left(y - \sum_{k=1}^{n} \langle y, z_k \rangle z_k \right) \perp z_j.$$

Dem.

$$\begin{aligned} \langle \left(y - \sum_{k=1}^{n} \langle y, z_k \rangle z_k \right), z_j \rangle &= \langle y, z_j \rangle - \left(\sum_{k=1}^{n} \langle y, z_k \rangle \right) \langle z_k, z_j \rangle \\ &= \langle y, z_j \rangle - \langle y, z_j \rangle = 0. \end{aligned}$$

□

Corolario 1.1. *La diferencia entre un vector y, y su mejor aproximación mediante combinaciones lineales de los $x_1, x_2, \ldots, x_n$, es siempre ortogonal a todos los x_j.*

Dem. Es inmediato a partir de la proposición anterior. □

Observación 1.1. En lenguaje geométrico, el anterior corolario establece que la distancia más corta desde un punto dado a la variedad es la longitud de un elemento perpendicular a la variedad.

Teorema 1.2. *Sea $a_1x_1 + a_2x_2 + \ldots + a_nx_n$ la mejor (en el sentido de minimizar la norma) de las combinaciones lineales que aproximan el vector y, supuesto que y no pertenece a la variedad engendrada por $\{x_k\}_{k=1}^n$ supuestos linealmente independientes. Entonces los coeficientes a_i son la solución del siguiente sistema de ecuaciones*

$$\begin{aligned} a_1\langle x_1, x_1 \rangle + a_2\langle x_2, x_1 \rangle + \ldots + a_n\langle x_n, x_1 \rangle &= \langle y, x_1 \rangle \\ a_1\langle x_1, x_2 \rangle + a_2\langle x_2, x_2 \rangle + \ldots + a_n\langle x_n, x_2 \rangle &= \langle y, x_2 \rangle \\ &\vdots \\ a_1\langle x_1, x_n \rangle + a_2\langle x_2, x_n \rangle + \ldots + a_n\langle x_n, x_n \rangle &= \langle y, x_n \rangle. \end{aligned} \tag{1}$$

Estas ecuaciones se conocen como sistema normal de ecuaciones.

Dem. Por el corolario previo

$$\langle y - (a_1x_1 + a_2x_2 + \ldots + a_nx_n), x_j \rangle = 0, \quad j = 1, 2, \ldots, n$$

cuando se desarrollan estas ecuaciones resulta el anterior sistema. □

[1] K-espacio vectorial, donde como es habitual $K = \mathbb{R}$ o $K = \mathbb{C}$.

Definición 1.1. *Dada una sucesión de vectores $x_1, x_2, \ldots, x_n$ en un espacio con producto interno, la matriz $n \times n$*

$$G = (\langle x_i, x_j \rangle) = \begin{pmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \ldots & \langle x_1, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \ldots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \ldots & \langle x_n, x_n \rangle \end{pmatrix}$$

se conoce como matriz de Gram de los vectores $x_1, x_2, \ldots, x_n$. Su determinante

$$g = g(x_1, x_2, \ldots, x_n) = |\langle x_i, x_j \rangle| = |\langle x_j, x_i \rangle|$$

se conoce como determinante de Gram, o gramiano, de dichos elementos.

Observación 1.2. La matriz de Gram es la transpuesta de la matriz de coeficientes de las ecuaciones normales. También es la matriz de la forma bilineal

$$\langle a_1 x_1 + a_2 x_2 + \ldots + a_n x_n, b_1 x_1 + b_2 x_2 + \ldots + b_n x_n \rangle = \sum_{i,j=1}^{n} a_i \bar{b}_j \langle x_i, x_j \rangle.$$

Observemos que $g(x_1, x_2, \ldots, x_n)$ es una función simétrica de sus argumentos. Si intercambiamos x_i por x_j eso significa intercambiar dos filas entre si y dos columnas entre si, y el determinante no varía.

Lema 1.1. *Sea $y_i = \sum_{j=1}^{n} a_{ij} x_j$, $i = 1, 2, \ldots, n$. Llamemos A a la matriz (a_{ij}) y $\widetilde{A}$ a su transpuesta conjugada $(\widetilde{a}_{ij})$. Entonces se cumplen*

$$\begin{aligned} G(y_1, y_2, \ldots, y_n) &= AG(x_1, x_2, \ldots, x_n)\widetilde{A} \\ g(y_1, y_2, \ldots, y_n) &= |\det A|^2 g(x_1, x_2, \ldots, x_n). \end{aligned}$$

Dem. Tenemos que

$$\langle x_k, y_i \rangle = \langle x_k, \sum_{j=1}^{n} a_{ij} x_j \rangle = \sum_{j=1}^{n} \overline{a}_{ij} \langle x_k, x_j \rangle$$

que matricialmente puede expresarse como:

$$\begin{pmatrix} \langle x_1, y_1 \rangle & \langle x_1, y_2 \rangle & \ldots & \langle x_1, y_n \rangle \\ \langle x_2, y_1 \rangle & \langle x_2, y_2 \rangle & \ldots & \langle x_2, y_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, y_1 \rangle & \langle x_n, y_2 \rangle & \ldots & \langle x_n, y_n \rangle \end{pmatrix} =$$

$$= \begin{pmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \ldots & \langle x_1, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \ldots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \ldots & \langle x_n, x_n \rangle \end{pmatrix} \begin{pmatrix} \overline{a}_{11} & \overline{a}_{21} & \ldots & \overline{a}_{n1} \\ \overline{a}_{12} & \overline{a}_{22} & \ldots & \overline{a}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \ldots & \overline{a}_{nn} \end{pmatrix} =$$

$$= G(x_1, x_2, \ldots, x_n)\widetilde{A}.$$

Además

$$
\begin{pmatrix}
a_{11} & a_{12} & \dots & a_{1n} \\
a_{21} & a_{22} & \dots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \dots & a_{nn}
\end{pmatrix}
\begin{pmatrix}
\langle x_1, y_1 \rangle & \langle x_1, y_2 \rangle & \dots & \langle x_1, y_n \rangle \\
\langle x_2, y_1 \rangle & \langle x_2, y_2 \rangle & \dots & \langle x_2, y_n \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle x_n, y_1 \rangle & \langle x_n, y_2 \rangle & \dots & \langle x_n, y_n \rangle
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\langle y_1, y_1 \rangle & \langle y_1, y_2 \rangle & \dots & \langle y_1, y_n \rangle \\
\langle y_2, y_1 \rangle & \langle y_2, y_2 \rangle & \dots & \langle y_2, y_n \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle y_n, y_1 \rangle & \langle y_n, y_2 \rangle & \dots & \langle y_n, y_n \rangle
\end{pmatrix} = G(y_{1,2}, \dots, y_n).
$$

Combinando estas dos ecuaciones obtenemos la primera identidad del lema. Tomando determinantes y observando que $|\widetilde{A}| = \overline{|A|}$ resulta la segunda. Un caso particular será

$$
g(\sigma_1 x_1, \sigma_2 x_2, \dots, \sigma_n x_n) = |\sigma_1|^2 |\sigma_2|^2 \dots |\sigma_n|^2 g(x_1, x_2, \dots, x_n).
$$

□

Teorema 1.3. *Sean $x_1, x_2, \dots, x_n \neq 0$. Entonces*

$$
0 \leq g(x_1, x_2, \dots, x_n) \leq \|x_1\|^2 \|x_2\|^2 \cdots \|x_n\|^2.
$$

El extremo inferior $g = 0$ se alcanza, si y solamente si, los elementos x_i no son linealmente independientes. El valor de la derecha se alcanza, si y solo si, los $\{x_j\}_{j=1}^n$ son ortogonales.

Dem. Supongamos, en primer lugar, que los x_j no son linealmente independientes. Entonces podemos encontrar constantes $a_1, a_2, \dots, a_n$ no todas nulas, tales que $a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0$. Supongamos que $a_j \neq 0$, y consideremos la transformación

$$
\begin{aligned}
y_1 &= x_1 \\
y_2 &= x_2 \\
&\vdots \\
y_{j-1} &= x_{j-1} \\
y_j &= a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0 \\
y_{j+1} &= x_{j+1} \\
&\vdots \\
y_n &= x_n
\end{aligned}
$$

Como $\langle y_j, y_i \rangle = \langle 0, y_i \rangle = 0$, resulta que $g(y_1, y_2, \dots, y_n) = 0$. Tenemos que

$$
|A| = \begin{vmatrix}
1 & 0 & \dots & 0 & \dots & 0 \\
0 & 1 & \dots & 0 & \dots & 0 \\
\vdots & \vdots & & \vdots & & \vdots \\
a_1 & a_2 & \dots & a_j & \dots & a_n \\
\vdots & \vdots & & \vdots & & \vdots \\
0 & 0 & \dots & 0 & \dots & 1
\end{vmatrix}.
$$

Desarrollando este determinante por la columna j-ésima, tendremos que

$$|A| = 0 + 0 + \ldots + a_j \cdot 1 + 0 + \ldots + 0 = a_j \neq 0.$$

Se sigue del Lema 1.1 que $g(x_1, x_2, \ldots, x_n) = 0$.

Supongamos seguidamente que los x_j son linealmente independientes. Entonces podemos[2] encontrar constantes a_{ij} tales que

$$z_k = a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kk}x_k, \quad \text{con} \quad a_{kk} > 0,$$

tales que los $\{z_k\}$ sean ortonormales. Por el lema anterior

$$1 = |\delta_{ij}| = g(z_1, z_2, \ldots, z_n) = g(x_1, x_2, \ldots, x_n)|A|^2$$

donde

$$|A| = \begin{vmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

Por tanto

$$g(x_1, x_2, \ldots, x_n) = \frac{1}{a_{11}^2} \cdot \frac{1}{a_{22}^2} \cdots \frac{1}{a_{nn}^2} > 0.$$

Luego $g = 0$ solo cuando las x_j no son linealmente independientes. Ahora probaremos que

$$\frac{1}{a_{kk}^2} \leq \|x_k\|^2.$$

Del proceso de ortogonalización de Gram-Schmidt sabemos que

$$\frac{1}{a_{kk}^2} = \|y_k\|^2 = \|x_k - \sum_{j=1}^{k-1} \langle x_k, z_j \rangle z_j\|^2 \leq \|x_k\|^2.$$

Si los x_j son ortogonales, entonces G es una matriz con los $\|x_j\|^2$ en la diagonal principal y ceros en el resto. En ese caso, $g(x_1, x_2, \ldots, x_n) = \|x_1\|^2\|x_2\|^2 \cdots \|x_n\|^2$. Supongamos recíprocamente que $g(x_1, x_2, \ldots, x_n) = \|x_1\|^2\|x_2\|^2 \cdots \|x_n\|^2$, tenemos que $g(x_1, x_2, \ldots, x_n) = \frac{1}{a_{11}^2} \cdot \frac{1}{a_{22}^2} \cdots \frac{1}{a_{nn}^2}$. Ahora como $1/a_{kk}^2 \leq \|x_k\|^2$ se sigue que $\|y_k\|^2 = \frac{1}{a_{kk}^2} = \|x_k\|^2$, $k = 1, 2, \ldots, n$. Pero entonces

$$\sum_{i=1}^{k-1} |\langle x_k, z_i \rangle|^2 = 0, \quad k = 1, 2, \ldots, n,$$

esto implica la ortogonalidad de la familia. □

Corolario 1.2. *(Desigualdad de Hadamard)*

Sea $D = (a_{ij})$ una matriz $n \times n$ con elementos complejos, entonces

$$|D|^2 \leq \prod_{k=1}^{n} (|a_{k1}|^2 + |a_{k2}|^2 + \ldots + |a_{kn}|^2). \tag{2}$$

[2] Los coeficientes que van apareciendo en el proceso de ortogonalización de Gram-Schmidt.

Si los elementos a_{ij} satisfacen $|a_{ij}| \leq M$, $i, j = 1, 2, \ldots, n$, entonces

$$|D| \leq M^n n^{n/2}.$$

Dem. Llamemos x_i al vector $(a_{i1}, a_{i2}, \ldots, a_{in})$. Introduzcamos el producto interno hermitiano en $\mathbb{C}^n$

$$\langle x_i, x_j \rangle = \sum_{k=1}^{n} a_{ik} \overline{a}_{jk}.$$

Si $\widetilde{D}$ designa la matriz transpuesta conjugada de D, $(\overline{a}_{ji})$, entonces

$$\begin{aligned} D\widetilde{D} &= \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \begin{vmatrix} \overline{a}_{11} & \overline{a}_{21} & \ldots & \overline{a}_{n1} \\ \overline{a}_{12} & \overline{a}_{22} & \ldots & \overline{a}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \ldots & \overline{a}_{nn} \end{vmatrix} \\ &= |\langle x_i, x_j \rangle| = g(x_1, x_2, \ldots, x_n) \leq \|x_1\|^2 \|x_2\|^2 \cdots \|x_n\|^2. \end{aligned}$$

Como $\|x_i\|^2 = \sum_{k=1}^n a_{ik}\overline{a}_{ik} = \sum_{k=1}^n |a_{ik}|^2$, y $|\overline{D}| = |\widetilde{D}|$, resulta (2). Si $|a_{ij}| \leq M$, tendremos que $|a_{k1}|^2 + |a_{k2}|^2 + \ldots + |a_{kn}|^2 \leq nM^2$, y por tanto $|D|^2 \leq n^n M^{2n}$. □

Ejemplo 1.1. Sean $f_i(t) \in C([a, b])$, $i = 1, 2, \ldots, n$. Estas funciones son linealmente independientes si y solo si el determinante de Gram $\left|\int_a^b f_i(t) f_j(t) dt\right| > 0$. Un resultado similar se obtiene para $f_i(t) \in L^2([a, b])$.

1.2 Aproximación en espacios con producto interno

En el teorema que veremos a continuación resolvemos el problema de calcular la distancia mínima de una variedad a un vector fijo y, externo a ella, y en el siguiente determinamos qué elemento de la variedad es que el está a distancia mínima de y. Finalmente obtenemos los sucesivos elementos de la base ortonormal utilizando la matriz de Gram.

Teorema 1.4. *Sean $x_1, x_2, \ldots, x_n$, vectores linealmente independientes, sea y un vector fuera del subespacio generado por ellos. Si*

$$\delta = \min_{a_i} \|y - (a_1 x_1 + a_2 x_2 + \ldots + a_n x_n)\|,$$

entonces

$$\delta^2 = \frac{g(x_1, x_2, \ldots, x_n, y)}{g(x_1, x_2, \ldots, x_n)}.$$

Dem. Llamemos v al vector $a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$ que minimiza la distancia a y. Entonces

$$\delta^2 = \|y - v\|^2 = \langle y - s, y - s \rangle = \langle y - v, y \rangle - \langle y - v, v \rangle.$$

Sabemos, ver Teorema 1.1, que $\langle y - v, v \rangle = 0$, por tanto

$$\delta^2 = \langle y - v, y \rangle = \langle y, y \rangle - \langle v, y \rangle$$

de donde

$$\langle v, y\rangle = \langle y, y\rangle - \delta^2. \tag{3}$$

Escribamos las ecuaciones normales en la forma siguiente y añadámosle como última ecuación (3) en forma desarrollada

$$\begin{array}{lcl} a_1\langle x_1, x_1\rangle + a_2\langle x_2, x_1\rangle + \ldots + a_n\langle x_n, x_1\rangle & -\langle y, x_1\rangle & = & 0 \\ & & \vdots & \\ a_1\langle x_n, x_1\rangle + a_2\langle x_n, x_1\rangle + \ldots + a_n\langle x_n, x_n\rangle & -\langle y, x_n\rangle & = & 0 \\ a_1\langle x_n, y\rangle + a_2\langle x_n, y\rangle + \ldots + a_n\langle x_n, y\rangle & +[\delta^2 - \langle y, y\rangle] & = & 0 \end{array}$$

Si introducimos $a_{n+1} = 1$, podemos suponer la última columna multiplicada por dicho valor. Tenemos un sistema homogeneo de $n+1$ ecuaciones con $n+1$ incógnitas, $a_1, a_2, \ldots, a_n, a_{n+1}(= 1)$, que posee una solución no trivial, luego el determinante de la matriz de coeficientes debe ser nulo, es decir

$$\begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle & 0 - \langle y, x_1\rangle \\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_n, x_2\rangle & 0 - \langle y, x_2\rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_n\rangle & 0 - \langle y, x_n\rangle \\ \langle x_1, y\rangle & \langle x_2, y\rangle & \ldots & \langle x_n, y\rangle & \delta^2 - \langle y, y\rangle \end{vmatrix} = 0.$$

En consecuencia

$$\begin{vmatrix} \langle x_1, x_1\rangle & \ldots & \langle x_n, x_1\rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \ldots & \langle x_n, x_n\rangle & 0 \\ \langle x_1, y\rangle & \ldots & \langle x_n, y\rangle & \delta^2 \end{vmatrix} = \begin{vmatrix} \langle x_1, x_1\rangle & \ldots & \langle x_n, x_1\rangle & \langle y, x_1\rangle \\ \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \ldots & \langle x_n, x_n\rangle & \langle y, x_n\rangle \\ \langle x_1, y\rangle & \ldots & \langle x_n, y\rangle & \langle y, y\rangle \end{vmatrix},$$

y desarrollando el primer determinante por la última columna resulta

$$\delta^2\, g(x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n, y).$$

□

Teorema 1.5. *Sean $x_1, x_2, \ldots, x_n$ vectores linealmente independientes. La solución al problema de hacer mínima la distancia a un vector dado y, que no esté en el subespacio generado por ellos, es decir*

$$\min_{a_i} \|y - (a_1x_1 + a_2x_2 + \ldots + a_nx_n)\|$$

viene dada por el vector $v = a_1x_1 + a_2x_2 + \ldots + a_nx_n$, tal que

$$v = \frac{-1}{g(x_1, x_2, \ldots, x_n)} \begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle & \langle y, x_1\rangle \\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_n, x_2\rangle & \langle y, x_2\rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_n\rangle & \langle y, x_n\rangle \\ x_1 & x_2 & \ldots & x_n & 0 \end{vmatrix}. \tag{4}$$

Además la diferencia $y - v$, vale

$$y - v = \frac{1}{g(x_1, x_2, \ldots, x_n)} \begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle & \langle y, x_1\rangle \\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_n, x_2\rangle & \langle y, x_2\rangle \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_n\rangle & \langle y, x_n\rangle \\ x_1 & x_2 & \ldots & x_n & y \end{vmatrix}. \tag{5}$$

Dem. A partir de las ecuaciones normales (1) y aplicando la regla de Cramer, tenemos

$$a_1 = \frac{1}{g(x_1, x_2, \ldots, x_n)} \begin{vmatrix} \langle y, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle \\ \langle y, x_2\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_n, x_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_n\rangle \end{vmatrix},$$

tendremos expresiones semejantes para los restantes a_k. Si desarrollamos el determinante (4) por la última fila, el coeficiente de x_1, es justamente $-a_1$, y análogamente con cada uno de los restantes coeficientes con lo que queda probada la primera afirmación.

