UNIVERSIDAD DE SEVILLA
FACULTAD DE MATEMÁTICAS

POLINOMIOS MATRICIALES ORTONORMALES

Memoria presentada por Pedro López Rodríguez para optar al grado de Doctor en Ciencias Matemáticas. Sevilla, octubre de 1995.

Antonio José Durán Guardeño
Director de la tesis

Pedro López Rodríguez

Este trabajo es debido al empeño y paciencia que durante varios años me ha dedicado el profesor Antonio José Durán Guardeño, a quién estoy sumamente agradecido.

También estoy muy agradecido al profesor Christian Berg (Universidad de Copenhague) por su hospitalidad y por sus valiosas sugerencias y constante apoyo.

A mis padres

INDICE

INTRODUCCIÓN

Una sucesión de polinomios ortogonales $p_n(t)$ $(n = 0, 1, 2, \dots)$ en la recta real con respecto a una medida positiva μ siempre satisface una relación de recurrencia de tres términos de la forma

$$tp_n(t) = a_{n+1}p_{n+1}(t) + b_n p_n(t) + a_n p_{n-1}(t), \tag{1}$$

con la condición inicial $p_{-1}(t) = 0$, y siendo a_n y b_n dados por

$$a_n = \int_{\mathbb{R}} tp_{n-1}(t)p_n(t)d\mu(t) \neq 0,$$

y

$$b_n = \int_{\mathbb{R}} tp_n^2(t)d\mu(t) \in \mathbb{R}.$$

El recíproco es también cierto: una sucesión de polinomios que verifique la relación de recurrencia dada por (1) con $a_{n+1} \neq 0$ y $b_n \in \mathbb{R}$ es siempre ortonormal con respecto a una medida positiva en la recta real. Este recíproco fue dado por Favard en 1935 ([**F**]), pero era conocido antes y aparece ya en los libros de Stone ([**S**, teorema 10.27, pag. 545-546]), Perron ([**P**, §36, teorema 4.6/4.7]) y Wintner ([**W**, §32 y §37]), quien atribuye el caso del soporte finito a E. Heine ([**H**, §108]). De forma elemental (careciendo del teorema de representación de Riesz) el resultado está en el trabajo de Stieltjes [**St**, §11] de 1894, y aparentemente el resultado era ya conocido por Chebyshev [**C**] cuando el soporte de la medida μ es finito.

Recientemente (a finales de los años 80) el estudio de un tipo de ortogonalidad más amplia que la definida por medidas positivas ha concentrado gran interés: son los llamados productos escalares de tipo Sobolev. Los polinomios ortogonales con respecto a un producto escalar

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

de tipo Sobolev discreto no verifican una relación de recurrencia de tres términos, pero sí otro tipo de relación más general donde aparece un número impar de términos mayor que tres. Sin embargo, no todas las familias de polinomios que verifican una relación de recurrencia de este tipo son ortogonales con respecto a un producto escalar del tipo Sobolev discreto.

Sorprendentemente, la relación de recurrencia de $(2N + 1)$ términos equivale a la ortogonalidad con respecto a una matriz de medidas, de la que los productos de Sobolev son un caso particular (véase [**D1**], [**D2**] y [**DV**]). Esta relación entre polinomios que satisfacen una fórmula de recurrencia de $(2N + 1)$ términos y polinomios matriciales ortogonales sobre la recta real ha dado impulso al estudio de estos últimos. Ahí está el origen de esta tesis.

Los polinomios matriciales ortogonales han sido estudiados en la última mitad de este siglo de manera esporádica: Krein obtuvo algunos resultados sobre problemas de momentos matriciales desde el punto de vista de la teoría de operadores ([**Kr**]). Más recientemente, en los años 80 han sido relacionados con la teoría de la dispersión por Geronimo ([**G1**]), y se ha establecido un análogo del teorema de Favard para fórmulas de recurrencia matriciales (Aptekarev-Nikishin, [**AN**]). Algunos resultados algebraicos y sobre los ceros fueron establecidos por Zhani ([**Z**]). Por último, durante los años de realización de esta tesis, algunos resultados sobre ceros y fórmulas de cuadratura han sido desarrollados por Sinap y Van Assche desde el punto de vista computacional ([**SV**]).

Sin embargo, sorprende que no se haya llevado a cabo un estudio sistemático de los polinomios matriciales ortogonales sobre la recta real, para determinar qué grado de analogía y diferencia existe con el caso escalar (sí hay un estudio más sistemático para polinomios matriciales ortogonales sobre la circunferencia unidad, véase [**DGK**] y [**G2**]). La intención de esta memoria es dar los primeros pasos en ese sentido. Concluimos esta introducción resumiendo su contenido:

En el capítulo 2 se estudian en detalle las topologías vaga y débil definidas sobre el espacio de las matrices de medidas definidas positivas, y se prueba que al igual que en el caso escalar, el conjunto de soluciones de un problema de momentos matricial es compacto y convexo para ambas topologías, que coinciden sobre dicho conjunto.

En el capítulo 3 se estudian las propiedades básicas de los ceros de una familia de polinomios matriciales ortonormales y la existencia de fórmulas de cuadratura asociadas. Los resultados guardan cierta analogía con el caso escalar, aunque el caracter matricial produce algunas diferencias. Así por ejemplo, mientras que los ceros de los polinomios matriciales ortogonales son como en el caso escalar reales, no son necesariamente simples, pudiendo darse multiplicidad, si bien esta está acotada por el tamaño de los polinomios matriciales. La prueba de los resultados dados es simple y se basa en las propiedades de la matriz de Jacobi asociada.

En el capítulo 4 se generaliza el teorema de Blumenthal, que describe el soporte de una medida positiva cuando los coeficientes de la relación de recurrencia de tres términos tienden a límites finitos, al caso matricial. Se considera una sucesión de polinomios escalares que satisfacen una relación de recurrencia de $(2N + 1)$ términos de modo que las sucesiones de recurrencia que aparecen en dicha relación sean convergentes, y se prueba que el soporte de la medida que ortonormaliza a la sucesión de polinomios matriciales asociados es un intervalo cerrado y quizás dos sucesiones tendiendo a los extremos de dicho intervalo.

En el capítulo 5 se definen los espacios $\mathcal{L}^p(\mu)$ asociados a una matriz de medidas μ definida positiva. Hasta donde sabemos, estos espacios sólo aparecen en la literatura para el caso $p = 2$ ([**R**], [**DS**]). Probamos que estos espacios son espacios de Banach (Hilbert en al caso $p = 2$) y que el dual de $\mathcal{L}^p(\mu)$ es $\mathcal{L}^q(\mu)$, siendo q el exponente conjugado de p. Las demostraciones dadas son más simples que las que aparecen en las referencias citadas (para $p = 2$). También probamos las inclusiones naturales entre estos espacios y damos un resultado básico de interpolación que es una generalización de un conocido teorema de Marcinkiewicz. Completamos el capítulo dando un resultado parcial sobre densidad de polinomios en $\mathcal{L}^1(\mu)$ que es una generalización del conocido teorema de Naimark ([**N**]), que caracteriza las medidas solución de un problema de momentos para las cuales los polinomios son densos en su espacio $\mathcal{L}^1(\mu)$ como aquellas que son extremales en dicho conjunto de soluciones.

Por último, en el capítulo 6 se estudian cuestiones de densidad para el problema de momentos matricial truncado y se describe con detalle el conjunto de matrices de medidas positivas de V_{2n-2} (conjunto de soluciones del problema de momentos matricial truncado de orden $2n-2$) para las cuales los polinomios hasta grado $n-1$ son densos en el correspondiente espacio $\mathcal{L}^2(\mu)$. Estas matrices de medidas son exactamente las extremales del conjunto V_{n-1}, y hay tantas como matrices A verifiquen que la matriz $A_n A$ es hermítica. Estas matrices de medidas son discretas y los puntos del soporte y los pesos se obtienen al desarrollar la expresión $(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda))$ en fracciones simples (P_n son los polinomios ortogonales asociados al problema de momentos y Q_n son los correspondientes polinomios de segunda especie). A diferencia del caso escalar, donde todas estas medidas alcanzan en todos sus puntos del soporte la máxima masa posible en V_{2n-2}, los pesos que los ceros soportan en el caso matricial no tienen porque ser los máximos, y sólo lo son cuando el correspondiente peso es una matriz invertible.

CAPÍTULO 1

PRELIMINARES

Una matriz de medidas μ de tamaño $N \times N$ sobre la recta real es una matriz cuadrada de tamaño $N \times N$ cuyas entradas son medidas complejas de Borel sobre la recta real, es decir

$$\mu = \begin{pmatrix} \mu_{1,1} & \cdots & \mu_{1,N} \\ \vdots & \ddots & \vdots \\ \mu_{N,1} & \cdots & \mu_{N,N} \end{pmatrix}.$$

Se dice que la matriz de medidas μ es definida positiva si para cualquier conjunto de Borel A la matriz numérica $\mu(A) = (\mu_{i,j}(A))_{i,j=1}^{N}$ es semidefinida positiva. Si μ es definida positiva, entonces las medidas $\mu_{i,i}$ que aparecen en la diagonal son todas medidas positivas, y el resto son medidas complejas, siendo $\mu_{i,j} = \overline{\mu_{j,i}}$. Todas las matrices de medidas que manejamos se suponen definidas positivas, salvo que se especifique lo contrario.

Para una matriz de medidas μ definida positiva se define el soporte de μ ($\mathrm{sop}(\mu)$) como el soporte de la medida traza de μ ($\tau\mu$), es decir

$$\mathrm{sop}(\mu) = \mathrm{sop}(\tau\mu) = \mathrm{sop}(\mu_{1,1} + \mu_{2,2} + \cdots + \mu_{N,N}). \tag{1.1}$$

Un polinomio matricial de grado n es una aplicación $P : \mathbb{C} \to \mathbb{C}^{N\times N}$ de la forma

$$P(t) = A_n t^n + A_{n-1} t^{n-1} + \cdots + A_1 t + A_0,$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

donde $A_0, A_1, \ldots, A_n$ son elementos de $\mathbb{C}^{N\times N}$ y $A_n \neq \theta$ (en lo sucesivo, θ representa la matriz nula, el tamaño de la misma queda determinado por el contexto). Este polinomio matricial se dice mónico cuando $A_n = I$, la matriz identidad.

El espacio de polinomios en la variable t con coeficientes en $\mathbb{C}^{N\times N}$ se denota por $\mathbb{C}^{N\times N}[t]$. Dado n un número natural, $\mathbb{C}_n^{N\times N}[t]$ denota el subespacio de $\mathbb{C}^{N\times N}[t]$ formado por todos los polinomios matriciales de grado menor o igual que n.

Obsérvese que mientras que definimos el polinomio $P^*(t)$ como

$$P^*(t) = \sum_{k=0}^{n} A_k^* t^k$$

tenemos que

$$P(t)^* = \left(\sum_{k=0}^{n} A_k t^k\right)^* = \sum_{k=0}^{n} A_k^* \bar{t}^k = P^*(\bar{t}).$$

Se dice que un punto t_0 es un cero del polinomio matricial $P(t)$ si $\det(P(t_0)) = 0$, es decir si t_0 es un cero del polinomio escalar $\det(P(t))$.

Se dice que una sucesión $(P_n)_{n=0}^{\infty}$ de polinomios matriciales con grado de P_n igual a n es ortogonal con respecto a una matriz de medidas definida positiva μ si

$$\int_{\mathbb{R}} P_n(t) d\mu(t) P_m^*(t) = \theta, \text{ para } n, m \geq 0, n \neq m,$$

y se dice que $(P_n)_{n=0}^{\infty}$ es ortonormal si

$$\int_{\mathbb{R}} P_n(t) d\mu(t) P_m^*(t) = \delta_{n,m} I, \text{ para } n, m \geq 0.$$

Obsérvese que si $(P_n)_n$ es una sucesión de polinomios matriciales ortonormales con respecto a μ y $(A_n)_n$ es una sucesión de matrices unitarias, es decir $A_n A_n^* = I$, entonces $(A_n P_n)_n$ es también una sucesión de polinomios matriciales ortonormales con respecto a μ.

Si $(P_n)_{n=0}^{\infty}$ es una sucesión de polinomios matriciales con grado de P_n igual a n y el coeficiente lider A_n de P_n es no singular para todo n, entonces cualquier polinomio matricial $P(t)$ de grado m puede expresarse como $P(t) = \sum_{k=0}^{m} \alpha_k P_k(t)$, siendo α_k ciertas matrices. La prueba es sencilla por inducción: si $P(t)$ es de grado 0, entonces $P(t)$ es una matriz C constante, es decir $P(t) = C$, y por tanto es posible expresar $P(t) = CA_0^{-1}P_0(t)$, pues $P_0(t) = A_0$. Supuesta ahora la propiedad para polinomios de grado $m-1$ y dado $P(t)$ de grado m, se tiene que $P(t)$ es de la forma $P(t) = B_m t^m + Q(t)$, siendo $Q(t)$ un polinomio de grado menor o igual que $m-1$. Así, $P(t) = B_m A_m^{-1} P_m(t) + R(t)$, siendo $R(t)$ otro polinomio de grado menor o igual que $m-1$ que por hipótesis de inducción puede expresarse como una combinación lineal de $P_0, \ldots, P_{m-1}$, teniéndose el resultado.

Se dice que una matriz de medidas μ con todos sus momentos $S_n = \int_{\mathbb{R}} t^n d\mu$ finitos es no degenerada si el único polinomio $P(t)$ de $\mathbb{C}^{N\times N}[t]$ que verifica

$$\int_{\mathbb{R}} P(t)d\mu(t)P^*(t) = \theta$$

es $P(t) = \theta$. De la definición es inmediato que para una matriz de medidas μ no degenerada se tiene que el primer momento S_0 es una matriz numérica definida positiva.

Para una matriz de medidas definida positiva μ con todos sus momentos finitos, son equivalentes el ser no degenerada y el tener una sucesión $(P_n)_{n=0}^{\infty}$ de polinomios ortonormales, siendo P_n de grado n y con coeficiente lider no singular, para todo n. En efecto, si tal sucesión existe, podemos escribir cualquier polinomio $P(t)$ como una combinación lineal finita de los polinomios P_n: $P = \sum_{k=0}^{n} \alpha_k P_k$, siendo todas las matrices α_k nulas sólo cuando $P = \theta$. En consecuencia

$$\int_{\mathbb{R}} P(t)d\mu(t)P^*(t) = \sum_{k=0}^{n} \alpha_k \alpha_k^*$$

que es distinto de cero si algún α_k no es θ.

Por otra parte, si μ no es degenerada, expresando $P(t) = \sum_{i=0}^{n} Y_i t^i$, la definición de no degenerada equivale a decir que para todo n natural la igualdad

$$\sum_{i,j=0}^{n} Y_i S_{i+j} Y_j^* = \theta$$

sólo es posible cuando todas las matrices cuadradas Y_i sean θ. Además, es claro que las matrices Y_i pueden sustituirse por vectores fila y_i y la definición es la misma.

Es posible definir un producto escalar sobre el espacio de polinomios $\mathbb{C}^{N\times N}[t]$ de la forma siguiente:

$$(P, Q) = \int_{\mathbb{R}} P(t) d\mu(t) Q^*(t).$$

Este producto escalar tiene las siguientes propiedades, que se verifican fácilmente:

(1) $(P, Q) = (Q, P)^*$

(2) Si $C_1, C_2 \in \mathbb{C}^{N\times N}$, entonces $(C_1P_1 + C_2P_2, Q) = C_1(P_1, Q) + C_2(P_2, Q)$

(3) $(tP, Q) = (P, tQ)$, para $t \in \mathbb{R}$.

Para un polinomio $P(t) = \sum_{i=0}^{n} \alpha_i t^i$ y un vector columna x en $\mathbb{C}^N$ se tiene que

$$x^*(P, P)x = \sum_{i,j=0}^{n} x^* \alpha_i S_{i+j} \alpha_j^* x$$

y por tanto por ser μ no degenerada se tiene que

$$\mathrm{Ker}(P, P) = \bigcap_{i=0}^{n} \mathrm{Ker}\alpha_i^*.$$

De esta forma, si un α_i es invertible, entonces (P, P) es invertible.

Ortonormalizando la sucesión $I, tI, t^2I, \ldots$ se obtiene una sucesión $P_n(t)$ de polinomios matriciales ortonormales con respecto a μ: llamamos

$R_0(t) = I$ y para $n \geq 1$,

$$R_n(t) = t^n I - \sum_{k=0}^{n-1} (t^n I, R_k)(R_k, R_k)^{-1} R_k(t).$$

Puesto que $(R_0, R_0) = S_0$ es invertible, podemos definir R_1 con esta fórmula, y si (R_k, R_k) es invertible para $0 \leq k \leq n-1$, podemos definir R_n. que verifica que (R_n, R_n) es invertible por ser el coeficiente de la potencia n la matriz identidad, de esta forma se tiene definido $R_n(t)$ por inducción.

Veamos ahora que si $i \neq j$, entonces $(R_i, R_j) = \theta$. En efecto,

$$(R_1, R_0) = (tI, R_0) - (tI, R_0)(R_0, R_0)^{-1}(R_0, R_0) = \theta.$$

Supuesto ahora que $(R_i, R_j) = \theta$, para $0 \leq i, j \leq n-1$, $i \neq j$, tenemos que para $0 \leq k \leq n-1$,

$$\begin{aligned}(R_n, R_k) &= (t^n I, R_k) - \sum_{i=0}^{n-1} (t^n I, R_i)(R_i, R_i)^{-1}(R_i, R_k) \\ &= (t^n I, R_k) - (t^n I, R_k)(R_k, R_k)^{-1}(R_k, R_k) = \theta,\end{aligned}$$

es decir, los polinomios R_n son ortogonales. Para obtener una sucesión ortonormal basta definir $P_n(t) = (R_n, R_n)^{-\frac{1}{2}} R_n(t)$, cuyo coeficiente líder es $(R_n, R_n)^{-\frac{1}{2}}$, que es no singular, y además es claro que $(P_n, P_m) = \delta_{n,m} I$. para $n, m \geq 0$.

Los polinomios escalares ortogonales en la recta real satisfacen una simple relación de recurrencia de tres términos de la forma

$$tp_n(t) = a_{n+1} p_{n+1}(t) + b_n p_n(t) + a_n p_{n-1}(t)$$

siendo $p_{-1} = 0$, $a_n \neq 0$ y $b_n \in \mathbb{R}$. El teorema de Favard prueba que esta relación de recurrencia es equivalente a la ortonormalidad (véase [**Ch.** cap. IV]).

Una relación de recurrencia similar, pero ahora con coeficientes matriciales, es válida para polinomios matriciales ortonormales. Si $(P_n)_n$ es una sucesión de polinomios matriciales ortonormales con respecto a una matriz de medidas μ definida positiva y no degenerada, es posible expresar el polinomio $tP_n(t)$ como una combinación lineal de los polinomios $P_0, \ldots, P_{n+1}$:

$$tP_n(t) = \sum_{i=0}^{n+1} \alpha_i P_i(t), \text{ donde } \alpha_i = (tP_n, P_i).$$

La propiedad (3) del producto escalar $(\cdot, \cdot)$ da que $(tP_n, P_i) = (P_n, tP_i) = \theta$, si $i + 1 \leq n - 1$, y por tanto $\alpha_i = \theta$ si $i \leq n - 2$, lo cual da

$$tP_n(t) = \alpha_{n+1}P_{n+1}(t) + \alpha_n P_n(t) + \alpha_{n-1}P_{n-1}(t).$$

Llamando $A_{n+1} = (tP_n, P_{n+1})$ y $B_n = (tP_n, P_n)$ se tiene que

$$\alpha_{n-1} = (tP_n, P_{n-1}) = (P_n, tP_{n-1}) = (tP_{n-1}, P_n)^* = A_n^*$$

y por tanto se obtiene la relación de recurrencia de tres términos

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_n P_n(t) + A_n^* P_{n-1}(t). \tag{1.2}$$

La matriz B_n es hermítica puesto que

$$B_n = (tP_n, P_n) = (P_n, tP_n) = (tP_n, P_n)^* = B_n^*.$$

Además, A_n es invertible, pues por ser el grado de $R_{n+1}(t) - tR_n(t)$ menor o igual que n se tiene que $(R_{n+1} - tR_n, R_{n+1}) = \theta$, de lo cual se deduce que $(tR_n, R_{n+1}) = (R_{n+1}, R_{n+1})$ y por tanto

$$\begin{aligned} A_{n+1} = (tP_n, P_{n+1}) &= (R_n, R_n)^{-\frac{1}{2}}(tR_n, R_{n+1})(R_{n+1}, R_{n+1})^{-\frac{1}{2}} \\ &= (R_n, R_n)^{-\frac{1}{2}}(R_{n+1}, R_{n+1})^{\frac{1}{2}} \end{aligned}$$

que es invertible. La descomposición polar de una matriz permite tomar A_n definida positiva.

Esta relación de recurrencia también caracteriza la ortonormalidad con respecto a una matriz de medidas, es decir, si una sucesión de polinomios matriciales $(P_n)_n$ verifica una relación de recurrencia de la forma (1.2), entonces existe una matriz de medidas μ definida positiva respecto de la cual los polinomios $(P_n)_n$ son ortonormales (véase [**AN**] o [**D2**]).

Dada $(P_n)_n$ una sucesión de polinomios matriciales ortonormales con respecto a μ, se definen los correspondientes polinomios de segunda clase $Q_n(t)$ como

$$Q_n(t) = \int_{\mathbb{R}} \frac{P_n(t) - P_n(x)}{t - x} d\mu(x), \quad n \geq 0.$$

Los polinomios de segunda clase $Q_n(t)$ también verifican la relación de recurrencia (1.2), con las condiciones iniciales $Q_0(t) = \theta$, $Q_1(t) = A_1^{-1}$.

Muy recientemente ha sido establecida una relación muy estrecha entre polinomios matriciales ortonormales (o polinomios matriciales verificando una relación de recurrencia de tres términos) y polinomios escalares verificando una relación de recurrencia de más términos (ver [**D1**], [**D2**] y [**DV**]). Así, en [**DV**, §2] se prueba el siguiente teorema:

Teorema A. *Supongamos que $p_n(t)$ $(n = 0, 1, 2, \cdots)$ es una sucesión de polinomios escalares que satisfacen la siguiente relación de recurrencia de $2N + 1$ términos*

$$t^N p_n(t) = c_{n,0} p_n(t) + \sum_{k=1}^{N} \left(\overline{c_{n,k}} p_{n-k}(t) + c_{n+k,k} p_{n+k}(t) \right), \tag{1.3}$$

donde $c_{n,0}$ $(n = 0, 1, 2, \cdots)$ es una sucesión real y $c_{n,k}$ $(n = 0, 1, 2, \cdots)$ son sucesiones complejas para $k = 1, \cdots, N$, con $c_{n,N} \neq 0$ para todo n y con las condiciones iniciales $p_k(x) = 0$ para $k < 0$ y p_k polinomios de *grado k dados, para $k = 0, \cdots, N-1$. Definimos la sucesión de polinomios matriciales $(P_n)_n$ como*

$$(1.4) \qquad P_n(t) = \begin{pmatrix} R_{N,0}(p_{nN})(t) & \dots & R_{N,N-1}(p_{nN})(t) \\ R_{N,0}(p_{nN+1})(t) & \dots & R_{N,N-1}(p_{nN+1})(t) \\ \vdots & \ddots & \vdots \\ R_{N,0}(p_{nN+N-1})(t) & \dots & R_{N,N-1}(p_{nN+N-1})(t) \end{pmatrix},$$

donde el operador $R_{N,m}$ ($m = 0, \cdots, N-1$) está definido por

$$R_{N,m}(p)(t) = \sum_n \frac{p^{(nN+m)}(0)}{(nN+m)!} t^n,$$

de modo que

$$p(t) = R_{N,0}(p)(t^N) + tR_{N,1}(t^N) + \cdots + t^{N-1} R_{N,N-1}(t^N).$$

Entonces la sucesión de polinomios matriciales definida por (1.4) es ortonormal en la recta real con respecto a una matriz de medidas definida positiva y satisface una relación de recurrencia de tres términos.

Recíprocamente, supongamos que $P_n = (P_{n,m,j})_{m,j=0}^{N-1}$ es una sucesión de polinomios ortonormales matriciales o equivalentemente que satisfacen una relación de recurrencia matricial de tres términos (sin pérdida de generalidad es posible suponer que el coeficiente líder de P_n es una matriz triangular inferior), entonces los polinomios escalares definidos por

$$p_{nN+m}(t) = \sum_{j=0}^{N-1} t^j P_{n,m,j}(t^N), \qquad (n \in \mathbb{N}, 0 \le m \le N-1),$$

satisfacen una relación de recurrencia de $2N+1$ términos de la forma (1.3).

Dada una sucesión de polinomios matriciales $(P_n)_n$ verificando la relación de recurrencia de tres términos (1.2):

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t)$$

con $P_0(t) = I$ y $P_{-1}(t) = \theta$, es posible suponer (véase la prueba del teorema anterior en [**DV**]) que las matrices A_n son matrices triangulares inferiores con $\det A_n \neq 0$ y $B_n^* = B_n$. Se tiene que la expresión de las matrices A_n y B_n, en términos de los coeficientes de recurrencia que aparecen en (1.3) es

$$A_n = \begin{pmatrix} c_{nN,N} & 0 & 0 & \dots & 0 \\ c_{nN,N-1} & c_{nN+1,N} & 0 & \dots & 0 \\ c_{nN,N-2} & c_{nN+1,N-1} & c_{nN+2,N} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{nN,1} & c_{nN+1,2} & c_{nN+2,3} & \dots & c_{nN+N-1,N} \end{pmatrix},$$

y

$$B_n = \begin{pmatrix} c_{nN,0} & c_{nN+1,1} & c_{nN+2,2} & \dots & c_{nN+N-1,N-1} \\ \overline{c_{nN+1,1}} & c_{nN+1,0} & c_{nN+2,1} & \dots & c_{nN+N-1,N-2} \\ \overline{c_{nN+2,2}} & \overline{c_{nN+2,1}} & c_{nN+2,0} & \dots & c_{nN+N-1,N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{c_{nN+N-1,N-1}} & \overline{c_{nN+N-1,N-2}} & \overline{c_{nN+N-1,N-3}} & \dots & c_{nN+N-1,0} \end{pmatrix}.$$

En [**D3**, §3] se prueba que para una familia de polinomios matriciales ortonormales $(P_n)_n$ verificando la fórmula de recurrencia de tres términos

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t),$$

si A es una matriz de tamaño $N \times N$ de modo que $A_nA = A^*A_n^*$, entonces la descomposición en fracciones simples

$$(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}R(\lambda) = \sum_{k=1}^{m} \frac{G_{n,k}}{\lambda - x_{n,k}} \tag{1.5}$$

es siempre posible si R es un polinomio matricial de grado menor o igual que $n-1$, independientemente de la multiplicidad de los ceros $x_{n,k}$ de $\det(P_n(\lambda) - AP_{n-1}(\lambda)) = 0$, y siendo $G_{n,k}$ determinadas matrices numéricas cuyo valor se da explícitamente:

$$G_{n,k} = \frac{1}{(\det(P_n(t) - AP_{n-1}(t)))^{(l_k)}(x_{n,k})} (\mathrm{Adj}(P_n(t) - AP_{n-1}(t)))^{(l_k-1)}(x_{n,k})R(x_{n,k}), \quad k = 1, \dots, m,$$

siendo l_k el orden de $x_{n,k}$ como cero del polinomio $P_n(\lambda) - AP_{n-1}(\lambda)$. En el caso particular $R(\lambda) = Q_n(\lambda) - AQ_{n-1}(\lambda)$, las matrices $G_{n,k}$ $(k = 1, \ldots, m)$ son semidefinidas positivas y el rango de $G_{n,k}$ coincide con la multiplicidad del cero $x_{n,k}$.

Además, si se define la matriz de medidas μ_n como

$$\mu_n = \sum_{k=1}^{m} G_{n,k}\delta_{x_{n,k}},$$

entonces μ_n verifica

$$\int_{\mathbb{R}} t^k d\mu_n = \int_{\mathbb{R}} t^k d\mu, \text{ para } k = 0, \ldots, 2n-2,$$

y

$$\int_{\mathbb{R}} \frac{d\mu_n(t)}{t-\lambda} = -(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)). \tag{1.6}$$

Para demostrar el teorema principal del capítulo 6 se hará uso de un resultado de Kochubei que aparece en [**K**]. Dado que este artículo está escrito en ucraniano, incluimos la demostración traducida para hacer la memoria más completa. Aprovechamos la ocasión para agradecer al profesor Alexander Aptekarev su colaboración para traducir este resultado.

Dada una matriz de medidas definida positiva μ con todos sus momentos finitos se denota por V_n el conjunto de matrices de medidas definidas positivas cuyos momentos hasta grado n coinciden con los de μ.

Lema K.

Si μ_0 es una matriz de medidas extremal en el conjunto V_{n-1}, entonces μ_0 tiene a lo sumo nN^2 puntos en el soporte.

Demostración

Supongamos por el contrario que μ_0 tuviera un número mayor o igual que nN^2+1 de puntos en el soporte. Entonces es posible escoger nN^2+1 intervalos disjuntos $\Delta_0, \ldots, \Delta_{nN^2}$ de modo que $\mu_0(\Delta_i) \neq \theta$, para $0 \leq i \leq nN^2$. Llamemos

$$\Delta_{nN^2+1} = \mathbb{R} \setminus \bigcup_{i=1}^{nN^2} \Delta_i.$$

Puesto que $\Delta_0 \subseteq \Delta_{nN^2+1}$ se tiene que $\mu_0(\Delta_{nN^2+1}) \neq \theta$.

