Universidad Nacional Autónoma
de México

Facultad de Ciencias

UN CRITERIO DE SOLUCIÓN AL PROBLEMA DE LOS MOMENTOS

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
ESTEBAN SÁNCHEZ GONZÁLEZ

TUTOR:
DR. LUIS ANTONIO RINCÓN SOLÍS

Ciudad Universitaria, Cd. Mx., 2017

1. Datos del alumno
Sánchez
González
Esteban
55 20 04 04 10
Universidad Nacional Autónoma de
México
Facultad de Ciencias
Matemáticas
304327942

2. Datos del tutor
Dr
Luis Antonio
Rincón
Solís

3. Datos del sinodal 1
Dra
María Asunción Begoña
Fernández
Fernández

4. Datos del sinodal 2
Dr
Gerardo
Sánchez
Licea

5. Datos del sinodal 3
Dr
Yuri
Salazar
Flores

6. Datos del sinodal 4
Dr
Fernando
Baltazar
Larios

Agradecimientos

Quiero agradecer al Dr. Luis Antonio Rincón Solís por haberme brindado el apoyo y la confianza para realizar este trabajo bajo su dirección, así como por el tiempo, las enseñazas e ideas que pudo compartir conmigo durante el periodo de desarrollo, ya que fueron un factor muy importarte para poder culminar.

A mis sinodales por su tiempo y dedicación en la revisión de mi trabajo.

También agradezco el apoyo y motivación de mi familia, en especial de mi madre Elodia y mis hermanas quienes siempre me brindaron los elementos necesarios para poder continuar con mis estudios.

Finalmente agradezco a la UNAM por darme la oportunidad de crecer tanto personalmente como académicamente y a los profesores de esta institución que motivaron mi interés en las Matemáticas.

Contenido

Introducción

En el presente trabajo se discutirá un criterio de solución al problema de los momentos de Hamburger. La afirmación de dicho problema es la siguiente:

Dada una sucesión de números reales $\{s_k\}_{k=0}^{\infty}$, se busca una función no decreciente $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ que satisfaga las ecuaciones

$$s_k = \int_{-\infty}^{\infty} u^k d\alpha(u), \quad \text{para} \quad k = 0, 1, 2, ...$$

Dicho de otra forma, se buscan condiciones para que una sucesión de números reales sea la sucesión de momentos de alguna función de distribución.

Este trabajo está basado principalmente en [1], utilizando conceptos previos de probabilidad y análisis matemático, así como herramientas de álgebra lineal.

El problema de los momentos se mencionó por primera vez en el año de 1894, cuando Thomas Jan Stieltjes (1856-1894) publicó el documento *"Recherches sur les fractions continues"*. En este trabajo, Stieltjes introdujo lo que ahora es conocido como la integral de Stieltjes con respecto a una función no decreciente con la cual definió el concepto de momento generalizado de orden k, como

$$\int_{0}^{\infty} u^k d\alpha(u), \quad \text{para} \quad k = 0, 1, 2, ...$$

Dicha integral también fue utilizada para resolver el siguiente problema al cual Stieltjes llamó *"el problema de los momentos"*. Dada una sucesión de números reales $s_0, s_1, s_2, ...$ se requiere encontrar una distribución de masa positiva α en el intervalo $[0, \infty)$ tal que $\{s_k\}_{k=0}^{\infty}$ sea la sucesión de momentos de α.

Una distribución de masa positiva α en el intervalo $[0, \infty)$ es una función no decreciente, tal que para cualesquiera $u_0 \geq 0$ y $u_1 > u_0$ el incremento $\alpha(u_1) - \alpha(u_0)$ representa la masa en el intervalo $[u_0, u_1)$. Por lo tanto la masa total en $[0, \infty)$ se puede escribir como

$$\lim_{v \longrightarrow \infty} [\alpha(v) - \alpha(0)] = \int_0^\infty d\alpha(u),$$

mientras que las integrales

$$\int_0^\infty u d\alpha(u), \quad \int_0^\infty u^2 d\alpha(u),$$

representan el momento estático y el momento de inercia de la distribución de masa en cuestión, respectivamente.

Para darle solución al problema de los momentos, Stieltjes se adentró en el estudio de fracciones continuas, pero lamentablemente murió el mismo año en que el documento con su trabajo fue publicado.

Anteriormente un problema relacionado al problema de los momentos fue desarrollado en 1873 por Pafnuty Lvovich Chebyshev (1821-1894), el cual fue estudiado y generalizado por Andrei Andreyevich Markov (1856-1922).

Una generalización al problema original de Stieltjes fue planteada y resuelta por Hans Ludwig Hamburger (1889-1956) en una serie de documentos de los cuales el primero se publicó en 1920: *"Über eine Erweiterung der Stieltjesschen Momentenproblems"*, o *"Una extensión de los problemas de momentos de Stieltjes"*. Dicha generalización extendía el soporte de α del intervalo $[0, \infty)$ a toda la recta real. Este problema ahora es conocido como el problema de los momentos de Hamburger. Como Stieltjes, Hamburger utilizó fracciones continuas para dar solución al problema.

Casi simultáneamente con Hamburger, el problema fue resuelto por otros matemáticos y avanzó en diferentes direcciones al darle solución por otros métodos. En particular, Rolf Herman Nevanlinna (1895-1980), Frigyes Riesz (1880-1956), Ernst David Hellinger (1883-1950) y Torsten Carleman (1892-1949), mostraron conexiones entre el problema de los momentos y muchas ramas del análisis, como el análisis funcional, la teoría de funciones y la teoría espectral de operadores.

En torno al problema de los momentos se pueden plantear una serie de preguntas con distintos enfoques. Con relación a la solubilidad cuestiones como

¿bajo qué condiciones tiene solución?, ¿qué condiciones debe de cumplir la sucesión de momentos? o ¿qué características debe tener la función de distribución? Acerca del número de soluciones interrogantes como, ¿el problema tiene solución?, ¿la solución encontrada es única?, o en caso de no serlo, ¿tiene una cantidad finita o infinita de soluciones?

En el trabajo [9] se puede apreciar una exposición orientada al número de soluciones del problema de los momentos y a la construcción de las mismas. El presente trabajo se desarrolla enfocado a la solubilidad del problema de los momentos, mostrando un criterio particular que lo resuelve. El contenido de las secciones de esta tesis es el siguiente:

En la primera sección se recuerdan conceptos de la teoría de transformaciones lineales y formas cuadráticas. Se define el concepto de sucesión definida positiva y se da una equivalencia a dicha definición.

En la segunda sección se define el funcional correspondiente a una sucesión. Se define el concepto de funcional positivo y se da una segunda equivalencia a la definición de sucesión definida positiva.

En la tercera sección se definen las matrices infinitas de Jacobi, el concepto de sucesión de polinomios ortonormales para una sucesión definida positiva y se da una sucesión particular de dichos polinomios. Además se muestra una relación de correspondencia uno a uno entre las sucesiones definidas positivas y las matrices infinitas de Jacobi.

En la cuarta sección se introduce otra sucesión de polinomios y se muestran propiedades que cumplen estos polinomios y los polinomios ortonormales. Se define el concepto de polinomio cuasiortogonal.

En la quinta sección se construye la fórmula de cuadratura, la cual tiene coeficientes que son importantes para construir la solución del problema de los momentos.

Finalmente, en la sexta sección, se muestra un criterio que da solución al problema de los momentos.

Además de las principales secciones ya mencionadas, se incluye un apéndice con tres secciones que tienen el siguiente contenido:

En la primera sección se define la integral de Riemann-Stieltjes y se muestran propiedades de la misma. En la segunda sección se muestra la fórmula de interpolación de Lagrange. En la tercera sección se desarrollan los teoremas de Helly.

El material expuesto en las secciones principales de esta tesis está basado principalmente de [1]. Además, para los apéndices se utilizó información de [3], [11] y [7]. Se realizó un particular esfuerzo para que las demostraciones que se tomaron de dichas referencias se presentaran de forma más detallada. También se desarrollaron casi completamente algunas pruebas que se presentan sólo como un bosquejo de idea en los libros mencionados.

1. Sucesiones definidas positivas

En esta sección se recordarán algunos conceptos de la teoría de transformaciones lineales y formas cuadráticas. En particular, se define un concepto importante para el desarrollo de este trabajo, las sucesiones definidas positivas. Se verá además una manera equivalente de definir este tipo de sucesiones.

A continuación se recuerda el concepto de transformación lineal para lo cual hay que establecer que dado un vector renglón $\bar{x} = (x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1}$, se tiene que el vector transpuesto denotado por

$$\bar{x}^{\,t} = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix},$$

es un vector columna que también pertenece a $\mathbb{R}^{n+1}$.