La siguiente identidad es obvia, sin más que desarrollar por la última columna

$$y = \frac{1}{g(x_1, x_2, \ldots, x_n)} \begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle & 0 \\ \langle x_1, x_2\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_n, x_2\rangle & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_n\rangle & 0 \\ x_1 & x_2 & \ldots & x_n & y \end{vmatrix} \tag{6}$$

calculando $y - v$ a partir de (6) y de (4) obtenemos (5). □

Corolario 1.3. *Sean $x_1, x_2, x_3, \ldots$ linealmente independientes y ortonormalizándolos de acuerdo con el esquema de Gram-Schmidt se llega a $z_1, z_2, z_3, \ldots$. Entonces*[3]

$$z_n = \frac{1}{\sqrt{g(x_1, x_2, \ldots, x_{n-1}) g(x_1, x_2, \ldots, x_n)}} \begin{vmatrix} \langle x_1, x_1\rangle & \langle x_2, x_1\rangle & \ldots & \langle x_n, x_1\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_1, x_n\rangle & \langle x_2, x_n\rangle & \ldots & \langle x_n, x_{n-1}\rangle \\ x_1 & x_2 & \ldots & x_n \end{vmatrix} \tag{7}$$

para $n > 1$, con $z_1 = x_1/\sqrt{g(x_1)}$. El coeficiente que multiplica a x_n en la expresión que nos da z_n, es

$$a_{nn} = \sqrt{\frac{g(x_1, x_2, \ldots, x_{n-1})}{g(x_1, x_2, \ldots, x_n)}}, \quad \textit{para} \quad n > 1. \tag{8}$$

Dem. Consideremos el problema de minimizar

$$\min_{a_i} \|x_n - (a_1 x_1 + a_2 x_2 + \ldots + a_{n-1} x_{n-1})\|.$$

[3]Ver por ejemplo (pág. 183 [Dav75])

Sabemos que la solución de este problema de mínimos cuadrados viene dada por

$$\sum_{k=1}^{n-1} \langle x_n, z_k \rangle z_k,$$

donde los $\{z_k\}$ son los $\{x_k\}$ ortonormalizados mediante Gram-Schmidt, en ese proceso se tiene que

$$y_n = x_n - \sum_{k=1}^{n-1} \langle x_n, z_k \rangle z_k, \quad \text{y luego se introduce} \quad z_n = \frac{y_n}{\|y_n\|}, \tag{9}$$

por lo que la solución del anterior problema de mínimo viene dada por $x_n - y_n$. Hemos partido de que $z_n = a_{nn} x_n + \ldots$, luego

$$\frac{z_n}{a_{nn}} = x_n + \ldots = x_n - y_n,$$

donde es obvio que $a_{n,n} = \|y_n\|$. Utilizando (5)

$$\begin{aligned} \frac{z_n}{\|y_n\|} &= x_n - (a_1 x_1 + a_2 x_2 + \ldots + a_{n-1} x_{n-1}) \\ &= \frac{1}{g(x_1, x_2, \ldots, x_{n-1})} \begin{vmatrix} \langle x_1, x_1 \rangle & \langle x_2, x_1 \rangle & \ldots & \langle x_n, x_1 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_1, x_{n-1} \rangle & \langle x_2, x_{n-1} \rangle & \ldots & \langle x_n, x_{n-1} \rangle \\ x_1 & x_2 & \ldots & x_n \end{vmatrix}. \end{aligned}$$

Por otro lado como el coeficiente $a_{nn} > 0$, podemos escribir

$$\frac{1}{a_{nn}} = \left\| \frac{z_n}{a_{nn}} \right\| = \min_{a_i} \|x_n - (a_1 x_1 + a_2 x_2 + \ldots + a_{n-1} x_{n-1})\|$$

y por el Teorema 1.4 resulta

$$\frac{1}{a_{nn}} = \sqrt{\frac{g(x_1, x_2, \ldots, x_{n-1}, x_n)}{g(x_1, x_2, \ldots, x_{n-1})}}.$$

□

Observación 1.3. Obsérvese que esta fórmula para nada utiliza que el producto interno tenga que ir asociado a una integral. Los resultados que se obtienen son, por tanto, más generales que los que se obtienen a partir de una medida. La introducción de los PO a partir de un funcional de momentos (método que sigue Chihara) cuando es definido positivo, es consistente con todo lo anterior, con la ventaja de que permite extensiones, que están vedadas dentro de la teoría de espacios con producto interno usual.

1.3 Otras propiedades de los determinantes de Gram

Teorema 1.6. *El determinante de Gram $g(x_1, x_2, \ldots, x_n)$, de los vectores $x_1, x_2, \ldots, x_n$ linealmente independientes, posee las siguientes propiedades.*

i) *g es simétrico respecto de sus argumentos.*

ii) $g(x_1, \dots, \sigma x_j, \dots, x_n) = |\sigma|^2 g(x_1, x_2, \dots, x_n)$

iii) $g(x_1, \dots, x_j + \sigma x_k, \dots, x_n) = g(x_1, \dots, x_n)$, *con* $j \neq k$

iv) $\sqrt{g(x_1' + x_1'', x_2, \dots, x_n)} \leq \sqrt{g(x_1', x_2, \dots, x_n)} + \sqrt{g(x_1'', x_2, \dots, x_n)}$

iv) $g(x_1, \dots, x_n) \leq g(x_1, \dots, x_n) g(x_{p+1}, \dots, x_n)$, *con* $1 \leq p < n$.

v) *La igualdad se alcanza en el apartado iv) si y solo si*

$$\langle x_i, x_j \rangle = 0, \quad 1 \leq i \leq p, \quad p+1 \leq j \leq n.$$

Dem. En las proposiciones anteriores hemos probado ya los apartados i) y ii). El apartado iii) es inmediato a partir de las propiedades elementales de los determinantes. iv). Podemos suponer que $x_2, x_3, \dots, x_n$ son independientes, en otro caso los dos miembros de iv) se anularían y la igualdad sería trivial. Ortonormalicemos $x_2, x_3, \dots, x_n$ y llamemos a los vectores normalizados $z_2, z_3, \dots, z_n$. Entonces[4] tendremos

$$\begin{aligned}
\frac{g^{1/2}(x_1' + x_1'', x_2, \dots, x_n)}{g^{1/2}(x_2, x_3, \dots, x_n)} &= \|x_1' + x_1'' - \sum_{k=2}^{n} \langle x_1' + x_1'', z_k \rangle z_k\| \\
&= \|x_1' - \sum_{k=1}^{n} \langle x_1', z_k \rangle x_k + x_1'' - \sum_{k=2}^{n} \langle x_1'', z_k \rangle z_k\| \\
&\leq \|x_1' - \sum_{k=2}^{n} \langle x_1', z_k \rangle z_k\| + \|x_1'' - \sum_{k=2}^{n} \langle x_1'', z_k \rangle z_k\| \\
&= \frac{g^{1/2}(x_1', x_2, \dots, x_n)}{g^{1/2}(x_2, \dots, x_n)} + \frac{g^{1/2}(x_1'', x_2, \dots, x_n)}{g^{1/2}(x_2, \dots, x_n)}.
\end{aligned}$$

Multiplicando esta desigualdad por el denominador resulta probada iv). Veamos v). Sea k satisfaciendo $1 \leq k < p$. Como hay más grados de libertad en el lado izquierdo se cumple que

$$\min_{a_j} \|x_k - (a_{k+1} x_{k+1} + \ldots + a_n x_n)\|^2 \leq \min_{b_j} \|x_k - (b_{k+1} x_{k+1} + \ldots + b_p x_p)\|^2.$$

De forma semejante

$$\min_{c_j} \|x_p - (c_{p+1} x_{p+1} + \ldots + c_n x_n)\|^2.$$

Por tanto por el Teorema 1.4

$$\frac{g(x_k, x_{k+1}, x_{k+2}, \dots, x_n)}{g(x_{k+1}, \dots, x_n)} \leq \frac{g(x_k, x_{k+1}, \dots, x_p)}{g(x_{k+1}, \dots, x_p)} \tag{10}$$

y

$$\frac{g(x_p, x_{p+1}, \dots, x_n)}{g(x_{p+1}, \dots, x_n)} \, g(x_{p+1}, \dots, x_n) \leq g(x_p). \tag{11}$$

[4] Obsérvese que esta propiedad es un análogo para gramianos de la desigualdad de Minkowski.

En particular, escribiendo para $k = 1, 2, \ldots, p-1$ en (10), tendremos

$$\begin{aligned}
\frac{g(x_1,\ldots,x_n)}{g(x_2,\ldots,x_n)} &\leq \frac{g(x_1,\ldots,x_p)}{g(x_2,\ldots,x_p)} \\
\frac{g(x_2,\ldots,x_n)}{g(x_3,\ldots,x_n)} &\leq \frac{g(x_2,\ldots,x_p)}{g(x_3,\ldots,x_p)} \\
&\vdots \\
\frac{g(x_p,\ldots,x_n)}{g(x_{p+1},\ldots,x_n)} &\leq g(x_p).
\end{aligned}$$

Multiplicando las anteriores igualdades resulta

$$\frac{g(x_1,\ldots,x_n)}{g(x_{p+1},\ldots,x_n)} \leq g(x_1,\ldots,x_p)$$

o lo que es lo mismo

$$g(x_1,\ldots,x_n) \leq g(x_1,\ldots,x_p)\; g(x_{p+1},\ldots,x_n). \tag{12}$$

Esta relación es una generalización de la desigualdad de Cauchy-Schwartz para gramianos.

Veamos por último que la igualdad en (12) se da si y solo en la anterior cadena de desigualdades se alcanza la igualdad, es decir, si y solamente si

$$\min_{c_j} \|x_p(c_{p+1}x_{p+1} + \ldots + c_n x_n)\|^2 = \|x_p\|^2, \tag{13}$$

y

$$\min_{a_j} \|x_k(a_{k+1}x_{k+1}+\ldots+a_n x_n)\|^2 = \min_{b_j} \|x_k-(b_{k+1}x_{k+1}+\ldots+b_p x_p)\|^2, \quad k = 1, 2, \ldots, p-1. \tag{14}$$

Ahora (13) se cumple, si y solo si,

$$\langle x_p, x_j\rangle = 0, \quad j = p+1, \ldots, n.$$

Haciendo $k = p-1$ en (14) tendremos

$$\min_{a_j} \|x_{p-1} - (a_p x_p + \ldots + a_n x_n)\|^2 = \min_{b_p} \|x_{p-1} - b_p x_p\|^2.$$

Por el mismo principio, esto ocurre si y solo si $x_{p-1} \perp x_{p+1}, x_{p+2}, \ldots, x_n$. Haciendo luego $k = p-2, p-3, \ldots, 1$ sucesivamente, llegamos a las condiciones de ortogonalidad de la tesis. $\square$

Observación 1.4. El determinante de Gram tiene una curiosa interpretación geométrica. Supongamos dados n vectores de $\mathbb{R}^n$, tales que $x_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \in \mathbb{R}^n$, $i = 1, 2, \ldots n$. Estos vectores son las aristas de cierto paralelepípedo n-dimensional (la generalización de un paralelogramo) cuyo volumen llamaremos $V = V(x_1, x_2, \ldots, x_n)$. Puede probarse que

$$V = \text{valor aboluto}\left(\begin{vmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nn} \end{vmatrix} \right). \tag{15}$$

Aunque elemental, la deducción de esta fórmula, está lejos de ser trivial, sin embargo aquí podemos obtenerla de otra manera.

Multiplicando el determinante de la transpuesta conjugada por (15) , resulta que

$$\sum_{j=1}^{n} \overline{x}_{ij}\; x_{jk} = \langle x_i, x_k \rangle$$

luego podemos escribir aplicando Binet-Cauchy

$$V^2 = \begin{vmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \dots & \langle x_1, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \dots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \dots & \langle x_n, x_n \rangle \end{vmatrix} = g(x_1, x_2, \dots, x_n).$$

En consecuencia

$$V(x_1, x_2, \dots, x_n) = \sqrt{g(x_1, x_2, \dots, x_n)}.$$

Se deduce (15) vía Teorema 1.4, si extendemos a dimensión superior lo que ocurre en dos o tres dimensiones. El volumen se obtiene multiplicando entre si las alturas, es decir

$$V(x_1, x_2, \dots, x_n) = \|x_1\| \cdot d(x_2; x_1) \cdot d(x_3; x_1, x_2) \cdots d(x_n; x_1, x_2, \dots, x_{n-1})$$

donde $d(x_i; x_1, \dots, x_{i-1})$ designa la distancia de x_i a la variedad lineal engendrada por $x_1, \dots, x_{i-1}$. Por tanto

$$d(x_i; x_1, x_2, \dots, x_{i-1}) = \sqrt{\frac{g(x_1, x_2, \dots, x_i)}{g(x_1, x_2, \dots, x_{i-1})}}$$

luego

$$V^2 = g(x_1) \frac{g(x_1, x_2)}{g(x_1)} \frac{g(x_1, x_2, x_3)}{g(x_1, x_2)} \cdots \frac{g(x_1, x_2, \dots, x_n)}{g(x_1, x_2, \dots, x_{n-1})} = g(x_1, \dots, x_n).$$

2 Aplicación a los P.O. por el camino algebraico

Con este título queremos indicar que no utilizaremos para los resultados que siguen nada relativo a la medida. Es decir todo es válido a partir de una matriz HDP, infinita sea o no de momentos. Es claro que en nuestro caso, la base *standard* de Π_M, espacio de los polinomios con producto interno dado por M, – que no es una base de Schauder– es $\{z^k\}_{k=0}^{\infty}$. En nuestro caso

$$x_1 = 1, \;\; x_2 = z, \;\; x_3 = z^2, \dots, x_{n+1} = z^n, \dots$$

y denominamos $\{p_n(z)\}_{n=1}^{\infty}$ a la base ortonormal SPON[5], es decir

$$z_1 = p_0(z), \;\; z_2 = p_1(z), \;\; z_3 = p_3(z), \dots, z_{n+1} = p_n(z), \dots$$

[5]Sucesión de polinomios ortonormales.

La matriz de Gram de orden $n+1$ es obviamente la sección de orden $n+1$ de la matriz infinita M,

$$M_{n+1} = \begin{pmatrix} \langle 1,1\rangle & \langle z,1\rangle & \dots & \langle z^n,1\rangle \\ \langle 1,z\rangle & \langle z,z\rangle & \dots & \langle z^n,z\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle 1,z^n\rangle & \langle z,z^n\rangle & \dots & \langle z^n,z^n\rangle \end{pmatrix} = \begin{pmatrix} c_{00} & c_{10} & \dots & c_{n0} \\ c_{01} & c_{11} & \dots & c_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{0n} & c_{1n} & \dots & c_{nn} \end{pmatrix}.$$

En teoría de P.O. se suele denominar

$$\Delta_n = |M_{n+1}| = g(x_1, x_2, \dots, x_{n+1}).$$

La sucesión de los PO normalizados se notará como $\{p_n(z)\}_{n=0}^{\infty}$, y a los correspondientes mónicos los representaremos –solo aquí– como $P_n(z)$. Asimismo el coeficiente $a_{n,n}$ –siguiendo la notación de la sección anterior– será

$$p_{n-1}(z) \equiv z_n = a_{n,n}x_n + a_{n,n-1}x_{n-1} + \dots + a_{n,1}x_1$$

se suele denotar como κ_{n-1} y a veces en el caso real como γ_{n-1}.

2.1 Obtención de los miembros de la SPON

Proposición 2.1. *Dada una matriz M infinita y HDP, la sucesión de $\{p_n(z)\}_{n=0}^{\infty}$, se obtiene mediante la expresión*

$$p_n(z) = \frac{1}{\sqrt{\Delta_{n-1}\Delta_n}} \begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} \\ c_{01} & c_{11} & \dots & c_{n,1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{0,n-1} & c_{1,n-1} & \dots & c_{n-1,n-1} \\ 1 & z & \dots & z^n \end{vmatrix},$$

siendo

$$\Delta_n = \begin{vmatrix} c_{00} & c_{10} & \dots & c_{n0} \\ c_{01} & c_{11} & \dots & c_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{0n} & c_{1n} & \dots & c_{nn} \end{vmatrix}.$$

Dem. Es una consecuencia directa de la fórmula (7) del Corolario 1.3. Basta con hacer $x_{n+1} = z^n$, $z_{n+1} = p_n(z)$, $n = 0, 1, 2, \dots$, y tener en cuenta que

$$\begin{pmatrix} \langle 1,1\rangle & \langle z,1\rangle & \dots & \langle z^n,1\rangle \\ \langle 1,z\rangle & \langle z,z\rangle & \dots & \langle z^n,z\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle 1,z^n\rangle & \langle z,z^n\rangle & \dots & \langle z^n,z^n\rangle \end{pmatrix} = \begin{pmatrix} c_{00} & c_{10} & \dots & c_{n0} \\ c_{01} & c_{11} & \dots & c_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{0n} & c_{1n} & \dots & c_{nn} \end{pmatrix}.$$

Nótese que los vectores que hemos llamado y_n en el Corolario 1.3, son los miembros de la SPOM (sucesion de P.O. mónicos). $\square$

Observación 2.1. Obsérvese que si trabajáramos con M traspuesta, la fila de potencias sucesivas de z tendría que ser columna para que no variara el determinante. En las matrices de Hankel o en las Toeplitz reales (por el hecho de ser simétricas) esa doble posibilidad no afecta al resultado. No es así en las complejas. Para ser más concretos: en el caso de matrices de Toeplitz complejas, la representación puede dar lugar a polinomios diferentes.

2.2 Propiedad de mínimo de los mónicos

Proposición 2.2. *El polinomio mónico de grado n con norma más pequeña es $P_n(z) = p_n(z)/\kappa_n$, es decir el mónico de grado n de la SPOM, y el valor de ese mínimo es justamente $1/\kappa_n$.*

DEM. Basta aplicar la fórmula (5) del Teorema 1.5, tras escoger $y = z^n$, nótese que según la demostración del Corolario 1.3 el mínimo toma el valor $1/a_{nn}$, en nuestro caso $a_{n+1,n+1} = \kappa_n$. □

2.3 Propiedad de mínimo de los núcleos y su expresión mediante determinantes

Definición 2.1. *Dada la SPON $\{p_n(z)\}$, respecto de un producto escalar dado por $M = (c_{ij})_{i,j=0}^{\infty}$, infinita y HDP, se denomina n-núcleo a la suma*[6]

$$K_n(x,y) = \sum_{j=0}^{n} \overline{p_j(x)} p_j(y). \tag{16}$$

Proposición 2.3. *El polinomio de grado n, verificando que $Q_n(z_0) = 1$, tiene norma mínima si y solo si, $Q_n(z) = K_n(z_0,z)/K_n(z_0,z_0)$, además el valor de ese mínimo al cuadrado es justamente $1/K_n(z_0,z_0)$.*

DEM. Es claro que cualquier polinomio $Q_n(z)$ podemos expresarlo en función de los polinomios de la SPON, es decir, existen constantes $c_k \in \mathbb{C}$, tales que

$$Q_n(z) = c_0 p_0(z) + c_1 p_1(z) + \ldots + c_n p_n(z).$$

La ventaja obvia de la base $\{p_n(z)\}$, es que se cumple

$$\|Q_n(z)\|^2 = \langle Q_n(z), Q_n(z)\rangle = \sum_{k=0}^{n} \bar{c}_j c_j = \sum_{k=0}^{n} |c_k|^2,$$

por la propia definición de ortonormalidad respecto del producto escalar dado por M. De acuerdo con la hipótesis

$$Q_n(z_0) = 1 = \sum_{k=0}^{n} c_k \, p_k(z_0).$$

[6] Hay distintas formas de definir los n-núcleos. Por ejemplo: $K_n(x,y) = \sum_{j=0}^{n} p_j(x)\overline{p_j(y)}$. Hemos optado por la de la pág. 377 de [Sze39] y pág. 38 [Chi78]

Aplicamos la desigualdad de Cauchy-Schwarz a la anterior igualdad y resulta

$$1 = \sum_{k=0}^{n} c_k \; p_k(z_0) \leq \sqrt{\sum_{k=0}^{n} |c_k|^2 \quad \sum_{k=0}^{n} |p_k(z_0)|^2}.$$

Hemos visto que $\|Q_n(z)\| = \sum_{k=0}^{n} |c_k|^2$. Luego elevando al cuadrado podemos escribir

$$1 \leq \|Q(z)\|^2 \sum_{k=0}^{n} |p_k(z_0)|^2 = \|Q(z)\|^2 K_n(z_0, z_0). \tag{17}$$

La igualdad[7] se alcanza, si y solo si,

$$c_k = A \, \overline{p_k(z_0)}, \quad k = 0, 1, \ldots, n.$$

Por hipótesis $Q_n(z_0) = 1$, por tanto se debe cumplir que

$$Q_n(z_0) = \sum_{k=0}^{n} c_k \, p_k(z_0) = \sum_{k=0}^{n} \left(A \, \overline{p_k(z_0)} \right) p_k(z_0) = 1 \quad \Rightarrow \quad A = \frac{1}{\sum_{k=0}^{n} |p_k(z_0)|^2}.$$

Luego

$$c_j = A \, \overline{p_j(z_0)} = \frac{\overline{p_j(z_0)}}{\sum_{k=0}^{n} |p_k(z_0)|^2}.$$

Por tanto

$$Q_n(z) = \sum_{j=0}^{n} c_j \; p_j(z) = \frac{\sum_{j=0}^{n} \overline{p_j(z_0)} p_j(z)}{\sum_{k=0}^{n} |p_k(z_0)|^2} = \frac{K_n(z_0, z)}{K_n(z_0, z_0)}.$$

Hemos visto en (17) que

$$K_n(z_0, z_0)^{-1} \leq \|Q_n(z)\|^2,$$

alcanzandose la igualdad –es decir el mínimo de la norma de Q_n –, si y solo si, $A = K_n(z_0, z_0)^{-1}$, cuando

$$Q_n(z) = \frac{K_n(z_0, z)}{K_n(z_0, z_0)},$$

y obviamente la norma de ese mínimo al cuadrado vale $1/K_n(z_0, z_0)$. □

La siguiente observación justifica la particular notación fila-columna para las matrices del producto escalar.