Dado un vector $(a_1, \ldots, a_{nN^2+1})$ en $\mathbb{R}^{nN^2+1}$ con $a_i \geq 0$ para $1 \leq i \leq nN^2+1$, es posible definir la matriz de medidas μ como

$$\mu(A) = \sum_{i=1}^{nN^2+1} a_i \mu_0(A \cap \Delta_i).$$

Llamemos W_{n-1} al conjunto de vectores $(a_1, \ldots, a_{nN^2+1})$ de $\mathbb{R}^{nN^2+1}$ tales que la matriz de medidas correspondiente μ pertenece a V_{n-1}.

Se tiene entonces que la matriz de medidas asociada al vector $(1, \ldots, 1)$ es μ_0 y por ser esta matriz de medidas extremal en V_{n-1} se tiene que $(1, \ldots, 1)$ es extremal en W_{n-1}.

El conjunto W_{n-1} está caracterizado por las ecuaciones

$$(1.7) \qquad \begin{cases} a_i \geq 0, \text{ para } 1 \leq i \leq nN^2+1 \\ \displaystyle\sum_{i=1}^{nN^2+1} a_i \int_{\Delta_i} t^k d\mu_0(t) = S_k, \text{ para } 0 \leq k \leq n-1, \end{cases}$$

que determinan un sistema de nN^2 ecuaciones en las nN^2+1 incógnitas a_i. Es posible por tanto escoger una solución $(h_1, \ldots, h_{nN^2+1})$ del sistema homogéneo (esto es, sustituyendo S_k por θ) de modo que $|h_i| < 1$ para $1 \leq i \leq nN^2+1$.

Definiendo $A_+ = (1+h_1, \ldots, 1+h_{nN^2+1})$ y $A_- = (1-h_1, \ldots, 1-h_{nN^2+1})$, es claro que A_+ y A_- verifican (1.7), y además $A = \frac{1}{2}(A_+ + A_-)$, por tanto A no es extremal en W_{n-1}, lo cual es una contradicción.

$\square$

En el capítulo 4 se hace uso de los polinomios de Chebyshev de primera y segunda clase. Los polinomios de Chebyshev de primera clase están definidos por

$$T_k(t) = \cos(k\mathrm{arcos}(t)), \text{ para } k \geq 0,$$

y son ortonormales en $[-1, 1]$ con respecto a la medida

$$\sqrt{\frac{2}{\pi}}(1 - t^2)^{-\frac{1}{2}} dt.$$

Los polinomios de Chebyshev de segunda clase están definidos por

$$U_k(t) = \frac{\mathrm{sen}((n+1)\mathrm{arcos}(t))}{\mathrm{arcos}(t)}, \text{ para } k \geq 0,$$

y son ortonormales en $[-1, 1]$ con respecto a la medida

$$\sqrt{\frac{2}{\pi}}(1 - t^2)^{\frac{1}{2}} dt.$$

Para más detalles véase [**Ri**].

CAPÍTULO 2

TOPOLOGÍAS EN LOS ESPACIOS DE MATRICES DE MEDIDAS

En este capítulo estudiamos algunas propiedades topológicas de los espacios de matrices de medidas que se usarán a lo largo de la memoria. Como conclusión se probará que el conjunto de soluciones de un problema de momentos matricial es, al igual que en el caso unidimensional, un conjunto compacto y convexo. Este último resultado garantizará la existencia de puntos extremales del conjunto de soluciones, lo que será de cierta utilidad cuando se estudie la densidad del espacio de polinomios en el espacio $\mathcal{L}^1$ de una matriz de medidas.

Denotamos por $\mathcal{M}_N$ el conjunto de matrices de medidas definidas positivas (ver preliminares, pag.1), y por $\mathcal{M}_N^*$ el conjunto de matrices de medidas definidas positivas que tienen momentos de cualquier orden y que son no degeneradas (ver preliminares, pag.3).

Dada $\mu \in \mathcal{M}_N^*$, consideramos el conjunto $[\mu]$ dado por

$$[\mu] = \Big\{ \nu \in \mathcal{M}_N^* \text{ tales que } \int_{\mathbb{R}} t^n \nu_{i,j} = \int_{\mathbb{R}} t^n d\mu_{i,j} \,, \quad \text{para todo } n \in \mathbb{N} \text{ y para todo } 1 \leq i, j, \leq N \Big\}$$

es decir el conjunto de matrices de medidas definidas positivas que tienen los mismos momentos que μ, o lo que es lo mismo, el conjunto de soluciones del problema de momentos definido por μ.

La topología vaga sobre $\mathcal{M}_N$ se define como la más fina para la cual todas las aplicaciones $\mu \to \int_{\mathbb{R}} f d\mu$ son continuas, donde $f \in \mathcal{C}_c(\mathbb{R})$ es arbitraria. $C_c(\mathbb{R})$ denota el conjunto de las funciones continuas de soporte compacto en $\mathbb{R}$.

La topología débil sobre $\mathcal{M}_N$ se define como la más fina para la cual todas las aplicaciones $\mu \to \int_{\mathbb{R}} f d\mu$ son continuas, donde $f \in C_b(\mathbb{R})$ es arbitraria. $C_b(\mathbb{R})$ denota el conjunto de las funciones continuas y acotadas en $\mathbb{R}$.

Puesto que $C_c(\mathbb{R})$ está incluido estrictamente en $C_b(\mathbb{R})$ es claro que la topología vaga es más fina, es decir tiene menos abiertos, que la topología débil.

Si $A \subseteq C_c(\mathbb{R})$ es un conjunto arbitrario que genera $C_c(\mathbb{R})$ entonces la topología vaga es también la más fina para la cual todas las aplicaciones $\mu \to \int_{\mathbb{R}} f d\mu$ son continuas, donde f recorre todo A. Esta observación es usada para $A = C_c^+(\mathbb{R})$.

Para $\mu_0 = (\mu_{0,i,j})_{1\le i,j\le N}$ los conjuntos

$$V_{f_1,\ldots,f_n;\epsilon}(\mu_0) = \Big\{\mu \in \mathcal{M}_N^* : \left| \int_{\mathbb{R}} f_p d\mu_{i,j} - \int_{\mathbb{R}} f_p d\mu_{0,i,j} \right| < \epsilon, \\ p = 1,\ldots,n,\ 1 \le i,j \le N\}$$

donde $f_1,\ldots,f_n \in C_c^+(\mathbb{R})$, $\epsilon > 0$ forman una base de entornos de μ_0 para la topología vaga. Análogamente se construye una base de entornos de μ_0 para la topología débil, tomando $f_0,\ldots,f_n$ en $C_b^+(\mathbb{R})$.

La topología vaga es Hausdorff porque si $\mu \neq \nu$, existe $f \in C_c^+(\mathbb{R})$ de modo que $\int_{\mathbb{R}} f d\mu \neq \int_{\mathbb{R}} f d\nu$. (Esto se tiene a partir de la unicidad en el teorema de representación de Riesz). Si se considera entonces ϵ tal que

$$\epsilon - \max_{1\le i,j\le N} \left| \int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\nu_{i,j} \right| > 0,$$

entonces $V_{f;\frac{\epsilon}{2}}(\mu) \cap V_{f;\frac{\epsilon}{2}}(\nu) = \emptyset$.

Como primer resultado establecemos la relación existente entre la convergencia débil y la vaga.

Teorema 2.1.

Dadas $(\mu_\alpha)_{\alpha\in A}$ *y* μ *en* $\mathcal{M}_N$ *finitas (es decir* $\tau\mu_\alpha(\mathbb{R}) < \infty$ *y* $\tau\mu(\mathbb{R}) < \infty$*) se tiene que* $\mu_\alpha \to \mu$ *débilmente si y sólo si* $\mu_\alpha \to \mu$ *vagamente y* $\lim_\alpha \tau\mu_\alpha(\mathbb{R}) = \tau\mu(\mathbb{R})$.

La demostración del teorema 2.1 sigue de los siguientes lemas.

Lema 2.2.

Sean $(\mu_\alpha)_{\alpha\in A}$, μ *en* $\mathcal{M}_N$, $\mu_\alpha \to \mu$ *vagamente y* $g : \mathbb{R} \to [0,\infty)$ *una función continua. Entonces*

a) Si $\limsup_{\alpha\in A} \int_{\mathbb{R}} gd\tau\mu_\alpha < \infty$ *entonces*

$$\int_{\mathbb{R}} gd\tau\mu < \infty \text{ y } \lim_\alpha \int_{\mathbb{R}} fd\mu_{\alpha,i,j} = \int_{\mathbb{R}} fd\mu_{i,j},$$

para cualquier $f \in o(g)$ *y para todo* $1 \leq i,j \leq N$.

b) Si $\lim_{\alpha\in A} \int_{\mathbb{R}} gd\tau\mu_\alpha = \int_{\mathbb{R}} gd\tau\mu < \infty$ *entonces*

$$\lim_\alpha \int_{\mathbb{R}} fd\mu_{\alpha,i,j} = \int_{\mathbb{R}} fd\mu_{i,j},$$

para cualquier $f \in O(g)$ *y para todo* $1 \leq i,j \leq N$.

Demostración

a) Escogemos $\phi_n \in C_c(\mathbb{R})$ verificando $\chi_{[-n,n]} \leq \phi_n \leq 1$. Tenemos entonces:

$$\begin{aligned}
\int_{\mathbb{R}} gd\tau\mu &= \lim_{n\to\infty} \int_{\mathbb{R}} g\phi_n d\tau\mu \\
&= \lim_{n\to\infty} \lim_{\alpha\in A} \int_{\mathbb{R}} g\phi_n d\tau\mu_\alpha \\
&\leq \limsup_\alpha \int_{\mathbb{R}} gd\tau\mu_\alpha < \infty.
\end{aligned}$$

Supongamos ahora que $f \in o(g)$ y $\epsilon > 0$ son dados. Existe un conjunto compacto K tal que $|f(x)| \leq \epsilon g(x)$, para cualquier $x \in K^C$. Sea

$\phi \in C_c^+(\mathbb{R})$ verificando $\chi_K \leq \phi \leq 1$. Tenemos entonces $|f|(1-\phi) \leq \epsilon g$ y por tanto:

$$\left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right|$$

$$\leq \left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f\phi d\mu_{i,j}\right| + \left|\int_{\mathbb{R}} f\phi d\mu_{i,j} - \int_{\mathbb{R}} f\phi d\mu_{\alpha,i,j}\right| + \left|\int_{\mathbb{R}} f\phi d\mu_{\alpha,i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right|$$

$$\leq \epsilon \int_{\mathbb{R}} g d|\mu_{i,j}| + \left|\int_{\mathbb{R}} f\phi d(\mu_{i,j} - \mu_{\alpha,i,j})\right| + \epsilon \int_{\mathbb{R}} g d|\mu_{\alpha,i,j}|$$
$$\leq \epsilon \int_{\mathbb{R}} g d\tau\mu + \left|\int_{\mathbb{R}} f\phi d(\mu_{i,j} - \mu_{\alpha,i,j})\right| + \epsilon \int_{\mathbb{R}} g d\tau\mu_\alpha.$$

Tomando ahora $\limsup_\alpha$ obtenemos

$$\limsup_\alpha \left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right| \leq \epsilon(c_1 + c_2)$$

y puesto que ϵ era arbitrario deducimos que

$$\limsup_\alpha \left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right| = 0$$

y por tanto

$$\lim_\alpha \int_{\mathbb{R}} f d\mu_{\alpha,i,j} = \int_{\mathbb{R}} f d\mu_{i,j}.$$

b) Supongamos ahora que $\lim_{\alpha \in A} \int_{\mathbb{R}} g d\tau\mu_\alpha = \int_{\mathbb{R}} g d\tau\mu < \infty$. Sea $f \in O(g)$ y $\epsilon > 0$ dado. Existe una constante C y un conjunto compacto K_0 de modo que $|f(x)| \leq Cg(x)$, para cualquier $x \in K_0^C$. Podemos elegir K un conjunto compacto, con $K_0 \subseteq K$ y siendo

$$\int_{\mathbb{R}} g d\tau\mu - \epsilon < \int_K g d\tau\mu$$

Si elegimos ahora $\phi \in C_c^+(\mathbb{R})$ tal que $\chi_K \leq \phi \leq 1$, entonces $\int_{\mathbb{R}} g(1-\phi)d\tau\mu < \epsilon$ y $|f|(1-\phi) \leq Cg(1-\phi)$.

Como antes, tenemos

$$\left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right|$$

$$\leq \left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f\phi d\mu_{i,j}\right| + \left|\int_{\mathbb{R}} f\phi d\mu_{i,j} - \int_{\mathbb{R}} f\phi d\mu_{\alpha,i,j}\right| + \left|\int_{\mathbb{R}} f\phi d\mu_{\alpha,i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right|$$

$$\begin{aligned}
&\leq \int_{\mathbb{R}} |f|(1-\phi)d|\mu_{i,j}| + \left|\int_{\mathbb{R}} f\phi d(\mu_{i,j} - \mu_{\alpha,i,j})\right| + C\int_{\mathbb{R}} g(1-\phi)d|\mu_{\alpha,i,j}| \\
&\leq \int_{\mathbb{R}} |f|(1-\phi)d\tau\mu + \left|\int_{\mathbb{R}} f\phi d(\mu_{i,j} - \mu_{\alpha,i,j})\right| + C\int_{\mathbb{R}} g(1-\phi)d\tau\mu_\alpha \\
&\leq C\epsilon + \left|\int_{\mathbb{R}} f\phi d(\mu_{i,j} - \mu_{\alpha,i,j})\right| + C\int_{\mathbb{R}} g(1-\phi)d\tau\mu_\alpha.
\end{aligned}$$

Tomando $\limsup_\alpha$ obtenemos

$$\limsup_\alpha \left|\int_{\mathbb{R}} f d\mu_{i,j} - \int_{\mathbb{R}} f d\mu_{\alpha,i,j}\right| \leq 2C\epsilon$$

porque

$$\lim_\alpha \int_{\mathbb{R}} g(1-\phi)d\tau\mu_\alpha = \int_{\mathbb{R}} g(1-\phi)d\tau\mu < \epsilon$$

por tanto el lema 2.2 está probado.

$\square$

El siguiente lema es consecuencia del anterior:

Lema 2.3.

Supongamos que $\mu_\alpha \to \mu$ *vagamente. Se tiene que*

a) Si $\limsup_\alpha \tau\mu_\alpha(\mathbb{R}) < \infty$ *entonces* $\tau\mu(\mathbb{R}) < \infty$ *y*

$$\lim_\alpha \int_{\mathbb{R}} f d\mu_{\alpha,i,j} = \int_{\mathbb{R}} f d\mu_{i,j},$$

para cualquier $f \in o(1)$ *y para todo* $1 \leq i,j \leq N$.

b) Si $\lim_\alpha \tau\mu_\alpha(\mathbb{R}) = \tau\mu(\mathbb{R})$ *entonces*

$$\lim_\alpha \int_{\mathbb{R}} f d\mu_{\alpha,i,j} = \int_{\mathbb{R}} f d\mu_{i,j},$$

para cualquier $f \in O(1)$ *y para todo* $1 \leq i,j \leq N$.

Demostración del teorema 2.1

Es claro que si $\mu_\alpha \to \mu$ débilmente entonces $\mu_\alpha \to \mu$ vagamente, y además, puesto que $\tau\mu_\alpha \to \tau\mu$ débilmente, tomando $f = 1$ en la definición de convergencia débil se tiene que $\tau\mu_\alpha(\mathbb{R}) \to \tau\mu(\mathbb{R})$.

Por otra parte, si $\mu_\alpha \to \mu$ vagamente y $\tau\mu_\alpha(\mathbb{R}) \to \tau\mu(\mathbb{R})$, de (b) del lema 2.3 y puesto que toda función de $C_b(\mathbb{R})$ es acotada y por tanto de $O(1)$, se deduce que $\mu_\alpha \to \mu$ débilmente.

$\square$

Es de interés saber si la topología vaga en $\mathcal{M}_N$ es metrizable, pues en tal caso las nociones topológicas pueden ser expresadas en términos de sucesiones en lugar de redes. Se tiene el siguiente teorema

Teorema 2.4. $\mathcal{M}_N$ *es metrizable para la topología vaga.*

La demostración no difiere mucho de la dada en [**B**] para el caso unidimensional. En los problemas 1.1 de la sección 7.7 y 2 de la sección 7.8 de [**B**] se prueba que dada una sucesión de funciones $\mathcal{D} = (\eta_n)_n$ en $C_c(\mathbb{R})$ que verifican:

(1) Para cada conjunto compacto $K \subseteq \mathbb{R}$ existe un entorno U relativamente compacto tal que cada función $f \in C_c(\mathbb{R})$ con

supp$f$ $\subseteq K$ es uniformemente aproximable en $\mathbb{R}$ por funciones de $\mathcal{D}$ cuyo soporte está en U.

(2) Para ese K y U, existe una η_n con $0 \leq \chi_K \leq \eta_n \leq 1$.

se tiene que una sucesión μ_n de medidas converge vagamente a μ si y sólo si

$$\lim_{n\to\infty} \int_{\mathbb{R}} \eta d\mu_n = \int_{\mathbb{R}} \eta d\mu, \quad \text{para cada} \quad \eta \in \mathcal{D}.$$

Además, si se define

$$\rho(\mu,\nu) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\left\{1, \left|\int_{\mathbb{R}} \eta_n d\mu - \int_{\mathbb{R}} \eta_n d\nu\right|\right\},$$

ρ es una métrica en $\mathcal{M}_1$ y la correspondiente topología es la topología vaga. De forma análoga se prueba que $\mathcal{M}_N$ es metrizable, definiendo para dos matrices de medidas $\mu = (\mu_{i,j})_{1\leq i,j\leq N}$, $\nu = (\nu_{i,j})_{1\leq i,j\leq N}$

$$\rho(\mu,\nu) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\left\{1, \max_{1\leq i,j\leq N} \left|\int_{\mathbb{R}} \eta_n d\mu_{i,j} - \int_{\mathbb{R}} \eta_n d\nu_{i,j}\right|\right\}.$$

ρ es una métrica en $\mathcal{M}_N$ y la topología correspondiente es la vaga.

A continuación probamos que el caracter definido positivo de las matrices de medidas se conserva por límite débil y vago:

Teorema 2.5.

Si $(\mu_n)_n$ son matrices de medidas en $\mathcal{M}_N$ finitas y $\mu = (\mu_{i,j})_{1\leq i,j\leq N}$ es una matriz de medidas finita, y se tiene que $\mu_n \to \mu$ vagamente, entonces μ es definida positiva. Es decir el caracter definido positivo se mantiene por límite vago.

Para la demostración usamos el siguiente lema.

Lema 2.6.

Dada $\mu = (\mu_{i,j})_{1\leq i,j\leq N}$ una matriz de medidas finita, las siguientes dos propiedades son equivalentes:

a) Para cada conjunto de Borel A la matriz numérica $(\mu_{i,j}(A))_{1\le i,j\le N}$ es semidefinida positiva.

b) Para cualquier $f \in C_c^+(\mathbb{R})$, la matriz numérica

$$\left(\int_{\mathbb{R}} f d\mu_{i,j}\right)_{1\le i,j\le N}$$

es semidefinida positiva.

Demostración

b)$\to$a) Dado un intervalo $(a,b) \subseteq \mathbb{R}$, consideramos f_n una sucesión creciente de funciones de $C_c^+(\mathbb{R})$ tales que $f_n \to \chi_{(a,b)}$ puntualmente. Consideramos además la descomposición de Hahn de la matriz de medidas μ: $\mu = \mu^1 - \mu^2 + i(\mu^3 - \mu^4)$, donde todas las medidas que aparecen en las matrices de medidas μ^i son medidas positivas. Puesto que todas estas medidas están acotadas por la medida $\tau\mu$, el teorema de la convergencia monótona da entonces que para $1 \le i,j \le N$,

$$\mu_{i,j}((a,b)) = \int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu_{i,j}(t)$$

$$= \int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu_{i,j}^1(t) - \int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu_{i,j}^2(t) + i\left(\int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu_{i,j}^3(t) - \int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu_{i,j}^4(t)\right)$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu_{i,j}^1(t) - \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu_{i,j}^2(t) + i\left(\lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu_{i,j}^3(t) - \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu_{i,j}^4(t)\right)$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d(\mu_{i,j}^1 - \mu_{i,j}^2 + i(\mu_{i,j}^3 - \mu_{i,j}^4))(t) = \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu_{i,j}(t)$$

y por tanto

$$\mu((a,b)) = \int_{\mathbb{R}} \chi_{(a,b)}(t) d\mu(t) = \lim_{n\to\infty} \int_{\mathbb{R}} f_n(t) d\mu(t).$$

Por ser todas las matrices numéricas $\int_{\mathbb{R}} f_n(t)d\mu(t)$ semidefinidas positivas se deduce que $\mu((a,b))$ también lo es.

Usando ahora la regularidad de μ se deduce que $\mu(A)$ es semidefinida positiva, para cualquier conjunto de Borel A.

Supongamos ahora a) y probemos b). Supongamos por el contrario que existe $f \in C_c^+(\mathbb{R})$ tal que la matriz numérica $\left(\int_{\mathbb{R}} f d\mu_{i,j}\right)_{1\leq i,j\leq N}$ no es semidefinida positiva, es decir existe un vector no nulo $c = (c_1, \ldots, c_N)$ tal que $c\left(\int_{\mathbb{R}} f d\mu_{i,j}\right) c^* \not\geq 0$. Llamemos

$$d = \operatorname{dist}\left\{c\left(\int_{\mathbb{R}} f d\mu_{i,j}\right) c^*, \{x \in \mathbb{R}, x \geq 0\}\right\} > 0.$$

Es claro que

$$d \geq \left|\Im c\left(\int_{\mathbb{R}} f d\mu_{i,j}\right) c^*\right| \tag{2.1}$$

Puesto que $f \in C_c^+(\mathbb{R})$, podemos encontrar una función simple $f_0 = \sum_{k=1}^{k_0} a_i \chi_{A_k}$, con $a_k \geq 0$, que aproxima a f de la forma siguiente

$$\|f - f_0\|_\infty \leq \frac{d}{2N^2 \max\limits_{1\leq i\leq N} |c_i|^2 \max\limits_{1\leq i,j\leq N} \|\mu_{i,j}\|}.$$

De la igualdad

$$c\left(\int_{\mathbb{R}} f d\mu_{i,j}\right) c^* = c\left(\int_{\mathbb{R}} (f - f_0) d\mu_{i,j}\right) c^* + c\left(\int_{\mathbb{R}} f_0 d\mu_{i,j}\right) c^*,$$

teniendo en cuenta que

$$\begin{aligned}\left|c\left(\int_{\mathbb{R}} (f - f_0) d\mu_{i,j}\right) c^*\right| &\leq 2N^2 \max_{1\leq i\leq N} |c_i|^2 \max_{1\leq i,j\leq N} \|\mu_{i,j}\| \|f - f_0\|_\infty \\ &\leq \frac{d}{2}\end{aligned}$$

y que

$$c\left(\int_{\mathbb{R}} f_0 d\mu_{i,j}\right)c^* = c\left(\sum_{k=1}^{k_0} a_k\mu_{i,j}(A_k)\right)c^* = \sum_{k=1}^{k_0} a_k c\mu_{i,j}(A_k)c^* \geq 0$$

se deduce que

$$\begin{aligned}\left|\Im c\left(\int_{\mathbb{R}} f d\mu_{i,j}\right)c^*\right| &= \left|\Im c\left(\int_{\mathbb{R}}(f-f_0)d\mu_{i,j}\right)c^*\right| \\ &\leq \left|c\left(\int_{\mathbb{R}}(f-f_0)d\mu_{i,j}\right)c^*\right| \leq \frac{d}{2}\end{aligned}$$

lo cual es una contradicción con (2.1).

□

Demostración del teorema 2.5

La demostración es similar a la del lema 2.6. Supongamos que $\mu_n \to \mu_0$ vagamente, siendo $\mu_n \in \mathcal{M}_N$. Tenemos que probar que la matriz de medidas μ_0 es definida positiva. Supongamos por el contrario que μ_0 no es definida positiva, entonces por el lema 2.5 existe $f \in C_c^+(\mathbb{R})$ tal que la matriz numérica $\left(\int_{\mathbb{R}} f d\mu_0\right)$ no es semidefinida positiva. Como hicimos antes, llamemos

$$d = \operatorname{dist}\left\{c\left(\int_{\mathbb{R}} f d\mu_0\right)c^*, \{x \in \mathbb{R}, x \geq 0\}\right\} > 0.$$

Es claro que

$$d \geq \left|\Im c\left(\int_{\mathbb{R}} f d\mu_0\right)c^*\right|. \tag{2.2}$$

Para cualquier número natural n tenemos

$$c\left(\int_{\mathbb{R}} f d\mu_0\right)c^* = c\left(\int_{\mathbb{R}} f d\mu_0 - \int_{\mathbb{R}} f d\mu_n\right)c^* + c\left(\int_{\mathbb{R}} f d\mu_n\right)c^* \tag{2.3}$$

y

$$\left| c\left(\int_{\mathbb{R}} f d\mu_0 - \int_{\mathbb{R}} f d\mu_n\right) c^* \right|$$
$$\leq \max_{1\leq i,j\leq N} \left| \int_{\mathbb{R}} f d\mu_{0,i,j} - \int_{\mathbb{R}} f d\mu_{n,i,j} \right| N^2 \max_{1\leq i\leq N} |c_i|^2.$$

Como $\mu_n \to \mu_0$ vagamente y $f \in C_c(\mathbb{R})$, existe un número natural n_0 tal que para cualquier $n \geq n_0$, tenemos

$$\max_{1\leq i,j\leq N} \left| \int_{\mathbb{R}} f d\mu_{0,i,j} - \int_{\mathbb{R}} f d\mu_{n,i,j} \right| \leq \frac{d}{2N^2 \max\limits_{1\leq i\leq N} |c_i|^2}$$

y por tanto

$$\left| c\left(\int_{\mathbb{R}} f d\mu_{0,i,j} - \int_{\mathbb{R}} f d\mu_{n,i,j}\right) c^* \right| \leq \frac{d}{2} \tag{2.4}$$

Ahora, de las igualdades (2.3) y (2.4), y puesto que

$$c\left(\int_{\mathbb{R}} f d\mu_n\right) c^* \geq 0$$

se deduce que para todo $n \geq n_0$

$$\left| \Im c\left(\int_{\mathbb{R}} f d\mu_0\right) c^* \right| = \left| \Im c\left(\int_{\mathbb{R}} f d\mu_{0,i,j} - \int_{\mathbb{R}} f d\mu_{n,i,j}\right) c^* \right|$$
$$\leq \left| c\left(\int_{\mathbb{R}} f d\mu_{0,i,j} - \int_{\mathbb{R}} f d\mu_{n,i,j}\right) c^* \right| \leq \frac{d}{2}$$

lo cual es una contradicción con (2.2).

□

Para finalizar el capítulo probamos que el conjunto de soluciones de un problema de momentos matricial es débil y vagamente compacto y convexo.

Teorema 2.7. *$[\mu]$ es un conjunto compacto y convexo en la topología débil y vaga que coinciden en $[\mu]$.*

Este resultado es consecuencia inmediata de los siguientes tres lemas.

Lema 2.8.

El conjunto $\{\mu \in \mathcal{M}_N : \tau\mu(\mathbb{R}) \leq 1\}$ es vagamente compacto.

Demostración

Dada $(\mu_n)_{n\in\mathrm{N}}$ en $\{\mu \in \mathcal{M}_N : \tau\mu(\mathbb{R}) \leq 1\}$, es posible descomponer $\mu_n = \mu_n^1 - \mu_n^2 + i(\mu_n^3 - \mu_n^4)$, siendo μ_n^i matrices de medidas en las que todas las componentes son medidas positivas. Puesto que $|\mu_{n,i,j}| \leq \tau\mu_n$ para todo n se deduce que todas estas medidas están acotadas por una constante fija, y puesto que el conjunto de medidas positivas $\{\mu \geq 0, \|\mu\| \leq a\}$ es vagamente compacto (véase por ejemplo [**B**]) podemos encontrar $\mu_0 = (\mu_{0,i,j})_{1\leq i,j\leq N}$ tal que $\mu_n \to \mu_0$ vagamente. Por el teorema 2.3, μ_0 es una matriz de medidas definida positiva .

Tenemos que probar ahora que $\tau\mu_0(\mathbb{R}) \leq 1$. Pero $\mu_n \to \mu_0$ vagamente implica que $\tau\mu_n \to \tau\mu_0$ vagamente. Consideremos ϕ_p una función en $C_c^+(\mathbb{R})$ tal que $\phi_p(x) = 1$, si $|x| \leq p$ y $\phi_p(x) = 0$, si $|x| \geq p+1$. Tenemos que:

$$
\begin{aligned}
\tau\mu_0(\mathbb{R}) &= \sup_{p\in\mathrm{N}} \int_{\mathbb{R}} \phi_p(x)\,d\tau\mu_0(x) \\
&= \sup_{p\in\mathrm{N}} \left\{ \lim_{\alpha\in A} \int_{\mathbb{R}} \phi_p(x)\,d\tau\mu_\alpha(x) \right\} \\
&\leq \sup_{p\in\mathrm{N}} \left\{ \lim_{\alpha\in A} \int_{\mathbb{R}} d\tau\mu_\alpha(x) \right\} \\
&= \sup_{p\in\mathrm{N}} \left\{ \lim_{\alpha\in A} \tau\mu_\alpha(\mathbb{R}) \right\} \\
&\leq 1.
\end{aligned}
$$

□

Lema 2.9. *Dada μ una matriz de medidas en $\mathcal{M}_N^*$, $[\mu]$ es un conjunto vagamente relativamente compacto.*

Demostración

Para cualquier $\nu \in [\mu]$, puesto que

$$\tau\nu(\mathbb{R}) = \int_{\mathbb{R}} d\tau\nu(x) = \tau\left(\int_{\mathbb{R}} d\nu(x)\right) = \tau\left(\int_{\mathbb{R}} d\mu(x)\right) = \tau\mu(\mathbb{R})$$

$\tau\nu(\mathbb{R})$ tiene el mismo valor positivo $\tau\mu(\mathbb{R})$. Por el lema 2.8, el conjunto $\{\nu \in \mathcal{M}_N : \tau\nu(\mathbb{R}) \leq \tau\mu(\mathbb{R})\}$ es vagamente compacto, y por tanto vagamente cerrado. Tenemos que $[\mu] \subseteq \{\nu \in \mathcal{M}_N : \tau\nu(\mathbb{R}) \leq \tau\mu(\mathbb{R})\}$, así que tomando cierre tenemos $\overline{[\mu]} \subseteq \{\nu \in \mathcal{M}_N : \tau\nu(\mathbb{R}) \leq \tau\mu(\mathbb{R})\}$. Ahora el conjunto $\overline{[\mu]}$ es un conjunto cerrado dentro de un compacto y por tanto compacto.