Definición 1.1. *Una transformación lineal es una función $T : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^{m+1}$ dada por la relación*

$$T(\bar{x}) = C\bar{x}^{\,t},$$

para una matriz C, a la cual se le conoce como matriz estándar de la transformación lineal T, y donde $\bar{x} = (x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1}$.

Sea I la matriz identidad. La función dada por la relación $T(\bar{x}) = I\bar{x}^t = \bar{x}^t$ para toda $\bar{x} \in \mathbb{R}^{n+1}$, llamada función identidad es un ejemplo de transformación lineal.

De la definición anterior se puede ver que una transformación lineal cumple la siguiente propiedad

$$T(\alpha\bar{x} + \bar{y}) = \alpha T(\bar{x}) + T(\bar{y}),$$

para cualesquiera $\bar{x}, \bar{y} \in \mathbb{R}^{n+1}$ y $\alpha \in \mathbb{R}$.

Definición 1.2. *Una transformación lineal se llama degenerada si el determinante de su matriz estándar es igual a cero, y no degenerada en el caso contrario.*

Por ejemplo, es sencillo verificar que las transformaciones lineales dadas por las relaciones

$$T_1(x,y) = \left(\begin{array}{cc} 1 & 2 \\ 0 & 1 \end{array} \right) \begin{pmatrix} x \\ y \end{pmatrix} = (x+2y, y)^t,$$

y

$$T_2(x,y) = \left(\begin{array}{cc} 2 & 2 \\ 1 & 1 \end{array} \right) \begin{pmatrix} x \\ y \end{pmatrix} = (2x+2y, x+y)^t,$$

son no degenerada y degenerada, respectivamente.

Definición 1.3. *Una forma cuadrática es una función* $q : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ *de la forma*

$$q(\bar{x}) = \sum_{i,k=0}^{n} a_{ik} x_i x_k,$$

en donde $\bar{x} = (x_0, x_1, ..., x_n) \in \mathbb{R}^{n+1}$ *y los coeficientes* a_{ik} *son números reales.*

La función $q(\bar{x}) = x^2 + xy + y^2$ definida de $\mathbb{R}^2$ a $\mathbb{R}$ es un ejemplo de forma cuadrática.

Toda forma cuadrática puede ser expresada de forma matricial de la siguiente manera

$$q(\bar{x}) = \begin{pmatrix} x_0 & x_1 & ... & x_n \end{pmatrix} \left(\begin{array}{cccc} a_{00} & a_{01} & ... & a_{0n} \\ a_{10} & a_{11} & ... & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & ... & a_{nn} \end{array} \right) \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Por ejemplo, para la forma cuadrática $q(\bar{x}) = x^2 + xy + y^2$ se tiene que

$$q(\bar{x}) = \begin{pmatrix} x & y \end{pmatrix} \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array} \right) \begin{pmatrix} x \\ y \end{pmatrix}.$$

La representación matricial de una forma cuadrática puede expresarse de una manera peculiar debido al siguiente resultado.

Teorema 1.4. *Toda forma cuadrática $q : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ puede ser expresada matricialmente de la forma*

$$q(\bar{x}) = \bar{x}\, A\, \bar{x}^{\,t},$$

donde A es una matriz simétrica llamada matriz asociada a la forma cuadrática.

Demostración. Dada la forma cuadrática $q(\bar{x}) = \sum_{i,k=0}^{n} a_{ik}x_ix_k$, se observan los pares de sumandos de la forma $a_{ik}x_ix_k$ y $a_{ki}x_kx_i$ notando que

$$\begin{aligned} a_{ik}x_ix_k + a_{ki}x_kx_i &= (a_{ik} + a_{ki})x_ix_k \\ &= \frac{(a_{ik} + a_{ki})}{2}x_ix_k + \frac{(a_{ik} + a_{ki})}{2}x_kx_i, \end{aligned}$$

por lo que la matriz de la forma cuadrática también puede ser expresada como la siguiente matriz simétrica

$$\begin{pmatrix} a_{00} & \dfrac{a_{01} + a_{10}}{2} & \dots & \dfrac{a_{0n} + a_{n0}}{2} \\ \dfrac{a_{10} + a_{01}}{2} & a_{11} & \dots & \dfrac{a_{1n} + a_{n1}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{a_{n0} + a_{0n}}{2} & \dfrac{a_{n1} + a_{1n}}{2} & \dots & a_{nn} \end{pmatrix},$$

$\square$

Se puede ver fácilmente que para una forma cuadrática su matriz asociada es única.

Como ejemplo se tiene que la forma cuadrática $q(\bar{x}) = x^2 + xy + y^2$, vista anteriormente, también se puede representar de forma matricial como

$$q(\bar{x}) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Si el vector $\bar{x} = (x_0, x_1, ..., x_n)$ está expresado en una base B de $\mathbb{R}^{n+1}$, se denotará como $[x]_B$, donde $[x]_B = [x_0, x_1, ..., x_n]$. Se puede hacer un cambio de base para expresar de manera diferente una forma cuadrática con lo cual se obtiene lo siguiente.

Sea B' una base de $\mathbb{R}^{n+1}$ diferente de B y P la matriz del cambio de base de B a B', entonces se tiene que

$$[x]_B = [x]_{B'} P,$$

para todo $[x]_B$ en $\mathbb{R}^{n+1}$, con lo cual se observa que se puede considerar un cambio de base como una transformación lineal. Así que sustituyendo en la expresión matricial de una forma cuadrática en su forma simétrica, obtenida en el teorema anterior, se tiene que

$$\begin{aligned} q(\bar{x}) &= \bar{x}\, A\, \bar{x}^t = [x]_B\, A\, [x]_B^t \\ &= ([x]_{B'} P) A (P^t [x]_{B'}^t) = [x]_{B'} (PAP^t) [x]_{B'}^t, \end{aligned}$$

en donde $A' = PAP^t$ es una matriz simétrica con la cual se expresa q en la base B'.

Es importante mencionar que al hacer el cambio de base de B a B', todo elemento en la nueva base B' proviene de un único elemento de la base B, además de que si $[x]_B \neq [0]_B$, entonces $[x]_B P^{-1} = [x]_{B'} \neq [0]_{B'} = [0]_B P^{-1}$.

El siguiente teorema muestra que se puede representar una forma cuadrática de una manera más sencilla.

Teorema 1.5. *Toda forma cuadrática $q : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ se puede expresar de tal forma que su matriz asociada sea una matriz diagonal.*

Demostración. Sea $q(\bar{x}) = [x]_B A [x]_B^t$ la forma matricial de la forma cuadrática q donde A es su matriz asociada. Como A es simétrica, admite una diagonalización ortogonal [5], es decir A es semejante a una matriz diagonal D, así que $D = PAP^{-1}$, donde P es ortogonal y es la matriz de cambio de base de B a B'. Al ser P ortogonal, $P^{-1} = P^t$, por lo que

$$D = PAP^{-1} = PAP^t,$$

con lo cual se tiene que A y D son congruentes.

Luego se tiene que en la base B' la forma cuadrática puede expresarse de la siguiente forma

$$\begin{aligned} q(\bar{x}) &= [x]_B \, A \, [x]_B^t = ([x]_{B'} P) A (P^t [x]_{B'}^t) \\ &= [x]_{B'} \, D \, [x]_{B'}^t = \lambda_0 y_0^2 + \cdots + \lambda_n y_n^2, \end{aligned}$$

donde

$$[x]_{B'} = (y_0, \ldots, y_n) = \bar{y} \quad \text{y} \quad \lambda_0, ..., \lambda_n,$$

son los valores propios [5] de la matriz A.

□

Por el teorema anterior toda forma cuadrática puede ser expresada como $q(\bar{x}) = \lambda_0 y_0^2 + \cdots + \lambda_n y_n^2$, donde $(y_0, ..., y_n) = [x]_{B'} = [x]_B P^{-1}$ y $\lambda_0, \lambda_1, ..., \lambda_n$ son los valores propios de su matriz asociada.

Es sencillo verificar que para la forma cuadrática $q(\bar{x}) = x^2 + xy + y^2$ que se ha visto anteriormente, su representación diagonal es la siguiente

$$q(\bar{x}) = [x]_{B'} \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix} [x]_{B'}^t,$$

donde $[x]_{B'} = [x]_B P^{-1}$ y

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

Se puede efectuar una clasificación de las formas cuadráticas de acuerdo al signo que toman los vectores bajo éstas, de la siguiente forma.

Definición 1.6. *Sea* $q(\bar{x}) = \sum_{i,k=0}^{n} a_{ik}x_i x_k$ *una forma cuadrática. Se dice que*

a) *q es definida positiva* $\Longleftrightarrow$ *para toda* $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\},\ q(\bar{x}) > 0,$

b) *q es definida negativa* $\Longleftrightarrow$ *para toda* $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\},\ q(\bar{x}) < 0,$

c) *q es indefinida* $\Longleftrightarrow$ *existen* $\bar{x}_1, \bar{x}_2 \in \mathbb{R}^{n+1}$ *tales que* $q(\bar{x}_1) > 0$ *y* $q(\bar{x}_2) < 0$.