Observación 2.2.

La notación históricamente usual para la matriz del producto escalar proviene de su origen como matrices de momentos y es, como ya hemos visto, contraria a la habitual fila-columna, es decir

$$M = \begin{pmatrix} c_{0,0} & c_{1,0} & c_{2,0} & \cdots \\ c_{0,1} & c_{1,1} & c_{2,1} & \cdots \\ c_{0,2} & c_{1,2} & c_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad c_{jk} = \int z^j \overline{z}^k = \langle z^j, z^k \rangle_M.$$

[7]La desigualdad de Cauchy-Schwarz para suma de productos complejos con resultado real.

El producto escalar sobre un espacio vectorial complejo debe verificar $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$, y $\langle u, \beta v\rangle = \overline{\beta}\langle u, v\rangle$, por lo que el producto de dos polinomios de $p_1(z) = a_0 + a_1 z$, por $q_1(z) = b_0 + b_1 z$ se podría expresar matricialmente de las dos posibles formas que señalamos a continuación

$$
\begin{aligned}
\langle p_1(z), q_1(z)\rangle_M = & \int (a_0 + a_1 z)(\overline{b_0 + b_1 z}) d\mu(z) = \\
& = a_0\overline{b}_0 c_{0,0} + a_0\overline{b}_1 c_{0,1} + a_1\overline{b}_0 c_{1,0} + a_1\overline{b}_1 c_{1,1} = \\
& = (\overline{b}_0, \overline{b}_1, 0, \ldots) \begin{pmatrix} c_{0,0} & c_{1,0} & \cdots \\ c_{0,1} & c_{1,1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ 0 \\ \vdots \end{pmatrix} = \\
& = (a_0, a_1, 0, \ldots) \begin{pmatrix} c_{0,0} & c_{0,1} & \cdots \\ c_{1,0} & c_{1,1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \overline{b}_0 \\ \overline{b}_1 \\ 0 \\ \vdots \end{pmatrix}.
\end{aligned}
$$

Observación 2.3. Antes de probar la expresión del núcleo mediante determinantes, y con carácter puramente algebraico, es preciso recordar que los n-núcleos tienen capacidad reproductora, i.e. para cualquier $Q(x)$, polinomio de grado a lo sumo n, se cumplen

$$
\begin{aligned}
\langle Q(x), K_n(x, y)\rangle &= Q(y), \\
\langle K_n(x, y), Q(y)\rangle &= \overline{Q(x)}.
\end{aligned}
$$

Obviamente $\langle K_n(x, y), Q(x)\rangle = \overline{Q(y)}$, etc., además el n-núcleo es único, ya que si otra función de dos variables $H_n(x, y)$ tuviera capacidad reproductora se cumplirían

$$
\begin{aligned}
\langle H_n(x, z), K_n(x, y)\rangle &= H_n(y, z), \\
\langle H_n(x, z), K_n(x, y)\rangle &= \overline{K_n(z, y)} = K_n(y, z).
\end{aligned}
$$

En consecuencia $K_n(y, z) = H_n(y, z)$, $\forall y, z \in \mathbb{C}$.

Proposición 2.4. *Se cumple*[8] *que*

$$
\begin{aligned}
K_n(x, y) &= \sum_{j=0}^{n} \overline{p_j(x)} p_j(y) = -\frac{1}{\Delta_n} \begin{vmatrix} c_{00} & c_{10} & \ldots & c_{n,0} & 1 \\ c_{01} & c_{11} & \ldots & c_{n,1} & \overline{x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \ldots & c_{n,n} & \overline{x}^n \\ 1 & y & \ldots & y^n & 0 \end{vmatrix} \\
&= (1, y, \ldots, y^n) M_{n+1}^{-1} \begin{pmatrix} 1 \\ \overline{x} \\ \vdots \\ \overline{x}^n \end{pmatrix}.
\end{aligned}
$$

[8] Se denomina complemento Schur (pág. 46 [Lan85]) o identidad de Magnus (pág. 52 [Bre80], a la siguiente fórmula. Si A y D son matrices cuadradas y $|A| \neq 0$, entonces:

$$
\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D - CA^{-1}B|\,|A|.
$$

Dem. Introduzcamos la función de dos variables

$$H_n(x,y) = -\frac{1}{\Delta_n}\begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} & 1 \\ c_{01} & c_{11} & \dots & c_{n,1} & \overline{x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \dots & c_{n,n} & \overline{x}^n \\ 1 & y & \dots & y^n & 0 \end{vmatrix}.$$

Resulta inmediato que

$$\langle x^m, H_n(x,y)\rangle = -\frac{1}{\Delta_n}\begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} & c_{m0} \\ c_{01} & c_{11} & \dots & c_{n,1} & c_{m1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \dots & c_{n,n} & c_{mn} \\ 1 & y & \dots & y^n & 0 \end{vmatrix}.$$

Restamos de la última columna la columna $(m+1)$–ésima y resulta

$$\langle x^m, H_n(x,y)\rangle = -\frac{1}{\Delta_n}\begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} & 0 \\ c_{01} & c_{11} & \dots & c_{n,1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \dots & c_{n,n} & 0 \\ 1 & y & \dots & y^n & -y^m \end{vmatrix} = y^m.$$

De modo análogo

$$\langle H_n(x,y), y^m\rangle = -\frac{1}{\Delta_n}\begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} & 1 \\ c_{01} & c_{11} & \dots & c_{n,1} & \overline{x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \dots & c_{n,n} & \overline{x}^n \\ c_{0m} & c_{1m} & \dots & c_{n,m} & 0 \end{vmatrix}.$$

Restando de la última fila la fila $(m+1)$–ésima tendremos que

$$\langle H_n(x,y), y^m\rangle = -\frac{1}{\Delta_n}\begin{vmatrix} c_{00} & c_{10} & \dots & c_{n,0} & 1 \\ c_{01} & c_{11} & \dots & c_{n,1} & \overline{x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \dots & c_{n,n} & \overline{x}^n \\ 0 & 0 & \dots & 0 & -\overline{x}^m \end{vmatrix} = \overline{x}^m.$$

Esta función de dos variables puede extenderse por linealidad a cualquier polinomio. En otras palabras $H_n(x,y)$ tiene capacidad reproductora, y por lo visto en la observación previa debe coincidir con $K_n(x,y)$. La segunda parte de la afirmación se explica en el pie de página correspondiente. □

La propiedad de mínimo de los núcleos podemos verla ahora desde este punto de vista.

Corolario 2.1. *El polinomio*

$$Q_n(y) = \frac{K_n(z_0, y)}{K_n(z_0, z_0)},$$

cumple obviamente $Q_n(z_0) = 1$ *y además tiene norma mínima, que justamente es*

$$\left\| \frac{K_n(z_0, y)}{K_n(z_0, z_0)} \right\|^2 = \frac{1}{K_n(z_0, z_0)}.$$

DEM. Si en la fórmula (5) del Teorema 1.5, escogemos $n := n + 1$, y hacemos $x_k := y^{k-1}$ e $y := K_n(x, y)$, resulta que la última columna toma el valor

$$\langle y, x_{j+1} \rangle = \langle K_n(x, y), y^j \rangle = \overline{x}^j, \quad j = 0, 1, \ldots, n$$

y queda la fórmula

$$K_n(x, y) = -\frac{1}{\Delta_n} \begin{vmatrix} c_{00} & c_{10} & \ldots & c_{n,0} & 1 \\ c_{01} & c_{11} & \ldots & c_{n,1} & \overline{x} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \ldots & c_{n,n} & \overline{x}^n \\ 1 & y & \ldots & y^n & 0 \end{vmatrix}.$$

Si dividimos ambos términos por $K_n(z_0, z_0)$ y particularizamos $x := z_0$, resulta

$$Q_n(y) = \frac{K_n(z_0, y)}{K_n(z_0, z_0)},$$

que obviamente cumple $Q_n(z_0) = 1$ y además su norma al cuadrado será

$$\langle \frac{K_n(z_0, y)}{K_n(z_0, z_0)}, \frac{K_n(z_0, y)}{K_n(z_0, z_0)} \rangle = \frac{K_n(z_0, z_0)}{K_n(z_0, z_0)^2} = \frac{1}{K_n(z_0, z_0)}.$$

□

Definición 2.2. *Notaremos con* $\Delta_n^{(m)}$ *al determinante de orden* $n + 1$*, tras eliminar* m *filas y* m *columnas en la matriz* M*.*

Ejemplo 2.1. Nótese que de la anterior expresión para los núcleos se deduce inmediatamente que desarrollando por la última columna y luego por la última fila resulta

$$K_n(0, 0) = -\frac{1}{\Delta_n} \begin{vmatrix} c_{00} & c_{10} & \ldots & c_{n,0} & 1 \\ c_{01} & c_{11} & \ldots & c_{n,1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n} & c_{1,n} & \ldots & c_{n,n} & 0 \\ 1 & 0 & \ldots & 0 & 0 \end{vmatrix} = -\frac{1}{\Delta_n}(-1)^{2n+1}\Delta_{n-1}^{(1)} = \frac{\Delta_{n-1}^{(1)}}{\Delta_n}.$$

Donde

$$\Delta_{n-1}^{(1)} = \begin{vmatrix} c_{11} & c_{21} & \ldots & c_{n,1} \\ c_{12} & c_{22} & \ldots & c_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1,n} & c_{2,n} & \ldots & c_{n,n} \end{vmatrix}.$$

Nótese que si M es de Toeplitz $\Delta_{n-1}^{(1)} = \Delta_{n-1}$ y se tiene que $K_n(0, 0) = \kappa_n^2$. De hecho M es de Toeplitz, si y solo si, $K_n(0, 0) = \kappa_n^2$, $\forall n \in \mathbb{Z}_+$.

Cuando se trabaja con matrices de Toeplitz se suele notar $\varphi_n(z) = p_n(z)$ y $\Phi_n(z) = P_n(z)$, es decir respectivamente a los normalizados y a los mónicos. En ese contexto se introduce también el denominado polinomio reverso.

Definición 2.3. *Dado un polinomio $q_n(z) = \sum_{k=}^{n} a_k z^k$, polinomio como mucho de grado n, llamaremos reverso de ese polinomio y se notará como $q_n^*(z)$ al polinomio $q_n^*(z) = \sum_{k=0}^{n} \overline{a}_k z^{n-k} = z^n \overline{q}_n(1/z)$, donde $\overline{q}_n(z)$ es el polinomio que resulta de conjugar los coeficientes de $q_n(z)$. Obviamente $[q_n^*(z)]^* = q_n(z)$.*

Sabemos que el polinomio con coeficiente conductor unidad, que hace mínima su norma al cuadrado, es el mónico de la SPOM. Su norma coincide[9] con la de su reverso–que por tanto también es de norma mínima– pero su reverso pertenece, además, a aquellos polinomios $q_n(z)$ de grado n que verifican $q_n(0) = 1$. El polinomio solución que minimiza la norma al cuadrado de los polinomios con esa propiedad vimos que era justamente $K_n(z,0)/K_n(0.0)$. Ambos polinomios con norma mínima coinciden, de ahí que los mínimos que alcanzan valgan

$$\frac{1}{K_n(0,0)} = \frac{1}{\kappa_n^2}.$$

En cuanto a los polinomios, lo que ocurre es que el reverso del mónico coincide con el polinomio $K_n(z,0)/K_n(0,0)$, de donde, como $K_n(0,0) = \kappa_n^2$, resulta que

$$\left(\frac{\varphi_n(z)}{\kappa_n}\right)^* = \frac{K_n(z,0)}{K_n(0,0)} \quad \Rightarrow \quad K_n(z,0) = \kappa_n \varphi_n(z)^*. \tag{18}$$

Hemos probado el siguiente corolario.

Corolario 2.2. *Si M es Toeplitz se cumplen.*

i) $K_n(0,0) = \kappa_n^2$.

ii) $K_n(z,0) = \kappa_n \varphi_n(z)^*$.

3 Construcción de la matriz D

3.1 Propiedades. Generalización de la matriz de Jacobi

Sea $M = (c_{ij})_{i,j=0}^{\infty}$, matriz hermitiana infinita, definida positiva,

$$M = \begin{pmatrix} c_{0,0} & c_{1,0} & \cdots & c_{n-1,0} & c_{n,0} & \cdots \\ c_{0,1} & c_{1,1} & \cdots & c_{n-1,1} & c_{n,1} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ c_{0,n-1} & c_{1,n-1} & \cdots & c_{n-1,n-1} & c_{n,n-1} & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \end{pmatrix}.$$

[9] Las matrices de Toeplitz finitas de orden n tienen la propiedad de que operando con la contraidentidad I_n^c, resulta $I_n^c T_n I_n^c = \overline{T}_n$. La conclusión es que $\|p_n\| = \|p_n^*\|$.

En lo que sigue llamaremos M' a la matriz que resulta de eliminar la primera columna en la matriz M. Las matrices finitas M_n y M'_n serán las secciones de orden n de M y M' respectivamente, *i.e.*, las restricciones a las primeras n filas y columnas.

Proposición 3.1. *Sea $\{Q_n(z)\}$ la S.P.O. asociada*[10] *a M, entonces se cumple que las n raíces de $Q_n(z)$ son los autovalores de la matriz $M_n^{-1}M'_n$.*

Dem. Partimos de la conocida expresión para una base ortogonal obtenida por el procedimiento de Gram-Schmidt de la base canónica $\{1, z, z^2, z^3, \ldots\}$,

$$Q_n(z) = \begin{vmatrix} c_{0,0} & c_{1,0} & \cdots & c_{n-1,0} & c_{n,0} \\ c_{0,1} & c_{1,1} & \cdots & c_{n-1,1} & c_{n,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{0,n-1} & c_{1,n-1} & \cdots & c_{n-1,n-1} & c_{n,n-1} \\ 1 & z & \ldots & z^{n-1} & z^n \end{vmatrix}.$$

Si en este determinante multiplicamos la primera columna por z, le restamos la segunda, y dividimos todo el determinante por z para que no varíe resulta

$$Q_n(z) = \frac{1}{z} \begin{vmatrix} zc_{0,0} - c_{1,0} & c_{1,0} & \cdots & c_{n-1,0} & c_{n,0} \\ zc_{1,0} - c_{1,1} & c_{1,1} & \cdots & c_{n-1,1} & c_{n,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ zc_{0,n-1} - c_{1,n-1} & c_{1,n-1} & \cdots & c_{n-1,n-1} & c_{n,n-1} \\ 0 & z & \ldots & z^{n-1} & z^n \end{vmatrix}.$$

repitiendo este proceso $n-1$ veces (hasta la penúltima columna) queda

$$\begin{aligned} Q_n(z) &= \frac{1}{z^n} \begin{vmatrix} zc_{0,0} - c_{1,0} & zc_{1,0} - c_{1,1} & \cdots & zc_{n-1,1} - c_{n,0} & c_{n,0} \\ zc_{0,1} - c_{1,1} & zc_{1,1} - c_{2,1} & \cdots & zc_{n-1,2} - c_{n,1} & c_{n,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ zc_{0,n-1} - c_{1,n} & zc_{1,n-1} - c_{2,n-1} & \cdots & zc_{n-1,n-1} - c_{n,n-1} & c_{n,n-1} \\ 0 & 0 & \ldots & 0 & z^n \end{vmatrix} \\ &= \begin{vmatrix} zc_{0,0} - c_{1,0} & zc_{1,0} - c_{1,1} & \cdots & zc_{n-1,1} - c_{n,0} \\ zc_{0,1} - c_{1,1} & zc_{1,1} - c_{2,1} & \cdots & zc_{n-1,2} - c_{n,1} \\ \vdots & \vdots & \ddots & \vdots \\ zc_{0,n-1} - c_{1,n} & zc_{1,n-1} - c_{2,n-1} & \cdots & zc_{n-1,n-1} - c_{n,n-1} \end{vmatrix}. \end{aligned}$$

La expresión anterior puede encontrarse en la pág. 27 de [Sze39]. Llamando[11] M_n y M'_n a las matrices correspondientes, podemos escribir la anterior relación como $Q_n(z) = |zM_n - M'_n|$. Luego los ceros de $Q_n(z)$ son aquellos valores de z que resuelven

$$|zM_n - M'_n| = 0. \tag{19}$$

[10] Por comodidad no es ni la de los mónicos ni la de los normalizados, serían los mónicos multiplicados por Δ_{n-1}.

[11] Los sucesivos determinantes se suelen determinar Δ_n, de acuerdo con esa notación tenemos que $\Delta_n = |M_{n+1}|$ y $\Delta'_n = |M'_{n+1}|$.

Si suponemos que $|M_n| \neq 0$, $\forall n$, podemos multiplicar $Q_n(z) = |zM_n - M_n'| = 0$, por $|M_n^{-1}|$ y tendremos

$$|M_n^{-1}|\, Q_n(z) = |M_n^{-1}|\, |zM_n - M_n'| = 0.$$

aplicando Binet-Cauchy y agrupando el producto de determinantes resulta que

$$|M_n^{-1}|\, Q_n(z) = |zI_n - M_n^{-1}M_n'| = 0.$$

Es evidente que los valores de z que hacen nulo el polinomio $Q_n(z)$ son a su vez los autovalores de la matriz $M_n^{-1}M_n'$. $\square$

Observación 3.1. Evidentemente en la proposición anterior no se ha utilizado para nada que sea definida positiva ni tampoco que sea hermitiana. Lo único que hace falta es que la matriz M sea fuertemente regular, es decir que $|M_n| \neq 0$, $\forall n$. Esto permitiría extender este método a matrices de Hankel complejas, no necesariamente definidas positivas, pero si fuertemente regulares.

Analicemos ahora la matriz $M_n^{-1}M_n'$, en lo que sigue probaremos que es una matriz de Fröbenius, es decir tiene la estructura que señalamos en la siguiente proposición, y corresponde a los coeficientes del polinomio mónico P_n de la sucesión.