□

Lema 2.10. *Dada μ una matriz de medidas en $\mathcal{M}_N^*$, el conjunto $[\mu]$ es un conjunto vagamente cerrado.*

Demostración

Si $\mu_n \to \mu$ vagamente, entonces por el teorema 2.4 $\mu \geq 0$, y teniendo en cuenta que todas las matrices de medidas μ_n tienen los mismos momentos, deducimos que $\limsup_n \int_{\mathbb{R}} t^{2n_0} d\tau\mu_n < \infty$ y por tanto para cualquier polinomio $p(t) \geq 0$ tenemos que $\limsup_n \int_{\mathbb{R}} p(t) d\tau\mu_n < \infty$.

Dado un número natural n_0 consideramos el polinomio

$$p(t) = (1 + t^{2n_0})(1 + t^2).$$

Es claro que $t^{n_0} \in o(p(t))$ y por tanto el lema 2.2 da

$$\lim_n \int_{\mathbb{R}} t^{n_0} d\mu_{n,i,j} = \int_{\mathbb{R}} t^{n_0} d\mu_{i,j} \quad \text{para todo} \quad 1 \leq i, j \leq N$$

y por tanto los momentos de μ son los mismos que los de μ_n.

□

Demostración del teorema 2.7

Las topologías vaga y débil coinciden en $[\mu]$ pues si μ_α y μ pertenecen a $[\mu]$, y $\mu_\alpha \to \mu$ vagamente, al ser los momentos de $\tau\mu_\alpha$ igual a los de $\tau\mu$, la condición $\tau\mu(\mathbb{R})_\alpha \to \tau\mu(\mathbb{R})$ se satisface trivialmente, y esto implica que $\mu_\alpha \to \mu$ según el teorema 2.1.

De los lemas 2.9 y 2.10 se sigue que $[\mu]$ es compacto. Por último, es claro que $[\mu]$ es convexo, pues si μ_1 y μ_2 pertenecen a $[\mu]$, todas las matrices de medidas del segmento $t\mu_1 + (1-t)\mu_2$, $t \in [0,1]$ están en $[\mu]$.

□

CAPÍTULO 3

CEROS Y FÓRMULA DE CUADRATURA

En este capítulo estudiamos las propiedades básicas de los ceros de una familia de polinomios ortogonales matriciales y la existencia de fórmulas de cuadratura asociadas. Veremos que guarda una sugestiva analogía con el caso escalar, aunque el carácter matricial introduce importantes perturbaciones. Estas propiedades se deducirán de la fórmula de recurrencia matricial a través de la matriz de Jacobi asociada. Como aplicación, las usamos para derivar una nueva demostración de la versión matricial del teorema de Favard.

Partiremos pues de una sucesión de polinomios matriciales de tamaño $N \times N$ que satisfacen una relación de recurrencia de tres términos de la forma

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t) \tag{3.1}$$

siendo $P_0(t) = I$ y $P_{-1}(t) = \theta$, donde $P_n(t)$ son polinomios matriciales con coeficientes en $\mathbb{C}^{N\times N}$ y los coeficientes de recurrencia A_{n+1}, B_n son también matrices de tamaño $N \times N$, que verifican $\det A_k \neq 0$ y $B_k^* = B_k$. Como consecuencia de la fórmula de recurrencia P_n es un polinomio de grado n con coeficiente lider no singular.

La sucesión de polinomios $(P_n)_n$ induce un producto escalar para el que son ortonormales: en efecto, si P y Q son dos polinomios matriciales de tamaño $N \times N$, podemos escribir

$$P(t) = \sum_l \Gamma_l P_l(t) \quad \text{y} \quad Q(t) = \sum_l \Lambda_l P_l(t)$$

y se define

$$< P, Q >_{(P_n)} = \sum_l \Gamma_l \Lambda_l^*$$

que es un producto escalar para el que la sucesión $(P_n)_n$ es ortonormal.

Recordemos que, por definición, un punto x_0 es un cero del polinomio matricial $P(x)$ si $\det(P(x_0)) = 0$, es decir, x_0 es un cero del polinomio escalar $\det(P(x))$, y la multiplicidad de x_0 como cero de $P(x)$ es la que tenga como cero de $\det(P(x))$.

Consideremos la matriz de Jacobi asociada a los polinomios P_n, esto es, la matriz infinita que se obtiene al situar las matrices $(A_k)_k$, $(B_k)_k$ y $(A_k^*)_k$, que aparecen en la relación de recurrencia (3.1) en las tres diagonales centrales de la matriz J:

$$J = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^* & B_1 & A_2 & & \\ & A_2^* & B_2 & A_3 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{3.2}$$

La N-matriz de Jacobi juega un papel fundamental en el estudio de los ceros y fórmula de cuadratura para los polinomios matriciales $(P_n)_n$, más aún, el teorema principal de esta sección es una consecuencia del siguiente lema.

Lema 3.1. *Para $n \in \mathbb{N}$, los ceros del polinomio matricial $P_n(t)$ son los mismos que los del polinomio $\det(tI_{nN} - J_{nN})$ (con la misma multiplicidad), donde I_{nN} es la matriz identidad de tamaño nN y J_{nN} es la N-matriz de Jacobi truncada al tamaño nN:*

$$J_{nN} = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^* & B_1 & A_2 & & \\ & \ddots & \ddots & \ddots & \\ & & A_{n-2}^* & B_{n-2} & A_{n-1} \\ & & & A_{n-1}^* & B_{n-1} \end{pmatrix}$$

Para probar el lema 3.1, necesitamos el siguiente lema que es interesante por sí mismo. Para una matriz cuadrada dada $A = (a_{i,j})_{1 \leq i,j \leq N}$, denotamos por $\mathrm{Adj}(A)$ la matriz de tamaño $N \times N$ cuya entrada (i,j) es el valor del determinante de la matriz obtenida al eliminar en A la fila i y la columna j multiplicado por $(-1)^{i+j}$. Esta matriz verifica la propiedad

$$A\mathrm{Adj}(A) = \mathrm{Adj}(A)A = \det(A)I.$$

Lema 3.2. *Sea $A(t)$ un polinomio matricial de tamaño $N \times N$ y sea a un cero de $A(t)$. Llamamos*

$$R(a, A) = \{v \in \mathbb{C}^N : A(a)v = \theta\},$$

esto es, los autovectores a la derecha de la matriz $A(a)$ asociados a cero.

Si $dim(R(a, A)) = p$, entonces

$$\Big(Adj\big(A(t)\big)\Big)^{(l)}(a) = \theta, \quad \text{para } l = 0, \dots, p-2,$$

y a es un cero de $A(t)$ de multiplicidad al menos p.

Demostración del lema 3.2:

Consideremos las siguientes notaciones. Llamamos $A_{i,j}(t)$ al polinomio matricial de tamaño $(N-1)\times(N-1)$ obtenido de $A(t)$ quitando la i-ésima fila y la j-ésima columna. Denotamos por $r_{i,j}(t) = \det(A_{i,j}(t))$, es decir, el menor correspondiente a la entrada (i,j) del polinomio matricial $A(t)$. Salvo un signo, $r_{i,j}$ es la entrada (j,i) del polinomio matricial $\mathrm{Adj}(A(t))$. Tenemos que probar que $r_{i,j}^{(l)}(a) = 0$ para $l = 0, \dots, p-2$.

Llamamos $A_{i,j,(m_1,k_1),\dots,(m_n,k_n)}(t)$ al polinomio matricial obtenido derivando k_d ($k_d \geq 1$) veces la columna m_d ($d = 1, \dots, n \leq N$) de $A_{i,j}$. Tenemos entonces la siguiente expresión para $r_{i,j}^{(l)}(a)$:

$$r_{i,j}^{(l)}(a) = \sum_{\sum_{d=1}^{n} k_d = l} \alpha_{(m_1,k_1),\dots,(m_n,k_n)} \det\left(A_{i,j,(m_1,k_1),\dots,(m_n,k_n)}(a)\right),$$

para ciertos enteros no negativos $\alpha_{(m_1,k_1),\ldots,(m_n,k_n)}$. El resultado

$$\Big(\mathrm{Adj}\big(A(t)\big)\Big)^{(l)}(a)=\theta, \quad \text{para } l=0,\ldots,p-2$$

se tiene si probamos que para $0\le l\le p-2$, $\sum_{d=1}^{n}k_d=l$ implica

$$\det\big(A_{i,j,(m_1,k_1),\ldots,(m_n,k_n)}(a)\big)=0.$$

Consideremos el subespacio U de $R(a,A_{i,j})$ definido por

$$u\in U \text{ si y sólo si } u_{m_1}=\cdots=u_{m_n}=0,$$

donde u_m denota la componente m-ésima del vector u. De la definición de $A_{i,j}$, es claro que

$$\dim(R(a,A_{i,j}))\ge \dim(R(a,A))-1.$$

Tenemos entonces que

$$\begin{aligned}\dim(U)&\ge \dim(R(a,A_{i,j}))-n\ge \dim(R(a,A))-n-1=p-n-1\\ &=p-1-\sum_{d=1}^{n}1\ge p-1-\sum_{d=1}^{n}k_d=p-1-l\ge 1.\end{aligned}$$

Por tanto, podemos tomar $u\in U$, $u\ne\theta$.

La matriz $A_{i,j,(m_1,k_1),\ldots,(m_n,k_n)}(a)$ difiere de $A_{i,j}(a)$ sólo en las columnas $m_1,\ldots,m_n$. El vector u tiene estas componentes iguales a 0 y puesto que $u\in R(a,A_{i,j})$ se sigue que

$$A_{i,j,(m_1,k_1),\ldots,(m_n,k_n)}(a)u=A_{i,j}(a)u=\theta.$$

De $u\ne\theta$, tenemos que

$$\det(A_{i,j,(m_1,k_1),\ldots,(m_n,k_n)}(a))=0.$$

Derivando la fórmula $\mathrm{Adj}(A(t))A(t) = \det A(t)I$ y teniendo en cuenta lo que ya hemos probado, obtenemos que

$$\Big(\mathrm{Adj}\big(A(t)\big)\Big)^{(l)}(a)A(a) = \left(\det A(t)\right)^{(l)}(a)I, \quad \text{para } l = 0, \cdots, p-1.$$

Puesto que $A(a)$ es singular, deducimos que $\left(\det A(t)\right)^{(l)}(a) = 0$, $l = 0, \cdots, p-1$, y por tanto a es un cero de $A(t)$ de multiplicidad al menos p.

□

Probamos ahora el lema 3.1

Demostración del lema 3.1:

Sea a un autovalor de la matriz J_{nN}, y sea v un autovector ($v \neq \theta$) de esta matriz correspondiente al autovalor a. Escribimos v como un vector columna:

$$v = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{pmatrix}, \tag{3.3}$$

donde $v_i \in \mathbb{C}^N$. La ecuación $J_{nN}v = av$ puede escribirse

$$\begin{aligned} B_0 v_0 + A_1 v_1 &= a v_0, \\ A_1^* v_0 + B_1 v_1 + A_2 v_2 &= a v_1, \\ &\vdots \\ A_{n-1}^* v_{n-2} + B_{n-1} v_{n-1} &= a v_{n-1}, \end{aligned}$$

lo cual a partir de la relación de recurrencia de tres términos para P_n y usando que A_n es no singular da sucesivamente

$$\begin{aligned} v_1 &= P_1(a) v_0, \\ v_2 &= P_2(a) v_0, \\ &\vdots \\ v_{n-1} &= P_{n-1}(a) v_0, \\ \theta &= P_n(a) v_0. \end{aligned}$$

Esto prueba que $v_0 \neq \theta$ (en caso contrario $v = \theta$), y por tanto $P_n(a)$ es singular, es decir a es un cero de P_n.

Si V_a es el espacio de autovectores a la derecha de la matriz J_{nN} para el autovalor a, entonces $W_a = \{v_0 : v \in V_a\}$ es un subespacio de $\mathbb{C}^N$ de la misma dimensión que V_a, y $P_n(a)v_0 = \theta$, para $v_0 \in W_a$. Esto prueba que $W_a \subset R(a, P_n)$, donde con $R(a, P_n)$ denotamos el espacio de autovectores a la derecha de la matriz $P_n(a)$ para el autovalor 0.

Recíprocamente, si $P_n(a)v_0 = \theta$ para algún $v_0 \in \mathbb{C}^N$, y definimos $v_k = P_k(a)v_0$, $k = 1, \cdots, n-1$, entonces v definido por (3.3) es un autovector de J_{nN} correspondiente a a. Esto prueba que $W_a = R(a, P_n)$. Entonces, concluimos que la dimensión de $R(a, P_n)$ es también la multiplicidad de a como autovalor de J_{nN}. Del lema 3.2 deducimos que a es un cero de $P_n(t)$ de multiplicidad al menos la multiplicidad de a como autovalor de J_{nN}.

Los ceros de P_n son los ceros de $\det P_n$. Puesto que las matrices $(A_k)_k$ son triangulares inferiores y no singulares, de la relación de recurrencia de tres términos se sigue que $\det P_n$ es un polinomio de grado justamente nN. La matriz J_{nN} es una matriz de tamaño $nN \times nN$ y por tanto $\det(tI_{nN} - J_{nN})$ es un polinomio de grado justamente nN, así que del resultado probado antes, se sigue que los ceros del polinomio matricial P_n coinciden con (y tienen la misma multiplicidad que) los del polinomio $\det(tI_{nN} - J_{nN})$. Más aún, la matriz

$$A_a = \begin{pmatrix} P_0(a) \\ P_1(a) \\ \vdots \\ P_{n-1}(a) \end{pmatrix} \tag{3.4}$$

define una biyección entre $R(a, P_n)$ y el subespacio de autovectores de la matriz J_{nN} asociados al autovalor a.

$\square$

Ahora podemos enunciar y demostrar el teorema principal de este capítulo

Teorema 3.3.

(1) *Los ceros de P_n tienen una multiplicidad no mayor que N. Además P_n tiene nN ceros (contando las multiplicidades) y todos los ceros son reales ($n \in \mathbb{N}$).*

(2) *Si a es un cero de multiplicidad p de P_n, entonces* $\text{rango}(P_n(a)) = N - p$.

(3) *Si llamamos $x_{n,k}$ ($k = 1, \cdots, nN$) a los ceros de P_n ordenados en orden creciente (y teniendo encuenta sus multiplicidades), entonces*

$$x_{n+1,k} \leq x_{n,k} \leq x_{n+1,k+N} \quad \text{para } k = 1, \cdots, nN.$$

(4) *Si a es un cero de P_n y P_{n+1}, entoncés $P_n(a)$ y $P_{n+1}(a)$ no tienen ningún autovector asociado a cero común.*

(5) *Si $x_{n,k}$ es un cero de P_n de multiplicidad N, entonces $P_n(a) = \theta$. Más aún, cada valor complejo de $x_{n,k}^{\frac{1}{N}}$ es un cero de los N polinomios escalares consecutivos $p_{nN}(t), \cdots, p_{nN+N-1}(t)$ (para la definición de los polinomios escalares asociados a P_n véase los preliminares). En este caso el número real $x_{n,k}$ no puede ser un cero del polinomio matricial P_{n+1}.*

(6) *Fórmula de cuadratura. Cada cero $x_{n,k}$ (que se repetirá de acuerdo con su multiplicidad) del polinomio matricial P_n ($k = 1, \cdots, nN$) tiene asociada una matriz $B_{n,k}$ de tamaño $N \times N$ semidefinida positiva y de rango 1 de modo que*

$$< P, Q >_{(P_n)} = \sum_{k=1}^{nN} P(x_{n,k}) B_{n,k} Q^*(x_{n,k}) \tag{3.5}$$

para P, Q polinomios matriciales cualesquiera verificando $\text{gr}(P) + \text{gr}(Q) \leq 2n - 1$, *donde $< \cdot, \cdot >_{(P_n)}$ denota el producto escalar matricial definido por los $(P_n)_n$.*

La demostración de estos resultados es muy simple, y se deriva del hecho de ser los ceros de P_n autovalores de la matriz hermítica J_{nN}.

Otras fórmulas de cuadratura para polinomios ortogonales matriciales han sido encontradas recientemente y por otros métodos por Sinap y Van Assche (véase [**SV**]). La demostración dada aquí para la fórmula (3.5) debería compararse con la dada por ellos.

Debemos reseñar que en un trabajo aún por aparecer, Antonio J. Durán ha completado el teorema 3.3 dando nuevas propiedades de los ceros de los polinomios ortogonales matriciales y una expresión cerrada para los pesos de cuadratura. Estos resultados han sido usados para extender el conocido teorema de Markov a una sucesión de polinomios ortogonales matriciales (véase [**D3**]).

Demostración del teorema 3.3:

Del lema 3.1, y teniendo en cuenta que la matriz J_{nN} es una matriz hermítica de tamaño $nN \times nN$ de $(2N+1)$ diagonales, se tiene (1) del teorema 3.3.

Hemos probado en el lema 3.1 que la multiplicidad de a coincide con la dimensión de $R(a, P_n)$, el espacio de autovectores a la derecha de la matriz $P_n(a)$ asociados al autovalor 0. Y la dimensión de este espacio es justamente $N - \text{rango}(P_n(a))$.

Para probar (3) del teorema 3.3, es suficiente observar que la matriz J_{nN} se obtiene de $J_{(n+1)N}$ eliminando las últimas N filas y columnas, de este modo el principio de inclusión para matrices hermíticas [**HJ**, p. 189] da las propiedades de separación para los ceros de P_n.

Para probar (4), obsérvese que si $P_n(a)$ y $P_{n+1}(a)$ tuvieran un autovector común v, la fórmula de recurrencia implicaría que a es un cero de P_{n-1}, y entonces v sería también autovector de $P_{n-1}(a)$ asociado a 0. Procediendo sucesivamente llegaríamos a que a es un cero de P_0, que es una contradicción porque P_0 es no singular.

Probamos ahora (5) del teorema 3.3. Si $x_{n,k}$ es un cero de P_n de multiplicidad N, deducimos del punto (2) de este teorema que $R(x_{n,k}, P_n)$ tiene dimensión N, es decir, cada vector u de $\mathbb{C}^N$ es un autovector de $P_n(x_{n,k})$ asociado al autovalor 0, así, $P_n(x_{n,k}) = \theta$. Si suponemos que $x_{n,k}$ es también un cero del polinomio matricial P_{n+1}, tendríamos que necesariamente $P_n(x_{n,k})$ y $P_{n+1}(x_{n,k})$ tendrían un autovector común, en contra de (4).

El resto de (5) se sigue de la relación existente entre los polinomios ortogonales matriciales y los polinomios escalares que satisfacen una relación de recurrencia de (2N+1) términos (ver preliminares).

Por último probamos (6), es decir, la fórmula de cuadratura. Así, consideremos un cero $x_{n,l+1}$ de P_n de multiplicidad $m \leq N$. Podemos suponer que $x_{n,l+1} = x_{n,l+2} = \cdots = x_{n,l+m}$. Ya probamos que la matriz $A_{x_{n,l+1}}$ (véase (3.4)) establece una biyección entre el subespacio $R(x_{n,l+1}, P_n)$ y el subespacio de autovectores de la matriz J_{nN} asociados al autovalor $x_{n,l+1}$. Ahora, elegimos una base $\{v_{l+1}, \cdots, v_{l+m}\}$ en $R(x_{n,l+1}, P_n)$ de modo que los vectores $A_{x_{n,l+1}}v_{l+1}, \cdots, A_{x_{n,l+1}}v_{l+m}$ formen una base ortonormal del subespacio de autovectores asociados al autovalor $x_{n,l+1}$ de la matriz J_{nN}. Procediendo de esta forma para cada cero del polinomio matricial P_n, y puesto que los autovectores correspondientes a diferentes autovalores son ortogonales, obtenemos una base ortonormal

$$\left\{A_{x_{n,1}}v_1, \cdots, A_{x_{n,nN}}v_{nN}\right\}$$

en $\mathbb{C}^{nN}$. Si usamos estos vectores como las columnas de una matriz B, es inmediato que esta matriz de tamaño $nN \times nN$ es unitaria. La fórmula de cuadratura para los polinomios $(P_n)_n$ esta implícitamente contenida en esta propiedad de la matriz B, como se probará a continuación.

De acuerdo con la definición de las matrices $A_{x_{n,l}}$ (véase (3.4)), podemos escribir la matriz B como

$$B = (P_k(x_{n,l})v_l)_{\substack{k=0,\cdots,n-1,\\ l=1,\cdots,nN}} \tag{3.6}$$

donde $P_k(x_{n,l})v_l$ es una caja de tamaño $N \times 1$. Ahora, consideramos la matriz C de tamaño $nN^2 \times nN$ definida por

$$C = \begin{pmatrix} v_1 & \theta & \theta & \dots & \theta \\ \theta & v_2 & \theta & \dots & \theta \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta & \theta & \theta & \dots & v_{nN} \end{pmatrix}.$$

Si consideramos la matriz CC^* dividida en cajas de tamaño $N \times N$, podemos escribirla como una matriz diagonal por cajas

$$CC^* = \begin{pmatrix} B_{n,1} & & \\ & \ddots & \\ & & B_{n,nN} \end{pmatrix},$$

donde las matrices $B_{n,k}$ están definidas por $B_{n,k}(i,j) = v_{k,i}\bar{v}_{k,j}$, $i,j = 1,\cdots,N$. Así, estas matrices son semidefinidas positivas y de rango 1. Según la definición de la matriz B (véase (3.6)) y C, la condición $BB^* = I$ puede escribirse como

$$\begin{pmatrix} P_0(x_{n,1}) & \dots & P_0(x_{n,nN}) \\ \vdots & \ddots & \vdots \\ P_{n-1}(x_{n,1}) & \dots & P_{n-1}(x_{n,nN}) \end{pmatrix} CC^* \begin{pmatrix} P_0^*(x_{n,1}) & \dots & P_{n-1}^*(x_{n,1}) \\ \vdots & \ddots & \vdots \\ P_0^*(x_{n,nN}) & \dots & P_{n-1}^*(x_{n,nN}) \end{pmatrix} = I.$$

Si separamos las matrices

$$\begin{pmatrix} P_0(x_{n,1}) & \dots & P_0(x_{n,nN}) \\ \vdots & \ddots & \vdots \\ P_{n-1}(x_{n,1}) & \dots & P_{n-1}(x_{n,nN}) \end{pmatrix},$$

en el producto anterior, en cajas de tamaño $N \times nN$, a partir de la definición de las matrices $B_{n,k}$ $(k = 1,\cdots,nN)$, tenemos

$$\sum_{i=1}^{nN} P_k(x_{n,i})B_{n,i}P_l^*(x_{n,i}) = \delta_{k,l}I_{N\times N} \tag{3.7}$$

para $0 \le k,l \le n-1$.

Por definición de los vectores $(v_i)_{i=1,\cdots,nN}$, tenemos que $P_n(x_{n,i})v_i = v_i^* P_n^*(x_{n,i}) = \theta$, $i = 1, \cdots, nN$. Esto da

$$(3.8) \qquad (P_n(x_{n,1}), \cdots, P_n(x_{n,nN}))\, C = \theta \quad \text{y} \quad C^* \begin{pmatrix} P_n^*(x_{n,1}) \\ \vdots \\ P_n^*(x_{n,nN}) \end{pmatrix} = \theta.$$

Así, para $k = 0, \cdots, n-1$ tenemos que

$$(3.9) \qquad \begin{cases} (P_k(x_{n,1}), \cdots, P_k(x_{n,nN}))\, CC^* \begin{pmatrix} P_n^*(x_{n,1}) \\ \vdots \\ P_n^*(x_{n,nN}) \end{pmatrix} = \theta, \\ (P_n(x_{n,1}), \cdots, P_n(x_{n,nN}))\, CC^* \begin{pmatrix} P_k^*(x_{n,1}) \\ \vdots \\ P_k^*(x_{n,nN}) \end{pmatrix} = \theta \end{cases}.$$

De (3.7) y (3.9), obtenemos

$$\sum_{i=1}^{nN} P_k(x_{n,i}) B_{n,i} P_l^*(x_{n,i}) = \delta_{k,l} I_{N\times N}$$

para $k = 0, \cdots, n-1$ y $l = 0, \cdots, n$, o $k = 0, \cdots, n$ y $l = 0, \cdots, n-1$. La ortonormalidad de los polinomios $(P_n)_n$ da que

$$\int_{\mathbb{R}} P_k(t) dM(t) P_l^*(t) = \delta_{k,l} I_{N\times N} = \sum_{i=1}^{nN} P_k(x_{n,i}) B_{n,i} P_l^*(x_{n,i})$$

para $k = 0, \cdots, n-1$ y $l = 0, \cdots, n$, o $k = 0, \cdots, n$ y $l = 0, \cdots, n-1$. Es decir, la fórmula de cuadratura (3.5) para los polinomios $(P_k)_{k=0}^{n-1}$, $(P_l)_{l=0}^{n}$ o $(P_k)_{k=0}^{n}$, $(P_l)_{l=0}^{n-1}$. Por linealidad, la fórmula de cuadratura se tiene para todos los polinomios P, Q que satisfagan $\mathrm{gr}(P)+\mathrm{gr}(Q) \leq 2n-1$.

$\square$

Observación 3.4.

Procediendo como en la demostración del lema 3.1 es posible probar que los ceros de ciertas perturbaciones de los polinomios matriciales P_n son también reales, y también es posible dar una fórmula de cuadratura a partir de los mismos. Sea A una matriz de tamaño $N \times N$. Si $A_n A = A^* A_n^*$, entonces

(1) Los ceros de $P_n - AP_{n-1}$ tienen una multiplicidad no mayor que N. Además, $P_n - AP_{n-1}$ tiene nN ceros (contando las multiplicidades) y todos los ceros son reales ($n \in \mathbb{N}$).

(2) Si a es un cero de multiplicidad p de $P_n - AP_{n-1}$, entonces $\operatorname{rango}(P_n(a) - AP_{n-1}(a)) = N - p$.

(3) Fórmula de cuadratura. Cada cero $x_{n,k}$ del polinomio matricial $P_n - AP_{n-1}$ ($k = 1, \ldots, nN$) tiene asociada una matriz $B_{n,k}$ de tamaño $N \times N$ semidefinida positiva y de rango 1 de modo que

$$< P, Q >_{(P_n)} = \sum_{k=1}^{nN} P(x_{n,k}) B_{n,k} Q^*(x_{n,k}) \tag{3.9}$$

para P, Q polinomios matriciales cualesquiera verificando $\operatorname{gr}(P) + \operatorname{gr}(Q) \leq 2n - 2$.

Estas perturbaciones de los polinomios $(P_n)_n$ jugarán un papel esencial para describir las medidas extremales correspondientes a un problema de momentos matricial truncado (véase capítulo 6).

Demostración

Sea a un cero del polinomio matricial $P_n - AP_{n-1}$, y sea v_0 un autovector de la matriz numérica $P_n(a) - AP_{n-1}(a)$ asociado a 0. Definamos $v_k = P_k(a)v_0$, $k = 1, \ldots, n-1$, y v como en (3.3). Entonces, usando la relación de recurrencia de tres términos (3.1), es fácil mostrar que v es un autovector de la matriz

$$\begin{pmatrix} B_0 & A_1 & & & \\ A_1^* & B_1 & A_2 & & \\ & \ddots & \ddots & \ddots & \\ & & A_{n-2}^* & B_{n-2} & A_{n-1} \\ & & & A_{n-1}^* & B_{n-1} + A_n A \end{pmatrix}$$

asociado a a. Puesto que $A_n A$ es hermítica a debe ser real. El resto de (1) y (2) se tiene igual que en la demostración del lema 3.3.

La demostración de la fórmula de cuadratura es igual que en el teorema 3.3, salvo que para estas perturbaciones no se tiene la fórmula (3.8) y por tanto la igualdad sólo se da para polinomios P y Q tales que $\mathrm{gr}(P)+\mathrm{gr}(Q) \leq 2n-2$.

$\square$

La parte final del capítulo está dedicada a la construcción de una medida ortogonalizante para la sucesión $(P_n)_n$, usando la fórmula de cuadratura. Esto proporciona una nueva demostración para la versión matricial del teorema de Favard (para otras demostraciones usando diferentes técnicas véase [**AN**], o [**D2**]).

A partir de las matrices semidefinidas positivas $(B_{n,k})_{n,k}$ que aparecen en la fórmula de cuadratura (3.9), podemos definir una sucesión de medidas discretas y definidas positivas como

$$\mu_n = \sum_{k=1}^{nN} B_{n,k} \delta_{x_{n,k}},$$

donde $x_{n,k}$ $(k = 1, \cdots, nN)$ son los ceros de P_n. En analogía con el caso escalar, una medida ortogonalizante para los polinomios matriciales $(P_n)_n$ puede obtenerse como un punto límite de estas matrices de medidas.

Para obtener este punto límite bastaría aplicar el lema 2.8, teniendo en cuenta que la topología vaga es metrizable. Haremos sin embargo aquí otra construcción ligeramente distinta.

En primer lugar, necesitamos recordar algunos resultados bien conocidos. Para $a, b \in \mathbb{R}^+$, $a < b$, denotamos por $\mathcal{C}([-a,b], \mathbb{C}^{N\times N})$ el espacio de las funciones continuas del intervalo $[-a,b]$ con valores en el espacio vectorial del matrices complejas de tamaño $N \times N$. El espacio dual $\mathcal{C}'([-a,b], \mathbb{C}^{N\times N})$ es el espacio de matrices de tamaño $N \times N$ cuyas componentes son medidas de Borel. Es claro que la bola unidad de $\mathcal{C}'([-a,b], \mathbb{C}^{N\times N})$ es débil compacta (teorema de Banach Alaoglu).