Estos son algunos casos particulares de las formas cuadráticas, ya que en general pueden tener otros comportamientos, pero no son relevantes para el desarrollo del presente trabajo.

Como ejemplos de las definiciones anteriores se tiene que

a) $q(\bar{x}) = x_0^2 + x_1^2$ es una forma cuadrática definida positiva,

b) $q(\bar{x}) = -x_0^2 - x_1^2$ es una forma cuadrática definida negativa y

c) $q(\bar{x}) = x_0^2 - x_1^2$ es una forma cuadrática indefinida.

Ahora se verá una caracterización de las formas cuadráticas definidas positivas.

Teorema 1.7. *Una forma cuadrática* $q : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ *es definida positiva si y sólo si todos los valores propios de su matriz asociada son positivos.*

Demostración.

$\Rightarrow$) Supóngase que q es definida positiva. Entonces $q(\bar{x}) > 0$ para cualquier $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\}$, por lo que para el vector $\bar{x}_i = \bar{e}_i\, P$, el i-ésimo vector de la base canónica multiplicado por la matriz del cambio de base, se tiene que

$$q(\bar{x}_i) = (\bar{e}_i\, P)A(\bar{e}_i\, P)^t = \bar{e}_i\,(PAP^t)\,\bar{e}_i^t = \bar{e}_i\, D\, \bar{e}_i^t = \lambda_i > 0,$$

para toda $i = 0, 1, ..., n$.

$\Leftarrow$) Supóngase que $\lambda_i > 0$ para toda $i = 0, 1, ..., n$. Entonces para todo $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\}$, con $\bar{x} = [x]_B = [x]_{B'}P$, donde $[x]_{B'} = \bar{y} \neq \bar{0}$, se tiene que

$$q(\bar{x}) = [x]_B\, A\, [x]_B^t = ([x]_{B'}P)A(P^t[x]_{B'}^t) = [x]_{B'}\, D\, [x]_{B'}^t = \sum_{i=0}^{n} \lambda_i y_i^2 > 0.$$

$\square$

La forma cuadrática definida positiva $q(\bar{x}) = x_0^2 + x_1^2$ se puede escribir como

$$q(\bar{x}) = \begin{pmatrix} x_0 & x_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix},$$

en donde se observa que sus valores propios son positivos.

El siguiente teorema muestra otra característica de las formas cuadráticas definidas positivas, para lo cual primero se recordará la definición de menor principal de una matriz y se verá una proposición que tiene que ver con el signo del determinante de la matriz asociada a una forma cuadrática.

Definición 1.8. *Dada una matriz $A = (a_{ij})$ de $n+1$ filas y $n+1$ renglones, los menores principales de orden $1, 2, ..., n$ de esa matriz son los determinantes*

$$a_{00}, \quad \begin{vmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{vmatrix}, \; \cdots, \; \begin{vmatrix} a_{00} & a_{01} & ... & a_{0n} \\ a_{10} & a_{11} & ... & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & ... & a_{nn} \end{vmatrix}.$$

Proposición 1.9. *Sea una forma cuadrática q y sea una transformación lineal no degenerada T. Entonces el signo del determinante de la matriz asociada a q es el mismo que el de la matriz asociada a la composición $q(T)$.*

Demostración. Sean $q(\bar{x}) = \bar{x}\, A\, \bar{x}^{\,t}$ y $T(\bar{x}) = C\, \bar{x}^{\,t}$ una forma cuadrática y una transformación lineal no degenerada, respectivamente. Entonces

$$q(T(\bar{x})) = (\bar{x}C^{\,t})A(C\,\bar{x}^{\,t}) = \bar{x}(C^{\,t}A\,C)\bar{x}^{\,t},$$

así que después de aplicar la transformación lineal se obtiene una forma cuadrática cuya matriz asociada es $C^t A\, C$, pero como $|C^t| = |C|$, se tiene que

$$|C^{\,t}AC| = |C^{\,t}||A||C| = |A||C|^2,$$

con lo cual se observa que el signo del determinante de la matriz asociada a q se mantiene.

$\square$

Teorema 1.10. *Una forma cuadrática es definida positiva si y sólo si los menores principales de su matriz asociada son positivos.*

Demostración. Se procede por inducción sobre m, el número de variables de la forma cuadrática. Para $m = 1$ se tiene que $q(\bar{x}) = a_{00}x_0^2$ es una forma cuadrática definida positiva si y sólo si el único menor principal $a_{00} > 0$.

Supóngase que el resultado se cumple por inducción para $m = n$. Se verá que se cumple para $m = n + 1$.

Sea $q(\bar{x}) = \sum_{i,k=0}^{n} a_{ik}x_i x_k$ una forma cuadrática con $n + 1$ variables. Dado que se puede representar con una matriz asociada simétrica entonces

$$
\begin{aligned}
q(\bar{x}) &= \sum_{i=0}^{n}\sum_{k=0}^{n} a_{ik}x_i x_k = \sum_{i=0}^{n}\left(\sum_{k=0}^{n-1} a_{ik}x_i x_k + a_{in}x_i x_n\right) \\
&= \sum_{i=0}^{n}\sum_{k=0}^{n-1} a_{ik}x_i x_k + \sum_{i=0}^{n} a_{in}x_i x_n \\
&= \sum_{k=0}^{n-1}\left(\sum_{i=0}^{n-1} a_{ik}x_i x_k + a_{nk}x_n x_k\right) + \sum_{i=0}^{n} a_{in}x_i x_n \qquad (1.1) \\
&= \sum_{i,k=0}^{n-1} a_{ik}x_i x_k + \sum_{k=0}^{n-1} a_{nk}x_n x_k + \sum_{i=0}^{n-1} a_{in}x_i x_n + a_{nn}x_n^2 \\
&= \sum_{i,k=0}^{n-1} a_{ik}x_i x_k + 2\sum_{i=0}^{n-1} a_{in}x_i x_n + a_{nn}x_n^2,
\end{aligned}
$$

donde el primero de estos tres términos representa la forma cuadrática de n variables

$$
r(x_0, x_1, ..., x_{n-1}) = \sum_{i,k=0}^{n-1} a_{ik}x_i x_k,
$$

para la cual sus menores principales coinciden con los menores principales de la forma cuadrática q, excepto por el último de q. Después de notar lo anterior se continuará con el paso inductivo.

$\Rightarrow$) Supóngase que q es definida positiva. Entonces r también es definida positiva, ya que si existe $(x_0, x_1, ..., x_{n-1}) \neq \bar{0}$ tal que $r(x_0, x_1, ..., x_{n-1}) \leq 0$, se tiene que $q(x_0, x_1, ..., x_{n-1}, 0) = r(x_0, x_1, ..., x_{n-1}) \leq 0$, lo cual es una contradicción.

Luego, por hipótesis de inducción, todos los menores principales de r son positivos, así que sólo falta comprobar que el último menor principal de q también es positivo. Pero esto se cumple, pues por el Teorema 1.5 el último menor tiene el mismo signo que el producto de todos los valores propios de su matriz asociada, los cuales son positivos por el Teorema 1.7.

$\Leftarrow$) Supóngase ahora que todos los menores principales de la forma cuadrática q son positivos. Entonces los menores principales de r también son positivos, pues coinciden con los de q, y por la hipótesis de inducción r es definida positiva.

Luego por el Teorema 1.5 se puede escribir a r como

$$r(\bar{x}) = \sum_{i=0}^{n-1} \lambda_i {y'_i}^2 = \sum_{i=0}^{n-1} (\sqrt{\lambda_i} y'_i)^2 = \sum_{i=0}^{n-1} y_i^2,$$

pues $\lambda_i > 0$ para toda $i = 0, 1, ..., n-1$, y haciendo $\sqrt{\lambda_i} y'_i = y_i$.