Proposición 3.2. *Se cumple que*

$$M_n^{-1}M_n' = \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{n-1} \\ 0 & 0 & \dots & 1 & a_n \end{pmatrix}.$$

Dem. En efecto fraccionemos M_n en cajas del siguiente modo

$$M_n = \left(\begin{array}{c|ccc} c_{0,0} & c_{1,0} & \dots & c_{n-1,0} \\ \hline c_{0,1} & c_{1,1} & \dots & c_{n-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{0,n-1} & c_{1,n-1} & \dots & c_{n-1,n-1} \end{array}\right) = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix},$$

es decir

$$\begin{aligned} Q_{11} &= c_{0,0} \\ Q_{12} &= (c_{1,0}, c_{2,0}, \dots, c_{n-1,0}) \\ Q_{21} &= (c_{0,1}, c_{0,2}, \dots, c_{0,n-1})^t \\ Q_{22} &= (c_{ij})_{i,j=1}^{n-1}. \end{aligned}$$

La matriz M_n' puede escribirse utilizando algunas de las cajas anteriores en la forma

$$M_n' = \begin{pmatrix} Q_{12} & S \\ Q_{22} & T \end{pmatrix}$$

donde

$$\begin{array}{rcl} S & = & (c_{n,0}) \\ T & = & (c_{n,1}, c_{n,2}, \dots, c_{n,n-1})^t. \end{array}$$

Si invertimos por cajas M_n, de modo que

$$M_n^{-1} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

se cumplirá que $M_n^{-1} M_n = I_n$, lo que traducido a las cajas será:

$$\begin{array}{rcl} R_{11}Q_{11} + R_{12}Q_{21} & = & 1 \\ R_{11}Q_{12} + R_{12}Q_{22} & = & (0, 0, \dots, 0)_{n-1} \\ R_{21}Q_{11} + R_{22}Q_{21} & = & (0, 0, \dots, 0)_{n-1} \\ R_{21}Q_{12} + R_{22}Q_{22} & = & I_{n-1}. \end{array}$$

Por tanto al efectuar el producto $M_n^{-1} M_n'$ aparecerán algunas cajas que sabemos cuanto valen. Tendremos

$$\begin{array}{rcl} M_n^{-1} M_n' & = & \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} Q_{12} & S \\ Q_{22} & T \end{pmatrix} = \\ & = & \begin{pmatrix} R_{11}Q_{12} + R_{12}Q_{22} & R_{11}S + R_{12}T \\ R_{21}Q_{12} + R_{22}Q_{22} & R_{21}S + R_{22}T \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{n-1} \\ 0 & 0 & \dots & 1 & a_n \end{pmatrix}. \end{array}$$

□

Corolario 3.1. *Si con* $P_n(z) = a_{n,0} + a_{n,1}z + a_{n,2}z^2 + \ldots + a_{n,n-1}z^{n-1} + z^n$, *denotamos el polinomio mónico de la S.P.O. entonces se cumple que*

$$P_n(z) = |zI_n - F_n|,$$

con

$$F_n = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_{n,0} \\ 1 & 0 & \dots & 0 & -a_{n,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -a_{n,n-2} \\ 0 & 0 & \dots & 1 & -a_{n,n-1} \end{pmatrix}.$$

Dem. Esto es evidente ya que

$$\begin{array}{rcl} P_n(z) & = & |M_n^{-1}|\, Q_n(z) = |zI_n - M_n^{-1} M_n'| \\ & = & |zI_n - F_n|, \end{array}$$

y sabemos que la matriz compañera[12] de un polinomio mónico dado siempre toma la forma arriba señalada. □

[12] Aquella que tiene como polinomio característico dicho polinomio.

Definición 3.1. *Sea* $M = (c_{ij})_{i,j=0}^{\infty}$ *una matriz compleja HDP infinita. Si* $M_n = T_n T_n^H$ *es la*[13] *descomposición de Choleski de* M_n*, por ser* M_n *matriz HDP, entonces* T_n *es invertible. Definimos*

$$D_n = T_n^{-1} M_n' T_n^{-H}, \qquad n = 1, 2, 3, \ldots$$

Dada una matriz A_n *de orden* n*, sea* $k \in \mathbb{N}, k < n$*. Llamaremos* $(A_n)_k$ *a la matriz que resulta de tomar en* A_n *las* k *primeras filas y columnas.*

Proposición 3.3. *Se cumple que*

$$D_n = T_n^{-1} M_n' T_n^{-H} = T_n^H F_n T_n^{-H},$$

y además

$$P_n(z) = |I_n z - D_n|, \qquad n = 1, 2, 3, \ldots$$

Dem. Hemos visto en la proposición anterior que

$$M_n^{-1} M_n' = F_n, \tag{20}$$

Como $M_n = T_n T_n^H$, resulta que $M_n^{-1} = T_n^{-H} T_n^{-1}$, substituyendo en (20) tenemos que

$$T_n^{-H} T_n^{-1} M_n' = F_n.$$

Multipliquemos por la izquierda por T_n^H y por la derecha por T_n^{-H}, resulta

$$T_n^H T_n^{-H} T_n^{-1} M_n' T_n^{-H} = T_n^H F_n T_n^{-H},$$

y simplificando queda finalmente

$$D_n = T_n^{-1} M_n' T_n^{-H} = T_n^H F_n T_n^{-H}.$$

El segundo apartado es inmediato ya que la matriz D_n es semejante a la F_n, y por tanto tienen el mismo polinomio característico. $\square$

Proposición 3.4. *La matriz* D_n *es una matriz de Hessenberg superior (es decir* $d_{ij} = 0$ *si* $i > j + 1$*). Además la subdiagonal verifica*

$$d_{i,i-1} = \sqrt{\frac{|M_i|\,|M_{i-2}|}{|M_{i-1}|^2}}, \qquad 2 \leq i \leq n,$$

con $|M_0| = 1$ *y* $n \geq 2$

[13] Sea $a \in \mathbb{C}$ tal que $|a| = 1$, se cumple que $a\overline{a} = 1$, la descomposición de Choleski no es única en cuanto que $M_n = (aT_n)(aT_n)^H$. Esta falta de unicidad no tendrá consecuencias para nosotros ya que esos posibles factores se cancelan en el momento de definir la matriz D_n, por lo que para nuestras necesidades podemos considerarla única. Utilizaremos la notación T^H en vez de la más familiar T^*, en particular cuando tengamos que invertir y al abordar las transformaciones de semejanza.

Dem. La matriz T_n es triangular inferior, como es bien sabido T_n^{-1} también es triangular inferior. De modo análogo T_n^H es triangular superior y en consecuencia también lo es T_n^{-H}. En la descomposición de Choleski de M_n, la diagonal de T_n, $\{t_{ii}\}_{i=1}^n$, es la misma que la de T_n^H, la diagonal de T_n^{-H} será $\{1/t_{ii}\}_{i=1}^n$.

Desarrollando la igualdad $D_n = T_n^H F_n T_n^{-H}$ tendremos

$$D_n = \begin{pmatrix} t_{11} & \square & \dots & \square \\ 0 & t_{22} & \dots & \square \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ 0 & 1 & \dots & 0 & a_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & a_n \end{pmatrix} \begin{pmatrix} 1/t_{11} & \square & \dots & \square \\ 0 & 1/t_{22} & \dots & \square \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/t_{nn} \end{pmatrix}.$$

Efectuando el producto de estas tres matrices queda

$$D_n = \begin{pmatrix} \square & \square & \dots & \square & \square \\ t_{22}/t_{11} & \square & \dots & \square & \square \\ 0 & t_{33}/t_{22} & \dots & \square & \square \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & t_{nn}/t_{n-1,n-1} & \square \end{pmatrix},$$

luego D_n es una matriz de Hessenberg (triangular+1) superior. Si efectuamos la descomposición de Choleski y de Gauss-Banachiewicz simultáneamente[14] de M_n resulta

$$M_n = T_n T_n^H = V_n E_n V_n^H,$$

con $E_n = (\delta_{ij} e_{i-1})_{i,j=1}^n$, donde como sabemos $e_{i-1} = |M_i|/|M_{i-1}|$ y $|M_0| = 1$.

Al ser M_n definida positiva podemos hablar de la raíz positiva de E_n y tendremos que

$$T_n T_n^H = V_n \sqrt{E_n} \sqrt{E_n} V_n^H.$$

Donde $V_n = (v_{ij})_{i,j=1}^n$ es triangular inferior, verificando además $v_{ii} = 1$, $i = 1, 2, \dots, n$. La unicidad con esta condición está asegurada y conduce a que $T_n = V_n \sqrt{E_n}$ y por tanto $t_{ii} = \sqrt{e_{i-1}}$. Substituyendo queda

$$d_{i,i-1} = \frac{\sqrt{e_{i-1}}}{\sqrt{e_{i-2}}} = \sqrt{\frac{|M_i|\,|M_{i-2}|}{|M_{i-1}|^2}}, \qquad 2 \le i \le n,$$

con $|M_0| = 1$ y $n \ge 2$.

$\square$

Observación 3.2. Que M_n es submatriz de M_{n+1} es una obviedad ya que en el trasfondo la matriz infinita M existe. Esto no es cierto en general cuando se obtienen matrices que dependen de n. Por ejemplo es claro que salvo en casos muy especiales

$$(A_n^{-1})_k \ne A_k^{-1}.$$

[14] En realidad subordinamos la descomposición de Choleski a la de Gauss-Banachiewicz para evitar los problemas de unicidad antes mencionados. Es decir cuando llegue el momento escogeremos entre todas las posibilidades de descomposición de Choleski (para nosotros es lo mismo) aquella que verifica $T_n = V_n\sqrt{E_n}$.

Uno de esos casos es justamente el de las matrices triangulares invertibles. Es decir se cumple que

$$(T_n^{-1})_k = T_k^{-1}, \qquad 1 \le k \le n.$$

Esta propiedad es válida para las triangulares inferiores y las superiores.

Proposición 3.5. *Se cumple que*

$$D_n = (D_{n+1})_n, \qquad n = 2, 3, 4, \ldots$$

Dem. Por construcción $M_n = (M_{n+1})_n$ y $M_n' = (M_n')_n$. Por otro lado $M_{n+1} = T_{n+1}T_{n+1}^H$. Por ser triangulares T_n es submatriz de T_{n+1}, mientras que T_n^H lo es de T_{n+1}^H. La misma triangularidad dice que T_n^{-1} es submatriz de T_{n+1}^{-1}, y del mismo modo T_n^{-H} es submatriz de T_{n+1}^{-H}. Con la notación introducida $T_n^{-1} = (T_{n+1}^{-1})_n$ y $T_n^{-H} = (T_{n+1}^{-H})_n$. Se cumple que

$$D_n = T_n^{-1} M_n' T_n^{-H}, \qquad D_{n+1} = T_{n+1}^{-1} M_{n+1}' T_{n+1}^{-H}.$$

Tenemos que $M_n' = (M_{n+1}')_n$, y por la triangularidad de T_{n+1}^{-1} y T_{n+1}^{-H} resulta que $D_n = (D_{n+1})_n$. □

Observación 3.3. Este resultado es importante porque dada M infinita HDP, podemos hablar de la matriz infinita de Hessenberg D asociada a ella.

Proposición 3.6. *Si M es una matriz de Hankel definida positiva, entonces la matriz D es la tridiagonal de Jacobi correspondiente a M.*

Dem. Si M es de Hankel y definida positiva, entonces es real (ver ejercicio 3.7 de la pág. 18 de [Chi78]). Es decir M_n es una matriz simétrica y real. Las matrices de la descomposición de Choleski son reales. Por tanto podemos escribir

$$M_n = T_n T_n^t.$$

Por otro lado si M_n es de Hankel y simétrica también lo es M_n', es decir $M_n' = (M_n')^t$. Tenemos ahora que

$$D_n = T_n^{-1} M_n' T_n^{-t}.$$

Calculemos D_n^t, resulta

$$\begin{aligned} D_n^t &= (T_n^{-1} M_n' T_n^{-t})^t \\ &= T_n^{-1} (M_n')^t T_n^{-t} = T_n^{-1} M_n' T_n^{-t} = D_n. \end{aligned}$$

Luego D_n es simétrica, como era de Hessenberg, ahora necesariamente es tridiagonal y simétrica.

Resulta inmediato comprobar, desarrollando por la última columna

$$P_n(z) = |I_n z - D_n| = \begin{vmatrix} z - d_{11} & -d_{12} & \ldots & 0 & 0 \\ -d_{12} & z - d_{22} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & z - d_{n-1,n-1} & -d_{n,n-1} \\ 0 & 0 & \ldots & -d_{n,n-1} & z - d_{n,n} \end{vmatrix},$$

que obtenemos la fórmula de recurrencia a trés términos

$$P_n(z) = (z - d_{nn})P_{n-1}(z) - d^2_{n,n-1}P_{n-2}(z),$$

que como sabemos permite construir la tridiagonal de Jacobi. Luego esta tridiagonal es justamente la matriz J, es decir si M es de Hankel y HDP entonces $D = J$. □

Proposición 3.7. *Si $\alpha \in \mathbb{R}$, $\alpha \neq 0$, y llamamos $D_n(M_{n+1})$ a la matriz D_n que se obtiene a partir de la matriz*[15] *M_{n+1}. Entonces se cumple que*

$$D_n(\alpha M_{n+1}) = D_n(M_{n+1}).$$

Dem. En efecto

$$\alpha M_n = (\sqrt{\alpha}T_n)(\sqrt{\alpha}T_n)^H,$$

luego

$$\begin{aligned} D_n(\alpha M_{n+1}) &= (\sqrt{\alpha}T_n)^{-1}(\alpha M'_n)(\sqrt{\alpha}T_n)^{-H} = \frac{1}{\alpha}T_n^{-1}(\alpha M'_n)T_n^{-H} \\ &= T_n^{-1}M'_nT_n^{-H} = D_n(M_{n+1}). \end{aligned}$$

□

3.2 La descomposición LR y la matriz $\widetilde{D}$ con unos en la subdiagonal

Observación 3.4. Si la matriz M no es hermitiana definida positiva no es posible efectuar la descomposición de Choleski de sus secciones, no obstante si M es fuertemente regular podemos todavía descomponer M_n en producto de triangular inferior por triangular superior o descomposición LR. Esta descomposición es única si exigimos $r_{k,k} = 1$, $k = 1, 2, \ldots, n$.

En ese caso introducimos una matriz que denominaremos $\widetilde{D}$ para distinguirla de la D cuyas propiedades veremos en lo que sigue.

Definición 3.2. *Sea M una matriz fuertemente regular, y tal que para todo n, $M_n = L_nR_n$ sea la descomposición LR de M_n con la condición de que $r_{ii} = 1$, $i = 1, 2, \ldots, n$. Llamaremos*

$$\widetilde{D}_n = L_n^{-1}M'_nR_n^{-1}, \qquad n = 1, 2 \ldots$$

Proposición 3.8. *Se cumplen las siguientes propiedades:*

i) $P_n(z) = |I_nz - \widetilde{D}_n|$, $n = 1, 2, \ldots$.

ii) $\widetilde{D}_n$ *es una matriz de Hessenberg superior y* $d_{i,i-1} = 1$, $i = 2, 3, \ldots n$.

iii) $(\widetilde{D}_{n+1})_n = \widetilde{D}_n$.

[15] Para poder construirla hace falta M_n y M'_n, además de M_n necesitamos la columna $(c_{n,0}, c_{n,1}, \ldots, c_{n,n-1})^t$ de M_{n+1}, no es preciso conocer, sin embargo, el elemento $c_{n,n}$ de M_{n+1}.

iv) Si M es HDP entonces se cumple que

$$D_n = \sqrt{E_n}\,\widetilde{D}_n(\sqrt{E_n})^{-1},$$

con $E_n = (\delta_{ij}e_{i-1})_{i,j=1}^n$, con $e_{i-1} = |M_i|/|M_{i-1}|$ y $|M_0| = 1$.

Dem.

i) Sabemos que $F_n = M_n^{-1}M'_n$, como $M_n = L_nR_n$ y ambas son invertibles resulta que $M_n^{-1} = R_n^{-1}L_n^{-1}$, de donde

$$F_n = R_n^{-1}L_n^{-1}M'_n,$$

multiplicando a la izquierda por R_n y por R_n^{-1} a la derecha queda finalmente

$$\widetilde{D}_n = L_n^{-1}M'_nR_n^{-1} = R_nF_nR_n^{-1}.$$

La matriz $\widetilde{D}_n$ es semejante a F_n luego ambas tienen el mismo polinomio característico, como el de F_n es justamente $\widetilde{P}_n(z)$ se cumple que $\widetilde{P}_n(z) = |I_nz - \widetilde{D}_n|$.

Las propiedades ii) y iii) son análogas a las de las Proposiciones 3.4 y 3.5, aunque en el caso de la primera la subdiagonal este formada por unos.

iv) Finalmente utilizando de nuevo la descomposición de Gauss-Banachiewicz de las secciones de M, que es una matriz infinita HDP, tenemos

$$M_n = V_nE_nV_n^H,$$

con $v_{ii} = 1$, $i = 1, 2, \ldots, n$. Como es bien sabido $E_n = (\delta_{ij}e_{i-1})_{i,j=1}^n$ con $e_{i-1} = |M_i|/|M_{i-1}|$ y $|M_0| = 1$.

Es claro que si $M_n = L_nR_n$ con $r_{ii} = 1$, $i = 1, 2, \ldots, n$, resulta que $V_n^H = R_n$ y además $T_n^H = \sqrt{E_n}R_n = \sqrt{E_n}V_n^H$. Es decir $V_n^H = (\sqrt{E_n})^{-1}T_n^H$. Hemos visto que $\widetilde{D}_n = R_nF_nR_n^{-1}$, substituyendo R_n, como E_n es una matriz diagonal, real y con elementos positivos, cumple que $(\sqrt{E_n})^H = \sqrt{E_n}$, y tendremos que

$$\begin{aligned}
\widetilde{D}_n &= R_nF_nR_n^{-1} = V_n^HF_nV_n^{-H} \\
&= (\sqrt{E_n})^{-1}T_n^HF_nT_n^{-H}\sqrt{E_n} \\
&= (\sqrt{E_n})^{-1}D_n\sqrt{E_n}.
\end{aligned}$$

Despejando finalmente D_n queda

$$D_n = \sqrt{E_n}\,\widetilde{D}_n(\sqrt{E_n})^{-1}.$$

□

4 Ejemplos

4.1 Medida de Lebesgue en $\mathbb{T}$

Uno de los ejemplos más simples consiste en tomar como M la matriz identidad. En este caso

$$M_n = T_n = T_n^H = T_n^{-1} = T_n^{-H} = I_n, \qquad \forall n.$$

y como

$$M' = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

resulta, tras aplicar $D_n = T_n^{-1} M_n' T_n^{-H}$ que

$$M = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad \Longrightarrow \qquad D = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Si nos preguntamos si la matriz M es de momentos. La respuesta es que si. Consideremos la medida de Lebesgue sobre la circunferencia unidad, es decir

$$\begin{aligned} c_{jk} &= \int_{\mathbb{T}} z^j \overline{z}^k |dz| = \int_0^{2\pi} [e^{i\theta}]^j [e^{-i\theta}]^k d\theta \\ &= \int_0^{2\pi} [e^{i\theta}]^{j-k} d\theta = 2\pi \delta_{jk}. \end{aligned}$$

Vemos que la matriz de momentos es la matriz identidad multiplicada por la constante 2π (que no afecta al cálculo de la D). Resultados menos obvios nos dicen que los polinomios ortogonales respecto de esa medida, que son $\{z_k\}_{k=0}^{\infty}$, no forman un sistema completo, es decir $\overline{\Pi} \neq L^2_\mu(\mathbb{T})$, pero esta medida $\{z^k\}_{k=0}^{\infty}$ si es una base de Schauder.

El operador definido en ℓ^2 por la matriz D (el shift-right) es un operador subnormal (tiene una extensión[16] normal), y su espectro es $\sigma(D) = \{z | z \in \mathbb{C} \wedge |z| \leq 1\}$. En este caso el soporte de la medida es $\mathbb{T} = \{z | z \in \mathbb{C} \wedge |z| = 1\}$. Vemos que ambos no coinciden, $\sigma(D) \subset \text{supp}(\mu)$. El espectro de una extensión normal es más pequeño que el del correspondiente operador subnormal, (ver artículo [Atz75]). El problema se produce por la incompletitud de $\overline{\Pi}$.

[16] Una extensión normal, no necesariamente la mínima, es $V = \begin{pmatrix} U & W \\ 0 & U^* \end{pmatrix}$ donde $W = \begin{pmatrix} 1 & 0 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$, siendo U el shift-right y U^* el shift-left respectivamente. Es claro que $V : \ell^2 \bigoplus \ell^2 \longrightarrow \ell^2 \bigoplus \ell^2$ y se cumple que $VV^* = V^*V$, ver prob. 195 de [Hal80]. La mínima extensión normal se suele denominar men($\cdot$).