Ahora, para $a, b \in \mathbb{R}^+$, consideramos la matriz de medidas

$$\mu_{n_{|[-a,b]}} = \sum_{\substack{k=1 \\ x_{n,k} \in [-a,b]}}^{nN} B_{n,k} \delta_{x_{n,k}}.$$

La norma de una matriz de medidas $\mu = (\mu_{i,j})_{1 \leq i,j \leq N}$ se define como

$$\|\mu\| = \sum_{i,j=1}^{N} \|\mu_{i,j}\| = \sum_{i,j=1}^{N} |\mu_{i,j}(\mathbb{R})|.$$

Todas las matrices de medidas μ_n tienen igual el primer momento. Además, sabemos que todas las matrices de medidas definidas positivas están acotadas por su traza. Por ello, cualesquiera que sean $n \in \mathbb{N}$ y $a, b \in \mathbb{R}^+$ tenemos que

$$\begin{aligned}\|\mu_{n_{|[-a,b]}}\| = \sum_{i,j=1}^{N} \mu_{n,i,j_{|[-a,b]}} &\leq \sum_{i,j=1}^{N} \|\mu_{n,i,j}\| \\ &= \sum_{i,j=1}^{N} |\mu_{n,i,j}|(\mathbb{R}) \leq N^2 \tau \mu_n(\mathbb{R}).\end{aligned}$$

$N^2 \tau \mu_n(\mathbb{R})$ no depende de n por ser $\tau \mu_n(\mathbb{R})$ la traza del primer momento de μ_n, que no depende de n. Es decir, tenemos que cualesquiera que sean $n \in \mathbb{N}$ y $a, b \in \mathbb{R}^+$, las medidas $\mu_{n_{|[-a,b]}}$ están en una bola de $C'([-a,b], \mathbb{C}^{N\times N})$ de radio constante, y por tanto podemos aplicar el teorema de Banach Alaoglu.

Dadas dos sucesiones crecientes $(a_k)_k, (b_k)_k$, para las cuales $a_k, b_k \to +\infty$ cuando $k \to +\infty$, llevando a cabo un proceso diagonal y teniendo en cuenta la débil compacidad de la bola unidad del espacio $C'([-a,b], \mathbb{C}^{N\times N})$, obtenemos una sucesión creciente de enteros no negativos $(n_m)_m$, y para cada $k \in \mathbb{N}$ una matriz de medidas $\mu^{(k)} \in C'([-a_k, b_k], \mathbb{C}^{N\times N})$ tal que para todo $k \in \mathbb{N}$, μ_{n_m} converge débilmente a $\mu^{(k)}$ en $C'([-a_k, b_k], \mathbb{C}^{N\times N})$, es decir

$$\lim_{m\to\infty} \int_{[-a_k,b_k]} f(t)\, d\mu_{n_m}(t) = \int_{[-a_k,b_k]} f(t)\, d\mu^{(k)}(t). \tag{3.11}$$

para toda $f \in C([-a_k, b_k], \mathbb{C}^{N\times N})$. Además es claro que también podemos obtener $\mu^{(k)} = \mu^{(k')}$ en $[-a_k, b_k]$ para $k \leq k'$.

Puesto que las matrices de medidas μ_{n_m} son definidas positivas, los resultados probados en el capítulo 2 (véase teorema 2.5) prueban que las matrices $\mu^{(k)}$ son también definidas positivas, $k \in \mathbb{N}$.

Dado un conjunto de Borel $A \subseteq \mathbb{R}$, se define

$$\mu(A) = \lim_{k\to\infty} \mu^{(k)}(A \cap [-a_k, b_k]).$$

Debemos probar que este límite existe para cada conjunto de Borel A. Obsérvese que puesto que $\mu^{(k)} = \mu^{(k+1)}$ en $[-a_k, b_k]$, se tiene que

$$\begin{aligned}\mu^{(k+1)}&(A \cap [-a_{k+1}, b_{k+1}]) \\ &= \mu^{(k)}(A \cap [-a_k, b_k]) + \mu^{(k+1)}(A \cap [-a_{k+1}, b_{k+1}] \setminus [-a_k, b_k])\end{aligned}$$

y por tanto la sucesión de matrices numéricas $\mu^{(k)}(A \cap [-a_k, b_k])$ es creciente. Además, esta sucesión está acotada, pues

$$\begin{aligned}\mu^{(k)}(A \cap [-a_k, b_k]) &\leq \mu^{(k)}([-a_k, b_k]) \\ &= \lim_{m\to\infty} \mu_{n_m}([-a_k, b_k]) \\ &\leq \lim_{m\to\infty} \mu_{n_m}(\mathbb{R})\end{aligned}$$

que es una constante pues las matrices μ_{n_m} tienen igual el primer momento. Así, $\mu(A)$ está bien definida.

Además, por ser las matrices $\mu^{(k)}$ definidas positivas, μ es también definida positiva.

Ahora, probamos que

$$\int_{\mathbb{R}} t^n I \, d\mu(t) = \lim_{m\to\infty} \int_{\mathbb{R}} t^n I \, d\mu_{n_m}(t), \quad \text{para todo } n \in \mathbb{N}, \tag{3.12}$$

donde I es la matriz identidad de tamaño N. Para un $n \in \mathbb{N}$ fijo, la fórmula de cuadratura (3.9) da para $n_m, n_{m'} > n$ que

$$\int_{\mathbb{R}} t^n I \, d\mu_{n_m}(t) = \int_{\mathbb{R}} t^n I \, d\mu_{n_{m'}}(t).$$

Si llamamos A_n a esta matriz, debemos probar que $A_n = \int_{\mathbb{R}} t^n I\, d\mu(t)$. Sea $\|\cdot\|_2$ la norma espectral definida por

$$\|A\|_2 = \max\left\{\sqrt{\lambda} : \lambda \text{ es un autovalor de } A^*A\right\}.$$

Para $n_m > n$, tenemos

$$\left\|\int_{-a_k}^{b_k} t^n I d\mu(t) - A_n\right\|_2 = \left\|\int_{-a_k}^{b_k} t^n I d\mu(t) - \int_{\mathbb{R}} t^n I d\mu_{n_m}(t)\right\|_2$$

$$\leq \left\|\int_{-a_k}^{b_k} t^n I d\mu(t) - \int_{-a_k}^{b_k} t^n I d\mu_{n_m}(t)\right\|_2 + \left\|\int_{-\infty}^{-a_k} t^n I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^n I d\mu_{n_m}(t)\right\|_2 .$$

Dado $\epsilon > 0$, puesto que $\mu = \mu^{(k)}$ en $[-a_k, b_k]$, para n_m suficientemente grande, (3.11) da

$$\left\|\int_{-a_k}^{b_k} t^n I d\mu(t) - \int_{-a_k}^{b_k} t^n I d\mu_{n_m}(t)\right\|_2 < \epsilon$$

Ahora, si tomamos un entero no negativo l tal que $2l > n$, tenemos que

$$|t^n| = \left|\frac{t^{2l}}{t^{2l-n}}\right| \leq \left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n} t^{2l} \quad \text{para } t \notin [-a_k, b_k].$$

Por tanto, tenemos las siguientes desigualdades de matrices (como es usual $A \leq B$ si $B - A$ es semidefinida positiva)

$$-\left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n}\left(\int_{-\infty}^{-a_k} t^{2l} I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^{2l} I d\mu_{n_m}(t)\right)$$

$$\leq \int_{-\infty}^{-a_k} t^n I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^n I d\mu_{n_m}(t)$$

$$\leq \left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n}\left(\int_{-\infty}^{-a_k} t^{2l} I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^{2l} I d\mu_{n_m}(t)\right).$$

Como $t^{2l} \geq 0$, para $t \in \mathbb{R}$, y la matriz de medidas μ_{n_m} es positiva definida, se sigue que la matriz

$$\left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n} \left(\int_{-\infty}^{-a_k} t^{2l} I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^{2l} I d\mu_{n_m}(t)\right)$$

es también definida positiva. Por la definición de la norma espectral $\|\cdot\|_2$, obtenemos

$$\begin{aligned}
&\left\| \int_{-\infty}^{-a_k} t^n I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^n I d\mu_{n_m}(t) \right\|_2 \\
&\qquad \leq \left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n} \left\| \int_{-\infty}^{-a_k} t^{2l} I d\mu_{n_m}(t) + \int_{b_k}^{\infty} t^{2l} I d\mu_{n_m}(t) \right\|_2 \\
&\qquad \leq \left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n} \| \int_{\mathbb{R}} t^{2l} d\mu_{n_m}(t) \|_2 \\
&\qquad = \left(\frac{1}{\min\{a_k, b_k\}}\right)^{2l-n} \|A_{2l}\|_2
\end{aligned}$$

Tomando $k \to \infty$, se tiene (3.12).

De (3.12) y de la fórmula de cuadratura (3.9), se sigue que los polinomios $(P_n)_n$ son ortonormales son respecto a la matriz de medidas definida positiva μ. Obsérvese que μ puede no ser única puesto que hemos seleccionado sólo un posible límite débil de la sucesión de medidas discretas dada por la fórmula de cuadratura.

CAPÍTULO 4

EL TEOREMA DE BLUMENTHAL PARA POLINOMIOS MATRICIALES ORTOGONALES

En este capítulo generalizamos a polinomios matriciales ortogonales el teorema de Blumenthal: si los coeficientes de la relación de recurrencia asociada a una sucesión de polinomios ortogonales escalares en la recta real tienden a límites finitos, digamos $\lim_{n\to\infty} a_n = a$ y $\lim_{n\to\infty} b_n = b$, entonces el soporte de la medida ortogonalizante es el intervalo $[b-2a, b+2a]$, y posiblemente dos sucesiones que tienden a los extremos del intervalo.

Comenzamos por una sucesión de polinomios que satisfacen una relación de recurrencia de $(2N+1)$ términos de la forma

$$t^N p_n(t) = c_{n,0} p_n(t) + \sum_{k=1}^{N} \left(\overline{c_{n,k}} p_{n-k}(t) + c_{n+k,k} p_{n+k}(t)\right), \tag{4.1}$$

donde $c_{n,0}$ $(n = 0, 1, 2, \cdots)$ es una sucesión de números reales y $c_{n,k}$ $(n = 0, 1, 2, \cdots)$ son sucesiones complejas para $k = 1, \cdots, N$, con $c_{n,N} \neq 0$ para todo n y con las condiciones iniciales $p_k(x) = 0$ si $k < 0$ y p_k polinomios de grado k dados, si $k = 0, \cdots, N-1$. Suponemos que los coeficientes de esta relación de recurrencia tienden a límites finitos. Denotamos estos límites por $c_0, \cdots, c_N$:

$$c_k = \lim_{n\to\infty} c_{n,k}, \quad k = 0, \cdots, N.$$

Puesto que la sucesión $(c_{n,0})_n$ es real, el número c_0 es también real.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

Es conveniente plantear el problema en términos matriciales. Para ello consideramos los polinomios matriciales $(P_n)_n$ asociados a los polinomios escalares $(p_n)_n$ como se describió en los preliminares. Estos polinomios matriciales satisfacen una relación de recurrencia de tres términos de la forma

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t) \tag{4.2}$$

con $P_0(t) = I$ y $P_{-1}(t) = \theta$, donde los coeficientes de recurrencia A_{n+1}, B_n son también matrices de tamaño $N \times N$.

Teniendo en cuenta la relación, dada en los preliminares, existente entre los coeficientes matriciales de la relación de recurrencia de $(P_n)_n$ y los coeficientes escalares en la relación de $(2N + 1)$ términos verificada por $(p_n)_n$, la hipótesis de la convergencia de las sucesiones que aparecen en la fórmula de recurrencia (4.1) es equivalente a lo siguiente: los coeficientes matriciales $(A_k)_k$, $(B_k)_k$ tienden a las matrices límite $A = (A_{i,j})_{i,j=1,\cdots,N}$, $B = (B_{i,j})_{i,j=1,\cdots,N}$, que satisfacen

(4.3)
$$\begin{cases} \text{para } i < j, \quad A_{i,j} = 0, \\ \text{para } m = 0, \cdots, N-1, \quad A_{i+m,i} = A_{j+m,j},\ i,j = 1, \cdots, N-m, \\ \text{para } m = 0, \cdots, N-1, \quad B_{i+m,i} = B_{j+m,j},\ i,j = 1, \cdots, N-m, \\ \text{y para } m = 2, \ldots, N, \quad B_{1,m} = A_{N-m+2,1} \end{cases}$$

es decir, la matriz límite A es triangular inferior, A y B tienen el mismo valor en cada diagonal (son matrices de Toeplitz finitas) y el valor de la i-ésima diagonal superior de B es el mismo que el de la $(N + 2 - i)$-ésima diagonal inferior de A.

Tenemos entonces el siguiente teorema

Teorema 4.1. *Sea $(P_n)_n$ una sucesión de polinomios matriciales verificando una relación de recurrencia como (4.2), y supongamos que las sucesiones de matrices $(A_n)_n$ y $(B_n)_n$ convergen respectivamente a las matrices límite A y B, verificando estas (4.3). Definamos el polinomio* *trigonométrico*

$$t(x) = \left(\sum_{k=1}^{N} 2\Re(c_k)T_k(cos(x))\right) - \left(\sum_{k=1}^{N} 2sen(x)\Im(c_k)U_{k-1}(cos(x))\right),$$

donde $(T_k)_k$ *y* $(U_k)_k$ *son los polinomios de Chebyshev de primera y segunda clase, respectivamente. Sea* μ *la matriz de medidas definida positiva respecto de la cual los polinomios matriciales* $(P_n)_n$ *son ortonormales. Entonces el soporte de* μ *(ver (1.1) para la definición del soporte de* μ*) es el intervalo compacto*

$$\left[c_0 + \inf_{x\in[-\pi,\pi]} t(x), c_0 + \sup_{x\in[-\pi,\pi]} t(x)\right]$$

y, posiblemente, dos sucesiones de números reales fuera de este intervalo que tienden a los extremos del mismo. Más concretamente

$$\left[c_0 + \inf_{x\in[-\pi,\pi]} t(x), c_0 + \sup_{x\in[-\pi,\pi]} t(x)\right] \subset sop(\mu),$$

y para cada $\epsilon > 0$ *el conjunto*

$$sop(\mu) \setminus \left[c_0 + \inf_{x\in[-\pi,\pi]} t(x) - \epsilon, c_0 + \sup_{x\in[-\pi,\pi]} t(x) + \epsilon\right]$$

es finito.

Demostración

Procederemos en dos pasos. Primero reduciremos el problema al cálculo del espectro de un operador definido en el espacio de Hardy H^2 y segundo calcularemos explícitamente este espectro.

PRIMER PASO

Consideramos la N-matriz de Jacobi asociada a los polinomios matriciales $(P_n)_n$, es decir

$$J = \begin{pmatrix} B_0 & A_1 & & & \\ A_1^* & B_1 & A_2 & & \\ & A_2^* & B_2 & A_3 & \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

donde

$$A_n = \begin{pmatrix} c_{nN,N} & 0 & 0 & \cdots & 0 \\ c_{nN,N-1} & c_{nN+1,N} & 0 & \cdots & 0 \\ c_{nN,N-2} & c_{nN+1,N-1} & c_{nN+2,N} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{nN,1} & c_{nN+1,2} & c_{nN+2,N} & \cdots & c_{nN+N-1,N} \end{pmatrix},$$

y

$$B_n = \begin{pmatrix} c_{nN,0} & c_{nN+1,1} & c_{nN+2,2} & \cdots & c_{nN+N-1,N-1} \\ \overline{c_{nN+1,1}} & c_{nN+1,0} & c_{nN+2,1} & \cdots & c_{nN+N-1,N-2} \\ \overline{c_{nN+2,2}} & \overline{c_{nN+2,1}} & c_{nN+2,0} & \cdots & c_{nN+N-1,N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{c_{nN+N-1,N-1}} & \overline{c_{nN+N-1,N-2}} & \overline{c_{nN+N-1,N-3}} & \cdots & c_{nN+N-1,0} \end{pmatrix}.$$

Las sucesiones en las diagonales de esta matriz infinita, es decir $(c_{n+k,k})_n$, $k = 0, \cdots, N$ y $(\overline{c_{n,k}})_n$, $k = 1, \cdots, N$, son por hipótesis sucesiones convergentes, con límites $c_k = \lim_{n\to\infty} c_{n,k}$, $k = 0, \cdots, N$. Esto implica que el operador asociado a la matriz J en el espacio de Hilbert ℓ^2 y definido por

$$J : \ell^2 \to \ell^2$$
$$J((a_n)_n) = (a_n)_n J,$$

es acotado.

En [**D2**, Sect. 3], se muestra como obtener una matriz de medidas ortogonalizante para $(P_n)_n$ a partir de la resolución de la identidad de cualquier extensión autoadjunta del operador J. Puesto que el operador J es acotado, J tiene una única extensión autoadjunta, y por tanto la matriz de medidas que ortogonaliza a los polinomios matriciales $(P_n)_n$ es única, y además su soporte coincide con el espectro de J. Así, vamos a determinar este espectro. Procedemos como en [**MNV**, Sect. 4].

Usando el siguiente teorema de H. Weyl (véase por ejemplo [**RN**, §134, p. 367]), sustituimos la matriz J por una más simple.

Teorema (Weyl). *Sean T y U operadores autoadjuntos acotados en un espacio de Hilbert, siendo U compacto. Entonces el espectro esencial de T y de $T+U$ coinciden.*

El espectro esencial de un operador se define como el conjunto de puntos límite de su espectro. Llamemos J_0 a la matriz infinita de $(2N+1)$ diagonales cuyo valor en cada diagonal es el valor límite de la correspondiente diagonal de la matriz J:

$$J_0 = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_N & 0 & \cdots \\ \overline{c_1} & c_0 & c_1 & \cdots & c_{N-1} & c_N & \cdots \\ \overline{c_2} & \overline{c_1} & c_0 & \cdots & c_{N-2} & c_{N-1} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \overline{c_N} & \overline{c_{N-1}} & \overline{c_{N-2}} & \cdots & c_0 & c_1 & \cdots \\ 0 & \overline{c_N} & \overline{c_{N-1}} & \cdots & \overline{c_1} & c_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Puesto que las diagonales de la matriz infinita de $(2N+1)$ diagonales $J - J_0$ tienden a cero, el operador definido por esta matriz es compacto, y por tanto J y J_0 tienen el mismo espectro esencial.

El teorema 4.1 seguirá si probamos que el espectro de J_0 es el intervalo compacto

$$\left[c_0 + \inf_{x\in[-\pi,\pi]} t(x), c_0 + \sup_{x\in[-\pi,\pi]} t(x)\right].$$

Podemos suponer que $c_N \neq 0$. Si no, tendríamos que determinar el espectro de un operador definido por una matriz infinita de $(2M+1)$ diagonales, con $M < N$, en lugar del de una con $(2N+1)$ diagonales.

Consideremos el operador J_0 actuando sobre el espacio de Hardy H^2, es decir, el espacio de Hilbert de las funciones analíticas f en el disco unidad D, $f(z) = \sum_{j=0}^{\infty} a_j z^j$, con los coeficientes de Taylor $(a_j)_j$ en ℓ^2, dotado de la norma

$$\|f\| = \left(\sum_{j=0}^{\infty} |a_j|^2 \right)^{\frac{1}{2}}.$$

Teniendo en cuenta el isomorfismo

$$(a_j)_{j=0}^{\infty} \longmapsto \sum_{j=0}^{\infty} a_j z^j$$

entre ℓ^2 y H^2, la imagen por J_0 de una función f de H^2 se obtiene multiplicando la matriz J_0 por la derecha por el vector cuyas componentes son los coeficientes del desarrollo de f y expresando después el resultado en función de f. De esta forma se obtiene la siguiente expresión para el operador J_0:

$$(J_0 f)(z) = c_N z^N f(z) + \cdots + c_1 z f(z) + c_0 f(z)$$

$$+\overline{c_1}\frac{f(z)}{z} + \cdots + \overline{c_N}\frac{f(z)}{z^N}$$

$$-f(0)\left(\frac{\overline{c_1}}{z} + \cdots + \frac{\overline{c_N}}{z^N}\right) - f'(0)\left(\frac{\overline{c_2}}{z} + \cdots + \frac{\overline{c_N}}{z^{N-1}}\right) - \cdots - \frac{f^{(N-1)}(0)}{(N-1)!}\frac{\overline{c_N}}{z}.$$

SEGUNDO PASO Cálculo del espectro de J_0.

Recordemos que un número complejo λ pertenece al espectro de J_0 si el operador $J_0 - \lambda I$ no tiene una inversa acotada (I es el operador identidad) en H^2.

Consideremos la ecuación $(J_0 - \lambda I)f = g$, donde $g \in H^2$ está dada. Esta ecuación puede escribirse como

$$\left(c_N z^N + \cdots + c_1 z + c_0 - \lambda + \frac{\overline{c_1}}{z} + \cdots + \frac{\overline{c_N}}{z^N}\right) f(z)$$

$$- f(0)\left(\frac{\overline{c_1}}{z} + \cdots + \frac{\overline{c_N}}{z^N}\right) - \cdots - \frac{f^{(N-1)}(0)}{(N-1)!}\frac{\overline{c_N}}{z} = g(z)$$

Resolviendo la ecuación para $f(z)$ obtenemos:
(4.4)

$$f(z) = \frac{z^N g(z) + f(0)\left(\overline{c_1}z^{N-1} + \cdots + \overline{c_N}\right) + \cdots + \frac{f^{(N-1)}(0)}{(N-1)!}\overline{c_N}z^{N-1}}{c_N z^{2N} + \cdots + c_1 z^{N+1} + (c_0 - \lambda)z^N + \overline{c_1}z^{N-1} + \cdots + \overline{c_N}}.$$

Llamemos $p_\lambda(z)$ al polinomio del denominador, esto es,

$$p_\lambda(z) = c_N z^{2N} + \cdots + c_1 z^{N+1} + (c_0 - \lambda)z^N + \overline{c_1}z^{N-1} + \cdots + \overline{c_N}.$$

El siguiente lema será la clave para determinar cuándo la ecuación (4.4) define un operador acotado en H^2.

Lema 4.2. *Para cada g en H^2 la función f dada por (4.4) queda unívocamente determinada por la función g si y sólo si exactamente N de las raíces de $p_\lambda(z)$ están dentro del disco unidad $D = \{z : |z| < 1\}$.*

Demostración

Obsérvese que las $2N$ raíces del polinomio $p_\lambda(z)$ son de la forma $x_1, \ldots, x_N, \dfrac{1}{\overline{x_1}}, \cdots \dfrac{1}{\overline{x_N}}$.

Supongamos primero que ninguna de las raíces $x_1, \ldots, x_N, \dfrac{1}{\overline{x_1}}, \cdots \dfrac{1}{\overline{x_N}}$ tiene módulo 1. Entonces N de estas raíces están dentro de D. Podemos suponer que estas son $x_1, \ldots, x_N$. La función f definida por (4.4) es analítica en D si y sólo si $x_1, \cdots, x_N$ son también ceros del numerador en (4.4). Esto da un sistema lineal de ecuaciones con las N incógnitas $f(0), \ldots, \dfrac{f^{(N-1)}(0)}{(N-1)!}$, siendo el número de ecuaciones igual al número de raíces diferentes en $x_1, \ldots, x_N$.

Caso 1. Si todas estas raíces son diferentes tenemos un sistema lineal cuadrado cuyo determinante es

$$\det\begin{pmatrix} \overline{c_1}x_1^{N-1}+\cdots+\overline{c_N} & \overline{c_2}x_1^{N-1}+\cdots+\overline{c_N}x_1 & \dots & \overline{c_N}x_1^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{c_1}x_N^{N-1}+\cdots+\overline{c_N} & \overline{c_2}x_N^{N-1}+\cdots+\overline{c_N}x_N & \dots & \overline{c_N}x_N^{N-1} \end{pmatrix}.$$

Para simplificarlo, se suma en primer lugar la N-ésima columna (es decir la última) multiplicada por $-\dfrac{\overline{c_i}}{\overline{c_N}}$ a la columna i-ésima, para $1 \leq i \leq N-1$. En un segundo paso se suma la columna $(N-1)$-ésima multiplicada por $-\dfrac{\overline{c_{i+1}}}{\overline{c_N}}$ a la columna i-ésima, para $1 \leq i \leq N-2$. En el paso $N-1$ y último se suma la segunda columna multiplicada por $-\dfrac{\overline{c_{N-1}}}{\overline{c_N}}$ a la primera. Después de estas simplificaciones el determinante queda reducido a

$$\overline{c_N}^N \det\begin{pmatrix} 1 & x_1 & \dots & x_1^{N-1} \\ 1 & x_2 & \dots & x_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & \dots & x_N^{N-1} \end{pmatrix},$$

que es un determinante de Vandermonde con valor $\overline{c_N}^N \displaystyle\prod_{1\leq i<j\leq N}(x_j - x_i)$.

Si todas las raíces $x_1, \dots, x_N$ son diferentes, entonces para cualquier g dada en H^2 podemos determinar unívocamente la correspondiente f en H^2.

Caso 2. Veamos ahora que ocurre si la multiplicidad de alguna de las raíces $x_1, \dots, x_N$ es mayor que 1. Supongamos que tenemos p raíces diferentes $x_1, \dots, x_p$ con multiplicidades $N_1, \dots, N_p$, respectivamente, y $N_1 + \cdots + N_p = N$. En este caso, para compensar las raíces de $p_\lambda(z)$ forzamos que el numerador de (4.4) tenga un cero de orden justamente N_m en x_m, (para $m = 1, \cdots, p$). Entonces de nuevo obtenemos un sistema lineal cuadrado con las mismas incógnitas que antes.

Teniendo en cuenta que la derivada de orden j del numerador de (4.4) es

$$\left(z^N g(z)\right)^{(j)} + f(0)\left(\frac{(N-1)!}{(N-j-1)!}\overline{c_1}z^{N-j-1} + \cdots + j!\overline{c_{N-j}}\right) + \ldots$$
$$\cdots + \frac{f^{(N-1)}(0)}{(N-1)!}\frac{(N-1)!}{(N-j-1)!}\overline{c_N}z^{N-j-1}, \quad 0 \le j \le N,$$

el determinante de este sistema queda

$$\det\begin{pmatrix} \overline{c_1}x_1^{N-1} + \cdots + \overline{c_N} & \ldots & \overline{c_N}x_1^{N-1} \\ (N-1)\overline{c_1}x_1^{N-2} + \cdots + \overline{c_{N-1}} & \ldots & (N-1)\overline{c_N}x_1^{N-2} \\ \vdots & \ddots & \vdots \\ \frac{(N-1)!}{(N-N_1)!}\overline{c_1}x_1^{N-N_1} + \cdots + (N_1-1)!\overline{c_{N-N_1+1}} & \ldots & \frac{(N-1)!}{(N-N_1)!}\overline{c_N}x_1^{N-N_1} \\ \vdots & \ddots & \vdots \\ \overline{c_1}x_p^{N-1} + \cdots + \overline{c_N} & \ldots & \overline{c_N}x_p^{N-1} \\ (N-1)\overline{c_1}x_p^{N-2} + \cdots + \overline{c_{N-1}} & \ldots & (N-1)\overline{c_N}x_p^{N-2} \\ \vdots & \ddots & \vdots \\ \frac{(N-1)!}{(N-N_p)!}\overline{c_1}x_p^{N-N_1} + \cdots + (N_p-1)!\overline{c_{N-N_p+1}} & \ldots & \frac{(N-1)!}{(N-N_p)!}\overline{c_N}x_p^{N-N_p} \end{pmatrix}.$$

Después de llevar a cabo simplificaciones análogas se convierte en

$$\overline{c_N}^N \det\begin{pmatrix} 1 & x_1 & \ldots & x_1^{N_1} & \ldots & x_1^{N_p} & \ldots & x_1^{N-1} \\ 0 & 1 & \ldots & N_1x_1^{N_1-1} & \ldots & N_px_1^{N_p-1} & \ldots & (N-1)x_1^{N-2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & N_1! & \ldots & \frac{N_p!}{(N_p-N_1+1)!}x_1^{N_p-N_1+1} & \ldots & \frac{(N-1)!}{(N-N_1)!}x_1^{N-N_1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_p & \ldots & x_p^{N_1} & \ldots & x_p^{N_p} & \ldots & x_p^{N-1} \\ 0 & 1 & \ldots & N_1x_p^{N_1-1} & \ldots & N_px_p^{N_p-1} & \ldots & (N-1)x_p^{N-2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & N_p! & \ldots & \frac{(N-1)!}{(N-N_p)!}x_p^{N-N_p} \end{pmatrix}.$$

que podemos expresar en términos de derivadas parciales del determinante de Vandermonde:

$$\overline{c_N}^N \frac{\partial^{1+\cdots+N_1-1+\cdots+1+\cdots+N_p-1}}{\partial y_2 \cdots \partial y_{N_1}^{N_1-1} \cdots \partial y_{N-N_p+1} \cdots \partial y_N^{N_p-1}} \left(\det \begin{pmatrix} 1 & y_1 & \ldots & y_1^{N-1} \\ 1 & y_2 & \ldots & y_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & y_N & \ldots & y_N^{N-1} \end{pmatrix} \right)$$

con

$$y_1 = \cdots = y_{N_1} = x_1, \quad \ldots \quad y_{N-N_p+1} = \cdots = y_N = x_p$$

esto es

$$\overline{c_N}^N \frac{\partial^{1+\cdots+N_1-1+\cdots+1+\cdots+N_p-1}}{\partial y_2 \cdots \partial y_{N_1}^{N_1-1} \cdots \partial y_{N-N_p+1} \cdots \partial y_N^{N_p-1}} \left(\prod_{1 \le i < j \le N} (y_j - y_i) \right)$$

con

$$y_1 = \cdots = y_{N_1} = x_1, \quad \ldots \quad , y_{N-N_p+1} = \cdots = y_N = x_p.$$

Observemos que podemos expresar

$$\prod_{1 \le i < j \le N} (y_j - y_i)$$

$$= \prod_{1 \le i < j \le N_1} (y_j - y_i) \prod_{N_1+1 \le i < j \le N_1+N_2} (y_j - y_i)$$
$$\cdots \prod_{N-N_p+1 \le i < j \le N} (y_j - y_i) \prod_{(i,j) \text{ resto}} (y_j - y_i).$$

Al derivar con respecto a las variables y_i aparecen varios sumandos. Para que al realizar las sustituciones $y_1 = \cdots = y_{N_1} = x_1, \ldots, y_{N-N_p+1} = \cdots = y_N = x_p$ uno de estos sumandos no se anule, los p productos primeros de los $p+1$ del desarrollo anterior deben aparecer derivados con respecto a todas sus variables el máximo número posible de veces, pues en caso contrario al realizar la sustitución quedaría 0. De este modo el valor del determinante es

$$\overline{c_N}^N (N_1 - 1)! \ldots (N_p - 1)! \prod_{(i,j) \text{ resto}} (y_j - y_i)$$

con

$$y_1 = \cdots = y_{N_1} = x_1, \quad \ldots \quad y_{N-N_p+1} = \cdots = y_N = x_p,$$

que no es cero.