Sustituyendo esta representación de r en la última expresión de (1.1) y tomando $x_n = y_n$, se obtiene que

$$q(\bar{x}) = \sum_{i=0}^{n-1} y_i^2 + 2 \sum_{i=0}^{n-1} b_{in} y_i y_n + b_{nn} y_n^2,$$

para algunas constantes b_{in}. Luego, como

$$y_i^2 + 2b_{in} y_i y_n = (y_i + b_{in} y_n)^2 - b_{in}^2 y_n^2,$$

se puede sustituir esta expresión en la ecuación anterior, por lo que

$$q(\bar{x}) = \sum_{i=0}^{n-1} \left[(y_i + b_{in} y_n)^2 - b_{in}^2 y_n^2 \right] + b_{nn} y_n^2,$$

y aplicando la siguiente transformación lineal no degenerada

$$z_i = y_i + b_{in} y_n, \quad z_n = y_n,$$

con $i = 0, 1, ..., n-1$, se tiene que

$$
\begin{aligned}
q(\bar{x}) &= \sum_{i=0}^{n-1}(z_i^2 - b_{in}^2 z_n^2) + b_{nn} z_n^2 \\
&= \sum_{i=0}^{n-1} z_i^2 - \sum_{i=0}^{n-1} b_{in}^2 z_n^2 + b_{nn} z_n^2 \\
&= \sum_{i=0}^{n-1} z_i^2 + \left(b_{nn} - \sum_{i=0}^{n-1} b_{in}^2 \right) z_n^2 \\
&= \sum_{i=0}^{n-1} z_i^2 + c z_n^2.
\end{aligned}
$$

Así, q es definida positiva si c es positivo. Nótese que c es el determinante de la matriz asociada a la forma cuadrática del lado derecho de la última igualdad. Entonces por la proposición previa al teorema, esta forma cuadrática tiene el mismo signo que el determinante de q, la cual por hipótesis tiene determinante positivo.

□

Nuevamente para la forma cuadrática definida positiva

$$
q(\bar{x}) = \begin{pmatrix} x_0 & x_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix},
$$

se puede ver que sus menores principales son positivos.

Con las nociones anteriores es posible definir, finalmente, el concepto de sucesión definida positiva, lo cual se hace a continuación.

Definición 1.11. *Una sucesión definida positiva es una sucesión infinita de números reales $\{s_k\}_{k=0}^{\infty}$ tal que todas las formas cuadráticas*

$$
\sum_{i,k=0}^{n} s_{i+k} x_i x_k, \quad n = 0, 1, 2, ..., \tag{1.2}
$$

son definidas positivas.

Se puede ver que las formas cuadráticas de la forma (1.2) tienen una matriz asociada simétrica. Esto es debido a la expresión de sus coeficientes, y se aprecia mejor en la representación matricial de las formas cuadráticas

$$\sum_{i,k=0}^{n} s_{i+k} x_i x_k = \begin{pmatrix} x_0 & x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dada la definición, no es sencillo mostrar que una sucesión particular es definida positiva. A pesar de ello se pueden dar varios ejemplos de sucesiones definidas positivas como la sucesión de números reales dada por $0!, 1!, 2!, 3!, \dots$ Esto se dará como consecuencia del resultado principal que se estudia en este trabajo.

A continuación se muestra una equivalencia de la definición de sucesión definida positiva.

Teorema 1.12. *Una sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva si y sólo si todos los determinantes*

$$D_n = \begin{vmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{vmatrix}, \tag{1.3}$$

con $n = 0, 1, 2, \dots$, son positivos.

Demostración.

$\Rightarrow$) Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Entonces para todo n se tiene que la forma cuadrática $\sum_{i,k=0}^{n} s_{i+k} x_i x_k$ es definida positiva, así que por el Teorema 1.10 el menor

$$D_n = \begin{vmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{vmatrix},$$

es positivo.

$\Leftarrow$) Si todos los determinantes D_n son positivos para toda $n \geq 0$, entonces para cualquier $n \geq 0$, la forma cuadrática $\sum_{i,k=0}^{n} s_{i+k}x_ix_k$ cumple que todos los menores principales de su matriz asociada son positivos. Luego por el Teorema 1.10, la forma cuadrática $\sum_{i,k=0}^{n} s_{i+k}x_ix_k$ es definida positiva para todo $n \geq 0$, por lo tanto la sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva.

$\square$

En el apéndice de este documento se agrega como ejemplo el cálculo de los determinantes D_n hasta $n = 6$ para algunas sucesiones definidas positivas. Pero, dada una sucesión definida positiva, resulta complicado determinar de manera general que estos determinantes son positivos, por lo que se omiten los ejemplos.

Se continuará trabajando con las sucesiones definidas positivas en las secciones posteriores. Además este concepto es fundamental para mostrar la solución al problema de los momentos.

2. El funcional correspondiente a una sucesión

A continuación se define el funcional correspondiente a una sucesión de números reales, no necesariamente definida positiva. Más adelante se demostrará una propiedad de dicho funcional con la cual se muestra otra equivalencia de la definición de sucesión definida positiva.

Definición 2.1. *Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión de números reales y sea $R(\lambda) = p_0 + p_1\lambda + \cdots + p_n\lambda^n$ un polinomio. Se define el funcional $\mathcal{S}$ correspondiente a la sucesión como*

$$\mathcal{S}\{R(\lambda)\} = p_0 s_0 + p_1 s_1 + \cdots + p_n s_n.$$

Además de la notación $\mathcal{S}\{R(\lambda)\}$ también se usará la notación $\mathcal{S}_\lambda\{R(\lambda)\}$.

Este funcional es lineal en el sentido algebraico ya que para una constante $c \in \mathbb{R}$ y dos polinomios arbitrarios $R_1(\lambda)$ y $R_2(\lambda)$ se cumple que

$$\mathcal{S}\{cR_1(\lambda) + R_2(\lambda)\} = c\,\mathcal{S}\{R_1(\lambda)\} + \mathcal{S}\{R_2(\lambda)\}.$$

Definición 2.2. *Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Su funcional correspondiente $\mathcal{S}$ es positivo si para cualquier polinomio $R(\lambda) \geq 0$ no idénticamente cero, se tiene que $\mathcal{S}\{R\} > 0$.*

Ahora se verá un lema que ayudará a la demostración de la segunda equivalencia de la definición de sucesión definida positiva.

Lema 2.3. *Sea $R(\lambda) \geq 0$ un polinomio no idénticamente cero. Entonces $R(\lambda)$ tiene grado par, es decir, grado $2(n+1)$ para alguna $n \in \mathbb{N}$. Además se puede expresar como*

$$R(\lambda) = [A(\lambda)]^2 + [B(\lambda)]^2,$$

donde $A(\lambda)$ y $B(\lambda)$ son polinomios de grado no mayor a $n+1$, de la siguiente forma,

$$A(\lambda) = \sum_{k=0}^{n+1} x_k \lambda^k, \quad B(\lambda) = \sum_{k=0}^{n+1} y_k \lambda^k,$$

para ciertos coeficientes reales x_k y y_k.

Demostración. Todo polinomio $R(\lambda)$ que satisface la condición $R(\lambda) \geq 0$ no idénticamente cero para todo λ real, tiene grado par, pues si fuera de grado impar, tomaría valores negativos. Entonces el grado de $R(\lambda)$ es $2(n+1)$, para algún $n \in \mathbb{N}$.

Así, para obtener la expresión anunciada de $R(\lambda)$ primero se debe observar que por el Teorema Fundamental del Álgebra se puede escribir al polinomio como

$$R(\lambda) = c\prod_{k=0}^{2n+1}(\lambda - z_i),$$

donde $c > 0$ es el coeficiente del término principal y z_i es raíz del polinomio $R(\lambda)$, para $i = 0, 1, ..., 2n + 1$.

Considérese a z_0, la primera raíz del polinomio $R(\lambda)$. Entonces se puede escribir al polinomio como

$$R(\lambda) = c\,(\lambda - z_0)\prod_{k=1}^{2n+1}(\lambda - z_i).$$

Si $z_0 \in \mathbb{R}$, como $R(\lambda) \geq 0$ se cumple que

$$\prod_{k=1}^{2n+1}(\lambda - z_i) \geq 0 \quad \text{si} \quad \lambda > z_0,$$

$$\text{y} \quad \prod_{k=1}^{2n+1}(\lambda - z_i) \leq 0 \quad \text{si} \quad \lambda < z_0,$$

por lo que $\prod_{k=1}^{2n+1}(\lambda - z_i)$ cambia de signo en z_0. Por lo tanto, z_0 es nuevamente raíz del polinomio $R(\lambda)$. Procediendo sucesivamente y dado que $R(\lambda)$ tiene grado par, se concluye que toda raíz real tiene multiplicidad par.

Además recordando que si un número complejo es raíz de un polinomio, su conjugado también lo es, se puede escribir finalmente a $R(\lambda)$ de la forma

$$R(\lambda) = c\prod_{k=0}^{n}(\lambda - \alpha_k - i\beta_k)(\lambda - \alpha_k + i\beta_k),$$

donde $\beta_k \geq 0$, y $\alpha_k \in \mathbb{R}$. Esto incluye el caso cuando alguna de las raíces tiene multiplicidad mayor a 1.