4.2 Un producto escalar sin medida asociada

Consideremos la matriz[17] HDP $M = (c_{ij})_{i,j=0}^{\infty}$ tal que $c_{i,j} = \min(i,j) + 1$, es decir

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots \\ 1 & 2 & 2 & 2 & \dots \\ 1 & 2 & 3 & 3 & \dots \\ 1 & 2 & 3 & 4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Se prueba trivialmente por inducción que $|M_n| = 1$, $n = 1, 2, \ldots$. La descomposición de Choleski en este caso es sencilla, vamos a efectuarla para las submatrices 4×4 por razones de espacio

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Por otro lado

$$T_4^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \qquad M_4' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 \\ 2 & 3 & 4 & 4 \end{pmatrix}.$$

Tenemos que

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 2 & 3 & 3 & 3 \\ 2 & 3 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

La matriz D en este caso es

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

La afirmación que la matriz M no es de momentos se prueba facilmente si observamos que la diagonal principal de M no verifica la desigualdad de Cauchy-Schwartz en el siguiente sentido. Si la matriz fuera de momentos existiría un recinto Ω y una medida $\mu(z)$ tales

[17] No es más que un caso particular de las matrices que podemos denominar *left upper corner* $M(\{a_n\}, \ulcorner) = (a_{\min\{i,j\}})_{i,j=1}^{\infty}$, que son definidas positivas si la sucesion $\{a_n\}$ es monónota creciente.

que

$$\begin{aligned}
c_{nn} &= \int_\Omega z^n \overline{z}^n d\mu(z) = \int_\Omega |z|^{2n} d\mu(z) = \\
&= \int_\Omega |z|^{n-1}|z|^{n+1} d\mu(z) \le \sqrt{\int_\Omega (|z|^{n-1})^2 d\mu(z) \quad \int_\Omega (|z|^{n+1})^2 d\mu(z)} = \\
&= \sqrt{c_{n-1,n-1}}\ \sqrt{c_{n+1,n+1}}.
\end{aligned}$$

Es decir si la matriz es de momentos se tiene que cumplir

$$c_{n,n}^2 \le c_{n-1,n-1}\ c_{n+1,n+1}, \qquad n = 0, 1, 2, \ldots$$

Lo que no ocurre en este caso, ya que $c_{n-1,n-1}c_{n+1,n+1} = (n-1)(n+1) = n^2-1 < n^2 = c_{n,n}^2$.

4.3 Una D no acotada. Medida de Laguerre

Consideremos la función de densidad $w(x) = x^\alpha e^{-x}$ en $[0, +\infty)$, con $\alpha = 0$ genera los P.O. de Laguerre más sencillos, tendremos que los momentos son

$$S_n = \int_0^\infty x^n e^{-x} dx = \Gamma(n+1) = n!.$$

En este caso es más cómoda la descomposición LR y la obtención de $\widetilde{D}_5$ en vez de D_5, para evitar operaciones irracionales. Lo efectuamos a título de ejemplo para matrices 5×5

$$\begin{pmatrix} 1 & 1 & 2 & 6 & 24 \\ 1 & 2 & 6 & 24 & 120 \\ 2 & 6 & 24 & 120 & 720 \\ 6 & 24 & 120 & 720 & 5040 \\ 24 & 120 & 720 & 5040 & 40320 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 4 & 4 & 0 & 0 \\ 6 & 18 & 36 & 36 & 0 \\ 24 & 96 & 288 & 576 & 576 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 & 6 & 24 \\ 0 & 1 & 4 & 18 & 96 \\ 0 & 0 & 1 & 9 & 72 \\ 0 & 0 & 0 & 1 & 16 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Efectuemos el cálculo $L_5^{-1} M_5' R_5^{-1}$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & -1 & \frac{1}{4} & 0 & 0 \\ -\frac{1}{6} & \frac{1}{2} & -\frac{1}{4} & \frac{1}{36} & 0 \\ \frac{1}{24} & \frac{-1}{6} & \frac{1}{8} & -\frac{1}{36} & \frac{1}{576} \end{pmatrix} \begin{pmatrix} 1 & 2 & 6 & 24 & 120 \\ 2 & 6 & 24 & 120 & 720 \\ 6 & 24 & 120 & 720 & 5040 \\ 24 & 120 & 720 & 5040 & 40320 \\ 120 & 720 & 5040 & 40320 & 362880 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 & -6 & 24 \\ 0 & 1 & -4 & 18 & -96 \\ 0 & 0 & 1 & -9 & 72 \\ 0 & 0 & 0 & 1 & -16 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

y tenemos

$$\widetilde{D}_5 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 4 & 0 & 0 \\ 0 & 1 & 5 & 9 & 0 \\ 0 & 0 & 1 & 7 & 16 \\ 0 & 0 & 0 & 1 & 9 \end{pmatrix} \text{ y } D_5 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 2 & 0 & 0 \\ 0 & 2 & 5 & 3 & 0 \\ 0 & 0 & 3 & 7 & 4 \\ 0 & 0 & 0 & 4 & 9 \end{pmatrix}.$$

La diagonal de la tridiagonal vemos que resulta como esperábamos $2n+1$, mientras que la superdiagonal es n, ver la tabla sobre los coeficientes de los casos clásicos en las págs. 217-221, de [Chi78].

4.4 Un ejemplo con matriz M que desconocemos si es o no de momentos

Tomemos la matriz de Césaro utilizada[18] en el método utilizado para acelerar convergencia de ciertas sucesiones

$$C_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \text{ y construimos } M_4 = C_4 C_4^t = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/2 & 1/3 & 1/4 \\ 1/3 & 1/3 & 1/3 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}.$$

Es inmediato que $|C_n| = 1/n!$. A partir de estas matrices construimos

$$M_n = (\frac{1}{\max(i,j)+1})_{i,j=0}^{\infty}, \qquad \text{con } |M_n| = \frac{1}{(n!)^2}.$$

Es simétrica, real y definida positiva[19]. Calculemos $\widetilde{D}_4$ y D_4. La descomposición LR resulta ser

$$\begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/2 & 1/3 & 1/4 \\ 1/3 & 1/3 & 1/3 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/4 & 0 & 0 \\ 1/3 & 1/6 & 1/9 & 0 \\ 1/4 & 1/8 & 1/12 & 1/16 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 0 & 1 & 2/3 & 1/2 \\ 0 & 0 & 1 & 3/4 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Calculemos $\widetilde{D}_4 = L_4^{-1} M_4' R_4^{-1}$,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 4 & 0 & 0 \\ 0 & -6 & 9 & 0 \\ 0 & 0 & -12 & -16 \end{pmatrix} \begin{pmatrix} 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/3 & 1/4 & 1/5 \\ 1/4 & 1/4 & 1/4 & 1/5 \end{pmatrix} \begin{pmatrix} 1 & -1/2 & 0 & 0 \\ 0 & 1 & -2/3 & 0 \\ 0 & 0 & 1 & -3/4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

que tras operar queda

$$\widetilde{D}_4 = \begin{pmatrix} 1/2 & 1/12 & 1/36 & 1/80 \\ 1 & 2/12 & 2/36 & 2/80 \\ 0 & 1 & 3/36 & 3/80 \\ 0 & 0 & 1 & 4/80 \end{pmatrix}.$$

Observamos que

$$\widetilde{d}_{ij} = \begin{cases} \dfrac{i}{j^2(1+j)} & \text{si} \quad i \leq j, \\ 1 & \text{si} \quad i = j+1, \\ 0 & \text{si} \quad i > j+1. \end{cases}$$

También es inmediato obtener los polinomios

$$P_n(z) = (n+1)z^n - nz^{n-1}.$$

[18]Sorprendentemente puede probarse que el operador $C : \ell^2 \longrightarrow \ell^2$ es subnormal. Esta subnormalidad en nada afecta al problema de momentos asociado, dado que es D quien tiene que ser subnormal no uno de los factores de M.

[19]Un caso particular de la matrices *lower right corner* $M(\{a_n\}, \lrcorner)$, que se prueba son definidas positivas si la sucesión es monótona decriente estricta.

Utilizando el apartado iv) de la Proposición 3.8, como $|M_n| = 1/(n!)^2$, tenemos $e_n = |M_n|/|M_{n-1}| = [1/(n!)^2]/[1/((n-1)!)^2] = 1/n^2$.

$$D_n = \begin{pmatrix} 1 & & & \\ & 1/2 & & \\ & & 1/3 & \\ & & & 1/4 \end{pmatrix} \begin{pmatrix} 1/2 & 1/12 & 1/36 & 1/80 \\ 1 & 2/12 & 2/36 & 2/80 \\ 0 & 1 & 3/36 & 3/80 \\ 0 & 0 & 1 & 4/80 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{pmatrix}$$

y operando queda finalmente

$$D_4 = \begin{pmatrix} 1/2 & 2/12 & 3/36 & 4/80 \\ 1/2 & 2/12 & 3/36 & 4/80 \\ 0 & 2/3 & 3/36 & 4/80 \\ 0 & 0 & 3/4 & 4/80 \end{pmatrix},$$

que resulta ser

$$d_{ij} = \begin{cases} \dfrac{j}{j^2(1+j)} & \text{si} \quad i \leq j, \\ \dfrac{j}{1+j} & \text{si} \quad i = j+1, \\ 0 & \text{si} \quad i > j+1. \end{cases}$$

M si cumple la condición de Cauchy-Schwartz, pero eso no garantiza que la matriz sea de momentos. Averiguar si es o no subnormal o incluso hiponormal no parece demasiado trivial.

4.5 Un ejemplo construido a partir de un operador hiponormal que no es subnormal

Se sabe que el operador $T = 2U + U^*$, es decir

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 2 & 0 & 1 & 0 & \dots \\ 0 & 2 & 0 & 1 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

operando en ℓ^2 es hiponormal, pero no subnormal. Probar que es hiponormal es sencillo ya que

$$T^*T = \begin{pmatrix} 4 & 0 & 2 & 0 & \dots \\ 0 & 5 & 0 & 2 & \dots \\ 2 & 0 & 5 & 0 & \dots \\ 0 & 2 & 0 & 5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad TT^* = \begin{pmatrix} 1 & 0 & 2 & 0 & \dots \\ 0 & 5 & 0 & 2 & \dots \\ 2 & 0 & 5 & 0 & \dots \\ 0 & 2 & 0 & 5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

luego $T^*T - TT^* \geq 0$. Sin embargo T^2 no es hiponormal, que puede hacerse como ejercicio. En consecuencia T^2 no es subnormal. Hay un resultado (ver prob. 209 [Hal80]) que afirma que si W es subnormal también lo es W^2. Luego si T^2 no es subnormal tampoco lo es T. Y estamos en la situación descrita. La matriz M generada a partir de T va a verificar Cauchy-Schwartz por ser hiponormal (ver [Tor93]), sera hermitiana definida positiva, pero no de momentos. El ejemplo queda incompleto ya que no hemos dado una regla de formación de M. Puede obtenerse como ejercicio. Aquí calculamos M mediante la fórmula (26), que más adelante veremos. Por ejemplo de orden 5×5 es

$$M_5 = \begin{pmatrix} 1 & 0 & 2 & 0 & 8 \\ 0 & 4 & 0 & 16 & 0 \\ 2 & 0 & 20 & 0 & 112 \\ 0 & 16 & 0 & 128 & 0 \\ 8 & 0 & 112 & 0 & 896 \end{pmatrix}.$$

Se cumplirá que $\lim_{n\to\infty} \sqrt{c_{n+1,n+1}/c_{n,n}} = k$. Es una conjetura trivial: $k = 3$.

4.6 Medida planar de Lebesgue en el disco unidad

Conocemos la fórmula de Green[20] que nos relaciona las integrales dobles con las de línea

$$\iint_G f\overline{g'}dm = \frac{1}{2i}\oint_{\partial G} f\overline{g}dz.$$

Aplicándola al producto interno $\langle z^p, z^q \rangle$ resulta, si consideramos la medida de Lebesgue planar en el círculo

$$\langle z^p, z^q \rangle = \iint_G z^p \overline{z}^q dm = \frac{1}{2i(q+1)}\oint_{\partial G} z^p \overline{z}^{q+1} dz.$$

Tomando ahora $G = \{z : |z| < 1\}$ haciendo $z = e^{i\theta}$ tendremos que $dz = ie^{i\theta}d\theta$, luego

$$\begin{aligned} c_{p,q} &= \frac{1}{2i(q+1)}\int_0^{2\pi} e^{ip\theta}e^{-i(q+1)\theta}ie^{i\theta}d\theta \\ &= \frac{1}{2i(q+1)}\int_0^{2\pi} e^{i(p-q)\theta}d\theta = \begin{cases} 0 & \text{si} \quad p \neq q \\ \frac{\pi}{q+1} & \text{si} \quad p = q. \end{cases} \end{aligned}$$

Luego

$$M = \begin{pmatrix} \pi & 0 & 0 & \dots \\ 0 & \pi/2 & 0 & \dots \\ 0 & 0 & \pi/3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

[20] $\oint f(x+yi)\overline{g(x+yi)}(dx+idy) = \oint (f(x+yi)\overline{g(x+yi)})dx + (f(x+iy)\overline{g(x+yi)}i)dy$. Llamaremos $P(x,y) = f(x+yi)\overline{g(x+yi)}$, $Q(x,y) = f(x+iy)\overline{g(x+yi)}i$. El teorema de Green en su versión tradicional dice $\oint_{\partial G} P(x,y)dx + Q(x,y)dy = \iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy$, tenemos que $\frac{\partial Q}{\partial x} = if'(z)\overline{g(z)} + if(z)\overline{g'(z)}$, $\frac{\partial P}{\partial y} = f'(z)\overline{g(z)}i + f(z)\overline{g'(z)}\overline{i}$, luego finalmente $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 2if(z)\overline{g'(z)}$.

Resulta la matriz

$$
\begin{aligned}
D &= \begin{pmatrix} \sqrt{1/\pi} & 0 & 0 & \dots \\ 0 & \sqrt{2/\pi} & 0 & \dots \\ 0 & 0 & \sqrt{3/\pi} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \dots \\ \pi/2 & 0 & 0 & \dots \\ 0 & \pi/3 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \sqrt{1/\pi} & 0 & 0 & \dots \\ 0 & \sqrt{2/\pi} & 0 & \dots \\ 0 & 0 & \sqrt{3/\pi} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 & \dots \\ \sqrt{1/2} & 0 & 0 & \dots \\ 0 & \sqrt{2/3} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.
\end{aligned}
$$

Es conocido (ver [Con81] pág. 156) que en ℓ^2 los únicos shifts con peso hiponormales son aquellos en que la sucesión es monótona creciente con límite ($\sup |\alpha_n| = \|D\|$). En este caso $\{\sqrt{n/(n+1)}\}_{n=1}^{\infty}$ cumple esa condición, lo que era presumible ya que existe solución al problema de los momentos y la subnormalidad implica la hiponormalidad. Curiosamente para este shift, la matriz D, a diferencia de lo que ocurría con el shift-right U, no es cuasinormal.

4.7 Otro ejemplo sobre la circunferencia unidad

Consideremos la matriz de Toeplitz

$$
M_n = \begin{pmatrix} 2 & 1 & 0 & 0 & \dots \\ 1 & 2 & 1 & 0 & \dots \\ 0 & 1 & 2 & 1 & \dots \\ 0 & 0 & 1 & 2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.
$$

Es fácil probar por inducción que $|M_n| = n + 1$. Luego es definida positiva. En consecuencia el problema de los momentos sobre la circunferencia tiene solución y siempre es un problema determinado. Además

$$
\kappa_n = \sqrt{\frac{|M_{n-1}|}{|M_n|}} = \sqrt{\frac{n}{n+1}} \longrightarrow 1 < +\infty,
$$

luego se cumple la condición de Szegö y el sistema no es completo, luego la matriz D no es unitaria (normal). La matriz D tiene que ser subnormal y por tanto hiponormal, probarlo directamente no es aparentemente trivial, aunque obtuviésemos su expresión explícita en función de n, toma la forma

$$
D = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{12} & -\frac{\sqrt{90}}{60} & \dots \\ \frac{\sqrt{3}}{2} & \frac{1}{6} & -\frac{\sqrt{2}}{12} & \frac{\sqrt{30}}{60} & \dots \\ 0 & \frac{\sqrt{8}}{3} & \frac{1}{12} & -\frac{\sqrt{15}}{60} & \dots \\ 0 & 0 & \frac{\sqrt{15}}{4} & \frac{1}{20} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.
$$

Veremos más adelante cómo se calculan explícitamente sus elementos, utilizando el hecho de que $\Phi_n(0) = (-1)^n/(n+1)$.

Se prueba, ver página 48 de estas notas, que $D^H D = I$, pero en general $DD^H \neq I$, porque en ese caso sería unitaria y no lo es. Salvo que falle la condición de Szegö, que en este caso concreto no falla ya que $\sum_{n=0}^{\infty} |\Phi_n(0)|^2 < +\infty$.

5 Cuatro resultados sencillos

5.1 Autovalores y autovectores de D_n^t

Proposición 5.1. *Si z_{nk} es una raíz de $p_n(z)$ (aparte claro está de ser autovalor de D_n y de D_n^t) se cumple que*

$$z_{nk} \begin{pmatrix} p_0(z_{nk}) \\ p_1(z_{nk}) \\ \vdots \\ p_{n-1}(z_{nk}) \end{pmatrix} = D_n^t \begin{pmatrix} p_0(z_{nk}) \\ p_1(z_{nk}) \\ \vdots \\ p_{n-1}(z_{nk}) \end{pmatrix},$$

es decir el autovector correspondiente a z_{nk} es

$$(p_0(z_{nk}), p_1(z_{nk}), \ldots, p_{n-1}(z_{nk}))^t.$$

Dem. Escribamos con detalle el determinante

$$P_n(z) = |I_n z - D_n|,$$

tendremos

$$P_n(z) = \begin{vmatrix} z-d_{11} & -d_{12} & -d_{13} & \ldots & -d_{1,n-2} & -d_{1,n-1} & -d_{1,n} \\ -d_{21} & z-d_{22} & -d_{23} & \ldots & -d_{2,n-2} & -d_{2,n-1} & -d_{2,n} \\ 0 & -d_{32} & z-d_{33} & \ldots & -d_{3,n-2} & -d_{3,n-1} & -d_{3,n} \\ 0 & 0 & -d_{43} & \ldots & -d_{4,n-2} & -d_{4,n-1} & -d_{4,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -d_{n-1,n-2} & z-d_{n-1,n-1} & -d_{n-1,n} \\ 0 & 0 & 0 & \ldots & 0 & -d_{n,n-1} & z-d_{n,n} \end{vmatrix}.$$

Desarrollando por la última fila reiteradamente

$$\begin{aligned} P_n(z) &= (z-d_{n,n})P_{n-1}(z) \\ &+ [d_{n,n-1}](-d_{n-1,n})P_{n-2}(z) \\ &+ [d_{n,n-1}d_{n-1,n-2}](-d_{n-2,n})P_{n-3}(z) \\ &+ [d_{n,n-1}d_{n-1,n-2}d_{n-2,n-3}](-d_{n-3,n})P_{n-4}(z) \\ &\vdots \\ &+ [d_{n,n-1}d_{n-1,n-2}\ldots d_{21}](-d_{1,n})P_0(z). \end{aligned} \tag{21}$$