De nuevo, dada g cualquiera en H^2, podemos determinar unívocamente la correspondiente f en H^2.

Por otro lado, si λ es tal que alguna de las raíces $x_1, \ldots, x_N$ del polinomio $p_\lambda(z)$ tiene módulo 1, entonces debemos compensar las raíces del denominador en D y también las que están sobre el círculo unidad T, de modo que la función f definida por (4.4) pertenezca a H^2. Pero en este caso obtenemos un sistema lineal de ecuaciones con N incógnitas y más de N ecuaciones (recordemos que las raíces de $p_\lambda(z)$ son $x_1, \cdots, x_N$, $\frac{1}{\overline{x_1}}, \cdots, \frac{1}{\overline{x_N}}$). La matriz de coeficientes para este sistema es como antes (pero no una matriz cuadrada), y por tanto el rango de esta matriz es N.

Sin embargo, teniendo en cuenta que los términos independientes dependen del valor de g y algunas de sus derivadas en los puntos $x_1, \ldots, x_N$, es claro que una elección apropiada de la función g da una matriz aumentada para este sistema justo con rango $N+1$, y por tanto, para esta función g, la ecuación (4.4) no tiene una solución en H^2. Para este λ, el operador $J_0 - \lambda I$ no tiene una inversa acotada en H^2.

□

De este lema y del teorema de la aplicación abierta se deduce que para un λ dado, la ecuación (4.4) define un operador acotado en H^2 si y sólo si el polinomio

$$p_\lambda(z) = c_N z^{2N} + \cdots + c_1 z^{N+1} + (c_0 - \lambda) z^N + \overline{c_1} z^{N-1} + \cdots + \overline{c_N}$$

tiene exactamente N raíces dentro del disco unidad $D = \{z : |z| < 1\}$. Por tanto el espectro de J_0 es el conjunto de los λ para los cuales $p_\lambda(z)$ tiene al menos una raíz en $T = \{|z| = 1\}$.

Pero decir esto es lo mismo que decir que la ecuación

$$c_0 - \lambda = -c_N z^N - \overline{c_N} \frac{1}{z^N} - \cdots - c_1 z - \overline{c_1} \frac{1}{z} \tag{4.5}$$

tiene una solución en T. Si escribimos

$$h(z) = -c_N z^N - \overline{c_N}\frac{1}{z^N} - \cdots - c_1 z - \overline{c_1}\frac{1}{z},$$

entonces la ecuación (4.5) tiene una solución en T si y sólo si

$$\inf_{z\in T} h(z) \leq c_0 - \lambda \leq \sup_{z\in T} h(z).$$

Por tanto deducimos que el espectro de J_0 es el intervalo compacto

$$I = \left[c_0 - \sup_{z\in T} h(z),\ \ c_0 - \inf_{z\in T} h(z)\right].$$

Pero de acuerdo con la definición de los polinomios de Chebyshev de primera y segunda clase, concluimos que para $z = e^{ix} = \cos(x) + i\mathrm{sen}(x)$. $x \in [-\pi, \pi]$

$$\begin{aligned}
h(z) &= -c_N z^N - \overline{c_N}\frac{1}{z^N} - \cdots - c_1 z - \overline{c_1}\frac{1}{z} = \\
&= -\sum_{k=1}^{N} c_k(\cos(kx) + i\mathrm{sen}(kx)) - \sum_{k=1}^{N} \overline{c_k}(\cos(kx) - i\mathrm{sen}(kx)) = \\
&= -\sum_{k=1}^{N}(c_k + \overline{c_k})\cos(kx) - i\sum_{k=1}^{N}(c_k - \overline{c_k})\mathrm{sen}(kx) = \\
&= -\left(\sum_{k=1}^{N} 2\Re(c_k)T_k(\cos(x))\right) + \left(\sum_{k=1}^{N} 2\mathrm{sen}(x)\Im(c_k)U_{k-1}(\cos(x))\right) = \\
&= -t(x)
\end{aligned}$$

Y por tanto la prueba del Teorema 4.1 está acabada.

□

La hipótesis hecha sobre la convergencia de las sucesiones en la fórmula de recurrencia de $(2N + 1)$ términos impone algunas restricciones a las matrices límite (ver (4.3)) cuando planteamos el problema en términos matriciales.

Parece natural intentar extender el teorema 4.1 eliminando estas restricciones. Para ello habría que considerar la hipótesis, natural para el problema matricial, de la convergencia de las sucesiones de matrices $(A_n)_n$ y $(B_n)_n$ a matrices cuelesquiera A y B sin la condición (4.3).

Teniendo en cuenta los resultados de Van Assche y Geronimo ([**GV**]) sobre el soporte de medidas ortogonalizantes de polinomios verificando fórmulas de recurrencia con sucesiones de coeficientes finitamente periódicas, es posible conjeturar lo siguiente:

Conjetura.

Sea $(P_n)_n$ una sucesión de polinomios matriciales verificando una relación de recurrencia como (4.2), y supongamos que las sucesiones de matrices $(A_n)_n$ y $(B_n)_n$ convergen respectivamente a las matrices límite A y B. Sea μ la matriz de medidas definida positiva respecto de la cual los polinomios matriciales $(P_n)_n$ son ortonormales. Entonces el soporte de μ es una unión finita de intervalos y posiblemente un número finito de sucesiones que tienden a los extremos de los intervalos.

A continuación damos un ejemplo mostrando que en una hipotética extensión del teorema 4.1 no es posible esperar que el soporte de μ sea esencialmente un intervalo compacto.

Ejemplo.

Recordemos que se denomina clase de Nevai $M(a, b)$ al conjunto de medidas positivas con momentos finitos tales que los coeficientes de la relación de recurrencia de tres términos que verifican sus polinomios ortogonales

$$tp_n(t) = a_{n+1}p_{n+1}(t) + b_n p_n(t) + a_n p_{n-1}(t)$$

verifican

$$\lim_{n\to\infty} a_n = \frac{a}{2} \quad \text{y} \quad \lim_{n\to\infty} b_n = b.$$

Sean μ_1 y μ_2 dos medidas positivas en las clases de Nevai $M(a,b)$ y $M(a',b')$, respectivamente, donde $a \neq a'$, y $b \neq b'$. Llamemos $(p_{n,1})_n$, $(p_{n,2})_n$ a los polinomios ortonormales asociados a μ_1 y μ_2 respectivamente.

Definimos la matriz de medidas definida positiva ν como

$$\nu = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

Si escribimos las relaciones de recurrencia para $(p_{n,1})_n$, $(p_{n,2})_n$ como

$$a_{n+1,1}p_{n+1,1}(t) + b_{n,1}p_{n,1}(t) + a_{n,1}p_{n-1,1}(t) = tp_{n,1}(t),$$
$$a_{n+1,2}p_{n+1,2}(t) + b_{n,2}p_{n,2}(t) + a_{n,2}p_{n-1,2}(t) = tp_{n,2}(t),$$

entonces es claro que los polinomios matriciales ortogonales con respecto a ν son

$$P_k = \begin{pmatrix} p_{k,1} & 0 \\ 0 & p_{k,2} \end{pmatrix}$$

y que los coeficientes matriciales de recurrencia para estos polinomios matriciales son

$$A_k = \begin{pmatrix} a_{k,1} & 0 \\ 0 & a_{k,2} \end{pmatrix}, \qquad B_k = \begin{pmatrix} b_{k,1} & 0 \\ 0 & b_{k,2} \end{pmatrix}.$$

Estos coeficientes matriciales tienden a las matrices

$$A = \begin{pmatrix} \frac{a}{2} & 0 \\ 0 & \frac{a'}{2} \end{pmatrix}, \qquad B = \begin{pmatrix} b & 0 \\ 0 & b' \end{pmatrix}.$$

Pero el soporte de ν es $[b-a, b+a] \cup [b'-a', b'+a']$ y, quizás, dos sucesiones tendiendo a los extremos de estos intervalos. Tomando $b+a < b'-a'$, el soporte de ν no es un intervalo, aunque sí, esencialmente, una unión finita de intervalos.

CAPÍTULO 5

LOS ESPACIOS $\mathcal{L}_n^p$ DE UNA MATRIZ DE MEDIDAS

Definiremos en este capítulo los espacios de funciones donde de manera natural los polinomios matriciales serán usados para aproximar. Se trata de los espacios $\mathcal{L}_n^p(\mu)$, $1 \leq p \leq \infty$, $n \geq 1$ (n será el némero de filas de una función f en el espacio $\mathcal{L}_n^p(\mu)$). Salvo el espacio $\mathcal{L}_1^2(\mu)$, que fue definido en la década de los 60 (**[R]**, **[DS]**), el resto de los espacios $\mathcal{L}_n^p(\mu)$ no parece haber sido considerado hasta el momento. Probaremos que estos espacios son de Banach (Hilbert para $p = 2$) y que el dual de $\mathcal{L}_n^p(\mu)$ es $\mathcal{L}_n^q(\mu)$, con p y q exponentes conjugados. Para el caso $p = 2$, las demostraciones son nuevas y más simples que las ya conocidas (**[R]**, **[DS]**).

Completamos el capítulo con un resultado sobre densidad del espacio de polinomios matriciales en $\mathcal{L}_n^1(\mu)$. Este resultado se puede considerar una extensión del clásico debido a Naimark (**[N]**) para el caso escalar. No es sin embargo exhaustivo por lo que el caso general queda como problema abierto.

Sea $\mu = (\mu_{i,j})_{i,j=1}^N$ una matriz de medidas definida positiva. Teniendo en cuenta la desigualdad $0 \leq \mu \leq (\tau\mu)I$, se tiene que cada medida $\mu_{i,j}$ es absolutamente continua con respecto a la medida traza $\tau\mu$. Así, las derivadas de Radon-Nikodym $m_{i,j} = \dfrac{d\mu_{i,j}}{d\tau\mu}$ están bien definidas salvo en un conjunto de medida nula para la traza. La matriz de funciones

$$M = (m_{i,j})_{i,j=1}^N = \left(\frac{d\mu_{i,j}}{d\tau\mu}\right)_{1 \leq i,j \leq N}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

se denomina derivada de μ respecto a su traza. M está formada por funciones medibles e integrables con respecto a $\tau\mu$, y para cualquier conjunto de Borel A se tiene la igualdad

$$\int_A M(t)d\tau\mu(t) = \mu(A).$$

En ([R]) se prueba la desigualdad $0 \leq M(t) \leq I$, que se verifica en casi todo t $(\tau\mu)$.

Dado p, $1 \leq p < \infty$, y n un número natural fijo, definimos el espacio $\mathcal{L}_n^p(\mu)$ como el conjunto de matrices de funciones de tamaño $n \times N$ $f : \mathbb{R} \to M_{n\times N}(\mathbb{C})$ tales que $\tau(f(t)M(t)^{\frac{2}{p}}f(t)^*)^{\frac{1}{2}} \in L^p(\tau\mu)$, y definimos

$$\|f\|_{p,\mu} = \|\tau(f(t)M(t)^{\frac{2}{p}}f(t)^*)^{\frac{1}{2}}\|_{p,\tau\mu} = \left(\int_{\mathbb{R}} \tau(f(t)M(t)^{\frac{2}{p}}f(t)^*)^{\frac{p}{2}}d\tau\mu\right)^{\frac{1}{p}}.$$

y si $p = \infty$, decimos que $f \in \mathcal{L}_n^\infty(\mu)$ si $\tau(f(t)f(t)^*)^{\frac{1}{2}} \in L^\infty(\tau\mu)$ y entonces definimos

$$\|f\|_{\infty,\mu} = \|\tau(f(t)f(t)^*)^{\frac{1}{2}}\|_{\infty,\tau\mu}.$$

En ambos casos, identificamos $f(t)$ con $g(t)$ si $(f(t) - g(t))M(t)^{\frac{2}{p}} = \theta$.

De la definición de $\mathcal{L}_n^\infty(\mu)$ se deduce fácilmente que una función $f = (f_{i,j})_{1\leq i\leq n, 1\leq j\leq N}$ pertenece al espacio $\mathcal{L}_n^\infty(\mu)$ si y sólo si cada componente $f_{i,j}$ pertenece al espacio $L^\infty(\tau\mu)$.

Sin embargo, para los espacios $\mathcal{L}_n^p(\mu)$, con $p > 1$, esta propiedad no se cumple. Vemos a continuación un ejemplo de esto.

Ejemplo

Consideremos la matriz de medidas

$$\mu = \frac{1}{2}\begin{pmatrix} 1 & t-1 \\ t-1 & 1 \end{pmatrix} \chi_{[0,1]}(t)dm$$

donde dm denota la medida de Lebesgue. Por ser el determinante de $M(t)$ igual a $\frac{1}{4}t(2-t)$, que no es negativo en $[0,1]$, la matriz de medidas μ es definida positiva. Su traza es $\chi_{[0,1]}(t)dm$ es decir la medida de Lebesgue en $[0,1]$.

La descomposición en valores singulares de la matriz $M(t)$ en $[0,1]$ es

$$M(t) = \frac{1}{2}P\begin{pmatrix} 2-t & 0 \\ 0 & t \end{pmatrix}P^*, \quad \text{siendo} \quad P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

y por tanto

$$M(t)^{\frac{2}{p}} = \frac{1}{2^{\frac{2}{p}}}P\begin{pmatrix} (2-t)^{\frac{2}{p}} & 0 \\ 0 & t^{\frac{2}{p}} \end{pmatrix}P^*.$$

Para la función $f(t) = \left(\frac{1}{t^{\frac{1}{p}}}, \frac{1}{t^{\frac{1}{p}}}\right)$ tenemos que

$$\begin{aligned}
&\left[f(t)M(t)^{\frac{2}{p}}f(t)^*\right]^{\frac{p}{2}} \\
&= \left[\frac{1}{2^{\frac{2}{p}}}\frac{1}{2}\begin{pmatrix} \frac{1}{t^{\frac{1}{p}}} & \frac{1}{t^{\frac{1}{p}}} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} (2-t)^{\frac{2}{p}} & 0 \\ 0 & t^{\frac{2}{p}} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{t^{\frac{1}{p}}} \\ \frac{1}{t^{\frac{1}{p}}} \end{pmatrix}\right]^{\frac{p}{2}} \\
&= \left[\frac{1}{2}\frac{1}{2^{\frac{p}{2}}}\begin{pmatrix} 0 & \frac{2}{t^{\frac{1}{p}}} \end{pmatrix}\begin{pmatrix} (2-t)^{\frac{2}{p}} & 0 \\ 0 & t^{\frac{2}{p}} \end{pmatrix}\begin{pmatrix} 0 \\ \frac{2}{t^{\frac{1}{p}}} \end{pmatrix}\right]^{\frac{p}{2}} = 2^{\frac{p}{2}-1},
\end{aligned}$$

que por ser constante es integrable con respecto a $\tau\mu = \chi_{[0,1]}(t)dm$. Así, $f \in \mathcal{L}_1^p(\mu)$, para $p > 1$. Sin embargo, es claro que ninguna de las componentes de $f(t)$ pertenece al espacio $L^p(\tau\mu)$.

$\square$

Obsérvese que si $f = (f_{i,j})_{1\leq i\leq n, 1\leq j\leq N}$, y llamamos f_i a las filas de f, tenemos que para $p > 1$

$$\begin{aligned}
\tau(f(t)M(t)^{\frac{2}{p}}f(t)^*)^{\frac{p}{2}} &= \left(\sum_{i=1}^n f_i(t)M(t)^{\frac{2}{p}}f_i(t)^*\right)^{\frac{p}{2}} \\
&= \left(\sum_{i=1}^n \|f_i(t)M(t)^{\frac{1}{p}}\|_E^2\right)^{\frac{p}{2}},
\end{aligned}$$

que es la norma $\dfrac{2}{p}$ en $\mathbb{R}^N$ del vector

$$\left(\|f_1(t)M(t)^{\frac{1}{p}}\|_E^p, \ldots, \|f_n(t)M(t)^{\frac{1}{p}}\|_E^p\right)$$

y donde $\|\cdot\|_E$ denota la norma 2 en $\mathbb{R}^N$.

Usando ahora que las normas $\dfrac{2}{p}$ y 1 son equivalentes en $\mathbb{R}^N$, existen dos constantes positivas C_1 y C_2 tales que

$$\begin{aligned} C_1 \sum_{i=1}^{n} \|f_i(t)M(t)^{\frac{1}{p}}\|_E^p \leq & \left(\sum_{i=1}^{n} \|f_i(t)M(t)^{\frac{1}{p}}\|_E^2\right)^{\frac{p}{2}} \\ \leq & C_2 \sum_{i=1}^{n} \|f_i(t)M(t)^{\frac{1}{p}}\|_E^p, \end{aligned}$$

y de esta fórmula se obtiene integrando que

$$C_1 \sum_{i=1}^{n} \|f_i\|_{p,\mu}^p \leq \|f\|_{p,\mu}^p \leq C_2 \sum_{i=1}^{n} \|f_i\|_{p,\mu}^p.$$

Esta propiedad es interesante pues reduce el estudio de los espacios de matrices de funciones $\mathcal{L}_n^p(\mu)$ al de vectores de funciones $\mathcal{L}_1^p(\mu)$. Por ejemplo para estudiar cuestiones de densidad en los espacios $\mathcal{L}_n^p(\mu)$: dada una matriz de medidas definida positiva, el espacio de polinomios $\mathbb{C}^{n\times N}[x]$ es denso en $\mathcal{L}_n^p(\mu)$ si y sólo si el espacio de polinomios $\mathbb{C}^N[x]$ es denso en el espacio $\mathcal{L}_1^p(\mu)$, pues si es posible aproximar en norma p todas las filas de una función f por polinomios, entonces es posible aproximar en norma p la función f por un polinomio, y viceversa. Al espacio $\mathcal{L}_1^p(\mu)$ lo denotamos por $\mathcal{L}^p(\mu)$.

Probamos en primer lugar que $\|\cdot\|_{p,\mu}$ es una norma en $\mathcal{L}_n^p(\mu)$. Necesitamos varios lemas técnicos, que incluimos al no haberlos encontrado en la bibliografía:

Lema 5.1. *Si M es una matriz numérica semidefinida positiva, entonces define un producto escalar en $M_{n\times N}(\mathbb{C})$ dado por $(a,b)=\tau(aMb^*)$. Más aún, $\|a\|_M=(a,a)^{\frac{1}{2}}$ es una seminorma en $M_{n\times N}(\mathbb{C})$.*

Demostración

(a) Para probar la desigualdad de Cauchy-Schwartz, llamemos λ_i, $1\leq i\leq N$ a los autovalores de M y v_i, $1\leq i\leq N$ a una base de $\mathbb{C}^N$ formada por autovectores de M, tales que v_i es un autovector asociado a λ_i, $1\leq i\leq N$, y $v_iv_j^*=\delta_{i,j}$.

Dados a y b en $M_{n\times N}(\mathbb{C})$, podemos expresar

$$a=\begin{pmatrix}\sum_{i=1}^{N}a_{1,i}v_i\\ \vdots\\ \sum_{i=1}^{N}a_{n,i}v_i\end{pmatrix}\text{ y } b=\begin{pmatrix}\sum_{i=1}^{N}b_{1,i}v_i\\ \vdots\\ \sum_{i=1}^{N}b_{n,i}v_i\end{pmatrix}$$

donde $a_{j,i}$ y $b_{j,i}$ son números complejos. Tenemos entonces

$$\tau(aMb^*)=\sum_{j=1}^{n}\sum_{i=1}^{N}a_{j,i}\overline{b_{j,i}}\lambda_i$$

y

$$\|a\|_M\|b\|_M=\left(\sum_{j=1}^{n}\sum_{i=1}^{N}|a_{j,i}|^2\lambda_i\right)^{\frac{1}{2}}\left(\sum_{j=1}^{n}\sum_{i=1}^{N}|b_{j,i}|^2\lambda_i\right)^{\frac{1}{2}}$$

por tanto es suficiente aplicar la desigualdad de Cauchy-Schwartz a los vectores $(a_{1,1}\lambda_1^{\frac{1}{2}},\ldots,a_{1,N}\lambda_N^{\frac{1}{2}},\ldots\ldots,a_{n,1}\lambda_1^{\frac{1}{2}},\ldots,a_{n,N}\lambda_N^{\frac{1}{2}})$ y $(b_{1,1}\lambda_1^{\frac{1}{2}},\ldots,b_{1,N}\lambda_N^{\frac{1}{2}},\ldots\ldots,b_{n,1}\lambda_1^{\frac{1}{2}},\ldots,b_{n,N}\lambda_N^{\frac{1}{2}})$.

(b) Para la desigualdad triangular, tenemos

$$\|a+b\|^2=\|a\|^2+\|b\|^2+(a,b)+(b,a)$$

y

$$(\|a\|+\|b\|)^2=\|a\|^2+\|b\|^2+2\|a\|\|b\|,$$

por tanto basta observar que

$$(a,b)+(b,a) \leq |(a,b)+(b,a)| \leq |(a,b)|+|(b,a)| \leq 2\|a\|\|b\|.$$

$\square$

A partir de ahora, si $a \in M_{n\times N}(\mathbb{C})$, M es una matriz numérica semidefinida positiva y $p \geq 1$, usamos la siguiente notación: $\|a\|_{M,p} = \tau(aM^{\frac{2}{p}}a^*)^{\frac{1}{2}}$. Como consecuencia del lema 5.1, $\|\cdot\|_{M,p}$ es una seminorma en $M_{n\times N}(\mathbb{C})$.

Lema 5.2. *Si M es una matriz semidefinida positiva, $p \geq 1$ y $\dfrac{1}{p}+\dfrac{1}{q}=1$, entonces para cualesquiera vectores a, b en $M_{n\times N}(\mathbb{C})$ tenemos que*

$$|\tau(aMb^*)| \leq \|a\|_{M,p}\|b\|_{M,q}.$$

Demostración

La demostración es similar a (a) del lema anterior, aplicando la desigualdad de Cauchy-Schwartz a los vectores
$(a_{1,1}\lambda_1^{\frac{1}{p}},\ldots,a_{1,N}\lambda_N^{\frac{1}{p}},\ldots\ldots,a_{n,1}\lambda_1^{\frac{1}{p}},\ldots,a_{n,N}\lambda_N^{\frac{1}{p}})$
y $(b_{1,1}\lambda_1^{\frac{1}{q}},\ldots,b_{1,N}\lambda_N^{\frac{1}{q}},\ldots\ldots,b_{n,1}\lambda_1^{\frac{1}{q}},\ldots,b_{n,N}\lambda_N^{\frac{1}{q}})$.

$\square$

Lema 5.3. *(Desigualdad de Hölder). Si $1 \leq p < \infty$, $\frac{1}{p}+\frac{1}{q}=1$, $f \in \mathcal{L}_n^p(\mu)$ y $g \in \mathcal{L}_n^q(\mu)$, entonces $\tau(f(t)M(t)g(t)^*) \in L^1(\tau\mu)$, y*

$$\|\tau(fMg^*)\|_{1,\tau\mu} \leq \|f\|_{p,\mu}\|g\|_{q,\mu}.$$

Demostración

Si $f=\theta$ o $g=\theta$, la desigualdad es trivial. Así que podemos suponer $f\neq\theta$ y $g\neq\theta$.

Supongamos primero que $1<p<\infty$. Entonces $1<q<\infty$ y tenemos la desigualdad $cd \leq \dfrac{c^p}{p}+\dfrac{d^q}{q}$ para $c,d\in\mathbb{R}$, $c,d\geq 0$.

Para cualquier $t \in \mathbb{R}$, ponemos

$$c = \frac{\|f(t)\|_{M,p}}{\|f\|_{p,\mu}} \text{ y } d = \frac{\|g(t)\|_{M,q}}{\|g\|_{q,\mu}}.$$

De la desigualdad anterior tenemos que

$$\frac{\|f(t)\|_{M,p}\|g(t)\|_{M,q}}{\|f\|_{p,\mu}\|g\|_{q,\mu}} \leq \frac{\|f(t)\|_{M,p}^p}{p\|f\|_{p,\mu}^p} + \frac{\|g(t)\|_{M,q}^q}{q\|g\|_{q,\mu}^q},$$

y usando el lema 5.2,

$$|\tau(f(t)M(t)g(t)^*)| \leq \|f(t)\|_{M,p}\|g(t)\|_{M,q}.$$

De ambos hechos tenemos que

$$\frac{|\tau(f(t)M(t)g(t)^*)|}{\|f\|_{p,\mu}\|g\|_{q,\mu}} \leq \frac{\|f(t)\|_{M,p}^p}{p\|f\|_{p,\mu}^p} + \frac{\|g(t)\|_{M,q}^q}{q\|g\|_{q,\mu}^q}.$$

Ahora es claro que $\tau(fMg^*) \in L^1(\tau\mu)$, y, más aún, integrando, obtenemos

$$\frac{\|\tau(fMg^*)\|_{1,\tau\mu}}{\|f\|_{p,\mu}\|g\|_{q,\mu}} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

que da el resultado.

Para el caso $p = 1$, $q = \infty$, es suficiente observar que para cualquier número real t,

$$|\tau(f(t)M(t)g(t)^*)| \leq \|f(t)\|_{M,1}\|g(t)\|_E \leq \|f(t)\|_{M,1}\|g\|_{\infty,\tau\mu},$$

de lo cual deducimos que $\tau(fMg^*) \in L^1(\tau\mu)$ y que

$$\|\tau(fMg^*)\|_{1,\tau\mu} \leq \|f\|_{1,\mu}\|g\|_{\infty,\mu}.$$

□

Estamos ya en condiciones de probar que $\|\cdot\|_{p,\mu}$ es una norma en $\mathcal{L}_n^p(\mu)$.

Teorema 5.4. $\|\cdot\|_{p,\mu}$ *es una norma en* $\mathcal{L}_n^p(\mu)$.

Demostración

Ya sabemos que $\|f\|_p = 0$ si y sólo si $f = 0$, para $1 \leq p \leq \infty$, porque identificamos $f(t)$ con $g(t)$ si $((f(t) - g(t))M(t) = 0$.

Probemos ahora la desigualdad de Minkowski: si $f, g \in \mathcal{L}_n^p(\mu)$, entonces $f + g \in \mathcal{L}_n^p(\mu)$ y $\|f + g\|_{p,\mu} \leq \|f\|_{p,\mu} + \|g\|_{p,\mu}$.

Si $p = \infty$, de $\|f(t) + g(t)\|_E \leq \|f(t)\|_E + \|g(t)\|_E$ en casi todo t deducimos que $\|f + g\|_{\infty,\mu} \leq \|f\|_{\infty,\mu} + \|g\|_{\infty,\mu}$.

Y si $1 \leq p < \infty$, obsérvese que

$$\begin{aligned}\|f(t) + g(t)\|_{M,p}^p \leq& (\|f(t)\|_{M,p} + \|g(t)\|_{M,p})^p \\ \leq& 2^{p-1}(\|f(t)\|_{M,p}^p + \|g(t)\|_{M,p}^p) \in L^1(\tau\mu)\end{aligned}$$

y por tanto $f + g \in \mathcal{L}_n^p(\mu)$. Más aún,

$$\begin{aligned}\|f(t) + g(t)\|_{M,p}^p &= \|f(t) + g(t)\|_{M,p}\|f(t) + g(t)\|_{M,p}^{p-1} \\ &\leq \|f(t)\|_{M,p}\|f(t) + g(t)\|_{M,p}^{p-1} + \|g(t)\|_{M,p}\|f(t) + g(t)\|_{M,p}^{p-1}.\end{aligned}$$

Sabemos que $q = \dfrac{p}{p-1}$ y por tanto, $\|f(t) + g(t)\|_{M,p}^{p-1} \in L^q(\tau\mu)$. Ahora, usando la desigualdad de Hölder, obtenemos

$$\begin{aligned}&\int_{\mathbb{R}} \|f(t)\|_{M,p}\|f(t) + g(t)\|_{M,p}^{p-1} d\tau\mu \\ &\leq \|f\|_{p,\mu}\left(\int_{\mathbb{R}} (\|f(t) + g(t)\|_{M,p}^{p-1})^q d\tau\mu\right)^{\frac{1}{q}} = \|f\|_{p,\mu}\|f + g\|_{p,\mu}^{\frac{p}{q}}\end{aligned}$$

y análogamente,

$$\int_{\mathbb{R}} \|g(t)\|_{M,p}\|f(t) + g(t)\|_{M,p}^{p-1} d\tau\mu \leq \|g\|_{p,\mu}\|f + g\|_{p,\mu}^{\frac{p}{q}}.$$

Por tanto,

$$\|f + g\|_{p,\mu}^p \leq (\|f\|_{p,\mu} + \|g\|_{p,\mu})\|f + g\|_{p,\mu}^{\frac{p}{q}} = (\|f\|_{p,\mu} + \|g\|_{p,\mu})\|f + g\|_{p,\mu}^{p-1}$$

de lo cual deducimos que

$$\|f + g\|_{p,\mu} \leq \|f\|_{p,\mu} + \|g\|_{p,\mu}.$$

□

A continuación probamos que $\mathcal{L}_n^p(\mu)$ es un espacio de Banach.