Ahora se demostrará por inducción sobre n que se cumplen las igualdades

$$\prod_{k=0}^{n}(\lambda-\alpha_k-i\beta_k)=A(\lambda)+iB(\lambda),$$

$$\prod_{k=0}^{n}(\lambda-\alpha_k+i\beta_k)=A(\lambda)-iB(\lambda),$$

donde $A(\lambda)$ y $B(\lambda)$ son polinomios de grado no mayor a $n+1$. Para la primera igualdad, si $n=0$ se toma $A(\lambda)=\lambda-\alpha_0$ y $B(\lambda)=-\beta_0$. Ahora, suponiendo que se cumple la primera igualdad por inducción para $n=m$, es decir que

$$\prod_{k=0}^{m}(\lambda-\alpha_k-i\beta_k)=A(\lambda)+iB(\lambda),$$

se demostrará para $n=m+1$.

Usando la hipótesis inductiva se tiene que

$$\begin{aligned}\prod_{k=0}^{m}(\lambda-\alpha_k-i\beta_k)&=(A'(\lambda)+iB'(\lambda))(\lambda-\alpha_m-i\beta_m)\\&=\lambda A'(\lambda)-\alpha_m A'(\lambda)-i\beta_m A'(\lambda)\\&\quad+i\lambda B'(\lambda)-i\alpha_m B'(\lambda)+\beta_m B'(\lambda)\\&=\lambda A'(\lambda)-\alpha_m A'(\lambda)+\beta_m B'(\lambda)\\&\quad+i\left[\lambda B'(\lambda)-\beta_m A'(\lambda)-\alpha_m B'(\lambda)\right]\\&=A(\lambda)+iB(\lambda),\end{aligned}$$

con lo cual queda demostrado que se cumple la primera igualdad. La segunda igualdad se prueba de manera análoga.

Desarrollando la representación obtenida de $R(\lambda)$ se tiene que

$$
\begin{aligned}
R(\lambda) &= c\prod_{k=0}^{n}(\lambda-\alpha_k-i\beta_k)(\lambda-\alpha_k+i\beta_k)\\
&= c\,[A'(\lambda)+iB'(\lambda)][A'(\lambda)-iB'(\lambda)]\\
&= [\sqrt{c}\,A'(\lambda)]^2+[\sqrt{c}\,B'(\lambda)]^2\\
&= [A(\lambda)]^2+[B(\lambda)]^2.
\end{aligned}
$$

□

Se puede notar en la demostración del lema anterior que a partir de los coeficientes de cualquier polinomio $R(\lambda) \geq 0$ no idénticamente cero, se pueden obtener los coeficientes de los polinomios $A(\lambda)$ y $B(\lambda)$, e inversamente también es posible, es decir, que a partir de los coeficientes de ambos polinomios se pueden obtener los coeficientes del polinomio $R(\lambda)$.

A continuación se da la segunda equivalencia de la definición de sucesión definida positiva.

Teorema 2.4. *Una sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva si y sólo si su funcional correspondiente $\mathcal{S}$ es positivo.*

Demostración. Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión. Por el lema anterior se tiene que para un polinomio $R(\lambda) \geq 0$ no idénticamente cero

$$
\begin{aligned}
R(\lambda) &= [A(\lambda)]^2+[B(\lambda)]^2\\
&= \left[\sum_{k=0}^{n} x_k\lambda^k\right]^2+\left[\sum_{k=0}^{n} y_k\lambda^k\right]^2 \qquad (2.1)\\
&= \sum_{i,k=0}^{n}\lambda^{i+k}x_ix_k+\sum_{i,k=0}^{n}\lambda^{i+k}y_iy_k,
\end{aligned}
$$

por lo que aplicando el funcional correspondiente a la sucesión se tiene que

$$\mathcal{S}\{R(\lambda)\} = \sum_{i,k=0}^{n} s_{i+k}x_ix_k + \sum_{i,k=0}^{n} s_{i+k}y_iy_k.$$

Con esta información ahora se puede demostrar la equivalencia en el enunciado del teorema.

$\Rightarrow$) Supóngase que la sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva. Entonces las formas cuadráticas (1.2) son definidas positivas. Por la ecuación (2.1)

$$\mathcal{S}\{R(\lambda)\} = \sum_{i,k=0}^{n} s_{i+k}x_ix_k + \sum_{i,k=0}^{n} s_{i+k}y_iy_k > 0.$$

La desigualdad se verifica pues ambos términos son no negativos y por lo menos uno es estrictamente positivo.

$\Leftarrow$) Supóngase que el funcional $\mathcal{S}$ es positivo. Se verá que si $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\}$, entonces la forma cuadrática $\sum_{i,k=0}^{n} s_{i+k}x_ix_k$ es definida positiva.

Sea pues $\bar{x} \in \mathbb{R}^{n+1} \setminus \{\bar{0}\}$. Por la ecuación (2.1), tomando $A(\lambda) = B(\lambda) = \sum_{k=0}^{n} x_k\lambda^k$

$$R(\lambda) = \sum_{i,k=0}^{n} \lambda^{i+k}x_ix_k + \sum_{i,k=0}^{n} \lambda^{i+k}x_ix_k = 2\sum_{i,k=0}^{n} \lambda^{i+k}x_ix_k \geq 0,$$

así que $\sum_{i,k=0}^{n} \lambda^{i+k}x_ix_k \geq 0$.

Por lo que al ser $\mathcal{S}$ positivo se tiene que

$$\mathcal{S}\left\{\sum_{i,k=0}^{n} \lambda^{i+k}x_ix_k\right\} = \sum_{i,k=0}^{n} s_{i+k}x_ix_k > 0.$$

$\square$

Si la sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva, entonces también lo es la sucesión $\{as_k\}_{k=0}^{\infty}$ para cualquier $a > 0$. Además para una sucesión definida positiva se tiene que $s_0 > 0$, ya que la forma cuadrática $\sum_{i,k=0}^{0} s_{i+k}x_ix_k = s_0x_0^2$ debe ser definida positiva.

A continuación se verá un último concepto con el cual se trabajará en la siguiente sección.

Definición 2.5. *Se dice que una sucesión definida positiva* $\{s_k\}_{k=0}^{\infty}$ *es normalizada si* $s_0 = 1$.

El funcional correspondiente a una sucesión será de utilidad para la teoría mostrada en las siguientes secciones. Además la equivalencia de la definición de sucesión definida positiva ayudará en la demostración del resultado principal de este trabajo.

3. Asociación de sucesiones definidas positivas y matrices de Jacobi

En esta sección se mostrará una relación de correspondencia entre las sucesiones definidas positivas normalizadas y las matrices de Jacobi infinitas. Para lo anterior, a partir de una sucesión definida positiva normalizada se obtendrán polinomios con ciertas características y a partir de éstos se conseguirá una matriz de Jacobi infinita. Inversamente, a partir de una matriz de Jacobi infinita también se construirán ciertos polinomios por medio de los cuales se obtendrá una sucesión definida positiva normalizada.

A continuación se define un concepto importante de esta sección.

Definición 3.1. *Una matriz de Jacobi infinita es de la forma*

$$\begin{pmatrix} a_0 & b_0 & 0 & 0 & 0 & \dots \\ b_0 & a_1 & b_1 & 0 & 0 & \dots \\ 0 & b_1 & a_2 & b_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.1}$$

donde cada a_k es un número real y cada b_k es un número real positivo.

Se puede ver que las matrices de Jacobi infinitas son simétricas y tridiagonales.

Más adelante serán importantes los polinomios que cumplen con propiedades de ortonormalidad para una sucesión definida positiva de acuerdo a lo siguiente.

Definición 3.2. *Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Una sucesión de polinomios $\{P_k(\lambda)\}_{k=0}^{\infty}$ es ortonormal con respecto a la sucesión definida positiva si cumple las siguientes propiedades*

i) *$P_k(\lambda)$ es un polinomio exactamente de grado k, para cada $k = 0, 1, 2, ...,$ y su coeficiente principal es positivo.*

ii) *Se mantienen las siguientes relaciones de ortogonalidad*

$$\mathcal{S}\{P_i(\lambda)P_k(\lambda)\} = \delta_{ik} = \begin{cases} 0 & si \quad i \neq k, \\ 1 & si \quad i = k. \end{cases} \tag{3.2}$$

Dada una sucesión definida positiva $\{s_k\}_{k=0}^{\infty}$, se puede escribir una expresión explícita para los polinomios ortonormales en términos de determinantes, de la siguiente manera.

Proposición 3.3. *Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Entonces la sucesión de polinomios*

$$\begin{aligned} P_0(\lambda) &= 1, \\ P_k(\lambda) &= \frac{1}{\sqrt{D_{k-1}D_k}} \begin{vmatrix} s_0 & s_1 & \dots & s_k \\ s_1 & s_2 & \dots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_k & \dots & s_{2k-1} \\ 1 & \lambda & \dots & \lambda^k \end{vmatrix}, \end{aligned} \tag{3.3}$$

donde $k = 1, 2, 3, ...,$ es ortonormal con respecto a la sucesión definida positiva. Además, para una sucesión definida positiva normalizada, si se considera $D_{-1} = 1$, la expresión (3.3) se mantiene válida para $k = 0$.