Sabemos que

$$d_{i,i-1} = \frac{\kappa_{i-2}}{\kappa_{i-1}} = \sqrt{\frac{|M_i|\,|M_{i-2}|}{|M_{i-1}|^2}}, \qquad 2 \le i \le n,$$

con $|M_0| = 1$ y $n \ge 2$, se cumplirá

$$\begin{aligned}
\frac{\kappa_{n-2}}{\kappa_{n-1}} &= d_{n,n-1}\\
\frac{\kappa_{n-3}}{\kappa_{n-1}} = \frac{\kappa_{n-2}}{\kappa_{n-1}}\,\frac{\kappa_{n-3}}{\kappa_{n-2}} &= d_{n,n-1}\,d_{n-1,n-2}\\
\frac{\kappa_{n-4}}{\kappa_{n-1}} = \frac{\kappa_{n-2}}{\kappa_{n-1}}\,\frac{\kappa_{n-3}}{\kappa_{n-2}}\,\frac{\kappa_{n-4}}{\kappa_{n-3}} &= d_{n,n-1}\,d_{n-1,n-2}\,d_{n-2,n-3}\\
&\vdots\\
\frac{\kappa_0}{\kappa_{n-1}} = \frac{\kappa_{n-2}}{\kappa_{n-1}}\,\frac{\kappa_{n-3}}{\kappa_{n-2}}\,\frac{\kappa_{n-4}}{\kappa_{n-3}}\cdots\frac{\kappa_0}{\kappa_1} &= d_{n,n-1}\,d_{n-1,n-2}\,d_{n-2,n-3}\cdots d_{21}.
\end{aligned}$$

Luego substituyendo en (21) y teniendo en cuenta que $\|P_k(n)\| = 1/\kappa_n$, es decir $\kappa_n P_n(z) = p_n(z)$, resultará

$$\begin{aligned}
P_n(z) &= (z - d_{n,n})P_{n-1}(z)\\
&+ \frac{\kappa_{n-2}}{\kappa_{n-1}}(-d_{n-1,n})P_{n-2}(z)\\
&+ \frac{\kappa_{n-3}}{\kappa_{n-1}}(-d_{n-2,n})P_{n-3}(z)\\
&+ \frac{\kappa_{n-4}}{\kappa_{n-1}}(-d_{n-3,n})P_{n-4}(z)\\
&\vdots\\
&+ \frac{\kappa_0}{\kappa_{n-1}}(-d_{1,n})P_0(z).
\end{aligned} \tag{22}$$

Multiplicando por $-\kappa_{n-1}$ ambos miembros y haciendo $\kappa_j P_j(z) = p_j(z),\ j = 0, 1, \ldots n$ podemos escribir, despejando $zp_{n-1}(z)$, y teniendo en cuenta que $\kappa_{n-1}P_n(z) = \dfrac{\kappa_{n-1}}{\kappa_n}p_n(z)$

$$\begin{aligned}
zp_{n-1}(z) &= \frac{\kappa_{n-1}}{\kappa_n}p_n(z) + d_{nn}p_{n-1}(z) + d_{n-1,n}p_{n-2}(z) + d_{n-2,n}p_{n-3}(z)\\
&\quad + d_{n-3}p_{n-4}(z) + \ldots + d_{1n}p_0(z).
\end{aligned} \tag{23}$$

Como sabemos que $\kappa_{n-1}/\kappa_n = d_{n+1,n}$. El anterior desarrollo es válido para cualquier valor de n tomándolo para $n = 1, 2, \ldots$, y para el propio n tendremos finalmente

$$\begin{aligned}
zp_0(z) &= d_{11}p_0(z) + d_{21}p_1(z)\\
zp_1(z) &= d_{12}p_0(z) + d_{22}p_1(z) + d_{32}p_2(z)\\
zp_2(z) &= d_{13}p_0(z) + d_{23}p_1(z) + d_{33}p_2(z) + d_{43}p_3(z)\\
&\vdots\\
zp_{n-1}(z) &= d_{1n}p_0(z) + d_{2n}p_1(z) + d_{3n}p_2(z) + \ldots + d_{n,n}p_{n-1}(z) + d_{n+1,n}p_n(z).
\end{aligned} \tag{24}$$

Obsérvese que el elemento $d_{n+1,n}$ está en D_{n+1} pero no en D_n. La anterior expresión se puede escribir matricialmente

$$z\begin{pmatrix} p_0(z) \\ p_1(z) \\ p_2(z) \\ \vdots \\ p_{n-2}(z) \\ p_{n-1}(z) \end{pmatrix} = \begin{pmatrix} d_{11} & d_{21} & 0 & \dots & 0 & 0 \\ d_{12} & d_{22} & d_{32} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{1,n-2} & d_{2,n-2} & d_{3,n-2} & \dots & d_{n-1,n-2} & 0 \\ d_{1,n-1} & d_{2,n-1} & d_{3,n-1} & \dots & d_{n-1,n-1} & d_{n,n-1} \\ d_{1,n} & d_{2,n} & d_{3,n} & \dots & d_{n-1,n} & d_{n,n} \end{pmatrix} \begin{pmatrix} p_0(z) \\ p_1(z) \\ p_2(z) \\ \vdots \\ p_{n-2}(z) \\ p_{n-1}(z) \end{pmatrix} + d_{n+1,n} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} p_n(z). \tag{25}$$

Si ahora hacemos $z = z_{nk}$, cualquiera de las raíces de $p_n(z)$, el último sumando desaparece y queda

$$z_{nk}\begin{pmatrix} p_0(z_{nk}) \\ p_1(z_{nk}) \\ p_2(z_{nk}) \\ \vdots \\ p_{n-2}(z) \\ p_{n-1}(z_{nk}) \end{pmatrix} = \begin{pmatrix} d_{11} & d_{21} & 0 & \dots & 0 & 0 \\ d_{12} & d_{22} & d_{32} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ d_{1,n-2} & d_{2,n-2} & d_{3,n-2} & \dots & d_{n-1,n-2} & 0 \\ d_{1,n-1} & d_{2,n-1} & d_{3,n-1} & \dots & d_{n-1,n-1} & d_{n,n-1} \\ d_{1,n} & d_{2,n} & d_{3,n} & \dots & d_{n-1,n} & d_{n,n} \end{pmatrix} \begin{pmatrix} p_0(z_{nk}) \\ p_1(z_{nk}) \\ p_2(z_{nk}) \\ \vdots \\ p_{n-2}(z) \\ p_{n-1}(z_{nk}) \end{pmatrix},$$

es decir

$$z_{nk}\overline{v} = D_n^t\overline{v}.$$

Luego z_{nk} es un autovalor de D_n^t y el autovector correspondiente es

$$\overline{v} = (p_0(z_{nk}), p_1(z_{nk}), \dots, p_{n-1}(z_{nk}))^t.$$

□

Observación 5.1. Sabemos que los autovalores de D_n y D_n^t coinciden. Si tomamos conjugados en (25) tras particularizar en $z = z_{nk}$ resulta

$$\overline{z}_{nk}\begin{pmatrix} \overline{p_0(z_{nk})} \\ \overline{p_1(z_{nk})} \\ \vdots \\ \overline{p_{n-1}(z_{nk})} \end{pmatrix} = D_n^H \begin{pmatrix} \overline{p_0(z_{nk})} \\ \overline{p_1(z_{nk})} \\ \vdots \\ \overline{p_{n-1}(z_{nk})} \end{pmatrix}.$$

Los autovectores de A^H y A son complejos conjugados, pero aquí se trata de A^t y no podemos afirmar nada sobre ellos. También es interesante que observemos que

$$\sum_{k=0}^{n} |p_k(z_{nk})|^2 = K_n(z_{nk}, z_{nk}) = K_{n-1}(z_{nk}, z_{nk}) = \|\overline{v}\|^2.$$

Es decir la longitud al cuadrado del autovector correspondiente a z_{nk} es justamente el núcleo n-ésimo particularizado en dicha raíz. En el caso tridiagonal sabemos además que la constante de Christoffel p_{nk} asociada a z_{nk} vale justamente

$$p_{nk} = \frac{1}{K_n(z_{nk}, z_{nk})}.$$

Observación 5.1. Las fórmulas (24) se pueden resumir en la expresión

$$zp_{n-1}(z) = \sum_{k=0}^{n+1} d_{k,n} p_{k-1}(z), \quad n \geq 1.$$

que es la denominada de *fórmula de recurrencina larga* y que puede servir para introducir también la matriz D.

5.2 Obtención de M a partir de D

Proposición 5.2. *Dada $M = (c_{i,j})_{i,j=0}^{\infty}$ matriz HDP infinita y sea D la matriz de Hessenberg construida por el procedimiento anterior, entonces se cumple que*

$$c_{i,j} = \langle D_n^i e_0, D_n^j e_0 \rangle, \qquad 0 \leq i, j \leq n-1, \tag{26}$$

siendo $e_0^t = (1, 0, 0, 0, \ldots, 0)$.

Dem. Expresaremos los elementos de la base $\{z^n\}_{n=0}^{\infty}$, que no es de Schauder, en función de $\{p_n(z)\}_{n=0}^{\infty}$. Supondremos $c_{0,0} = 1$, lo que equivale, caso de que no sea así, a dividir D por $c_{0,0}$. Con eso conseguimos que $p_0(z) = 1$. Con e_0 indicaremos el vector columna $(1, 0, 0, \ldots, 0)^t$, de longitud m.

Tenemos que $p_0(z) = 1$, luego

$$1 = (1, 0, 0, \ldots, 0)_m \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix} = e_0^t \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix}$$

multiplicando ambos miembros de la igualdad anterior por z tendremos

$$\begin{aligned} z &= (1, 0, 0, \ldots, 0)_m \begin{pmatrix} zp_0(z) \\ zp_1(z) \\ \vdots \\ zp_{m-1}(z) \end{pmatrix} = e_0^t \left[D_m^t \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ d_{m+1,m} \end{pmatrix} p_m(z) \right] \\ &= e_0^t D_m^t \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix}, \end{aligned}$$

volviendo a multiplicar por z y utilizando de nuevo (25) obtendremos

$$\begin{aligned} z^2 &= e_0^t D_m^t \begin{pmatrix} zp_0(z) \\ zp_1(z) \\ \vdots \\ zp_{m-1}(z) \end{pmatrix} = e_0^t D_m^t \left[D_m^t \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_k \end{pmatrix} p_m(z) \right] \\ &= e_0^t (D_m^t)^2 \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix}, \quad \text{si } m \geq 3. \end{aligned}$$

En general tendremos

$$z^n = e_0^t (D_m^t)^n \begin{pmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{m-1}(z) \end{pmatrix}, \qquad \text{si } m \geq n+1.$$

Es decir el vector que nos propociona las componentes de z^n respecto de los polinomios normalizados será el vector fila $e_0^t(D_m^t)^n$, $0 \leq n \leq m-1$. Como para operar utilizamos la representación en vectores columna, será

$$e_n \equiv [e_0^t (D_m^t)^n]^t = D_m^n e_0, \quad 0 \leq n \leq m-1.$$

En este punto podemos volver a utilizar la notación usual para $e_0 \in \ell^2$, suponiéndolo con infinitas componentes. De modo análogo en lugar de tomar D_m^t, podemos considerar la matriz infinita D^t y seguiremos teniendo

$$e_n \equiv D^n e_0.$$

Sabemos que

$$\langle z^i, z^j \rangle_M = e_j^H M e_i = c_{i,j}.$$

Donde e_j y e_i son los vectores de las componentes de z^i y z^j respecto de la base *standard* $\{z^n\}_{n=0}^{\infty}$. Las componentes de z^i y z^j respecto de la base normalizada son, como hemos visto, los vectores $D^i e_0$ y $D^j e_0$. Es claro que expresando z^i y z^j respecto de la SPON la matriz del producto escalar será la matriz I, ya que

$$\langle p_n(z), p_m(z) \rangle = \delta_{n,m}.$$

Luego resultará

$$\begin{aligned} c_{j,k} &= \langle z^j, z^k \rangle_M = \overline{e_k^t}\ M\ e_j = \\ &= \overline{e_0^t (D^t)^k}\ I\ D^j e_0 = \overline{e}_0^t (D^H)^k\ I\ D^j e_0 = \\ &= (D^k e_0)^H\ I\ D^j e_0 = \langle D^j e_0, D^k e_0 \rangle. \end{aligned}$$

Si $j, k \leq n-1$, se verifica $\langle D^j e_0, D^k e_0 \rangle = \langle D_n^j e_0, D_n^k e_0 \rangle$, y tenemos el enunciado. $\square$

Observación 5.2. Nótese que la matriz M obtenida a partir de la proposición anterior esta *normalizada*. En el sentido de que $c_{00} = 1$. Si correspondiera a una medida sería de probabilidad.

5.3 Transformaciones de semejanza del soporte y/o espectro

Observación 5.2. A diferencia de lo que ocurre sobre la recta real en donde si desplazamos la distribución, respecto del origen, la matriz de momentos sigue siendo de Hankel, y no hay cambios esenciales en el planteamiento del problema de los momentos. En el caso complejo, por ejemplo para la circunferencia unidad, basta con que traslademos el centro de la circunferencia soporte a otro punto para que la matriz de momentos deje de ser de Toeplitz, o efectuemos una homotecia de centro el origen y razón $r \neq 1$.

Es interesante por tanto que obtengamos las matrices de paso que nos permitan, conocida una matriz de momentos, calcular la que resulta si modificamos la curva soporte mediante una transformación de semejanza.

Incluso si la matriz M no es matriz de momentos, pero si HDP, la transformación[21] $\varphi(z) = \alpha z + \beta$ puede interpretarse como una transformación de semejanza del espectro de la matriz D, la matriz M^φ que correspondería a la D^φ tras dicha transformación satisfaría las ecuaciones que calcularemos en lo que sigue.

Proposición 5.3. *Sea M_n la sección de orden n de la matriz de momentos de una cierta distribución $\sigma : \gamma \longrightarrow \mathbb{R}$, transformemos el recinto soporte $\varphi(\gamma) = \{\alpha z + \beta \mid z \in \gamma\}$, $\alpha, \beta \in \mathbb{C}$, con la condición de dejar invariante la distribución, es decir $\sigma^\varphi(\alpha z+\beta) = \sigma(z)$. Entonces la matriz de momentos transformada M_n^φ viene dada por*

$$M_n^\varphi = A_n^H\, M_n\, A_n, \qquad con \qquad A_n = (a_{i,j})_{i,j=1}^n = \begin{cases} \binom{j-1}{i-1}\alpha^{i-1}\beta^{j-i} & si \quad j \geq i \\ 0 & si \quad j < i, \end{cases} \tag{27}$$

donde con A^H queremos indicar la matriz transpuesta conjugada de A.

Dem. Resulta inmediato que

$$\begin{aligned} c_{kl}^\varphi &= \int_{\gamma^*} u^k \overline{u}^l d\sigma^*(u) = \int_\gamma (\alpha z+\beta)^k \overline{(\alpha z+\beta)}^l d\sigma^*(\alpha z+\beta) \\ &= \int_\gamma \left[\sum_{i=0}^k \binom{k}{i}\beta^{k-i}\alpha^i z^i\right] \left[\sum_{j=0}^l \binom{l}{j}(\overline{\beta})^{l-j}(\overline{\alpha})^j(\overline{z})^j\right] d\sigma(z) \\ &= \sum_{i=0}^k \sum_{j=0}^l \binom{k}{i}\binom{l}{j}\beta^{k-i}(\overline{\beta})^{l-j}\alpha^i(\overline{\alpha})^j\, c_{i,j} \end{aligned} \tag{28}$$

Efectuemos el producto $A_n^H M_n A_n = T_n = (t_{ij})_{ij=1}^n$, para ver si efectivamente se cumple el enunciado. Llamemos $B_n = A_n^H M_n = (b_{kp})_{k,p=1}^n$, con $t_{lk} = \sum_{q=1}^n b_{lq}a_{qk}$ y $b_{lq} = \sum_{p=1}^n \overline{a}_{pl} c_{q-1,p-1}$. Utilizando (27), tenemos

$$\begin{aligned} t_{lk} &= \sum_{p=1}^n \sum_{q=1}^n \overline{a}_{pl} a_{qk} c_{q-1,p-1} \\ &= \sum_{q=1}^k \sum_{p=1}^l \binom{l-1}{p-1}(\overline{\alpha})^{p-1}(\overline{\beta})^{l-p}\binom{k-1}{q-1}\alpha^{q-1}\beta^{k-q}c_{q-1,p-1}. \end{aligned} \tag{29}$$

[21] Es importante resaltar el orden: primero giro+homotecia, luego traslación.

Como la notación para los elementos de las matrices de momentos $M_n^\varphi = (c_{ij}^\varphi)_{ij=0}^n$ y $M_n = (c_{ij})_{i,j=0}^n$ está permutada y disminuida en una unidad para ambos índices, haciendo $q-1=i$, $p-1=j$, $k \equiv k+1$ y $l \equiv l+1$ en (29), resulta finalmente

$$c_{kl}^\varphi = t_{l+1,k+1} = \sum_{i=0}^{k}\sum_{j=0}^{l}\binom{l}{j}\binom{k}{i}(\overline{\alpha})^j\alpha^i(\overline{\beta})^{l-j}\beta^{k-i}c_{ij}.$$

Luego la expresión obtenida mediante el producto de matrices coincide con la obtenida en (28). $\square$

Observación 5.3. La matriz A, resulta ser de acuerdo con (27)

$$A = \begin{pmatrix} \binom{0}{0}\alpha^0\beta^0 & \binom{1}{0}\alpha^0\beta^1 & \binom{2}{0}\alpha^0\beta^2 & \binom{3}{0}\alpha^0\beta^3 & \cdots \\ 0 & \binom{1}{1}\alpha^1\beta^0 & \binom{2}{1}\alpha^1\beta^1 & \binom{3}{1}\alpha^1\beta^2 & \cdots \\ 0 & 0 & \binom{2}{2}\alpha^2\beta^0 & \binom{3}{2}\alpha^2\beta^1 & \cdots \\ 0 & 0 & 0 & \binom{3}{3}\alpha^3\beta^0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Ejemplo 5.1. Consideremos la medida de Lebesgue normalizada (de modo que $c_{00}=1$) sobre la circunferencia unidad.

$$c_{j,k} = \frac{1}{2\pi}\int_{-\pi}^{\pi}[e^{i\theta}]^j[e^{-i\theta}]^k d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i(j-k)\theta}d\theta = \delta_{jk}.$$

La matriz de momentos es en este caso la matriz unidad, que es obviamente de Toeplitz. Si hacemos $\alpha=1$ y $\beta=1$, transformamos el conjunto de puntos mediante $T(z)=z+1$,

$$T(\{z : |z|=1\}) = \{z+1 : |z|=1\} = \{u : |u-1|=1\},$$

resulta la circunferencia unidad centrada en $(1,0)$ con la medida de Lebesgue. La matriz de momentos pasa a ser mediante (27) la llamada matriz de Pascal. Por ejemplo para $n=5$ tendremos

$$\begin{aligned} M_5^\varphi &= \begin{pmatrix} 1&0&0&0&0\\1&1&0&0&0\\1&2&1&0&0\\1&3&3&1&0\\1&4&6&4&1 \end{pmatrix}\begin{pmatrix} 1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1 \end{pmatrix}\begin{pmatrix} 1&1&1&1&1\\0&1&2&3&4\\0&0&1&3&6\\0&0&0&1&4\\0&0&0&0&1 \end{pmatrix} = \\ &= \begin{pmatrix} 1&1&1&1&1\\1&2&3&4&5\\1&3&6&10&15\\1&4&10&20&35\\1&5&15&35&70 \end{pmatrix}. \end{aligned}$$

Observación 5.4. Es evidente que tomando determinantes en el triple producto anterior resulta un $|M_n^\varphi| = 1$, $n=1,2,\ldots$, lo que también se puede probar más trabajosamente por inducción.