Teorema 5.5. *Si $1 \leq p \leq \infty$, el espacio $\mathcal{L}_n^p(\mu)$ es un espacio de Banach, y si $p = 2$, $\mathcal{L}_n^2(\mu)$ es un espacio de Hilbert.*

Demostración

Dada $\mu = (\mu_{i,j})_{i,j=1}^N$ una matriz de medidas definida positiva y $M = (m_{i,j})_{i,j=1}^N$ su derivada de Radon-Nikodym , los autovalores $\lambda_i(t)$ asociados a $M(t)$ son reales y no negativos porque $M(t)$ es semidefinida positiva para cualquier número real t. Más aún, de $0 \leq M(t) \leq I$ deducimos que $0 \leq \lambda_i(t) \leq 1$, para cualquier autovalor $\lambda_i(t)$ de $M(t)$.

Llamemos $0 \leq \lambda_1(t) \leq \cdots \leq \lambda_N(t) \leq 1$ a los autovalores de $M(t)$ y $v_1(t), \ldots, v_N(t)$ a una base de $\mathbb{C}^N$ formada por autovectores de $M(t)$ asociados a $\lambda_1(t), \ldots, \lambda_N(t)$ respectivamente. Podemos elegir los autovectores verificando $v_i(t)v_j(t)^* = \delta_{i,j}$.

Dada una función $f : \mathbb{R} \longrightarrow M_{n \times N}(\mathbb{C})$, podemos expresarla de la forma siguiente:

$$f(t) = \begin{pmatrix} \displaystyle\sum_{i=1}^N \alpha_{1,i}(t)v_i(t) \\ \vdots \\ \displaystyle\sum_{i=1}^N \alpha_{n,i}(t)v_i(t) \end{pmatrix}$$

siendo $\alpha_{j,i}(t)$ funciones medibles, $1 \leq i \leq N$, $1 \leq j \leq n$.

Obsérvese que si para $M(t_0)$ los primeros k autovectores $v_1(t_0)$, $\ldots$, $v_k(t_0)$ son los asociados al autovalor 0, entonces los valores $\alpha_{j,1}(t_0), \ldots, \alpha_{j,k}(t_0)$ no son importantes porque a efectos de los espacios $\mathcal{L}_n^p(\mu)$, $1 \leq p \leq \infty$, $v_i(t_0)$ puede ser identificado con el vector nulo, $1 \leq i \leq k$.

Supongamos primero que $1 \leq p < \infty$. Según la definición, $f \in \mathcal{L}_n^p(\mu)$ si y sólo si $\tau(f(t)M(t)^{\frac{2}{p}} f(t)^*)^{\frac{1}{2}} \in L^p(\tau\mu)$.

Tenemos que

(5.1)

$$\tau(f(t)M(t)^{\frac{2}{p}}f(t)^*)^{\frac{p}{2}} = \left(\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)^{\frac{2}{p}}|\alpha_{j,i}(t)|^2|v_i(t)|_E^2\right)^{\frac{p}{2}}$$
$$= \left(\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)^{\frac{2}{p}}|\alpha_{j,i}(t)|^2\right)^{\frac{p}{2}},$$

que es la norma $\frac{2}{p}$ de la matriz de tamaño $n \times N$

$$\begin{pmatrix} \lambda_1(t)|\alpha_{1,1}(t)|^p & \dots & \lambda_N(t)|\alpha_{1,N}(t)|^p \\ \vdots & \ddots & \vdots \\ \lambda_1(t)|\alpha_{n,1}(t)|^p & \dots & \lambda_N(t)|\alpha_{n,N}(t)|^p \end{pmatrix}.$$

Por ser las normas $\frac{2}{p}$ y 1 equivalentes en $\mathbb{R}^{n\times N}$, existen dos constantes positivas C_1 y C_2 tales que

$$C_1\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)|\alpha_{j,i}(t)|^p \leq \left(\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)^{\frac{2}{p}}|\alpha_{j,i}(t)|^2\right)^{\frac{p}{2}}$$
$$\leq C_2\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)|\alpha_{j,i}(t)|^p$$

y por tanto, $f \in \mathcal{L}_n^p(\mu)$ es equivalente a $\lambda_i(t)^{\frac{1}{p}}\alpha_{j,i}(t) \in L^p(\tau\mu)$ $1 \leq i \leq N$, $1 \leq j \leq n$. Si consideramos el espacio

$$\Theta_p = \bigoplus_{i=1}^{N} L^p(\tau\mu\chi_{[\lambda_i\neq 0]}, \mathbb{C}^n)$$

cuyos elementos son matrices de funciones $(h_{j,i})_{1\leq j\leq n, 1\leq i\leq N}$ tales que $h_{j,i} \in L^p(\tau\mu\chi_{[\lambda_i\neq 0]})$, $\forall i,j$, podemos definir un operador $T_p : \mathcal{L}_n^p(\mu) \to \Theta_p$

de la forma siguiente: si

$$f(t) = \begin{pmatrix} \sum_{i=1}^{N} \alpha_{1,i}(t) v_i(t) \\ \vdots \\ \sum_{i=1}^{N} \alpha_{n,i}(t) v_i(t) \end{pmatrix},$$

definimos

$$T_p(f) = \begin{pmatrix} \lambda_1(t)^{\frac{1}{p}} \alpha_{1,1}(t) & \dots & \lambda_N(t)^{\frac{1}{p}} \alpha_{1,N}(t) \\ \vdots & \ddots & \vdots \\ \lambda_1(t)^{\frac{1}{p}} \alpha_{n,1}(t) & \dots & \lambda_N(t)^{\frac{1}{p}} \alpha_{n,N}(t) \end{pmatrix}.$$

Es claro que T_p es lineal y biyectivo entre $\mathcal{L}_n^p(\mu)$ y Θ_p , y si consideramos en el espacio Θ_p la norma $\|\cdot\|_{\Theta_p}$ dada por

$$\|h\|_{\Theta_p} = \|\|h(t)\|_E\|_{p,\tau\mu} = \left(\int_{\mathbb{R}} \tau(h(t)h(t)^*)^{\frac{p}{2}} d\tau\mu \right)^{\frac{1}{p}},$$

tenemos que T_p es una isometría:

$$\begin{aligned} \|T_p f\|_{\Theta_p} &= \left(\int_{\mathbb{R}} \|T_p f(t)\|_E^p d\tau\mu(t) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}} \left(\sum_{j=1}^{n} \sum_{i=1}^{N} \lambda_i(t)^{\frac{2}{p}} |\alpha_{j,i}(t)|^2 \right)^{\frac{p}{2}} d\tau\mu(t) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}} \tau(f(t)M(t)^{\frac{2}{p}} f(t)^*)^{\frac{p}{2}} d\tau\mu(t) \right)^{\frac{1}{p}} \\ &= \|f\|_{p,\mu}. \end{aligned}$$

Por tanto, tenemos que $\mathcal{L}_n^p(\mu)$ es isométrico a Θ_p.

Para $p = \infty$, de la expresión (5.1) deducimos que $f(t) \in \mathcal{L}_n^\infty(\mu)$ si y sólo si $|\alpha_{j,i}(t)|^2 \in L^\infty(\tau\mu)$, para $1 \le i \le N$, $1 \le j \le n$.

Ahora podemos considerar el espacio

$$\Theta_\infty = \bigoplus_{i=N}^{n} L^\infty(\tau\mu\chi_{[\lambda_i \neq 0]}, \mathbb{C}^n)$$

y definir el operador T_∞ como

$$T_\infty(f) = \begin{pmatrix} \alpha_{1,1}(t) & \dots & \alpha_{1,N}(t) \\ \vdots & \ddots & \vdots \\ \alpha_{n,1}(t) & \dots & \alpha_{n,N}(t) \end{pmatrix}.$$

que como antes es lineal y biyectivo. Si consideramos en el espacio Θ_∞ la norma $\|\cdot\|_{\Theta_\infty}$, dada por

$$\|h\|_{\Theta_\infty} = \|\|h(t)\|_E\|_{\tau\mu} = \sup_{t\in\mathbb{R}} \tau((h(t)h(t)^*)^{\frac{1}{2}},$$

tenemos que T_∞ es una isometría:

$$\begin{aligned} \|T_\infty f\|_{\Theta_\infty} &= \sup_{t\in\mathbb{R}} \|(T_\infty f)(t)\|_E \\ &= \sup_{t\in\mathbb{R}} \left(\sum_{j=1}^{n} \sum_{i=1}^{N} |\alpha_{j,i}(t)|^2 \right)^{\frac{1}{2}} \\ &= \sup_{t\in\mathbb{R}} \|\tau(f(t)f(t)^*)^{\frac{1}{2}}\| \\ &= \|f\|_{\infty,\mu}. \end{aligned}$$

por tanto $\mathcal{L}_n^\infty(\mu)$ es isométrico a Θ_∞.

Puesto que los espacios Θ_p son espacios de Banach para $1 \leq p \leq \infty$ (Hilbert si $p = 2$), tenemos que los espacios $\mathcal{L}_n^p(\mu)$ son espacios de Banach para $1 \leq p \leq \infty$ (Hilbert si $p = 2$).

□

Vemos a continuación como esta identificación permite calcular de una manera simple el dual de los espacios $\mathcal{L}_n^p(\mu)$.

Teorema 5.6. *Si $1 \leq p < \infty$ y q es su conjugado, es decir $\frac{1}{p} + \frac{1}{q} = 1$, $g \in \mathcal{L}^q_n(\mu)$ y definimos $\Lambda_g : \mathcal{L}^p_n(\mu) \longrightarrow \mathbb{C}$ como*

$$\Lambda_g f = \int_{\mathbb{R}} \tau(fMg^*)d\tau\mu, \quad \text{para todo} f \in \mathcal{L}^p_n(\mu) \tag{5.2}$$

entonces Λ es una aplicación lineal y continua, y $\|\Lambda\| \leq \|g\|_q$.

Además, si $\Lambda : \mathcal{L}^p_n(\mu) \longrightarrow \mathbb{C}$ es lineal y continua, entonces existe una única g en $\mathcal{L}^q_n(\mu)$ de modo que Λ se expresa como (5.2) y , $\|\Lambda\| = \|g\|_{q,\mu}$.

Si $\tau\mu$ es σ-finita, entonces este resultado es también cierto para $\mathcal{L}^\infty_n(\mu)$

Demostración

Es claro que Λ es lineal, y usando la desigualdad de Hölder

$$|\Lambda f| \leq \int_{\mathbb{R}} |fMg^*|d\tau\mu \leq \|f\|_{p,\mu}\|g\|_{q,\mu}.$$

Más aún,

$$\|\Lambda\| = \sup_{\|f\|_{p,\mu}\leq 1} |\Lambda f| \leq \|g\|_{q,\mu}.$$

Veamos ahora que para $1 \leq p < \infty$, cualquier aplicación lineal y continua de $\mathcal{L}^p_n(\mu)$ en $\mathbb{C}$ puede ser representada como en 5.2, para una cierta función única $g \in \mathcal{L}^q_n(\mu)$.

Supongamos primero que $1 < p < \infty$ y $\Lambda : \mathcal{L}^p_n(\mu) \to \mathbb{C}$ es una aplicación lineal y continua. T_p es una isometría, por tanto la composición $\Lambda T_p^{-1} : \Theta_p \to \mathbb{C}$ es también una aplicación lineal y continua. El espacio dual de Θ_p es isométrico a Θ_q, por tanto existe una función $G = (G_{j,i})_{1\leq j\leq n, 1\leq i\leq N} \in \Theta_q$ tal que para cualquier función $F = (F_{j,i})_{1\leq j\leq n,1\leq i\leq N} \in \Theta_p$ tenemos que

$$\Lambda T_p^{-1}(F) = \int_{\mathbb{R}} (\sum_{j=1}^{n}\sum_{i=1}^{N} F_{j,i}\overline{G_{j,i}})d\tau\mu,$$

y $\|G\|_{\Theta_q} = \|\Lambda T_p^{-1}\|$. Por tanto, para cualquier función $f \in \mathcal{L}^p_n(\mu)$, $f = (f_{j,i})_{1\leq j\leq n,1\leq i\leq N}$ tenemos que

$$\Lambda(f) = \Lambda T_p^{-1}(T_p f) = \int_{\mathbb{R}} (\sum_{j=1}^{n}\sum_{i=1}^{N} \lambda_i^{\frac{1}{p}} f_{j,i}\overline{G_{j,i}})d\tau\mu \tag{5.3}$$

Llamemos $g = T_q^{-1}(G) \in \mathcal{L}_n^q(\mu)$ es decir $G = T_q(g)$. (5.3) puede expresarse ahora como

$$\begin{aligned}\Lambda(f) &= \int_{\mathbb{R}} (\sum_{j=1}^{n}\sum_{i=1}^{N} \lambda_i^{\frac{1}{p}} f_{j,i}\overline{g_{j,i}}\lambda_i^{\frac{1}{q}})d\tau\mu \\ &= \int_{\mathbb{R}} (\sum_{j=1}^{n}\sum_{i=1}^{N} \lambda_i f_i\overline{g_i})d\tau\mu \\ &= \int_{\mathbb{R}} \tau(fMg^*)d\tau\mu,\end{aligned}$$

y además

$$\begin{aligned}\|\Lambda\| = \sup_{\|f\|_{p,\mu}=1} |\Lambda(f)| &= \sup_{\|F\|_{\Theta_p}=1} |\Lambda T_p^{-1}(F)| \\ &= \|\Lambda T_p^{-1}\| = \|G\|_{\Theta_q} = \|T_q(g)\|_{\Theta_q} = \|g\|_{q,\mu}.\end{aligned}$$

Así que el espacio dual de $\mathcal{L}_n^p(\mu)$ es isométrico a $\mathcal{L}_n^q(\mu)$.

Probamos ahora que si $\tau\mu$ es σ-finita, entonces el espacio dual de $\mathcal{L}_n^1(\mu)$ es isométrico a $\mathcal{L}_n^\infty(\mu)$. Supongamos que $\Lambda : \mathcal{L}_n^1(\mu) \to \mathbb{C}$ es una aplicación lineal y continua. Como en el caso anterior, $\Lambda T_1^{-1} : \Theta_1 \to \mathbb{C}$ es también una aplicación lineal y continua. Por ser $\tau\mu$ σ-finita, el espacio dual de Θ_1 es isométrico a Θ_∞ y por tanto existe una función $G = (G_{j,i})_{1\le j\le n, 1\le i\le N} \in \Theta_\infty$ tal que para cualquier función $F = (F_{j,i})_{1\le j\le n, 1\le i\le N} \in \Theta_1$ tenemos que

$$\Lambda T_1^{-1}(F) = \int_{\mathbb{R}} (\sum_{j=1}^{n}\sum_{i=1}^{N} F_{j,i}\overline{G_{j,i}})d\tau\mu,$$

y $\|G\|_{\Theta_\infty} = \|\Lambda T_1^{-1}\|$. Ahora, poniendo $g = T_\infty^{-1}(G) \in \mathcal{L}_n^\infty(\mu)$ tenemos como antes que para cualquier función $f \in \mathcal{L}_n^1(\mu)$

$$\Lambda(f) = \int_{\mathbb{R}} \tau(fMg^*)d\tau\mu$$

y además deducimos que

$$\begin{aligned}\|\Lambda\| = \sup_{\|f\|_{1,\mu}=1} |\Lambda(f)| &= \sup_{\|F\|_{\Theta_1}=1} |\Lambda T_1^{-1}(F)| \\ &= \|\Lambda T_1^{-1}\| = \|G\|_{\Theta_\infty} = \|T_1(g)\|_{\Theta_\infty} = \|g\|_{\infty,\mu}.\end{aligned}$$

Es decir, el espacio dual de $\mathcal{L}_n^1(\mu)$ es isométrico a $\mathcal{L}_n^\infty(\mu)$.

□

Finalizamos esta primera parte completando el estudio de los espacios $\mathcal{L}_n^p(\mu)$. Probaremos la continuidad de las inclusiones $\mathcal{L}_n^p(\mu) \subseteq \mathcal{L}_n^q(\mu)$ $(q > p)$ cuando $\tau\mu(\mathbb{R}) < \infty$ y el análogo al teorema de Marcinkiewicz.

Teorema 5.7. *Si μ es una matriz de medidas definida positiva tal que $\tau\mu(\mathbb{R}) < \infty$, entonces*

(a) $\mathcal{L}_n^\infty(\mu) \subseteq \mathcal{L}_n^1(\mu)$ y además la inclusión es continua, pues para toda función f de $\mathcal{L}_n^\infty(\mu)$ se tiene que

$$\|f\|_{1,\mu} \leq \tau\mu(\mathbb{R})\|f\|_{\infty,\mu}.$$

(b) Si $1 \leq p < q$, $\mathcal{L}_n^p(\mu) \subseteq \mathcal{L}_n^q(\mu)$ y también en este caso la inclusión es continua, pues para toda función f de $\mathcal{L}_n^p(\mu)$ se tiene que

$$\|f\|_{q,\mu} \leq \tau\mu(\mathbb{R})^{\frac{1}{q}-\frac{1}{p}}\|f\|_{p,\mu}.$$

Demostración

(a) Teniendo en cuenta que $0 \leq M^2 \leq M \leq I$ tenemos que dada una función f en $\mathcal{L}_n^\infty(\mu)$,

$$\begin{aligned}\|f\|_1 &= \int_{\mathbb{R}} \tau(fM^2f^*)^{\frac{1}{2}}\, d\tau\mu \\ &\leq \int_{\mathbb{R}} \tau(ff^*)^{\frac{1}{2}}\, d\tau\mu \\ &\leq \int_{\mathbb{R}} \|f\|_{\infty,\mu} \\ &= \tau\mu(\mathbb{R})\|f\|_{\infty,\mu}.\end{aligned}$$

(b) De $p > q$ y $0 \le M \le I$ se deduce que $0 \le M^{\frac{1}{q}} \le M^{\frac{1}{p}}$. Dada f una función en $\mathcal{L}_n^p(\mu)$, tenemos que la función $\tau(fM^{\frac{2}{p}}f^*)^{\frac{q}{2}} \in L^{\frac{p}{q}}(\tau\mu)$, y puesto que $\tau\mu(\mathbb{R}) < \infty$, la función constante $1 \in L^{\frac{p}{q-p}}(\tau\mu)$ ($\frac{p}{p-q}$ es el exponente conjugado de $\frac{p}{q}$). De la desigualdad para M y usando la desigualdad de Hölder en el caso unidimensional se deduce que

$$\begin{aligned}
\|f\|_{q,\mu}^q &= \int_{\mathbb{R}} \tau(fM^{\frac{2}{q}}f^*)^{\frac{q}{2}} d\tau\mu \\
&\le \int_{\mathbb{R}} \tau(fM^{\frac{2}{p}}f^*)^{\frac{q}{2}} d\tau\mu \\
&\le \left(\int_{\mathbb{R}} d\tau\mu\right)^{\frac{p-q}{p}} \left(\int_{\mathbb{R}} \tau(fM^{\frac{2}{p}}f^*)^{\frac{p}{2}} d\tau\mu\right)^{\frac{q}{p}} \\
&= \tau\mu(\mathbb{R})^{\frac{p-q}{p}} \|f\|_{p,\mu}^q.
\end{aligned}$$

Tomando ahora raíz q-ésima se obtiene la desigualdad buscada.

□

En el caso general, es decir, si $\tau\mu(\mathbb{R}) = \infty$, las inclusiones anteriores no son ciertas.

Teorema 5.8. *Si $\mu = (\mu_{i,j})_{1\le i,j\le N}$ es una matriz de medidas definida positiva y $1 \le p < q \le \infty$, entonces $\mathcal{L}_n^p(\mu) \cap \mathcal{L}_n^q(\mu) \subseteq \mathcal{L}_n^r(\mu)$, para todo r tal que $p < r < q$. Además, existe una constante positiva C tal que para toda función $f \in \mathcal{L}_n^p(\mu) \cap \mathcal{L}_n^q(\mu)$ se tiene que*

$$\|f\|_{r,\mu} \le C\|f\|_{p,\mu}^{\theta}\|f\|_{q,\mu}^{1-\theta},$$

siendo θ el número real entre 0 y 1 que verifica $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$.

Demostración

a) supongamos primero que $q < \infty$. Dada $f \in \mathcal{L}_n^p(\mu) \cap \mathcal{L}_n^q(\mu)$ la expresamos como

$$f(t) = \begin{pmatrix} \sum_{i=1}^{N} \alpha_{1,i}(t)v_i(t) \\ \vdots \\ \sum_{i=1}^{N} \alpha_{n,i}(t)v_i(t) \end{pmatrix}.$$

Tenemos que

$$\|f\|_{r,\mu}^r = \int_{\mathbb{R}} \tau(f(t)M(t)^{\frac{2}{r}} f(t)^*)^{\frac{r}{2}} d\tau\mu(t).$$

Obsérvese que

$$\tau(f(t)M(t)^{\frac{2}{r}} f(t)^*)^{\frac{r}{2}} = \left(\sum_{j=1}^{n}\sum_{i=1}^{N} \lambda_i(t)^{\frac{2}{r}} |\alpha_{j,i}(t)|^2\right)^{\frac{r}{2}},$$

que es la norma $\frac{2}{r}$ de la matriz de tamaño $n \times N$

$$(\lambda_i(t)|\alpha_{j,i}(t)|^r)_{1\leq j\leq n, 1\leq i\leq N}.$$

Por ser las normas $\frac{2}{r}$ y 1 equivalentes en $\mathbb{R}^{n\times N}$ existe una constante positiva C_1 de modo que

$$\tau(f(t)M(t)^{\frac{2}{r}} f(t)^*)^{\frac{r}{2}} \leq C_1 \sum_{j=1}^{n}\sum_{i=1}^{N} \lambda_i(t)|\alpha_{j,i}(t)|^r. \tag{5.4}$$

A continuación expresamos cada sumando $\lambda_i(t)|\alpha_{j,i}(t)|^r$ como

$$\lambda_i(t)|\alpha_{j,i}(t)|^r = \lambda_i(t)^{\frac{\theta r}{p}} |\alpha_{j,i}(t)|^{\theta r} \lambda_i(t)^{\frac{(1-\theta)r}{q}} |\alpha_{j,i}(t)|^{(1-\theta)r} \tag{5.5}$$

siendo θ el número real entre 0 y 1 que verifica $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$. Consideramos ahora $p' = \frac{p}{\theta r}$ y $q' = \frac{q}{(1-\theta)r}$, que son conjugados:

$$\frac{1}{p'} + \frac{1}{q'} = \frac{\theta r}{p} + \frac{(1-\theta)r}{q} = r\left(\frac{\theta}{p} + \frac{(1-\theta)}{q}\right) = r\frac{1}{r} = 1.$$

$\lambda_i(t)^{\frac{\theta r}{p}} |\alpha_{j,i}(t)|^{\theta r}$ pertenece al espacio $L^{p'}(\tau\mu)$ porque

$$\left(\lambda_i(t)^{\frac{\theta r}{p}} |\alpha_{j,i}(t)|^{\theta r}\right)^{p'} = \lambda_i(t)|\alpha_{j,i}(t)|^p$$

que pertenece a $L^1(\tau\mu)$ por ser $f \in \mathcal{L}_n^p(\mu)$. Análogamente se tiene que $\lambda_i(t)^{\frac{(1-\theta)r}{q}} |\alpha_{j,i}(t)|^{(1-\theta)r}$ pertenece al espacio $L^{q'}(\tau\mu)$ por ser $f \in \mathcal{L}_n^q(\mu)$.

Así, integrando en (5.4) y usando la desigualdad de Hölder se obtiene

(5.6) $\|f\|_{r,\mu}^r \leq$

$$C_1 \sum_{j=1}^{n} \sum_{i=1}^{N} \left(\int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^p d\tau\mu(t) \right)^{\frac{\theta r}{p}} \left(\int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^q d\tau\mu(t) \right)^{\frac{(1-\theta)r}{q}}$$

que es el producto escalar de las dos matrices

$$\left(\left(\int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^p d\tau\mu(t) \right)^{\frac{\theta r}{p}} \right)_{1 \leq j \leq n, 1 \leq i \leq N}$$

y

$$\left(\left(\int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^q d\tau\mu(t) \right)^{\frac{(1-\theta)r}{q}} \right)_{1 \leq j \leq n, 1 \leq i \leq N}$$

y por tanto menor o igual que el producto de la norma $p' = \frac{p}{\theta r}$ de la primera por la norma $q' = \frac{q}{(1-\theta)r}$ de la segunda, por ser p' y q' conjugados. Así:

$$\|f\|_{r,\mu}^r \leq C_1 \left(\sum_{j=1}^{n} \sum_{i=1}^{N} \int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^p d\tau\mu(t) \right)^{\frac{\theta r}{p}} \cdot$$

$$\cdot \left(\sum_{j=1}^{n} \sum_{i=1}^{N} \int_{\mathbb{R}} \lambda_i(t) |\alpha_{j,i}(t)|^q d\tau\mu(t) \right)^{\frac{(1-\theta)r}{q}}$$

$$= \left(\int_{\mathbb{R}} \left(\sum_{j=1}^{n} \sum_{i=1}^{N} \lambda_i(t) |\alpha_{j,i}(t)|^p \right) d\tau\mu(t) \right)^{\frac{\theta r}{p}} \cdot$$

$$\cdot \left(\int_{\mathbb{R}} \left(\sum_{j=1}^{n} \sum_{i=1}^{N} \lambda_i(t) |\alpha_{j,i}(t)|^q \right) d\tau\mu(t) \right)^{\frac{(1-\theta)r}{q}} .$$

$\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)|\alpha_{j,i}(t)|^p$ es la norma 1 de la matriz

$$(\lambda_i(t)|\alpha_{j,i}(t)|^p)_{1\leq j\leq n,1\leq i\leq N},$$

que es menor o igual que una constante positiva por la norma $\frac{2}{p}$ de la matriz. Análogamente para el segundo multiplicando, que es menor o igual que una constante positiva por la norma $\frac{2}{q}$ de la matriz $(\lambda_i(t)|\alpha_{j,i}(t)|^q)_{1\leq j\leq n,1\leq i\leq N}$. Por tanto

$$\|f\|_{r,\mu}^r \leq C_2\left(\int_{\mathbb{R}}\left(\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)^{\frac{2}{p}}|\alpha_{j,i}(t)|^2\right)^{\frac{p}{2}}d\tau\mu(t)\right)^{\frac{\theta r}{p}}\cdot$$

$$\cdot\left(\int_{\mathbb{R}}\left(\sum_{j=1}^{n}\sum_{i=1}^{N}\lambda_i(t)^{\frac{2}{q}}|\alpha_{j,i}(t)|^2\right)^{\frac{q}{2}}d\tau\mu(t)\right)^{\frac{(1-\theta)r}{q}}$$

$$= C_2\|f\|_{p,\mu}^{\theta r}\|f\|_{q,\mu}^{(1-\theta)r}$$

de lo cual se deduce que

$$\|f\|_{r,\mu} \leq C_3\|f\|_{p,\mu}^{\theta}\|f\|_{q,\mu}^{1-\theta}.$$

b) En el caso $q=\infty$ se tiene que $\frac{1}{r}=\frac{\theta}{q}$ y por tanto $\theta=\frac{p}{r}$, $p'=1$ y $q'=\infty$. La descomposición (5.5) es ahora

$$\lambda_i(t)|\alpha_{j,i}(t)|^r = \lambda_i(t)|\alpha_{j,i}(t)|^p|\alpha_{j,i}(t)|^{r-p}.$$

$\lambda_i(t)|\alpha_{j,i}(t)|^p$ está en $L^1(\tau\mu)$ por ser $f\in\mathcal{L}_n^p(\mu)$, y $|\alpha_{j,i}(t)|^{r-p}$ está en $L^\infty(\tau\mu)$ por ser $f\in\mathcal{L}_n^\infty(\mu)$.

La desigualdad (5.6) queda ahora

$$\|f\|_{r,\mu}^r \leq C_1\sum_{j=1}^{n}\sum_{i=1}^{N}\left(\int_{\mathbb{R}}\lambda_i(t)|\alpha_{j,i}(t)|^p d\tau\mu(t)\right)\sup_{t\in\mathbb{R}}|\alpha_{j,i}(t)|^{r-p}.$$

Razonando como antes se obtiene

$$\|f\|_{r,\mu}^r \leq C_2 \|f\|_{p,\mu}^p \sup_{t\in\mathbb{R}} \max_{1\leq j\leq n, 1\leq i\leq N} |\alpha_{j,i}(t)|^{r-p}.$$

Además, usando de nuevo la equivalencia de las normas en $\mathbb{R}^{n\times N}$,

$$\begin{aligned}\sup_{t\in\mathbb{R}} \max_{1\leq j\leq n, 1\leq i\leq N} |\alpha_{j,i}(t)|^{r-p} \leq & C_3 \left(\sup_{t\in\mathbb{R}} \left(\sum_{j=1}^{n} \sum_{i=1}^{N} |\alpha_{j,i}(t)|^2 \right)^{\frac{1}{2}} \right)^{r-p} \\ = & C_3 \|f\|_{\infty,\mu}^{r-p}.\end{aligned}$$

Tomando raíz r-ésima y puesto que $\frac{p}{r} = \theta$ y $\frac{r-p}{r} = 1 - \frac{p}{r} = 1 - \theta$ se obtiene que

$$\|f\|_{r,\mu} \leq C_4 \|f\|_{p,\mu}^{\theta} \|f\|_{\infty,\mu}^{1-\theta}.$$

□

En la segunda parte de este capítulo extendemos el teorema de Naimark sobre densidad de polinomios en $\mathcal{L}_n^1(\mu)$. El teorema de Naimark establece que los polinomios son densos en $L^1(\mu)$, μ medida positiva con momentos finitos, si y sólo si μ es un punto extremal del conjunto convexo $\{\nu \geq 0 : \int t^n d\nu = \int t^n d\mu, \ n \geq 0\}$.