Demostración. La primera propiedad se sigue del Teorema 1.12, por lo que sólo hace falta mostrar que la sucesión de polinomios cumple la segunda propiedad.

Se tiene que

$$P_k(\lambda)\lambda^j = \frac{1}{\sqrt{D_{k-1}D_k}} \begin{vmatrix} s_0 & s_1 & \dots & s_k \\ s_1 & s_2 & \dots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_k & \dots & s_{2k-1} \\ \lambda^j & \lambda^{j+1} & \dots & \lambda^{j+k} \end{vmatrix}.$$

Aplicando el funcional $\mathcal{S}$ de ambos lados de la ecuación anterior se obtiene

$$\mathcal{S}\{P_k(\lambda)\lambda^j\} = \frac{1}{\sqrt{D_{k-1}D_k}} \begin{vmatrix} s_0 & s_1 & \dots & s_k \\ s_1 & s_2 & \dots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_k & \dots & s_{2k-1} \\ s_j & s_{j+1} & \dots & s_{j+k} \end{vmatrix}.$$

Por independencia o dependencia lineal de los renglones en el determinante se tiene que

$$\begin{aligned} \mathcal{S}\{P_k(\lambda)\lambda^j\} &= \begin{cases} 0 & \text{si} \quad 0 \leq j \leq k-1, \\ \dfrac{1}{\sqrt{D_{k-1}D_k}}\, D_k & \text{si} \quad j = k, \end{cases} \\ &= \begin{cases} 0 & \text{si} \quad 0 \leq j \leq k-1, \\ \sqrt{\dfrac{D_k}{D_{k-1}}} & \text{si} \quad j = k. \end{cases} \end{aligned} \tag{3.4}$$

Por lo tanto, $\mathcal{S}\{P_k(\lambda)P_j(\lambda)\} = 0$, para $j = 0, 1, ..., k-1$.

Para ver que $\mathcal{S}\{P_k(\lambda)P_k(\lambda)\} = 1$, de la ecuación (3.3) se debe notar que

$$P_k(\lambda) \quad = \quad \frac{1}{\sqrt{D_{k-1}D_k}}(-1)^k \begin{vmatrix} 1 & \lambda & ... & \lambda^k \\ s_0 & s_1 & ... & s_k \\ s_1 & s_2 & ... & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_k & ... & s_{2k-1} \end{vmatrix}$$

por lo que

$$P_k(\lambda) = \frac{(-1)^k}{\sqrt{D_{k-1}D_k}} \left(1 \begin{vmatrix} s_1 & ... & s_k \\ s_2 & ... & s_{k+1} \\ \vdots & \ddots & \vdots \\ s_k & ... & s_{2k-1} \end{vmatrix} + \cdots + (-1)^k \lambda^k \begin{vmatrix} s_0 & ... & s_{k-1} \\ s_1 & ... & s_k \\ \vdots & \ddots & \vdots \\ s_{k-1} & ... & s_{2k-2} \end{vmatrix} \right),$$

entonces

$$\begin{aligned} P_k(\lambda) &= \frac{(-1)^{2k}}{\sqrt{D_{k-1}D_k}} \lambda^k D_{k-1} + R_{k-1}(\lambda) \\ &= \sqrt{\frac{D_{k-1}}{D_k}} \lambda^k + R_{k-1}(\lambda), \end{aligned} \tag{3.5}$$

donde $R_{k-1}(\lambda)$ es un polinomio de grado $k-1$, entonces

$$P_k(\lambda)P_k(\lambda) = P_k(\lambda)\sqrt{\frac{D_{k-1}}{D_k}} \lambda^k + P_k(\lambda)R_{k-1}(\lambda),$$

y finalmente por (3.4) se tiene que

$$\mathcal{S}\{P_k(\lambda)P_k(\lambda)\} = \sqrt{\frac{D_{k-1}}{D_k}} \sqrt{\frac{D_k}{D_{k-1}}} = 1.$$

□

A continuación se verá una relación uno a uno entre las sucesiones definidas positivas normalizadas y las matrices de Jacobi infinitas.

Proposición 3.4. *Existe una relación de correspondencia biunívoca entre las sucesiones definidas positivas normalizadas y las matrices de Jacobi infinitas.*

Demostración. Primero se mostrará que a partir de una sucesión definida positiva normalizada se puede generar una matriz de Jacobi infinita.

Sea pues $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva normalizada. Entonces por la proposición anterior se pueden construir los polinomios ortonormales $P_k(\lambda)$ con respecto a la sucesión definida positiva.

Es fácil verificar que con los polinomios $P_k(\lambda)$ se puede generar cualquier polinomio, es decir, que dado un polinomio arbitrario $R(\lambda)$ de grado $n \geq 0$, se puede escribir de la forma

$$R(\lambda) = \alpha_n P_n(\lambda) + \alpha_{n-1} P_{n-1}(\lambda) + \cdots + \alpha_0 P_0(\lambda),$$

para algunas constantes α_i, con $i = 0, 1, ..., n$. Entonces

$$\lambda P_k(\lambda) = a_{k,k+1} P_{k+1}(\lambda) + a_{k,k} P_k(\lambda) + a_{k,k-1} P_{k-1}(\lambda) + \cdots \qquad (3.6)$$

Por otro lado se tiene que

$$\lambda P_k(\lambda) = \lambda \left(\sqrt{\frac{D_{k-1}}{D_k}}\, \lambda^k + R_{k-1}(\lambda) \right) = \sqrt{\frac{D_{k-1}}{D_k}}\, \lambda^{k+1} + R_k(\lambda),$$

por lo que, comparando los coeficientes principales de $\lambda P_k(\lambda)$ se obtiene la siguiente igualdad

$$a_{k,k+1} \sqrt{\frac{D_k}{D_{k+1}}} = \sqrt{\frac{D_{k-1}}{D_k}},$$

así que

$$a_{k,k+1} = \frac{\sqrt{D_{k-1}D_{k+1}}}{D_k}.$$

Multiplicando ambos lados de (3.6) por $P_i(\lambda)$, con $i = 0, 1, ..., k$ y aplicando el funcional $\mathcal{S}$ se obtiene

$$\mathcal{S}\{\lambda P_k(\lambda)P_i(\lambda)\} = a_{k,k+1}\mathcal{S}\{P_{k+1}(\lambda)P_i(\lambda)\} + a_{k,k}\mathcal{S}\{P_k(\lambda)P_i(\lambda)\} + \cdots ,$$

además

$$\begin{aligned}
\mathcal{S}\{\lambda P_k(\lambda)P_i(\lambda)\} &= \mathcal{S}\left\{P_k(\lambda)\left(\sqrt{\frac{D_{i-1}}{D_i}}\lambda^{i+1} + R_i(\lambda)\right)\right\} \\
&= \sqrt{\frac{D_{i-1}}{D_i}}\,\mathcal{S}\{P_k(\lambda)\,\lambda^{i+1}\} + \mathcal{S}\{P_k(\lambda)R_i(\lambda)\} \\
&= \begin{cases} 0 & \text{si} \quad i = 0, 1, ..., k-2, \\ \sqrt{\dfrac{D_{i-1}D_k}{D_i D_{k-1}}} & \text{si} \quad i = k-1, \end{cases}
\end{aligned}$$

donde $R_i(\lambda)$ es un polinomio sin término constante. Por lo que, igualando estos dos resultados, se puede ver que

$$\begin{aligned}
a_{k,i} &= 0 \quad \text{si} \quad i = 0, 1, ..., k-2, \\
a_{k,k-1} &= \mathcal{S}\{\lambda P_k(\lambda)P_{k-1}(\lambda)\}, \\
a_{k,k} &= \mathcal{S}\{\lambda P_k(\lambda)P_k(\lambda)\}.
\end{aligned}$$