No es extraño[22] que

$$
\begin{aligned}
A_5^{-H}[(M^{\varphi})^{(1)}]_5 A_5^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 10 & 15 & 21 \\ 4 & 10 & 20 & 35 & 56 \\ 5 & 15 & 35 & 70 & 126 \\ 6 & 21 & 56 & 126 & 252 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ -1 & 3 & -3 & 1 & 0 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 10 & 15 & 21 \\ 4 & 10 & 20 & 35 & 56 \\ 5 & 15 & 35 & 70 & 126 \\ 6 & 21 & 56 & 126 & 252 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ 0 & 1 & -2 & 3 & -4 \\ 0 & 0 & 1 & -3 & 6 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.
\end{aligned}
$$

Análogamente

$$
\begin{aligned}
A_5^{-H}[(M^{\varphi})^{(2)}]_5 A_5^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 6 & 10 & 15 & 21 & 28 \\ 10 & 20 & 35 & 56 & 84 \\ 15 & 35 & 70 & 126 & 210 \\ 21 & 56 & 126 & 252 & 462 \\ 28 & 84 & 210 & 462 & 924 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 6 & 4 & 1 & 0 & 0 \\ 4 & 6 & 4 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \\ 0 & 1 & 4 & 6 & 4 \\ 0 & 0 & 1 & 4 & 6 \end{pmatrix}.
\end{aligned}
$$

Esto no es una casualidad. Si dada una matriz de momentos M consideramos[23] la matriz $M^{(m)}$ que resulta de eliminar en M las m primeras filas y las m primeras columnas, la función densidad $w(z)$ queda multiplicada por $z^m \overline{z}^m = |z|^{2m}$.

La matriz de Pascal es la matriz de momentos de la medida de Lebesgue normalizada pero no sobre la circunferencia unidad sino sobre $|u-1| = 1$, es decir

$$
\begin{aligned}
(c_{jk}^{\varphi})^{(m)} &= \frac{1}{2\pi} \int_{|u-1|=1} u^{j+m-1} \overline{u}^{k+m-1} |du| = \\
&= \frac{1}{2\pi} \int_{|z|=1} (z+1)^{j+m-1} \overline{(z+1)}^{k+m-1} |dz| = \\
&= \frac{1}{2\pi} \int_0^{2\pi} (e^{it}+1)^{j+m-1} (e^{-it}+1)^{k+m-1} dt.
\end{aligned}
$$

[22] Con la notación $A^{(m)}$ queremos indicar que en la matriz infinita A hemos eliminado las primeras m filas y columnas.

[23] Esta notación se utilizará también para los polinomios desplazados. No confundir $M^{(1)}$ con M', en la primera se eliminan primera fila y primera columna, en la segunda solo la primera columna.

Observación 5.5. Parece natural que nos preguntemos por la expresión de la matriz D^{φ} correspondiente a M^{φ} en función de la matriz D correspondiente a M. Se tiene el siguiente resultado.

Proposición 5.4. *Sean dos matrices M y M^{φ}, HDP infinitas y relacionadas mediante las expresiones (27), siendo $D_n = (d_{ij})_{ij=1}^{n}$ y $\widetilde{D}_n = (\widetilde{d}_{ij})_{ij=1}^{n}$, las matrices de orden n asociadas a M. Entonces las matrices D_n^{φ} y $\widetilde{D}_n^{\varphi}$ de orden n asociadas a M^{φ} vendrán dadas por*

$$D_n^{\varphi} = \begin{pmatrix} \alpha d_{11}+\beta & \frac{\alpha^2}{|\alpha|}d_{12} & \frac{\alpha^3}{|\alpha|^2}d_{13} & \frac{\alpha^4}{|\alpha|^3}d_{14} & \dots & \frac{\alpha^n}{|\alpha|^{n-1}}d_{1n} \\ |\alpha|d_{21} & \alpha d_{22}+\beta & \frac{\alpha^2}{|\alpha|}d_{23} & \frac{\alpha^3}{|\alpha|^2}d_{24} & \dots & \frac{\alpha^{n-1}}{|\alpha|^{n-2}}d_{2n} \\ 0 & |\alpha|d_{32} & \alpha d_{33}+\beta & \frac{\alpha^2}{|\alpha|}d_{34} & \dots & \frac{\alpha^{n-2}}{|\alpha|^{n-3}}d_{3n} \\ 0 & 0 & |\alpha|d_{43} & \alpha d_{44}+\beta & \dots & \frac{\alpha^{n-3}}{|\alpha|^{n-4}}d_{4n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \alpha d_{nn}+\beta \end{pmatrix} \tag{30}$$

y por

$$\widetilde{D}_n^{\varphi} = \begin{pmatrix} \alpha\widetilde{d}_{11}+\beta & \alpha^2\widetilde{d}_{12} & \alpha^3\widetilde{d}_{13} & \alpha^4\widetilde{d}_{14} & \dots & \alpha^n\widetilde{d}_{1n} \\ 1 & \alpha\widetilde{d}_{22}+\beta & \alpha^2\widetilde{d}_{23} & \alpha^3\widetilde{d}_{24} & \dots & \alpha^{n-1}\widetilde{d}_{2n} \\ 0 & 1 & \alpha\widetilde{d}_{33}+\beta & \alpha^2\widetilde{d}_{34} & \dots & \alpha^{n-2}\widetilde{d}_{3n} \\ 0 & 0 & 1 & \alpha\widetilde{d}_{44}+\beta & \dots & \alpha^{n-3}\widetilde{d}_{4n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \alpha\widetilde{d}_{nn}+\beta \end{pmatrix}. \tag{31}$$

Dem. Es un ejercicio que dejamos al lector. Puede verse en [Esc11]. □

Observación 5.3. Obviamente la matriz $\alpha D+I\beta$ tiene como espectro el espectro transformado de D, pero lo importante, aquí, es que al ser $\alpha \in \mathbb{C}$, $\alpha D+I\beta$ no tendría subdiagonal estrictamente positiva. El anterior resultado resuelve ese problema. Claramente D^{φ} es unitariamente semejante a $\alpha D + I\beta$.

Ejemplo 5.2. En el ejemplo de la matriz de Pascal hemos visto que $\alpha = 1$ y $\beta = 1$, luego en este caso y aplicando la fórmula (30) tenemos que

$$D = U = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad \Longrightarrow \qquad D^{\varphi} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 0 & 1 & 1 & 0 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Nótese que $\sigma(D^{\varphi})$ no es mas que $\sigma(U)$ trasladado a la derecha una unidad. Además $U+I$ es subnormal, lo que podemos probar directamente, sin utilizar el resultado que relaciona el que M sea de momentos con el que D sea subnormal. Si $p(z)$ es un polinomio y U el shift-right (que es subnormal) el operador $p(U)$ es subnormal (ver prob. 198 [Hal80]), si $N = \text{men}(U)$ (mínima extensión normal de U), tendremos que $p(N)$, que también es normal, sera la extensión de $p(U)$.

Observación 5.6. Es importante no utilizar equivocadamente los anteriores resultados ya que el orden en el que se efectuan las transformaciones es importante. Si queremos llevar los momentos de los P.O. de Tchebyschev de primera especie, del intervalo $[-1,1]$ al $[0,1]$, podemos seguir dos caminos, consistentes ambos en dos transformaciones, por ejemplo: Una traslación $\alpha = 1$ con $\beta = 1$ y luego una homotecia $\alpha = 1/2$ con $\beta = 0$, o bien una homotecia $\alpha = 1/2$ con $\beta = 0$ y luego una traslación $\alpha = 1$ con $\beta = 1/2$. En cada caso si se desea la matriz de momentos normalizada habrá que dividir por el c_{00} final.

5.4 Expresión paramétrica de D en el caso en que M sea de Toeplitz

Obtener ahora la fórmula de recurrencia para los mónicos en el caso de que M sea de Toeplitz es sencillo ya que resulta inmediato probar.

Corolario 5.1. *Se cumple que*

$$|\Phi_n(0)|^2 = 1 - \left(\frac{\kappa_{n-1}}{\kappa_n}\right)^2. \tag{32}$$

Asimismo

$$|\Phi_n(0)| < 1, \qquad \forall n \in \mathbb{N}. \tag{33}$$

Dem. A partir de (16) y expresando el polinomio normalizado en función del mónico queda

$$\kappa_n^2 - \kappa_{n-1}^2 = K_n(0,0) - K_{n-1}(0,0) = |\varphi_n(0)|^2 = \kappa_n^2\, |\Phi_n(0)|^2,$$

dividiendo por κ_n^2, tenemos (32).

Se cumple obviamente que $0 < (\kappa_{n-1}/\kappa_n)^2 \leq 1$. Luego necesariamente a partir de (32) se verifica que $0 \leq |\Phi_n(0)| < 1$, es decir (33). La consecuencia es que cuando M es de Toeplitz HDP, las raíces de los polinomios están todas dentro de la circunferencia de radio unidad. Propiedad muy conocida de la teoría de P.O. en la circunferencia.

Además observamos, lo que luego tendrá importantes[24] consecuencias, que la sucesión de números reales positivos $\{\kappa_n\}$ verifica $\kappa_{n-1} \leq \kappa_n$, $\forall n$, es decir, se trata de una sucesión monótona creciente. □

Teorema 5.1. *Si $\{\varphi_n(z)\}_{n=0}^{\infty}$ son polinomios ortonormales sobre la circunferencia unidad, entonces se verifican las siguientes relaciones:*

$$\begin{aligned} &i) \qquad \kappa_{n-1}\, z\, \varphi_{n-1}(z) &=& \kappa_n \varphi_n(z) - \varphi_n(0)\varphi_n^*(z), &(34)\\ &ii) \qquad \kappa_{n-1}\varphi_n(z) &=& \kappa_n\, z\, \varphi_{n-1}(z) + \varphi_n(0)\varphi_{n-1}^*(z). &(35) \end{aligned}$$

Dem.

i) Partimos del lado derecho de (18), es decir de

$$K_n(z,0) = \sum_{k=0}^{n} \varphi_k(z)\overline{\varphi_k(0)} = \kappa_n \varphi_n^*(z),$$

[24] Podrá darse que bien $\kappa_n \to \infty$, o bien que $\kappa_n \to \kappa$, cuando esto último ocurre se dice que se cumple la condición de Szegö, que implicará, como veremos, que la sucesión de polinomios no sean densos en $L^2_\nu(\mathbb{T})$.

de donde se deduce por la propia definición de n-núcleos que

$$\kappa_n\varphi_n^*(z) - \kappa_{n-1}\varphi_{n-1}^*(z) = K_n(z,0) - K_{n-1}(z,0) = \varphi_n(z)\overline{\varphi_n(0)}. \tag{36}$$

Recordemos que $q_n^*(z) = z^n\overline{q_n(z)}$, por tanto aplicando la definición de los reversos a la ecuación anterior queda

$$\kappa_n z^n\overline{\varphi_n(z)} - \kappa_{n-1}z^{n-1}\overline{\varphi_{n-1}(z)} = \varphi_n(z)\overline{\varphi_n(0)},$$

tomando conjugados, y multiplicando por z^n, queda

$$\kappa_n z^n\overline{z}^n\varphi_n(z) - \kappa_{n-1}z^n\overline{z}^{n-1}\varphi_{n-1}(z) = z^n\overline{\varphi_n(z)}\varphi_n(0),$$

se tendrá que $z^n\overline{z}^n = 1$, etc., y volviendo a aplicar la definición de reverso resultará

$$\kappa_n\varphi_n(z) - \kappa_{n-1}z\varphi_{n-1}(z) = \varphi_n(0)\varphi_n^*(z), \tag{37}$$

con lo que hemos probado (34).

ii) Veamos ahora que se verifica (35). En efecto, despejemos[25] $\varphi_n^*(z)$ en (37) y substituyamos en (36), quedará

$$\kappa_n\left(\frac{\kappa_n\varphi_n(z) - \kappa_{n-1}z\varphi_{n-1}(z)}{\varphi_n(0)}\right) - \kappa_{n-1}\varphi_{n-1}^*(z) = \varphi_n(z)\overline{\varphi_n(0)},$$

multiplicando por $\varphi_n(0)$ resulta

$$\kappa_n^2\varphi_n(z) - \kappa_n\kappa_{n-1}z\varphi_{n-1}(z) - \kappa_{n-1}\varphi_n(0)\varphi_{n-1}^*(z) = |\varphi_n(0)|^2\varphi_n(z).$$

Sabemos que $|\Phi_n(0)|^2 = 1 - \frac{\kappa_{n-1}^2}{\kappa_n^2}$, o lo que es lo mismo $|\varphi_n(0)|^2 = \kappa_n^2 - \kappa_{n-1}^2$, substituyendo este último valor en la anterior ecuación, cancelando $\kappa_n^2\varphi_n(z)$ en ambos términos, y eliminando en todos los factores κ_{n-1} podemos despejar $\kappa_{n-1}\varphi_n(z)$ y queda

$$\kappa_{n-1}\varphi_n(z) = \kappa_n\ z\ \varphi_{n-1}(z) + \varphi_n(0)\varphi_{n-1}^*(z),$$

que es lo que queríamos probar. $\square$

Corolario 5.2. *Se cumple la recurrencia para $n \geq 2$ siguiente*

$$z\varphi_{n-1}(z) = \frac{\kappa_{n-1}}{\kappa_n}\varphi_n(z) - \frac{\kappa_{n-1}}{\kappa_n}\frac{\varphi_n(0)}{\varphi_{n-1}(0)}\varphi_{n-1}(z) + \frac{\kappa_{n-2}}{\kappa_n}\frac{\varphi_n(0)}{\varphi_{n-1}(0)}z\varphi_{n-2}(z). \tag{38}$$

Dem. Reescribamos (34) para $n := n-1$ y despejemos $\varphi_{n-1}^*(z)$ tendremos

$$\varphi_{n-1}^*(z) = \frac{\kappa_{n-1}\varphi_{n-1}(z) - \kappa_{n-2}z\varphi_{n-1}(z)}{\varphi_{n-1}(0)}.$$

[25] Podría ocurrir que $\varphi_n(0) = 0$, y por tanto $z = 0$ sería también raíz de $\Phi_n(z)$, con lo que $\Phi_n(0) = 0$, por tanto $\kappa_{n-1}^2 = \kappa_n^2$, y ambas fórmulas de recurrencia coincidirían. Se tendría que $\varphi_n(z) = z\varphi_{n-1}(z)$, por lo que podemos suponer que $\varphi_n(0) \neq 0$.

Basta con despejar $z\varphi_{n-1}(z)$ en (35). Substituyamos este valor en (35) tendremos

$$\kappa_{n-1}\varphi_n(z) = \kappa_n \boxed{z\varphi_{n-1}(z)} + \varphi_n(0)\left(\frac{\kappa_{n-1}\varphi_{n-1}(z) - \kappa_{n-2}z\varphi_{n-1}(z)}{\varphi_{n-1}(0)}\right).$$

Despejando la expresión recuadrada tenemos que

$$z\varphi_{n-1}(z) = \frac{\kappa_{n-1}}{\kappa_n}\varphi_n(z) - \frac{\kappa_{n-1}}{\kappa_n}\frac{\varphi_n(0)}{\varphi_{n-1}(0)}\varphi_{n-1}(z) + \frac{\kappa_{n-2}}{\kappa_n}\frac{\varphi_n(0)}{\varphi_{n-1}(0)}z\varphi_{n-2}(z),$$

que es lo que queríamos probar. $\square$

Finalmente se tiene el teorema siguiente:

Teorema 5.2. *Dada la sucesión de polinomios ortogonales $\{\varphi_n(z)\}$, sobre la circunferencia unidad, se cumple que*

$$z\begin{pmatrix} \varphi_0(z) \\ \varphi_1(z) \\ \varphi_2(z) \\ \vdots \\ \varphi_{n-1}(z) \end{pmatrix} = \begin{pmatrix} -\Phi_1\overline{\Phi}_0 & \frac{\kappa_0}{\kappa_1} & 0 & \cdots & 0 & 0 \\ \frac{-\kappa_0}{\kappa_1}\Phi_2\overline{\Phi}_0 & -\Phi_2\overline{\Phi}_1 & \frac{\kappa_1}{\kappa_2} & \cdots & 0 & 0 \\ \frac{-\kappa_0}{\kappa_2}\Phi_3\overline{\Phi}_0 & \frac{-\kappa_1}{\kappa_2}\Phi_3\overline{\Phi}_1 & -\Phi_3\overline{\Phi}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{-\kappa_0}{\kappa_{n-1}}\Phi_n\overline{\Phi}_0 & \frac{-\kappa_1}{\kappa_{n-1}}\Phi_n\overline{\Phi}_1 & \frac{-\kappa_2}{\kappa_{n-1}}\Phi_n\overline{\Phi}_2 & \cdots & \frac{-\kappa_{n-2}}{\kappa_{n-1}}\Phi_n\overline{\Phi}_{n-2} & -\Phi_n\overline{\Phi}_{n-1} \end{pmatrix} \times$$

$$\times \begin{pmatrix} \varphi_0(z) \\ \varphi_1(z) \\ \varphi_2(z) \\ \vdots \\ \varphi_{n-1}(z) \end{pmatrix} + \frac{\kappa_{n-1}}{\kappa_n}\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \varphi_n(z) \end{pmatrix}$$

donde los Φ_j y $\overline{\Phi}_k$ están particularizados en el origen.

Dem. Escribiendo la expresión (38) para los valores de n siguientes: $n-1$, $n-2$, ..., 2 y 1 y sustituyendo sucesivamente cada una de estas relaciones en la expresión anterior se obtiene

$$\begin{aligned} z\varphi_{n-1}(z) &= \frac{\kappa_{n-1}}{\kappa_n}\varphi_n(z) + \left[\frac{\kappa_{n-2}^2 - \kappa_{n-1}^2}{\kappa_n\kappa_{n-1}}\right]\frac{\varphi_n(0)}{\varphi_{n-1}(0)}\varphi_{n-1}(z) \\ &\quad + \ldots + \left[\frac{\kappa_0^2 - \kappa_1^2}{\kappa_n\kappa_{n-1}}\right]\frac{\varphi_n(0)}{\varphi_1(0)}\varphi_1(z) + \left[\frac{0 - \kappa_1^2}{\kappa_n\kappa_{n-1}}\right]\frac{\varphi_n(0)}{\varphi_0(0)}\varphi_0(z). \end{aligned} \tag{39}$$

que es la fórmula de recurrencia larga; si comparamos con la fórmula (23), resulta que $d_{n+1,n} = \kappa_{n-1}/\kappa_n$, y si es $j \leq n$, que

$$d_{j,n} = \left[\frac{\kappa_{j-2}^2 - \kappa_{j-1}^2}{\kappa_n\kappa_{n-1}}\right]\frac{\varphi_n(0)}{\varphi_{j-1}(0)}.$$

Como se tiene que $\varphi_n(z) = \kappa_n\Phi_n(z)$ y $|\varphi_{j-1}(0)|^2 = \kappa_{j-1}^2|\Phi_{j-1}(0)|^2 = \kappa_{j-1}^2 - \kappa_{j-2}^2$, se obtiene de aquí que

$$d_{j,n} = \left[\frac{\kappa_{j-2}^2 - \kappa_{j-1}^2}{\kappa_n\kappa_{n-1}}\right]\frac{\kappa_n\Phi_n(0)\kappa_{j-1}\overline{\Phi_{j-1}(0)}}{\kappa_{j-1}^2 - \kappa_{j-2}^2} = \frac{-\kappa_{j-1}}{\kappa_{n-1}}\Phi_n(0)\overline{\Phi}_{j-1}(0),$$

es la expresión de los términos de la matriz D, siendo $j \leq n$. La relación (39) puede escribirse ahora como

$$\begin{aligned} z\varphi_{n-1}(z) &= \frac{\kappa_{n-1}}{\kappa_n}\varphi_n(z) - \Phi_n\overline{\Phi}_{n-1}\varphi_{n-1}(z) - \frac{\kappa_{n-2}}{\kappa_{n-1}}\Phi_n\overline{\Phi}_{n-2}\varphi_{n-2}(z) \\ & \quad \ldots - \frac{\kappa_1}{\kappa_{n-1}}\Phi_n\overline{\Phi}_1\varphi_1(z) - \frac{\kappa_0}{\kappa_{n-1}}\Phi_n\overline{\Phi}_0\varphi_0(z), \end{aligned} \tag{40}$$

donde los productos $\Phi_n\overline{\Phi}_{j-1}$ están particularizados en $z = 0$. Obteniéndose además la igualdad matricial del enunciado. $\square$

Observación 5.4. Transponiendo resulta, que la sección de orden n de la matriz D es

$$D_n = \begin{pmatrix} -\Phi_1\overline{\Phi}_0 & \frac{-\kappa_0}{\kappa_1}\Phi_2\overline{\Phi}_0 & \frac{-\kappa_0}{\kappa_2}\Phi_3\overline{\Phi}_0 & \ldots & \frac{-\kappa_0}{\kappa_{n-2}}\Phi_{n-1}\overline{\Phi}_0 & \frac{-\kappa_0}{\kappa_{n-1}}\Phi_n\overline{\Phi}_0 \\ \frac{\kappa_0}{\kappa_1} & -\Phi_2\overline{\Phi}_1 & \frac{-\kappa_1}{\kappa_2}\Phi_3\overline{\Phi}_1 & \ldots & \frac{-\kappa_1}{\kappa_{n-2}}\Phi_{n-1}\overline{\Phi}_1 & \frac{-\kappa_1}{\kappa_{n-1}}\Phi_n\overline{\Phi}_1 \\ 0 & \frac{\kappa_1}{\kappa_2} & -\Phi_3\overline{\Phi}_2 & \ldots & \frac{-\kappa_2}{\kappa_{n-2}}\Phi_{n-1}\overline{\Phi}_2 & \frac{-\kappa_2}{\kappa_{n-1}}\Phi_n\overline{\Phi}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\Phi_{n-1}\overline{\Phi}_{n-2} & \frac{-\kappa_{n-2}}{\kappa_{n-1}}\Phi_n\overline{\Phi}_{n-2} \\ 0 & 0 & 0 & \ldots & \frac{\kappa_{n-2}}{\kappa_{n-1}} & -\Phi_n\overline{\Phi}_{n-1} \end{pmatrix}. \tag{41}$$

Los productos $\Phi_n\overline{\Phi}_{j-1}$ están particularizados en $z = 0$.