Sea pues μ una matriz de medidas definida positiva y con momentos finitos para que los polinomios estén incluidos en $\mathcal{L}_n^1(\mu)$.

En el caso escalar la traza de μ coincide con μ y entonces la matriz de derivadas de Radon-Nikodym es $M(t) = 1$. Obsérvese que los autovalores de $M(t)$ se reducen en este caso a 1. En el caso matricial los autovalores de $M(t)$ tienen una estructura más complicada. Para extender el teorema de Naimark al caso matricial necesitamos como hipótesis adicional que los autovalores de $M(t)$, como ocurre en el caso escalar, no se aproximen a 0:

Teorema 5.9. *Dada una medida definida positiva $\mu = M(t)d\tau\mu$, llamamos $0 \leq \lambda_1(t) \leq \cdots \leq \lambda_N(t) \leq 1$ a los autovalores de la matriz*

$M(t)$ ordenados en orden creciente. Si existe $\epsilon > 0$ tal que $\lambda_1(t) \geq \epsilon$. para todo t del soporte de μ, entonces son equivalentes:

a) Los polinomios son densos en el espacio $\mathcal{L}_n^1(\mu)$.

b) μ es una medida extremal del conjunto $[\mu]$ (véase la definición de $[\mu]$ en el capítulo 2).

Demostración Basta probar el teorema para $n = 1$.

Para probar que a) implica b), supongamos que los polinomios no son densos en $\mathcal{L}^1(\mu)$ y probemos que μ no es extremal en $[\mu]$.

En tal caso, podemos encontrar un funcional no nulo Λ del espacio dual de $\mathcal{L}^1(\mu)$, por ejemplo de norma 1, de modo que $\Lambda(p) = 0$ para cualquier polinomio p. Por el teorema de dualidad de $\mathcal{L}^1(\mu)$ podemos representar este funcional con una función $g = (g_1, \ldots, g_N)$ de $\mathcal{L}^\infty(\mu)$ y de norma 1 como

$$\Lambda(f) = \int_{\mathbb{R}} f(t)M(t)g(t)^* d\tau\mu(t), \text{ para toda función } f \in \mathcal{L}^1(\mu).$$

Por ser $\Lambda(p) = 0$ para cualquier polinomio p deducimos que las medidas que aparecen en $M(t)g(t)^* d\tau\mu$ tienen todas momentos nulos, y además alguna de ellas no es cero pues de serlo todas tendríamos que $\Lambda \equiv 0$.

Escribamos $d(t)d\tau\mu$ para la parte real o imaginaria no nula de una de estas medidas, y supongamos que es la obtenida al multiplicar la fila i_0 de $M(t)$ por el vector $g(t)^*$. Obsérvese que

$$|d(t)| \leq |m_{i_0,1}(t)\overline{g_1(t)} + \cdots + m_{i_0,N}(t)\overline{g_N(t)}| \leq N$$

porque $|m_{i,j}| \leq 1$ y $|g_i| \leq 1$, para $0 \leq i, j \leq N$.

Descomponemos μ de la forma siguiente:

$$\begin{aligned}\mu =& \frac{1}{2}\left\{\left(M(t) + \frac{\epsilon}{2N}d(t)I\right)d\tau\mu + \left(M(t) - \frac{\epsilon}{2N}d(t)I\right)d\tau\mu\right\}\\ =& \frac{1}{2}(\mu_1 + \mu_2)\end{aligned}$$

Por la construcción, es claro que μ_1 y μ_2 tienen ambas los mismos momentos que μ. Veamos ahora que estas dos matrices de medidas son definidas positivas. Para ello basta ver que en cada punto t del soporte de μ las matrices numéricas $M(t) \pm \frac{\epsilon}{2N}d(t)I$ son semidefinidas positivas.

Dado un vector c cualquiera en $\mathbb{C}^N$ lo expresamos como $c = \sum_{i=1}^{N} c_i v_i$, donde $v_1, \ldots, v_n$ es una base ortonormalizada de autovectores de $\mathbb{C}^N$, v_i asociado a λ_i. Se tiene entonces que

$$\begin{aligned} c\left(M(t) \pm \frac{\epsilon}{2N}d(t)I\right)c^* =& \sum_{i=1}^{N} |c_i|^2 \lambda_i(t) \pm \frac{\epsilon}{2N} d(t) \sum_{i=1}^{N} |c_i|^2 \\ \geq& \epsilon \sum_{i=1}^{N} |c_i|^2 - \frac{\epsilon}{2N} N \sum_{i=1}^{N} |c_i|^2 \\ =& \frac{\epsilon}{2} \sum_{i=1}^{N} |c_i|^2 \geq 0 \end{aligned}$$

y por tanto ambas medidas son definidas positivas.

Para probar ahora que b) implica a) supongamos que μ no es extremal en el conjunto $[\mu]$ y probemos que entonces los polinomios no son densos en el espacio $\mathcal{L}^1(\mu)$.

Supongamos pues que μ admite la descomposición $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ donde μ_1 y μ_2 son dos matrices de medidas del conjunto $[\mu]$ ninguna de ellas igual a μ.

Puesto que $\mu_1 \leq 2\mu \leq 2\tau\mu I$ y $\mu_2 \leq 2\mu \leq 2\tau\mu I$ (estas desigualdades se entienden para cada conjunto de Borel), podemos expresar $\mu_1 = M_1(t)d\tau\mu$ y $\mu_2 = M_2(t)d\tau\mu$, siendo las matrices $M_1(t)$ y $M_2(t)$ las derivadas de Radon-Nikodym de las medidas μ_1 y μ_2 respecto de $\tau\mu$ respectivamente, es decir

$$\mu = M(t)d\tau\mu = \frac{1}{2}(M_1(t) + M_2(t))d\tau\mu.$$

Puesto que μ no es igual que μ_1 tenemos que $M(t)$ no es igual que $M_1(t)$ y por tanto existe un vector $e_{i_0} = (0, \ldots, 1, \ldots, 0)$, donde el 1 está en el lugar i_0 de modo que $M(t)e_{i_0}^* \neq M_1(t)e_{i_0}^*$.

Definimos los operadores T y T_1 sobre $\mathcal{L}^1(\mu)$ como

$$T(f) = \int_{\mathbb{R}} f(t)M(t)e_{i_0}^* d\tau\mu \quad \text{y} \quad T_1(f) = \int_{\mathbb{R}} f(t)M_1(t)e_{i_0}^* d\tau\mu.$$

Es claro que ambos son lineales y que T es continuo porque e_{i_0} es un elemento de $\mathcal{L}^\infty(\mu)$.

Para ver que T_1 es también continuo, observemos primero que por ser $M_1 \leq 2M \leq 2I$ se tiene que $M_1^2 \leq 4I$. Además, por ser $\lambda_1 \geq \epsilon$ se tiene que $M \geq \epsilon I$ y por tanto $M^2 \geq \epsilon^2 I$. Por tanto si llamamos $C = \frac{4}{\epsilon^2}$ tenemos que $CM^2 \geq 4I \geq M_1^2$. De esto y del lema 5.2 tenemos que

$$\begin{aligned}
|T_1(f)| &= \left| \int f(t)M_1(t)e_{i_0}^* d\tau\mu \right| \\
&\leq \int |f(t)M_1(t)e_{i_0}^*| d\tau\mu \\
&\leq \int (f(t)M_1^2(t)f(t)^*)^{\frac{1}{2}} (e_{i_0}M_1^2(t)e_{i_0}^*)^{\frac{1}{2}} d\tau\mu \\
&\leq \sqrt{2}C^{\frac{1}{2}} \int (f(t)M_1^2(t)f(t)^*)^{\frac{1}{2}} d\tau\mu \\
&= \sqrt{2}C\|f\|_{1,\mu}.
\end{aligned}$$

Por tanto si definimos $R = T - T_1$, R es un operador no nulo sobre $\mathcal{L}^1(\mu)$ y que se anula sobre los polinomios, por tanto los polinomios no son densos en $\mathcal{L}^1(\mu)$, como queríamos probar.

□

La existencia de puntos extremales está garantizada pues como se probó en el capítulo 2 el conjunto $[\mu]$ es compacto y convexo. La cuestión general de si son estos puntos extremales los que tienen el espacio de polinomios denso en sus espacios $\mathcal{L}^1(\mu)$, sin condiciones sobre la estructura de los autovalores de su matriz de derivadas de Radon-Nykodym, queda por tanto abierta.

CAPÍTULO 6

CUESTIONES DE DENSIDAD PARA EL PROBLEMA DE MOMENTOS MATRICIAL TRUNCADO

Sea $\nu = (\nu_{i,j})_{1\le i,j\le N} \in \mathcal{M}_N^*$, es decir ν es una matriz de medidas definida positiva con todos sus momentos finitos, y no degenerada (ver Preliminares, pag. 3 para la definición de no degenerada).

Para $n \ge 0$ denotamos por V_n el conjunto de matrices de medidas definidas positivas cuyos momentos hasta grado n inclusive son finitos (denotamos esta clase por $\mathcal{M}_{N,n}$) y coinciden con los de ν, es decir:

$$V_n = \Big\{\mu = (\mu_{i,j})_{1\le i,j\le N} \in \mathcal{M}_{N,n} : \int_{\mathbb{R}} t^p d\mu_{i,j} = \int_{\mathbb{R}} t^p d\nu_{i,j}, \\ \text{para } 0 \le p \le n \text{ y } 1 \le i,j \le N\Big\}.$$

Como es habitual denotamos por $(P_n)_n$ la sucesión de polinomios matriciales ortonormales asociados a ν, siendo $P_0(t) = I$ y $P_{-1}(t) = \theta$.

Estos polinomios satisfacen una relación de recurrencia de la forma

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \quad n \ge 0 \tag{6.1}$$

siendo $\det A_n \ne 0$ y $B_n^* = B_n$.

Debemos señalar que dado que sólo imponemos que las matrices de V_{2n-2} tengan momentos finitos hasta el grado $2n-2$, sólo tendremos asegurado, para $\mu \in V_{2n-2}$, que los polinomios de grado menor o igual que $n-1$ pertenecen al correspondiente $\mathcal{L}_N^2(\mu)$. En cualquier caso, los polinomios $(P_k)_{k=0,\dots,n-1}$ son ortonormales para cualquier medida de V_{2n-2}.

Se denota por $Q_n(t)$ la correspondiente sucesión de polinomios de segunda clase, que también verifican la relación de recurrencia (6.1), con condiciones iniciales $Q_0(t) = \theta$, $Q_1(t) = A_1^{-1}$.

En este capítulo caracterizamos las matrices μ de V_{2n-2} tales que los polinomios de grado menor o igual que $n-1$ son densos en $\mathcal{L}^2_N(\mu)$: μ es una matriz extremal (en el sentido de la convexidad) de V_{n-1}. También probamos un teorema de estructura para estas matrices de medidas: son las matrices de medidas discretas con masas en los ceros del polinomio $(P_n(\lambda)-AP_{n-1}(\lambda))$ y pesos los residuos de $(P_n(\lambda)-AP_{n-1}(\lambda))^{-1}(Q_n(\lambda)-AQ_{n-1}(\lambda))$ en dichos ceros, variando la matriz numérica A entre aquellas que verifican $A_nA = A^*A_n^*$

Antes de abordar el teorema necesitamos algunas fórmulas y resultados técnicos que agrupamos en dos lemas:

Lema 6.1.

Supongamos que $X(x) = (X_n(x))_{n=0}^{\infty}$ e $Y(y) = (Y_n(x))_{n=0}^{\infty}$ $(x, y \in \mathbb{C})$ son dos soluciones arbitrarias de la sucesión de recurrencia matricial de tres términos

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \quad n \geq 0$$

siendo $P_0(t) = I$ y $P_{-1}(t) = \theta$, $\det A_n \neq 0$ y $B_n^* = B_n$, y definamos

$$\omega_n = X_n^*(\overline{x})A_{n+1}Y_{n+1}(y) - X_{n+1}^*(\overline{x})A_{n+1}^*Y_n(y). \tag{6.2}$$

Se tiene entonces que

$$\omega_n = \omega_{n-1} + (y-\overline{x})X_n^*(\overline{x})Y_n(y) \tag{6.3}$$

y como consecuencia que

$$(y-\overline{x})\sum_{k=1}^{n-1} X_k^*(\overline{x})Y_k(y) = \omega_{n-1} - \omega_0. \tag{6.4}$$

Como caso particular de esta fórmula se deduce la fórmula de Christoffel-Darboux

(6.5)

$$P^*_{n-1}(u)A_nP_n(v) - P^*_n(u)A^*_nP_{n-1}(v) = (v-u)\sum_{k=0}^{n-1} P^*_k(u)P_k(v), \quad u, v \in \mathbb{C}.$$

con sus casos particulares

$$P^*_{n-1}(z)A_nP_n(z) - P^*_n(z)A^*_nP_{n-1}(z) = \theta, \quad z \in \mathbb{C} \tag{6.7}$$

y

$$P^*_{n-1}(z)A_nP'_n(z) - P^*_n(z)A^*_nP'_{n-1}(z) = \sum_{k=0}^{n-1} P^*_k(z)P_k(z), \quad z \in \mathbb{C}. \tag{6.8}$$

y la fórmula de Green

(6.9)

$$P^*_{n-1}(u)A_nQ_n(v) - P^*_n(u)A^*_nQ_{n-1}(v) = I + (v-u)\sum_{k=0}^{n-1} P^*_k(u)Q_k(v), \qquad u, v \in \mathbb{C}$$

Demostración

De la relación de recurrencia (6.1) se tiene que

$$A_{n+1}Y_{n+1}(y) = (y - B_n)Y_n(y) - A^*_nY_{n-1}(y)$$

y

$$X^*_{n+1}(\overline{x})A^*_{n+1} = X^*_n(\overline{x})(\overline{x} - B^*_n) - X^*_{n-1}(\overline{x})A_n$$

de donde se deduce que

$$\omega_n = X^*_n(\overline{x})[(y - B_n)Y_n(y) - A^*_nY_{n-1}(y)] - [X^*_n(\overline{x})(\overline{x} - B^*_n) - X^*_{n-1}(\overline{x})A_n]Y_n(y)$$

$$= yX_n^*(\overline{x})Y_n(y) - X_n^*(\overline{x})A_nY_{n-1}(y) - \overline{x}X_n^*(\overline{x})Y_n(y) + X_{n-1}^*(\overline{x})A_nY_n(y)$$
$$= (y-\overline{x})X_n^*(\overline{x})Y_n(y) + \omega_{n-1}.$$

La fórmula (6.4) es ahora inmediata sumando en (6.3) entre 1 y $n-1$.

Para probar (6.5), basta tomar $X_k = P_k(\overline{u})$, $Y_k = P_k(v)$. Se tiene entonces que

$$P_{n-1}^*(u)A_nP_n(v) - P_n^*(u)A_n^*P_{n-1}(v) - \omega_0 = (v-u)\sum_{k=1}^{n-1} P_k^*(u)P_k(v),$$

y puesto que

$$\begin{aligned}\omega_0 =& P_0^*(\overline{u})A_1P_1(v) - P_1^*(\overline{u})A_1^*P_0(v)\\ =& A_1A_1^{-1}(v-B_0) - (\overline{u}-B_0^*)A_1^{*-1}A_1^*\\ =& v-\overline{u}\end{aligned}$$

y $P_0^*(u)P_0^*(v) = I$ se deduce que

$$P_{n-1}^*(u)A_nP_n(v) - P_n^*(u)A_n^*P_{n-1}(v) = (v-u)\sum_{k=0}^{n-1} P_k^*(u)P_k(v).$$

La fórmula (6.7) es inmediata tomando $u = v = z$ en (6.5), (6.8) es inmediata también a partir de (6.5),y la demostración de (6.9) es similar a la de (6.5) tomando $X_k = Q_k(\overline{u})$ e $Y_k = P_k(v)$ y comprobando que $\omega_0 = -I$.

□

Lema 6.2. *Si A es una matriz de tamaño $N \times N$ tal que $A_nA = A^*A_n^*$, entonces*

$$(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$$
$$= (P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda))$$

para $\lambda \in \mathbb{C}$ que no sea un cero de $\det(P_n(\lambda) - AP_{n-1}(\lambda)) = 0$. En particular, para $\lambda \in \mathbb{R}$ que no sea un cero de $\det(P_n(\lambda) - AP_{n-1}(\lambda)) = 0$

se tiene que la matriz $(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$ *es hermítica.*

Demostración

Veamos primero que

$$\int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x) P_{n-1}^*(x) = A_n^{-1},$$

que será usado después. A partir de la fórmula de recurrencia de tres términos (6.1) tenemos que para $x \in \mathbb{R}$,

$$P_{n-1}^*(x) = (xP_n^*(x) - P_{n-1}^*(x)A_{n+1}^* - P_n^*(x)B_n^*)A_n^{-1}.$$

Usando ahora la ortonormalidad de P_n con respecto a cualquier polinomio de grado menor que n tenemos que

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x) P_{n-1}^*(x) \\
&= \int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x)(xP_n^*(x) - P_{n-1}^*(x)A_{n+1}^* - P_n^*(x)B_n^*)A_n^{-1} \\
&= \int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x) x P_n^*(x) A_n^{-1} \\
&= \int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x)(x - \lambda) P_n^*(x) A_n^{-1} \\
&= \int_{\mathbb{R}} (P_n(x) - P_n(\lambda)) d\mu(x) P_n^*(x) A_n^{-1} = A_n^{-1}.
\end{aligned}$$

Puesto que $A_n A = A^* A_n$ se tiene que $A_n^{-1} A^* = A A_n^{*\,-1}$, y por tanto que

$$\int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x) P_{n-1}^*(x) A^* = A \int_{\mathbb{R}} P_{n-1}(x) d\mu(x) \frac{P_n^*(\lambda) - P_n^*(x)}{\lambda - x}.$$

Usando esto y de nuevo la ortonormalidad de P_n con respecto a cualquier otro polinomio de grado menor que n tenemos que

$$(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda))$$

$$= (P_n(\lambda) - AP_{n-1}(\lambda))^{-1} \int_{\mathbb{R}} \frac{(P_n(\lambda) - AP_{n-1}(\lambda)) - (P_n(x) - AP_{n-1}(x))}{\lambda - x} d\mu(x)$$

$$= (P_n(\lambda) - AP_{n-1}(\lambda))^{-1} \int_{\mathbb{R}} \frac{(P_n(\lambda) - AP_{n-1}(\lambda)) - (P_n(x) - AP_{n-1}(x))}{\lambda - x} d\mu(x)$$

$$[(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*) - (P_n^*(x) - P_{n-1}^*(x)A^*)](P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1} -$$

$$-(P_n(\lambda) - AP_{n-1}(\lambda)) \int_{\mathbb{R}} \frac{P_n(\lambda) - P_n(x)}{\lambda - x} d\mu(x) P_{n-1}^*(x) A^* (P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$$

$$= (P_n(\lambda) - AP_{n-1}(\lambda))^{-1} \int_{\mathbb{R}} [(P_n(\lambda) - AP_{n-1}(\lambda)) - (P_n(x) - AP_{n-1}(x))] d\mu(x)$$

$$\frac{[(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*) - (P_n^*(x) - P_{n-1}^*(x)A^*)]}{\lambda - x} (P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1} -$$

$$-(P_n(\lambda) - AP_{n-1}(\lambda)) A \int_{\mathbb{R}} P_{n-1}(x) d\mu(x) \frac{P_n^*(\lambda) - P_n^*(x)}{\lambda - x} (P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$$

$$= \int_{\mathbb{R}} d\mu(x) \frac{(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*) - (P_n^*(x) - P_{n-1}^*(x)A^*)}{\lambda - x} (P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}$$

$$= (Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}.$$

□

Estamos ya en condiciones de establecer el principal teorema de este capítulo:

Teorema 6.3. *Para una matriz de medidas μ en V_{2n-2} las siguientes aserciones son equivalentes:*

(1) μ es una medida extremal del conjunto V_{n-1}.

(2) Los polinomios hasta grado $n-1$ son densos en el espacio $\mathcal{L}_N^2(\mu)$.

(3) Existe una matriz cuadrada A de tamaño $N \times N$ de modo que $A_n A = A^* A_n^*$ y $\mu = \sum_{i=1}^{m} G_i \delta_{x_i}$, donde x_i, $i = 1, \ldots, m$ son los distintos ceros del polinomio $(P_n(\lambda) - AP_{n-1}(\lambda))$ y G_i son las matrices hermíticas semidefinidas positivas que aparecen en la descomposición en fracciones

simples siguiente

$$(6.10) \qquad (P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)) = \sum_{i=1}^{m} \frac{G_i}{\lambda - x_i},$$

(ver (1.4) en los preliminares para la descomposición (6.10)).

Demostración

(1)→(2) Si μ es extremal en V_{n-1}, se prueba en [K] (ver preliminares. pag. 10) que μ es discreta y por tanto tiene la forma

$$\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$$

siendo q un cierto número natural, x_i números reales y G_i matrices numéricas semidefinidas positivas.

Razonamos por reducción al absurdo. Si los polinomios hasta el grado $n-1$ no son densos en el espacio $\mathcal{L}^2_N(\mu)$ entonces por el teorema de Hahn-Banach existe un operador no nulo Λ definido en $\mathcal{L}^2_N(\mu)$ de modo que se anula para cualquier polinomio de grado menor o igual que $n-1$. Por el teorema de dualidad podemos representar este operador mediante una única función $g = \sum_{i=1}^{q} A_i \delta_{x_i}$ de $\mathcal{L}^2_N(\mu)$. Para cualquier función f en $\mathcal{L}^2_N(\mu)$ el operador Λ está definido como

$$\Lambda(f) = \int_{\mathbb{R}} \tau(f(t)d\mu(t)g^*(t)) = \sum_{i=1}^{q} \tau(f(x_i)G_i A_i^*).$$

Por ser $\Lambda(p) = 0$ para cualquier polinomio p hasta grado $n-1$ tenemos que la matriz de medidas no necesariamente definida positiva

$$\mu_0 = \sum_{i=1}^{q} G_i A_i^* \delta_{x_i}$$

tiene momentos nulos hasta grado $n-1$.

Considerando

$$\mu_0^H = \sum_{i=1}^{q}(G_iA_i^* + A_iG_i)\delta_{x_i} = \sum_{i=1}^{q} H_i\delta_{x_i}$$

obtenemos una matriz de medidas hermítica con momentos nulos hasta grado $n-1$. Llamemos a_i al más pequeño autovalor no nulo de la matriz G_i (recordemos que G_i son semidefinidas positivas) y elijamos un número positivo C tal que

$$\frac{1}{C}\max_{1\leq i\leq q}\|H_i\|_2 < \min_{1\leq i\leq q} a_i.$$

Descomponemos la matriz de medidas μ de la forma siguiente

$$\mu = \frac{1}{2}\left(\mu + \frac{1}{C}\mu_0^H\right) + \frac{1}{2}\left(\mu - \frac{1}{C}\mu_0^H\right),$$

y probamos a continuación que las matrices de medidas $\mu \pm \frac{1}{C}\mu_0^H$ son definidas positivas, para lo cual es suficiente probar que las matrices numéricas $G_i \pm \frac{1}{C}H_i$ son semidefinidas positivas. Para ello, si v es un vector de $\mathbb{C}^N$ y pertenece a $\mathrm{Ker}G_i$, entonces

$$v\left(G_i \pm \frac{1}{C}H_i\right)v^* = v\left(G_i \pm \frac{1}{C}(G_iA_i^* + A_iG_i)\right) = 0$$

y si v no está en $\mathrm{Ker}G_i$ entonces

$$\begin{aligned} v\left(G_i \pm \frac{1}{C}H_i\right)v^* =& vG_iv^* \pm \frac{1}{C}vH_iv^* \\ \geq& \|v\|^2a_i - \frac{1}{C}\|H_i\|\|v\|^2 \\ \geq& \left(\min_{1\leq i\leq q} a_i - \frac{1}{C}\max_{1\leq i\leq q}\|H_i\|\right)\|v\|^2 > 0. \end{aligned}$$

Por tanto concluimos que la matriz de medidas μ no puede ser extremal en V_{n-1}, lo cual es una contradicción.

(2)→(1) Supongamos ahora que los polinomios hasta grado $n-1$ son densos en $\mathcal{L}_N^2(\mu)$ y por reducción al absurdo que μ no es extremal en el conjunto V_{n-1}, es decir μ puede descomponerse como $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$, siendo $0 < \alpha < 1$ y μ_1 y μ_2 dos matrices de medidas diferentes en V_{n-1}. Definimos entonces los operadores T y T_1 en $\mathcal{L}_N^2(\mu)$ como:

$$T(f) = \int_{\mathbb{R}} \tau(f d\mu I) \quad \text{y} \quad T_1(f) = \int_{\mathbb{R}} \tau(f d\mu_1 I).$$

Ambos son claramente lineales y T es continuo puesto que I está en $\mathcal{L}_N^2(\mu)$. Para T_1 tenemos:

$$\left| \int_{\mathbb{R}} \tau(f d\mu_1 I) \right| = \frac{1}{\alpha} \left| \int_{\mathbb{R}} \tau(f d(\mu - (1-\alpha)\mu_2) I) \right|$$

$$\leq \frac{1}{\alpha} \left(\int_{\mathbb{R}} \tau(f d(\mu - (1-\alpha)\mu_2) f^*) \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} \tau(I d(\mu - (1-\alpha)\mu_2) I) \right)^{\frac{1}{2}}$$

$$\leq \frac{1}{\alpha} \left(\int_{\mathbb{R}} \tau(f d\mu f^*) \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} d\tau \mu \right)^{\frac{1}{2}}$$

así que tomando $c = \frac{1}{\alpha} \left(\int_{\mathbb{R}} d\tau\mu \right)^{\frac{1}{2}}$ tenemos $|T_1(f)| \leq c \|f\|_{2,\mu}$, es decir T_1 es un operador continuo. Definiendo ahora $U = T - T_1$, U es un operador lineal y continuo definido en $\mathcal{L}_N^2(\mu)$ tal que U no es el operador nulo pero se anula sobre todos los polinomios hasta grado $n-1$, así que concluimos que los polinomios hasta grado $n-1$ no son densos en el espacio $\mathcal{L}_N^2(\mu)$.

Obsérvese que la demostración dada para (2)→(1) sirve también para un problema matricial de momentos no truncado, teniéndose entonces que si μ es una solución del problems de momentos y los polinomios son densos en el espacio $\mathcal{L}_N^2(\mu)$, entonces μ es extremal en el conjunto de soluciones de dicho problema de momentos.

Para probar (3)→(2) necesitamos dos lemas:

Lema 6.4. *Si $\mu = M(t)d\tau\mu$ es una matriz de medidas definida positiva y no degenerada y $(P_k)_{k=0}^{\infty}$ es una sucesión de polinomios matriciales ortonormales asociados, siendo $\mathrm{gr}(P_k) = k$ y con coeficiente lider no singular, entonces para cualquier f en $\mathcal{L}_N^2(\mu)$ y cada n natural, la mejor aproximación de f en $\mathcal{L}_N^2(\mu)$ por un polinomio de grado n es la dada por la serie de Fourier*

$$\sum_{k=0}^{n}(f, P_k)P_k$$

donde

$$(f, P_k) = \int_{\mathbb{R}} f(t)M(t)P_k^*(t)d\tau\mu(t).$$

Demostración

Cualquier polinomio matricial P de grado n puede expresarse como $P(t) = \sum_{k=0}^{n} \alpha_k P_k$, siendo α_k matrices en $\mathbb{C}^{N\times N}$. Para un polinomio $P(t) = \sum_{k=0}^{n} \alpha_k P_k$ tenemos que

$$\left\| f - \sum_{k=0}^{n} \alpha_k P_k \right\|_{2,\mu}^2 = \int_{\mathbb{R}} \tau \left(f - \sum_{k=0}^{n} \alpha_k P_k \right) M \left(f - \sum_{k=0}^{n} \alpha_k P_k \right)^* d\tau\mu$$

$$= \tau \left(\int_{\mathbb{R}} fMf^* d\tau\mu - \sum_{k=0}^{n} \int_{\mathbb{R}} fMP_k^* d\tau\mu\alpha_k^* - \sum_{k=0}^{n} \int_{\mathbb{R}} \alpha_k P_k M f^* d\tau\mu + \sum_{k=0}^{n} \alpha_k\alpha_k^* \right)$$

$$= \tau \left(\int_{\mathbb{R}} fMf^* d\tau\mu - \sum_{k=0}^{n} (f, P_k)\alpha_k^* - \sum_{k=0}^{n} \alpha_k (f, P_k)^* + \sum_{k=0}^{n} \alpha_k\alpha_k^* \right)$$

$$= \tau \left(\int_{\mathbb{R}} fMf^* d\tau\mu + \sum_{k=0}^{n} (\alpha_k - (f, P_k))(\alpha_k - (f, P_k))^* - \sum_{k=0}^{n} (f, P_k)(f, P_k)^* \right)$$

$$= \left\| f - \sum_{k=0}^{n} (f, P_k)P_k \right\|_{2,\mu}^2 + \sum_{k=0}^{n} \tau(\alpha_k - (f, P_k))(\alpha_k - (f, P_k))^*$$

que es mínimo cuando $\alpha_k = (f, P_k)$, es decir cuando se aproxima con la serie de Fourier.

$\square$

Lema 6.5.

Para una matriz de medidas definida positiva de la forma

$$\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$$

siendo q un cierto número natural, x_i números reales y G_i matrices numéricas semidefinidas positivas, son equivalentes

(a) $\overline{\mathbb{C}_{n-1}^{N\times N}[x]} = \mathcal{L}_N^2(\mu)$.

(b) $\displaystyle\sum_{i=1}^{q} \mathrm{rg}(G_i) = nN$.