También se puede escribir a $P_{k-1}(\lambda)$ en términos de los polinomios ortonormales, con lo que se tiene

$$\lambda P_{k-1}(\lambda) = a_{k-1,k}P_k(\lambda) + R_{k-1}(\lambda).$$

Multiplicando la ecuación anterior por $P_k(\lambda)$ y aplicando el funcional $\mathcal{S}$ se obtiene

$$\mathcal{S}\{\lambda P_{k-1}(\lambda)P_k(\lambda)\} = a_{k-1,k}.$$

De lo anterior se tiene que

$$a_{k,k-1} = a_{k-1,k},$$

además

$$\begin{aligned}
a_{k,k+1} &= \mathcal{S}\{\lambda P_k(\lambda)P_{k+1}(\lambda)\} \\
&= \mathcal{S}\left\{\lambda\left(\sqrt{\frac{D_{k-1}}{D_k}}\,\lambda^k + R_{k-1}(\lambda)\right)P_{k+1}(\lambda)\right\} \\
&= \mathcal{S}\left\{\sqrt{\frac{D_{k-1}}{D_k}}\,\lambda^{k+1}P_{k+1}(\lambda) + R_k(\lambda)P_{k+1}(\lambda)\right\} \\
&= \sqrt{\frac{D_{k-1}}{D_k}}\frac{1}{\sqrt{D_kD_{k+1}}}\,\mathcal{S}\left\{\lambda^{k+1}\begin{vmatrix} s_0 & s_1 & \dots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_k & s_{k+1} & \dots & s_{2k+1} \\ 1 & \lambda & \dots & \lambda^{k+1}\end{vmatrix}\right\} \\
&= \sqrt{\frac{D_{k-1}}{D_k}}\frac{1}{\sqrt{D_kD_{k+1}}}D_{k+1} \\
&= \frac{\sqrt{D_{k-1}D_{k+1}}}{D_k},
\end{aligned}$$

por lo tanto la expansión (3.6) tiene la forma

$$\lambda P_k(\lambda) = b_{k-1}P_{k-1}(\lambda) + a_kP_k(\lambda) + b_kP_{k+1}(\lambda),$$

con $k = 0, 1, 2, ...$, donde $b_{-1} = 0$ debido a que no hay polinomio ortonormal de grado -1 y

$$a_k = a_{k,k} = \mathcal{S}\{\lambda P_k(\lambda) P_k(\lambda)\}, \quad b_k = a_{k,k+1} = \frac{\sqrt{D_{k-1} D_{k+1}}}{D_k} > 0.$$

Cabe notar que

$$b_{k-1} = a_{k-1,k} = \frac{\sqrt{D_{k-2} D_k}}{D_{k-1}}.$$

Tomando $P_k(\lambda) = y_k$ se obtienen las siguientes ecuaciones

$$b_{k-1} y_{k-1} + a_k y_k + b_k y_{k+1} = \lambda y_k, \tag{3.7}$$

con $k = 0, 1, 2, ...$, y la condición inicial

$$(a_0 - \lambda) y_0 + b_0 y_1 = 0, \tag{3.8}$$

con las cuales se pueden encontrar sucesivamente los valores de y_k, para $k = 0, 1, 2, ...$, si el valor de y_0 está dado. De esta manera las ecuaciones (3.7) y (3.8) determinan la matriz de Jacobi infinita que se estaba buscando.

Ahora se tomará como punto de partida una matriz de Jacobi infinita (3.1). En este caso se pueden construir las ecuaciones (3.7) y tomar la condición inicial (3.8) y entonces poniendo $y_0 = 1 = P_0(\lambda)$, se pueden obtener los polinomios $y_k = P_k(\lambda)$, para $k = 1, 2, 3, ...$

Nuevamente se tiene que cualquier polinomio puede ser expresado en términos de esos polinomios.

Se va a construir el funcional $\mathcal{S}$ que debe ser lineal y su dominio debe ser la clase de todos los polinomios en una variable. Primero se supondrá que

$$\mathcal{S}\{P_i(\lambda) P_k(\lambda)\} = \delta_{ik},$$

es decir, se define el funcional para ciertos polinomios de forma especial.

Para definir el funcional para cualquier polinomio $R(\lambda)$ se debe notar que éste puede ser expresado como el producto de dos polinomios, es decir

$$R(\lambda) = A(\lambda)B(\lambda),$$

donde

$$A(\lambda) = \sum_{i=0}^{m} A_i P_i(\lambda), \quad B(\lambda) = \sum_{k=0}^{n} B_k P_k(\lambda),$$

pues con los polinomios $P_k(\lambda)$ se puede generar cualquier polinomio. Por lo que

$$\begin{aligned} R(\lambda) &= \left(\sum_{i=0}^{m} A_i P_i(\lambda)\right)\left(\sum_{k=0}^{n} B_k P_k(\lambda)\right) \\ &= \sum_{i=0}^{m}\sum_{k=0}^{n} A_i P_i(\lambda) B_k P_k(\lambda), \end{aligned}$$

de donde

$$\mathcal{S}\{R(\lambda)\} = \sum_{i=0}^{m}\sum_{k=0}^{n} A_i B_k \mathcal{S}\{P_i(\lambda)P_k(\lambda)\} = \sum_{j=0}^{r} A_j B_j,$$

para $r = \text{mín}\{m, n\}$.

Sin embargo para que esta definición sea correcta se debe verificar que el resultado que se obtiene es independiente de la forma en que el polinomio $R(\lambda)$ se divide en dos factores. Así, se tiene que mostrar que para dos polinomios arbitrarios $M(\lambda)$ y $N(\lambda)$ la siguiente igualdad se cumple,

$$\mathcal{S}\{A_1(\lambda)B_1(\lambda)\} = \mathcal{S}\{A_2(\lambda)B_2(\lambda)\},$$

en donde

$$A_1(\lambda) = \lambda M(\lambda), \quad B_1(\lambda) = N(\lambda),$$
$$A_2(\lambda) = M(\lambda), \quad B_2(\lambda) = \lambda N(\lambda).$$

Debido a la propiedad de linealidad, es suficiente realizar la prueba para $M(\lambda) = P_m(\lambda)$ y $N(\lambda) = P_n(\lambda)$. Pero en ese caso se tiene que

$$A_1(\lambda) = \lambda P_m(\lambda) = b_m P_{m+1}(\lambda) + a_m P_m(\lambda) + b_{m-1} P_{m-1}(\lambda),$$
$$B_2(\lambda) = \lambda P_n(\lambda) = b_n P_{n+1}(\lambda) + a_n P_n(\lambda) + b_{n-1} P_{n-1}(\lambda),$$

por las condiciones iniciales y entonces

$$\begin{aligned}
\mathcal{S}\{A_1(\lambda)B_1(\lambda)\} &= \mathcal{S}\{\lambda P_m(\lambda) P_n(\lambda)\} \\
&= \begin{cases} 0 & \text{si} \quad n < m-1 \text{ o } n > m+1, \\ b_{m-1} = b_n & \text{si} \quad n = m-1, \\ a_n & \text{si} \quad n = m, \\ b_{m+1} = b_{n-1} & \text{si} \quad n = m+1, \end{cases} \\
&= \mathcal{S}\{P_m(\lambda)\lambda P_n(\lambda)\} \\
&= \mathcal{S}\{A_2(\lambda)B_2(\lambda)\},
\end{aligned}$$

y de esa forma se obtiene la relación requerida.

Habiendo construido el funcional $\mathcal{S}$ se define la sucesión $\{s_k\}_{k=0}^{\infty}$ como

$$s_k = \mathcal{S}\{\lambda^k\} \quad \text{para} \quad k = 0, 1, 2, ..., \tag{3.9}$$

ahora se probará que esta sucesión es definida positiva. Nótese que para el polinomio

$$A(\lambda) = \sum_{i=0}^{m} x_i \lambda^i.$$

se tiene que

$$A^2(\lambda) = \left(\sum_{i=0}^{m} x_i \lambda^i \right) \left(\sum_{k=0}^{m} x_k \lambda^k \right)$$

$$= \sum_{i,k=0}^{m} x_i x_k \lambda^{i+k},$$

y por (3.9)

$$\mathcal{S}_\lambda \{ [A(\lambda)]^2 \} = \sum_{i,k=0}^{m} s_{i+k} x_i x_k.$$

El polinomio $A(\lambda)$ ahora será expresado en términos de los polinomios $P_k(\lambda)$

$$A(\lambda) = \sum_{i=0}^{m} \xi_i P_i(\lambda),$$

por lo que

$$A^2(\lambda) = \left(\sum_{i=0}^{m} \xi_i P_i(\lambda) \right) \left(\sum_{k=0}^{m} \xi_k P_k(\lambda) \right)$$

$$= \sum_{i,k=0}^{m} \xi_i \xi_k P_i(\lambda) P_k(\lambda),$$

y de la construcción del funcional $\mathcal{S}$ se tiene que

$$\mathcal{S} \{ [A(\lambda)]^2 \} = \sum_{i=0}^{m} \xi_i^2,$$

entonces

$$\sum_{i,k=0}^{m} s_{i+k}\, x_i x_k = \sum_{i=0}^{m} \xi_i^2, \tag{3.10}$$

y esta cantidad es positiva siempre que $A(\lambda)$ no sea idénticamente cero. Por lo tanto, cada matriz de Jacobi infinita genera una sucesión definida positiva. $\square$

Así, se ha probado la correspondencia deseada entre las matrices de Jacobi infinitas y las sucesiones definidas positivas normalizadas.

Los polinomios ortonormales introducidos en esta sección así como las ecuaciones (3.7) serán muy utilizados en este documento. Se trabajará con estos elementos para que más adelante se pueda construir la función de distribución asociada al problema de los momentos.