Proposición 5.5. *Si M es HDP infinita y D es la correspondiente matriz de Hessenberg, entonces*[26]

$$M \text{ es de Toeplitz} \iff D^H D = I.$$

Dem.

$$\begin{aligned} D^H D = I \;\Leftrightarrow\; (T^{-1}S_L T)(T^H S_R T^{-H}) = I \;\Leftrightarrow\; T^{-1}S_L M S_R T^{-H} = I \\ \Leftrightarrow\; S_L M S_R = TT^H = M \;\Leftrightarrow\; M \text{ es de Toeplitz} \;. \end{aligned}$$

Nótese que todas las operaciones con matrices infinita tienen sentido por ser T y T^H y sus inversas matrices triangulares así como S_R, cuya inversa por la izquierda es S_L. $\square$

Observación 5.5. Nótese que no afirmamos, en principio que $DD^H = I$. Si la medida fallara en satisfacer la condición de Szegö, es decir, si no se cumple $\sum_{k=}^{\infty} |\Phi_n(0)|^2 < \infty$ el sistema de P.O. sería completo, es decir $\overline{\Pi} = L^2_\mu(\mathbb{T})$, y D sería normal y por tanto unitaria, y si se cumpliría $D^H D = DD^H = I$.

6 El teorema de Szegö

No incluiremos la prueba del importante teorema de Szegö-Kolmogorov-Krein, debido a su extensión. Ni tampoco el teorema que prueba que si falla Szegö el sistema de polinomios es completo. Aunque los enunciaremos y pondremos diversos ejemplos. El teorema de Szegö afirma (ver cap. 3 de [Nik91]) lo siguiente:

[26] Ver ([Tom11])

Teorema 6.1. *Sea una medida $d\sigma(e^{i\theta})$ sobre la circunferencia unidad y tal que*

$$d\sigma(e^{i\theta}) = p(\theta)d\theta + d\nu(e^{i\theta}), \tag{42}$$

siendo $p(\theta)$ una función no negativa, sumable en $[0, 2\pi]$, y sea ν la suma de las componentes[27] *singulares y discretas de σ. Llamamos $\mathfrak{G}(p)$ a la media geométrica de la función $p(\theta)$ sobre la circunferencia unidad, es decir*

$$\mathfrak{G}(p(\theta)) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \ln(p(\theta))d\theta\right),$$

se cumple entonces que

$$\lim_n \frac{1}{\kappa_n^2} = \lim_n \frac{\Delta_{n+1}}{\Delta_n} = \lim_n \frac{1}{K_n(0,0)} = \mathfrak{G}(p).$$

Por otro lado tenemos (ver pág. 104 [Nik91]):

Teorema 6.2. *Si σ es una medida de la forma (42). El sistema de polinomios es completo en $L^2_\sigma([0, 2\pi])$, si y solo si la función de peso $p(\theta)$ falla en satisfacer la condición de Szegö.*

Corolario 6.1. *Si $L^2_\sigma = \overline{\Pi}$, el operador $S_\sigma : \overline{\Pi} \to \overline{\Pi}$, tal que $S_\sigma(p(z)) = zp(z)$, es normal*[28] *y por tanto lo es su representación matricial D, respecto de la base ortonormal.*

Ejemplo 6.1. Consideremos la medida $d\mu(\theta) = (1 + \cos\theta)/(2\pi)d\theta$, la matriz de momentos[29] es

$$M = \begin{pmatrix} 1 & 1/2 & 0 & 0\dots & \\ 1/2 & 1 & 1/2 & 0\dots & \\ 0 & 1/2 & 1 & 1/2\dots & \\ 0 & 0 & 1/2 & 1\dots & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Resulta inmediato por inducción que $\Delta_n = (n+1)/2^n$, por lo que

$$\lim_{n\to\infty} \frac{1}{\kappa_n^2} = \lim_{n\to\infty} \frac{\Delta_n}{\Delta_{n-1}} = \lim_{n\to\infty} \frac{(n+1)/2^n}{n/2^{(n-1)}} = \frac{1}{2}.$$

Por otro lado tenemos que puede, con cierto trabajo, calcularse que

$$\mathfrak{G}(1 + \cos\theta) = \exp\,(\frac{1}{2\pi}\int_0^{2\pi} \ln(1 + \cos\theta)d\theta = \exp\,(-\ln 2 - 2\pi i) = \frac{1}{2},$$

luego se obtiene la identidad esperada.

Observación 6.1. Este teorema se satisface también en el caso extremo en que $1/\infty = 0 = e^{-\infty}$. En ese caso falla la condición de Szegö, lo que equivale obviamente a que $K_n(0,0) = \sum_{k=0} |p_n(0)|^2 = +\infty$ que, por el teorema de comparación en el límite, es lo mismo que afirmar que $\sum_{k=0} |\Phi_k(0)|^2 = +\infty$ o que $\kappa_n \to \infty$.

[27] Es decir $\frac{d\nu(e^{i\theta})}{d\theta} = 0$, a.e. en $[(0, 2\pi)$.

[28] En este caso unitario, ya que S_σ coincide con $N_\sigma = L^2_\sigma \to L^2_\sigma$, siendo $N_\sigma(f) = zf$ y $N^H_\sigma(f) = \overline{z}f$ y por tanto $\|S_\sigma\|^2 = \|N_\sigma\|^2 = |z|^2 = 1$.

[29] La matriz es de Toplitz y su serie de Fourier es $1/2e^{-i\theta} + 1 + 1/2e^{-i\theta} = 1 + \cos\theta$, y dividimos por 2π para que resulte de probabilidad.

Ejemplo 6.2. Funciones tan simples como $p(\theta) = (\cos(m\theta) + |\cos(m\theta)|)/2$, con $m \in \mathbb{N}$, fallan la condición de Szegö porque su soporte no llena la circunferencia. Lo que también ocurre en los casos puramente atómicos no densos sobre la circunferencia.

Todo depende de la media geométrica de $p(\theta)$. Un ejemplo que no satisface la condición de Szegö y su soporte si llena la circunferencia es la función $p(\theta) = \exp(-1/|\cos(\theta)|)$, ya que es claro que la integral impropia vale $\int_0^{2\pi} \ln(\exp(-1/|\cos(\theta)|)d\theta = -\infty$, luego $\mathfrak{G}(p) = 0$.

Hay una condición suficiente[30] para saber si el soporte llena o no la circunferencia. Si $\Phi_n(0) \to 0$, podemos asegurar que el soporte llena la circunferencia.

Funciones como $p(\theta) = \cos(m\theta) + 1$, $m \in \mathbb{N}$, cuyo soporte llena la circunferencia y alcanza el cero m veces, podemos escoger m tan grande como deseemos, sin embargo siempre satisfacen la condición de Szegö. Incluso una familia de funciones como $p(\theta) = |\operatorname{sen}(\theta/2)|^m$, que se aproximan a 0 al aumentar $m \in \mathbb{N}$. Se calcula que para $m = 20$, $\kappa = 1024$; para $m = 30$, $\kappa = 32768$; para $m = 40$, $\kappa = 1048576$, etc., luego para todo m las funciones de esta familia satisfacen la condición de Szegö.

El ejemplo que ponemos a continuación, lo hemos construido a partir de los $\Phi_n(0)$. No conocemos la medida, pero es claro por lo arriba señalado, que llena la circunferencia, y sin embargo no satisface la condición de Szegö.

Ejemplo 6.3. Supongamos que $\Phi_n(0) = 1/\sqrt{n+1}$, resulta inmediato que falla la condición de Szegö, ya que $\sum_{n=0}^{\infty} |\Phi_n(0)|^2 = +\infty$, por lo que

$$D = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2\cdot3}} & -\frac{1}{\sqrt{3\cdot4}} & -\frac{1}{\sqrt{4\cdot5}} & -\frac{1}{\sqrt{5\cdot6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ \frac{\sqrt{1}}{\sqrt{2}} & -\frac{1}{\sqrt{2\cdot3}} & -\frac{1}{\sqrt{3\cdot4}} & -\frac{1}{\sqrt{4\cdot5}} & -\frac{1}{\sqrt{5\cdot6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & -\frac{1}{\sqrt{3\cdot4}} & -\frac{1}{\sqrt{4\cdot5}} & -\frac{1}{\sqrt{5\cdot6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ 0 & 0 & \frac{\sqrt{3}}{\sqrt{4}} & -\frac{1}{\sqrt{4\cdot5}} & -\frac{1}{\sqrt{5\cdot6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ 0 & 0 & 0 & \frac{\sqrt{4}}{\sqrt{5}} & -\frac{1}{\sqrt{5\cdot6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ 0 & 0 & 0 & 0 & \frac{\sqrt{5}}{\sqrt{6}} & \cdots & -\frac{1}{\sqrt{n(n+1)}} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \\ 0 & 0 & 0 & 0 & 0 & \cdots & \frac{n}{\sqrt{n+1}} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \end{bmatrix},$$

es unitaria, es decir $D^H D = DD^H$. Se puede verificar directamente que esta matriz es unitaria, basta utilizar repetidamente el trivial resultado de series $\displaystyle\sum_{n=m}^{\infty} \left[\frac{-1}{\sqrt{n(n+1)}}\right]^2 = \frac{1}{m}$.

Incidentalmente $S_R - D$ es una matriz tal que todas las diagonales tienden a cero, pero no define un operador compacto en ℓ^2. Nótese que $D + (S_R - D) = S_R$ y al ser D unitaria su índice en un punto del interior de la circunferencia es 0, y como difiere en un operador

[30]Probarlo requeriría algunos resultado no triviales como la desigualdad de Ullman, ver pág. 5 [Sta92] y nociones sobre la capacidad logarítmica de un conjunto.

compacto de S_R, tendría que tener el mismo índice que S_R (y el de S_R es -1 en los puntos de dentro) luego $S_R - D$ no define un operador compacto.

Agradecimiento Este trabajo ha sido realizado como parte del proyecto MTM 2016-80582-R (AEI/FEDER, U.E.) del Ministerio de Economía, Industria y Competitividad.

References

[Atz75] A. Atzmon, *A moment problem for positive measures on the unit disc*, *Pacific J. Mathematics*, **59**, 2, (1975).

[Bre80] C. Brezinski, "Padé-type Approximation and General Orthogonal Polynomials", Birkahuser Verlag, Bassel (1980).

[Chi78] Theodore S. Chihara, "An Introduction to Orthogonal Polynomials", Gordon and Breach, New York (1978).

[Con81] John B. Conway, "Subnormal operators", Pitman, London. (1981).

[Dav75] Philip J. Davis, "Interpolation and Approximation", Dover Publications Inc. New York (1975).

[Esc11] C. Escribano, A. Giraldo, M.A. Sastre, E. Torrano, *Computing the Hessenberg matrix associated with a self-similar measure*, Journal of Approximation Theory 163, Elsevier (2011), 49-64.

[Hal80] Paul R. Halmos, "A Hilbert Space Problem Book", Graduate Texts in Mathematics, Springer- Verlag, New York (1980).

[Lan85] P. Lancaster and M. Tismenetski, "Theory of Matrices", Academic Press,(1985).

[Nik91] E. M. Nikishin and V. N. Sorokin, "Rational Approximations and Orthogonality", Translations of Mathematical Monographs, Vol 92, A.M.S. (1991).

[Sta92] Herbert Stahl and Vilmos Totik, "General Orthogonal Polynomials", Cambridge University Press, Canada (1992).

[Sze39] Gabor Szegö, "Orthogonal Polynomials", American Mathematical Society, Coloquium Publications, Vol. 32, primera edición (1939), cuarta edición (1975).

[Tom11] V. Tomeo, E. Torrano, *Two applications of the subnormality of the Hessenberg matrix related to general orthogonal polynomials*, Linear Algebra and its Applications, Vol. 435, (2011), 2314-2320.

[Tor93] E. Torrano and R. Guadalupe, *On the moment problem in the bounded case*, Journal of Comp. and App. Math. Vol. 49 (1993).

Cuadernos de Trabajo
Facultad de Estudios Estadísticos

CT03/2018 **Las matemáticas en el cine**
Gloria Cabrera Gómez

CT02/2018 **Brecha madre-padre en el uso de las medidas de conciliación y su efecto sobre las carreras profesionales de las madres**
José Andrés Fernández Cornejo y Lorenzo Escot (Coordinadores)
Autores: José Andrés Fernández Cornejo, Lorenzo Escot, Eva María del Pozo, Sabina Belope Nguema, Cristina Castellanos Serrano, Miryam Martínez, Ana Bernabeu, Lorenzo Fernández Franco, Juan Ignacio Cáceres, María Ángeles Medina Sánchez

CT01/2018 **El campo de valores de una matriz y su aplicación a los polinomios ortogonales**
Venancio Tomeo Perucha y Emilio Torrano Giménez

CT02/2015 **Prospectiva del proceso electoral a Rector de la Universidad Complutense 2015. Resultados del sondeo de intención de voto en primera y segunda vuelta**
Lorenzo Escot Mangas, Eduardo Ortega Castelló, Lorenzo Fernández Franco y Kenedy Alba (Coordinadores)

CT01/2015 **The polemic but often decisive contribution of computer science to linguistic and statistical research on translation accuracy and efficiency**
Antonio Ñíguez Bernal

CT05/2014 **Las medidas estadísticas: ejercicios motivadores**
Almudena Pajares García y Venancio Tomeo Perucha

CT04/2014 **Jugando con la estadística (y la probabilidad)**
Gloria Cabrera Gómez

CT03/2014 **Análisis Estadístico de las Consultas a la Base de Datos de Ayudas e Incentivos a Empresas de la Dirección General de Industria y de la PYME.**
Adolfo Coello de Portugal Muñoz, Juana María Alonso Revenga

CT02/2014 **Values of games with weighted graphs**
E. González-Arangüena, C. Manuel; M. del Pozo

CT01/2014 **Estimación de la tasa de retorno de la carta del censa de los Estados Unidos a través del modelo de regresión lineal y técnicas de predicción inteligentes.**
José Luis Jiménez-Moro y Javier Portela García Miguel

CT03/2013 **Provisión de siniestros de incapacidad temporal utilizando análisis de supervivencia.**
Ana Crespo Palacios y Magdalena Ferrán Aranaz

CT02/2013 **Consumer need for touch and Multichannel Purchasing Behaviour.**
R. Manzano, M. Ferrán y D. Gavilán

CT01/2013 **Un método gráfico de comparación de series históricas en el mercado bursátil.**
Magdalena Ferrán Aranaz

CT03/2012 **Calculando la matriz de covarianzas con la estructura de una red Bayesiana Gaussiana**
Miguel A. Gómez-Villegas y Rosario Susi

CT02/2012	**What's new and useful about chaos in economic science.** *Andrés Fernández Díaz, Lorenzo Escot and Pilar Grau-Carles*
CT01/2012	**A social capital index** *Enrique González-Arangüena, Anna Khmelnitskaya, Conrado Manuel, Mónica del Pozo*
CT04/2011	**La metodología del haz de rectas para la comparación de series temporales.** *Magdalena Ferrán Aranaz*
CT03/2011	**Game Theory and Centrality in Directed Social Networks** *Mónica del Pozo, Conrado Manuel, Enrique González-Arangüena y Guillermo Owen.*
CT02/2011	**Sondeo de intención de voto en las elecciones a Rector de la Universidad Complutense de Madrid 2011** *L.Escot, E. Ortega Castelló y L. Fernández Franco (coords)*
CT01/2011	**Juegos y Experimentos Didácticos de Estadística y Probabilidad** *G. Cabrera Gómez y Mª.J. Pons Bordería*
CT04/2010	**Medio siglo de estadísticas en el sector de la construcción residencial** *M. Ferrán Aranaz*
CT03/2010	**Sensitivity to hyperprior parameters in Gaussian Bayesian networks.** *M.A. Gómez-Villegas, P. Main, H. Navarro y R. Susi*
CT02/2010	**Las políticas de conciliación de la vida familiar y laboral desde la perspectiva del empleador. Problemas y ventajas para la empresa.** *R. Albert, L. Escot, J.A. Fernández Cornejo y M.T. Palomo*
CT01/2010	**Propiedades exóticas de los determinantes** *Venancio Tomeo Perucha*
CT05/2009	**La predisposición de las estudiantes universitarias de la Comunidad de Madrid a auto-limitarse profesionalmente en el futuro por razones de conciliación** *R. Albert, L. Escot y J.A. Fernández Cornejo*
CT04/2009	**A Probabilistic Position Value** *A. Ghintran, E. González–Arangüena y C. Manuel*
CT03/2009	**Didáctica de la Estadística y la Probabilidad en Secundaria: Experimentos motivadores** *A. Pajares García y V. Tomeo Perucha*
CT02/2009	**La disposición entre los hombres españoles a tomarse el permiso por nacimiento. ¿Influyen en ello las estrategias de conciliación de las empresas?** *L. Escot, J.A. Fernández-Cornejo, C. Lafuente y C. Poza*
CT01/2009	**Perturbing the structure in Gaussian Bayesian networks** *R. Susi, H. Navarro, P. Main y M.A. Gómez-Villegas*
CT09/2008	**Un experimento de campo para analizar la discriminación contra la mujer en los procesos de selección de personal** *L. Escot, J.A. Fernández Cornejo, R. Albert y M.O. Samamed*
CT08/2008	**Laboratorio de Programación. Manual de Mooshak para el alumno** *D. I. de Basilio y Vildósola, M. González Cuñado y C. Pareja Flores*

CT07/2008 **Factores de protección y riesgo de infidelidad en la banca comercial**
J. Mª Santiago Merino

CT06/2008 **Multinationals and foreign direct investment: Main theoretical strands and empirical effects**
María C. Latorre

CT05/2008 **On the Asymptotic Distribution of Cook's distance in Logistic Regression Models**
Nirian Martín y and Leandro Pardo

CT04/2008 **La innovación tecnológica desde el marco del capital intelectual**
Miriam Delgado Verde, José Emilio Navas López, Gregorio Martín de Castro y Pedro López Sáez

CT03/2008 **Análisis del comportamiento de los indecisos en procesos electorales: propuesta de investigación funcional predictivo-normativa**
J. Mª Santiago Merino

CT02/2008 **Inaccurate parameters in Gaussian Bayesian networks**
Miguel A. Gómez-Villegas, Paloma Main and Rosario Susi

CT01/2008 **A Value for Directed Communication Situations.**
E. González–Arangüena, C. Manuel, D. Gómez, R. van den Brink

UNIVERSIDAD COMPLUTENSE
MADRID