Demostración

Para una matriz de medidas como esta (discreta y con soporte finito), es claro que todos los espacios $\mathcal{L}_N^p(\mu)$ son isomorfos, para $1 \leq p \leq \infty$, y cualquier función f de $\mathcal{L}_N^p(\mu)$ puede representarse como

$$f = \sum_{i=1}^{q} F_i \delta_{x_i}$$

donde F_i son matrices numéricas y x_i los puntos del soporte de μ. Puesto que identificamos f con θ si $F_i G_i = \theta$, para $i = 1, \ldots, q$, es claro que los F_i que dan todos los valores posibles de $f(x_i)$ como función de $\mathcal{L}_N^2(\mu)$ están en biyección con la imagen de la aplicación definida por G_i (que también denotamos por G_i), $G_i : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$, dada por $G_i(F) = FG_i$. Y teniendo en cuenta que dim $(\mathrm{Im} G_i) = N\mathrm{rg}(G_i)$, tenemos que

$$\dim(\mathcal{L}_N^2(\mu)) = \sum_{i=1}^{q} N\mathrm{rg}(G_i) = N \sum_{i=1}^{q} \mathrm{rg}(G_i). \tag{6.11}$$

(a)$\rightarrow$(b)

Si $\mathbb{C}_{n-1}^{N\times N}[x]$ es denso en $\mathcal{L}_N^2(\mu)$, podemos representar cualquier función f de $\mathcal{L}_N^2(\mu)$ como

$$f = \sum_{k=0}^{n-1} A_k P_k(t)$$

siendo A_k matrices de tamaño $N \times N$ cualesquiera; si

$$\sum_{k=0}^{n-1} A_k P_k(t) = \sum_{k=0}^{n-1} A'_k P_k(t),$$

basta usar la ortonormalidad de $P_0, \ldots, P_{n-1}$ para obtener inmediatamente que $A_k = A'_k$, para $k = 0, \ldots, n-1$.

De este modo, al variar las matrices A_k en $M_{N\times N}(\mathbb{C})$ se obtiene todo el espacio $\mathcal{L}^2_N(\mu)$ y por tanto deducimos que

$$\dim(\mathcal{L}^2_N(\mu)) = nN^2. \tag{6.12}$$

Teniendo en cuenta ahora (6.11) se tiene que

$$\sum_{i=1}^{q} \mathrm{rg}(G_i) = nN.$$

(b)$\rightarrow$(a)

Dados P, Q en $\mathbb{C}^{N\times N}_{n-1}[x]$, $P \neq Q$, podemos escribir

$$P = \sum_{k=0}^{n-1} A_k P_k, \quad Q = \sum_{k=0}^{n-1} B_k P_k$$

con $B_k \neq A_k$ para algún k entre 0 y $n-1$, por tanto $P \neq Q$ en $\mathcal{L}^2_N(\mu)$. Como $\mathbb{C}^{N\times N}_{n-1}[x] \subseteq \mathcal{L}^2_N(\mu)$, obtenemos que $\dim(\mathcal{L}^2_N(\mu)) \geq nN^2$.

Además, de (6.11) obtenemos

$$\dim(\mathcal{L}^2_N(\mu)) = N\left(\sum_{i=1}^{q} \mathrm{rg}(G_i)\right) = nN^2,$$

por tanto $\mathbb{C}^{N\times N}_{n-1}[x] = \mathcal{L}^2_N(\mu)$, y en consecuencia los polinomios hasta grado $n-1$ son densos en $\mathcal{L}^2_N(\mu)$.

$\square$

Volvamos ahora a la demostración del teorema, probando

(3)→(2)

Sea A una matriz numérica verificando $A_n A = A^* A_n^*$, y consideremos la descomposición en fracciones simples

$$(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)) = \sum_{i=1}^{m} \frac{G_i}{\lambda - x_i}.$$

Está probado en [**D3**] (ver preliminares, pag. 9) que el rango de G_i coincide con la multiplicidad de x_i como cero de $P_n(\lambda) - AP_{n-1}(\lambda)$. En consecuencia, $\sum_{i=1}^{q} \mathrm{rg}(G_i) = \sum_{i=1}^{q} \mathrm{multiplicidad}(x_i) = nN$. Del lema anterior deducimos que $(3) \to (2)$.

(2)→(3)

Supongamos ahora que $\mu \in V_{2n-2}$, y que $\mathbb{C}_{n-1}^{N\times N}[x]$ es denso en $\mathcal{L}_N^2(\mu)$. Por ser esta condición equivalente a la extremalidad de μ en V_{n-1}, en virtud de los resultados probados en [**K**], podemos expresar $\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$, con G_i matrices numéricas semidefinidas positivas y x_i números reales que vamos a determinar. Procedemos en varios pasos.

PRIMER PASO: Para $1 \leq i, j \leq q$, se tiene la siguiente fórmula

$$G_j(P_{n-1}^*(x_j)A_n P_n(x_i) - P_n^*(x_j)A_n^* P_{n-1}(x_i))G_i = \theta.$$

En virtud del lema 6.4, y puesto que por hipótesis $\overline{\mathbb{C}_n^{N\times N}[x]} = \mathcal{L}_N^2(\mu)$, para toda función f de $\mathcal{L}_N^2(\mu)$ tenemos que

$$f = \sum_{k=0}^{n-1} (f, P_k) P_k$$

en $\mathcal{L}_N^2(\mu)$, siendo (f, P_k) el coeficiente de Fourier de f asociado a P_k. es decir

$$
\begin{aligned}
0 &= \left\| f - \sum_{k=0}^{n-1} (f, P_k) P_k \right\|_{2,\mu}^2 \\
&= \sum_{i=1}^{q} \tau \left(f(x_i) - \sum_{k=0}^{n-1} (f, P_k) P_k(x_i) \right) G_i \left(f(x_i) - \sum_{k=0}^{n-1} (f, P_k) P_k(x_i) \right)^*
\end{aligned}
$$

lo cual equivale a que

$$
f(x_i) G_i = \sum_{k=0}^{n-1} (f, P_k) P_k(x_i) G_i, \quad \text{para } 1 \leq i \leq q.
$$

Para $\lambda \in \mathbb{C} \setminus \mathbb{R}$, llamamos $f_\lambda(t) = \dfrac{I}{t - \lambda}$. Puesto que f_λ es acotada, $f \in \mathcal{L}_N^2(\mu)$. Calculamos sus coeficientes de Fourier:

$$
\begin{aligned}
(f_\lambda, P_k) &= \int_{\mathbb{R}} \frac{I}{t - \lambda} d\mu(t) P_k^*(t) \\
&= \int_{\mathbb{R}} d\mu(t) \frac{P_k^*(t) - P_k^*(\lambda)}{(t - \lambda)} + \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} P_k^*(\lambda) \\
&= Q_k^*(\lambda) + \omega(\lambda) P_k^*(\lambda),
\end{aligned}
$$

siendo $\omega(\lambda)$ la transformada de Hilbert de f_λ, es decir

$$
\omega(\lambda) = \int_{\mathbb{R}} \frac{d\mu(t)}{t - \lambda} = \sum_{i=1}^{q} \frac{G_i}{x_i - \lambda}.
$$

Tenemos por tanto que para $1 \leq i \leq q$ y $\lambda \in \mathbb{C} \setminus \mathbb{R}$,

$$
\begin{aligned}
\frac{G_i}{x_i - \lambda} &= \sum_{k=0}^{n-1} (Q_k^*(\lambda) + \omega(\lambda) P_k^*(\lambda)) P_k(x_i) G_i \\
&= \sum_{k=0}^{n-1} Q_k^*(\lambda) P_k(x_i) G_i + \omega(\lambda) \sum_{k=0}^{n-1} P_k^*(\lambda) P_k(x_i) G_i
\end{aligned}
$$

lo cual da

$$\theta = \left[-I + (x_i - \lambda)\sum_{k=0}^{n-1} Q_k^*(\lambda)P_k(x_i)\right] G_i + \omega(\lambda)\left[(x_i - \lambda)\sum_{k=0}^{n-1} P_k^*(\lambda)P_k(x_i)\right] G_i$$

para $1 \leq i \leq q$ y $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Usando las fórmulas (6.5) y (6.9) se obtiene

$$\theta = \left[Q_{n-1}^*(\lambda)A_nP_n(x_i) - Q_n^*(\lambda)A_n^*P_{n-1}(x_i)\right] G_i + \omega(\lambda)\left[P_{n-1}^*(\lambda)A_nP_n(x_i) - P_n^*(\lambda)A_n^*P_{n-1}(x_i)\right] G_i \tag{6.13}$$

para $1 \leq i \leq q$ y $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Multiplicando (6.13) por $(x_j - \lambda)$ y tomando límite cuando λ tiende a j, para $j \neq i$, y teniendo en cuenta que $\omega(\lambda) = \sum_{i=1}^{q} \frac{G_i}{x_i - \lambda}$ se obtiene que para $1 \leq i, j \leq q$, $i \neq j$,

$$G_j(P_{n-1}^*(x_j)A_nP_n(x_i) - P_n^*(x_j)A_n^*P_{n-1}(x_i))G_i = \theta \tag{6.14}$$

para $1 \leq i, j \leq q$, $i \neq j$.

Puesto que $P_{n-1}^*(x_i)A_nP_n(x_i) - P_n^*(x_i)A_n^*P_{n-1}(x_i) = \theta$ (fórmula (6.7)), la igualdad (6.14) se tiene para todo $1 \leq i, j \leq q$.

SEGUNDO PASO: definición de la matriz A.

A es la matriz que representa a una aplicación lineal $A : \mathbb{C}^N \to \mathbb{C}^N$ que definimos como sigue:

Si v, w son vectores de $\mathbb{C}^N$ y $v = P_{n-1}(x_i)G_iw$, entonces definimos $Av = P_n(x_i)G_iw$, es decir, A está definida sobre la imagen de la aplicación que define la matriz $P_{n-1}(x_i)G_i$ al multiplicar por la derecha.

Probamos en primer lugar que A está bien definida, es decir que si un vector v está en la imagen de $P_{n-1}(x_p)G_p$ y de $P_{n-1}(x_q)G_q$, o está representado en la imagen de $P_{n-1}(x_p)G_p$ por dos vectores, entonces las dos definiciones coinciden. Supongamos pues que $v = P_{n-1}(x_p)G_p w_p = P_{n-1}(x_q)G_q w_q$, donde $p \neq q$, o si $p = q$, entonces $w_p \neq w_q$. Probemos que $P_n(x_p)G_p w_p = P_n(x_q)G_q w_q$.

En (6.14), poniendo $i = p$ y multiplicando a la derecha por w_p se obtiene
(6.15)
$$G_j P^*_{n-1}(x_j)A_n P_n(x_p)G_p w_p - G_j P^*_n(x_j)A^*_n v = \theta, \quad \text{para } 1 \leq j \leq q.$$

Poniendo en (6.14) $i = q$ y multiplicando a la derecha por w_q queda
(6.16)
$$G_j P^*_{n-1}(x_j)A_n P_n(x_q)G_q w_q - G_j P^*_n(x_j)A^*_n v = \theta, \quad \text{para } 1 \leq j \leq q.$$

Restando (6.15) y (6.16) se obtiene
(6.17)
$$G_j P^*_{n-1}(x_j)A_n (P_n(x_p)G_p w_p - P_n(x_q)G_p w_q) = \theta, \quad \text{para } 1 \leq j \leq q.$$

Multiplicando por $P_{n-1}(x_j)$ en (6.17) por la izquierda y sumando en j se obtiene

$$\left(\sum_{j=1}^{q} P_{n-1}(x_j)G_j P^*_{n-1}(x_j) \right) A_n (P_n(x_p)G_p w_p - P_n(x_q)G_p w_q) = \theta.$$

Teniendo ahora en cuenta que

$$\sum_{j=1}^{q} P_{n-1}(x_j)G_j P^*_{n-1}(x_j) = \int_{\mathbb{R}} P_{n-1}(t)d\mu(t)P^*_{n-1}(t) = I$$

y que A_n es invertible se deduce que $P_n(x_p)G_p w_p = P_n(x_q)G_p w_q$.

Veamos ahora que A está definido sobre todo $\mathbb{C}^N$. Llamemos T_i a la aplicación lineal de $\mathbb{C}^N$ en $\mathbb{C}^N$ representada por la matriz $P_{n-1}(x_i)G_i P^*_{n-1}(x_i)$, $1 \leq i \leq q$. A continuación probamos que la suma de las imágenes de las aplicaciones T_i es todo $\mathbb{C}^N$. Llamemos A_i a una base de $\mathrm{Im}(T_i)$. Si el espacio generado por los elementos de A_i $(1 \leq i \leq q)$, es decir $< A_1, \ldots, A_q >$ tiene dimensión menor o igual que $N-1$, existe un vector no nulo v ortonormal a $< A_1, \ldots, A_q >$ y por tanto perteneciente a $\mathrm{Ker}(T_i)$ para todo i, es decir $T_i v = \theta$ para todo i. Sumando en i se obtiene como antes que $v = \theta$ lo cual es absurdo.

Por tanto, cualquiera que sea el vector v en $\mathbb{C}^N$, existen vectores w_i en $\mathbb{C}^N$, $1 \leq i \leq q$, de modo que $v = \sum_{i=1}^{q} P_{n-1}(x_i) G_i P_{n-1}^*(x_i) w_i$. Llamando $v_i = P_{n-1}^*(x_i) w_i$, queda $v = \sum_{i=1}^{q} P_{n-1}(x_i) G_i v_i$, y por tanto la suma de las imágenes de los operadores T_i es todo $\mathbb{C}^N$, con lo cual la matriz A está definida en todo $\mathbb{C}^N$.

TERCER PASO: $A_n A = A^* A_n^*$.

Dado un vector cualquiera v en $\mathbb{C}^N$, lo expresamos como $v = \sum_{i=1}^{q} P_{n-1}(x_i) G_i v_i$, siendo v_i vectores en $\mathbb{C}^N$. Multiplicando en (6.14) a la derecha por v_i y sumando en i se obtiene que

$$G_j (P_{n-1}^*(x_j) A_n A - P_n^*(x_j) A_n^*) v = \theta,$$

para cualquier vector v en $\mathbb{C}^N$, por tanto se deduce que

$$G_j P_{n-1}^*(x_j) A_n A - G_j P_n^*(x_j) A_n^* = \theta.$$

Multiplicando ahora en esta fórmula a la izquierda por v_j^* y sumando se obtiene que

$$v^* A_n A - v^* A^* A_n^* = \theta$$

para cualquier vector v de $\mathbb{C}^N$, y por tanto $A_n A = A^* A_n^*$.

CUARTO PASO: Para $\lambda \in \mathbb{C} \setminus \mathbb{R}$,

$$\omega(\lambda) = -(P_n(\lambda) - A P_{n-1}(\lambda))^{-1} (Q_n(\lambda) - A Q_{n-1}(\lambda)).$$

Retomamos ahora la fórmula (6.13)

$$\theta = \left[Q_{n-1}^*(\lambda)A_nP_n(x_i) - Q_n^*(\lambda)A_n^*P_{n-1}(x_i)\right]G_i$$
$$+\,\omega(\lambda)\left[P_{n-1}^*(\lambda)A_nP_n(x_i) - P_n^*(\lambda)A_n^*P_{n-1}(x_i)\right]G_i.$$

para $1 \le i \le q$. Multiplicando a la derecha por w y llamando $v = P_{n-1}(x_i)G_iw$, tenemos que $Av = P_n(x_i)G_iw$. Teniendo en cuenta que todo vector de $\mathbb{C}^N$ se puede escribir como suma de vectores de la forma $P_{n-1}(x_i)G_iw$, $i = 1, \ldots, q$ (ver segundo paso), obtenemos que

$$\theta = \left[Q_{n-1}^*(\lambda)A_nA - Q_n^*(\lambda)A_n^*\right]v + \omega(\lambda)\left[P_{n-1}^*(\lambda)A_nA - P_n^*(\lambda)A_n^*\right]v$$

para todo v de $\mathbb{C}^N$ y todo $\lambda \in \mathbb{C}\setminus\mathbb{R}$. Como A_n es invertible deducimos que

$$\theta = \left[Q_{n-1}^*(\lambda)A_nAA_n^{*\,-1} - Q_n^*(\lambda)\right]A_n^*v$$
$$+\,\omega(\lambda)\left[P_{n-1}^*(\lambda)A_nAA_n^{*\,-1} - P_n^*(\lambda)\right]A_n^*v$$

para todo vector v de $\mathbb{C}^N$ y $\lambda \in \mathbb{C}\setminus\mathbb{R}$. Usando que $A_nA = A^*A_n^*$ se obtiene que

$$\theta = [Q_{n-1}^*(\lambda)A^* - Q_n^*(\lambda)] + \omega(\lambda)[P_{n-1}^*(\lambda)A^* - P_n^*(\lambda)], \quad \lambda \in \mathbb{C}\setminus\mathbb{R}$$

que despejando da

$$\omega(\lambda) = -(Q_n^*(\lambda) - Q_{n-1}^*(\lambda)A^*)(P_n^*(\lambda) - P_{n-1}^*(\lambda)A^*)^{-1}, \quad \lambda \in \mathbb{C}\setminus\mathbb{R}$$

y que gracias al lema 6.3 es

$$\omega(\lambda) = -(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda)), \quad \lambda \in \mathbb{C}\setminus\mathbb{R}.$$

Teniendo en cuenta que $w(\lambda) = \sum_{i=1}^{q} \frac{G_i}{x_i - \lambda}$ y la descomposición en fracciones simples de $(P_n(\lambda) - AP_{n-1}(\lambda))^{-1}(Q_n(\lambda) - AQ_{n-1}(\lambda))$ (ver preliminares, pag. 9), el teorema queda probado.

□

Conviene comparar el teorema 6.4 con su versión escalar. La representación de las medidas de V_{2n-2} extremales en V_{n-1} en términos de los polinomios $p_n - ap_{n-1}$ es casi trivial en el caso escalar, debido a que es posible despejar y obtener directamente $a = \dfrac{p_n(x_i)}{p_{n-1}(x_i)}$. Como se ha visto, la estructura matricial introduce complicaciones que hacen que el resultado para el caso matricial sea más laborioso de obtener, requiriendo nuevas ideas.

Estas diferencias de estructura entre el caso matricial y el escalar generan discrepancias importantes como se verá a continuación. En el caso escalar, los puntos del soporte de las medidas extremales soportan la máxima masa posible en V_{2n-2}; en efecto, en el caso escalar si $\mu \in V_{2n-2}$ es extremal en V_{n-1}, es bien conocido (véase por ejemplo [**A**]) que si $x_i \in \mathrm{sop}\mu$, entonces

$$\mu(\{x_i\}) = \frac{1}{\displaystyle\sum_{k=0}^{n-1} p_k^2(x_i)}$$

y $\mu(\{x_i\}) \geq \nu(\{x_i\})$, cualquiera que sea $\nu \in V_{2n-2}$.

Esta propiedad no es cierta en el caso matricial. En efecto. las medidas de V_{2n-2} extremales en V_{n-1} no tienen porqué soportar la máxima masa posible en V_{2n-2}; esto ocurrrirá sólo cuando la masa soportada corresponda a una matriz invertible, en cuyo caso el polinomio $P_n(\lambda) - AP_{n-1}(\lambda)$ tiene en ese punto un cero de multiplicidad máxima (N), y recíprocamente, en este caso la matriz A vendría dada por $A = P_n(a)P_{n-1}^{-1}(a)$, siendo a el punto que soporta la máxima masa. Dedicaremos la parte final del capítulo a demostrarlo:

Teorema 6.6. *Si $\mu \in V_{2n-2}$ es una matriz de medidas extremal en V_{n-1}. pongamos*

$$\mu = \sum_{i=1}^{q} G_i \delta_{x_i}$$

donde x_i son los ceros de $P_n(\lambda) - AP_{n-1}(\lambda)$, para cierta A tal que $A_n A = A^* A_n^*$, entonces las siguientes condiciones son equivalentes:

(1) *μ alcanza en x_{i_0} la máxima masa posible en V_{2n-2}, más concretamente,*

$$G_{i_0} = \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \right)^{-1}.$$

(2) *G_{i_0} es invertible.*

(3) *$P_n(\lambda) - AP_{n-1}(\lambda)$ tiene en x_{i_0} un cero de multiplicidad máxima (N).*

(4) *$P_{n-1}(x_{i_0})$ es invertible y $A = P_n(x_{i_0}) P_{n-1}^{-1}(x_{i_0})$.*

Demostración

Veamos en primer lugar que de (1) se deducen (2), (3) y (4). La máxima masa posible para $\mu \in V_{2n-2}$ en x_{i_0} es

$$\left(\sum_{k=0}^{n-1} P_k^*(\lambda_0) P_k(\lambda_0) \right)^{-1}.$$

puesto que

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) = \int_{\mathbb{R}} \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t) \right) d\mu(t) \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t) \right)^*$$

$$= \int_{\mathbb{R}\setminus\{x_{i_0}\}} \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t) \right) d\mu(t) \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(t) \right)^* +$$

$$+ \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \right) B \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \right)^*$$

donde se denota por B la masa de la matriz de medidas μ en x_{i_0}. Tenemos por tanto que

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \geq \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \right) B \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0}) P_k(x_{i_0}) \right)^*$$

y por tanto

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0}) \left[\left(\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0}) \right)^{-1} - B \right] \sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0})$$

de donde se deduce que

$$\left(\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0}) \right)^{-1} - B \geq 0.$$

Este hecho aparece también probado en [**Z**].

Así, si μ soporta en x_{i_0} la máxima masa posible, esta viene dada por

$$G_{i_0} = \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0}) \right)^{-1},$$

que es una matriz invertible por ser la inversa de una matriz invertible:

$$\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0}) \geq P_0^*(x_{i_0})P_0(x_{i_0}) = I.$$

La relación entre el rango de G_{i_0} y la multiplicidad de x_{i_0} (recuérdese que estos dos valores coinciden) nos dice que $P_n(\lambda) - AP_{n-1}(\lambda)$ tiene un cero de multiplicidad máxima en x_{i_0}, es decir de multiplicidad N. En este caso es fácil dar una expresión de la matriz A; en efecto, puesto que $P_n(\lambda) - AP_{n-1}(\lambda)$ tiene en x_{i_0} un cero de multiplicidad N, deducimos de la observación 3.4 (2) que $P_n(x_{i_0}) - AP_{n-1}(x_{i_0}) = \theta$, y por tanto que $P_n(x_{i_0}) = AP_{n-1}(x_{i_0})$. Si $P_{n-1}(x_{i_0})$ fuera singular se tendría que x_{i_0} sería un cero de $P_{n-1}(\lambda)$ y de $P_n(\lambda)$, y además $P_n(x_{i_0})$ y $P_{n-1}(x_{i_0})$ tendrían un autovector común asociado a 0, lo cual contradice el teorema 3.3 (4). En consecuencia, $P_{n-1}(x_{i_0})$ es invertible y por tanto $A = P_n(x_{i_0})P_{n-1}^{-1}(x_{i_0})$.

Para probar el recíproco vemos que (2) implica (1). Multiplicando (6.13) por $(x_i - \lambda)$, tomando límite cuando λ tiende a x_i y teniendo en cuenta que $\omega(\lambda) = \sum_{i=1}^{q} \frac{G_i}{x_i - \lambda}$ se obtiene que para $1 \leq i \leq q$

$$G_i(P_{n-1}^*(x_i)A_nP_n'(x_i) - P_n^*(x_i)A_n^*P_{n-1}'(x_i))G_i = G_i,$$

que usando (6.8) y particularizando para $i = i_0$ se trasforma en

$$G_{i_0} = G_{i_0}\left(\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0})\right)G_{i_0}.$$

Puesto que por hipótesis G_{i_0} es invertible es posible despejar G_{i_0} obteniéndose

$$G_{i_0} = \left(\sum_{k=0}^{n-1} P_k^*(x_{i_0})P_k(x_{i_0})\right)^{-1},$$

es decir, la máxima masa posible.

□

REFERENCIAS

[A]. N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, english translation, Oliver and Boyd, Edinburgo, 1965.

[AN]. A.I. Aptekarev y E.M. Nikishin, *The scattering problem for a discrete Sturm-Liouville operator*, Mat. USSR Sb. **49** (1984), 325-355.

[B]. Heinz Bauer, *Probability theory and elements of measure theory*, Academic Press, Nueva York, Londres, 1978.

[C]. P. L. Chebyshev, *Sur les fractions continues*, J. Math. Pures Appl. (2) **3** (1858), 203-230, Œuvres I, Chelsea Publ. Company, Nueva York.

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

[DGK]. Ph. Delsarte, Y. Genin e Y. Kamp, *Orthogonal polynomials matrices on the unit circle*, IEEE Trans. Circuits and Systems **25** (1978), 149-160.

[DS]. N. Dunford y J. T. Schwartz, *Linear Operators. Part II: Spectral theory. Self adjoint operators in Hilbert space*, Interscience publishers (a Division of John Wiley & Sons), Nueva York, Londres, 1983.

[D1]. A. J. Durán, *A generalization of Favard's theorem for polynomials satisfying a recurrence relation*, J. Approx. Theory **74** (1993), 83-109.

[D2]. A. J. Durán, *On orthogonal polynomials with respect to a positive definite matrix of measures*, Can. J. Math. **47** (1995), 88-112.

[D3]. A. J. Durán, *Markov's theorem for orthogonal matrix polynomials*, Can. J. Math. **Por aparecer**.

[DL1]. A. J. Durán y P. López-Rodríguez, *Orthogonal matrix polynomials: zeros and Blumenthal's theorem*, J. Approx. Theory **Por aparecer**.

[DL2]. A. J. Durán y P. López-Rodríguez, *The L^p spaces of a positive definite matrix of measures and density of matrix polynomials in L^1*, Preprint.

[DL3]. A. J. Durán y P. López-Rodríguez, *Density questions for the truncated matrix moment problem*, Preprint.

[DV]. A. J. Durán y W. Van Assche, *Orthogonal matrix polynomials and higher order recurrence relations*, Linear Algebra Appl. **Por aparecer**.

[G1]. J. S. Geronimo, *Scattering theory and matrix orthogonal polynomials on the real line*, Circuits Systems Signal Process. **1** (1982), 471-495.

[G2]. J. S. Geronimo, *Matrix orthogonal polynomials on the unit circle*, J. Math. Phys. **22(7)** (1981), 1359-1365.

[GV]. J. S. Geronimo y W. Van Assche, *Orthogonal polynomials with asymptotically periodic recurrence coefficients*, J. Approx. Theory **46** (1986), 251-283.

[F]. J. Favard, *Sur les polynomes de Tchebicheff*, C. R. Acad. Sci. Paris **200** (1935), 2052-2053.

[H]. E. Heine, *Handbuch der Kugelfunctionen*, 2. Anlage, vol. I, G. Reimer, Berlin, 1878.

[HJ]. R.A. Horn y C.A. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.

[K]. A. N. Kochubei, *The extreme points of the set of solutions of the truncated matrix moment problem*, (Ucraniano) Dopovïdï Akad. Nauk Ukraïn. RSR Ser. A (1972), 418-421.

[Kr]. M. G. Krein, *Fundamental aspects of the representation theory of Hermitian Operators with deficiency index* (m, m), Amer. Math. Soc. Transl. (2) Vol. 97, 1970 (Original en Ukrain. Mat. Z. 1 (1949), 3-66, 418-421.

[MNV]. A. Mate, P. Nevai y W. Van Assche, *The supports of measures associated with orthogonal polynomials and the spectra of the related self-adjoint operators*, Rocky Mt. J. Math. **21** (1991), 501-527.

[N]. M. A. Naimark, *Extremal spectral functions of a symmetric operator*, Bulletin de l'académie des sciences de l'URSS **11** (1947), 327-344.

[P]. O. Perron, *Die Lehre von den Kettenbrüchen*, vol. II, Teubner, Stuttgart, 1977.

[R]. M. Rosemberg, *The square-integrability of matrix-valued functions with respect to a non-negative hermitian measure*, Duke Math. J. **31** (1964), 291-298.

[Ri]. Theodore J. Rivlin, *Chebyshev Polynomials*, Second Edition, John Wiley & Sons, Inc., Nueva York, 1990.

[RN]. F. Riesz y B. Sz.-Nagy, *Functional analysis*, (traducción de la segunda edición francesa por L. F. Boron, original francés: *Leçons d'analyse fonctionelle*, Akadémiai Kiadó, Budapest, 1952) Ungar, Nueva York, 1955.

[S]. M. H. Stone, *Linear Transformations in Hilbert Space and Their Applications to Analysis*, Amer. Math. Soc. Colloq. Publ., vol. 15, Amer. Math. Soc., Providence, RI, 1932.

[St]. T. J. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse **9** (1895), A1-A47.

[SV]. A. Sinap y W. Van Assche, *Polynomial interpolation and Gaussian quadrature for matrix valued functions*, Linear Algebra Appl. **207** (1994), 71-114.

[V]. W. Van Assche, *Asymptotics for Orthogonal Polynomials*, Lecture Notes in Mathematics, Springer Verlag, Berlin, 1987.

[W]. Wintner, *Spektraltheorie der unendlichen Matrizen*, Hirzel, Leipzig, 1929.

[Z]. D. Zhani, *Problème des moments matriciels sur la droite: construction d'une famille de solutions et questions d'unicité*, Publications du Departement de Mathématiques, Université Claude-Bernard-Lyon I Nouvelle serie **21** (1984), 1-84.

UNIVERSIDAD DE SEVILLA

Reunido el Tribunal integrado por los abajo firmantes en el día de la fecha, para juzgar la Tesis Doctoral de D. PEDRO LOPEZ RODRIGUEZ titulada POLINOMIOS MATRICIALES ORTONORMALES

acordó otorgarle la calificación de APTO CUM LAUDE POR UNANIMIDAD

Sevilla, 13 de NOVIEMBRE 1.995

El Vocal, | El Vocal, | El Vocal,

El Presidente | El Secretario, | El Doctorado,