4. Propiedades de los polinomios ortonormales

En lo siguiente se introduce otra sucesión de polinomios que es solución de las ecuaciones (3.7) de la sección anterior, además se muestran propiedades que cumplen ambas sucesiones de polinomios. También se define el concepto de polinomio cuasiortogonal el cual será de utilidad en la siguiente sección.

Para la ecuación diferencial finita

$$b_{k-1}y_{k-1} + a_k y_k + b_k y_{k+1} = \lambda y_k, \quad k = 1, 2, 3, ..., \tag{4.1}$$

donde

$$a_k = \mathcal{S}\{\lambda P_k(\lambda) P_k(\lambda)\}, \quad b_k = \frac{\sqrt{D_{k-1}D_{k+1}}}{D_k} \quad \text{y} \quad b_{k-1} = \frac{\sqrt{D_{k-2}D_k}}{D_{k-1}},$$

anteriormente se introdujeron las soluciones $P_k(\lambda)$, las cuales satisfacen las condiciones iniciales

$$P_0(\lambda) = 1, \quad P_1(\lambda) = \frac{\lambda - a_0}{b_0}, \tag{4.2}$$

y por (3.5) se observa que su coeficiente principal es positivo.

Ahora se introducen otras soluciones las cuales se denotan por $Q_k(\lambda)$, para $k = 0, 1, 2, ...,$ y satisfacen las condiciones iniciales

$$Q_0(\lambda) = 0, \quad Q_1(\lambda) = \frac{1}{b_0}. \tag{4.3}$$

Así, $Q_k(\lambda)$ es un polinomio de grado exacto $k-1$. Estos polinomios cumplen las siguientes propiedades.

Proposición 4.1. *Sean $\{P_k(\lambda)\}_{k=0}^{\infty}$ y $\{Q_k(\lambda)\}_{k=0}^{\infty}$ las sucesiones de polinomios dadas por (4.2), (4.3) y soluciones de las ecuaciones (4.1). Entonces para $k = 0, 1, 2, ...$ se tiene que*

$$Q_k(\lambda) = \mathcal{S}_u \left\{ \frac{P_k(\lambda) - P_k(u)}{\lambda - u} \right\}. \tag{4.4}$$

Demostración. Se comprobará por inducción sobre k de la siguiente manera.

Para $k = 0$ se tiene que

$$\mathcal{S}_u \left\{ \frac{P_0(\lambda) - P_0(u)}{\lambda - u} \right\} = \mathcal{S}_u \left\{ \frac{1 - 1}{\lambda - u} \right\} = \mathcal{S}_u\{0\} = 0 = Q_0(\lambda).$$

Suponiendo que la igualdad (4.4) se cumple para $k = n$, se verá que se cumple para $k = n + 1$, es decir que

$$Q_{k+1}(\lambda) = \mathcal{S}_u \left\{ \frac{P_{k+1}(\lambda) - P_{k+1}(u)}{\lambda - u} \right\}.$$

Puesto que los polinomios $Q_k(\lambda)$ son solución de (4.1) se tiene para $k = n$ que

$$b_{n-1}Q_{n-1}(\lambda) + a_nQ_n(\lambda) + b_nQ_{n+1}(\lambda) = \lambda Q_n(\lambda).$$

Entonces por la hipótesis inductiva, la propiedad de linealidad del funcional $\mathcal{S}$ y que los polinomios $P_k(\lambda)$ también son solución de (4.1) se tiene que

$$\begin{aligned}
b_n Q_{n+1}(\lambda) &= \lambda Q_n(\lambda) - b_{n-1}Q_{n-1}(\lambda) - a_n Q_n(\lambda) \\
&= (\lambda - a_n)Q_n(\lambda) - b_{n-1}Q_{n-1}(\lambda) \\
&= (\lambda - a_n)\mathcal{S}_u\left\{\frac{P_n(\lambda) - P_n(u)}{\lambda - u}\right\} - b_{n-1}\mathcal{S}_u\left\{\frac{P_{n-1}(\lambda) - P_{n-1}(u)}{\lambda - u}\right\} \\
&= \mathcal{S}_u\left\{\frac{(\lambda - a_n)(P_n(\lambda) - P_n(u)) - b_{n-1}(P_{n-1}(\lambda) - P_{n-1}(u))}{\lambda - u}\right\} \\
&= \mathcal{S}_u\left\{\frac{b_n P_{n+1}(\lambda) - b_n P_{n+1}(u)}{\lambda - u}\right\} \\
&= b_n\mathcal{S}_u\left\{\frac{P_{n+1}(\lambda) - P_{n+1}(u)}{\lambda - u}\right\},
\end{aligned}$$

lo cual concluye la prueba.

$\square$

El siguiente resultado es análogo a la fórmula de Liouville-Ostrogradskii [4].

Proposición 4.2. *Sean* $\{P_k(\lambda)\}_{k=0}^{\infty}$ *y* $\{Q_k(\lambda)\}_{k=0}^{\infty}$ *las sucesiones de polinomios dadas por (4.2), (4.3) y soluciones de las ecuaciones (4.1). Entonces para toda* $k = 1, 2, 3, \ldots$ *se cumple que*

$$P_{k-1}(\lambda)Q_k(\lambda) - P_k(\lambda)Q_{k-1}(\lambda) = \frac{1}{b_{k-1}}. \tag{4.5}$$

Demostración. La prueba se hará por inducción sobre k.

Si $k = 1$ se tiene que

$$P_0(\lambda)Q_1(\lambda) - P_1(\lambda)Q_0(\lambda) = 1\left(\frac{1}{b_0}\right) - \left(\frac{\lambda - a_0}{b_0}\right)0 = \frac{1}{b_0}.$$

Suponiendo que la igualdad (4.5) se cumple para $k = n$, se verá que se cumple para $k = n + 1$, es decir que

$$P_n(\lambda)Q_{n+1}(\lambda) - P_{n+1}(\lambda)Q_n(\lambda) = \frac{1}{b_n}.$$

Como los polinomios $Q_k(\lambda)$ son solución de las ecuaciones (4.1) se tiene que

$$b_{n-1}Q_{n-1}(\lambda) + a_nQ_n(\lambda) + b_nQ_{n+1}(\lambda) = \lambda Q_n(\lambda),$$

por lo que

$$Q_{n+1}(\lambda) = \frac{(\lambda - a_n)Q_n(\lambda) - b_{n-1}Q_{n-1}(\lambda)}{b_n},$$

y de la misma forma se cumple que

$$P_{n+1}(\lambda) = \frac{(\lambda - a_n)P_n(\lambda) - b_{n-1}P_{n-1}(\lambda)}{b_n},$$

pues los polinomios $P_k(\lambda)$ también son solución de las ecuaciones (4.1). Así que, usando la hipótesis de inducción,

$$\begin{aligned}
P_n(\lambda)Q_{n+1}(\lambda) - P_{n+1}(\lambda)Q_n(\lambda) &= P_n(\lambda)\left(\frac{(\lambda - a_n)Q_n(\lambda) - b_{n-1}Q_{n-1}(\lambda)}{b_n}\right) \\
&\quad - \left(\frac{(\lambda - a_n)P_n(\lambda) - b_{n-1}P_{n-1}(\lambda)}{b_n}\right)Q_n(\lambda) \\
&= \frac{b_{n-1}P_{n-1}(\lambda)Q_n(\lambda) - b_{n-1}Q_{n-1}(\lambda)P_n(\lambda)}{b_n} \\
&= \frac{b_{n-1}(P_{n-1}(\lambda)Q_n(\lambda) - Q_{n-1}(\lambda)P_n(\lambda))}{b_n} \\
&= \left(\frac{b_{n-1}}{b_n}\right)\left(\frac{1}{b_{n-1}}\right) = \frac{1}{b_n}.
\end{aligned}$$

□

Ahora se demostrará un resultado análogo a la fórmula de Green [2].

Proposición 4.3. *Sean y_k soluciones de (4.1) con el parámetro λ. Sean z_k soluciones de la misma ecuación pero con el parámetro μ. Ambas soluciones para $k = 0, 1, 2, ...$ Entonces se tiene que*

$$(\mu - \lambda)\sum_{k=m}^{n-1} y_k z_k = b_{n-1}(y_{n-1}z_n - y_n z_{n-1}) - b_{m-1}(y_{m-1}z_m - y_m z_{m-1}), \quad (4.6)$$

Demostración. Se tiene que

$$\begin{aligned}
(\mu - \lambda)\sum_{k=m}^{n-1} y_k z_k &= \sum_{k=m}^{n-1} \mu y_k z_k - \lambda y_k z_k \\
&= \sum_{k=m}^{n-1} (b_{k-1}z_{k-1} + a_k z_k + b_k z_{k+1})y_k \\
&\quad - (b_{k-1}y_{k-1} + a_k y_k + b_k y_{k+1})z_k \\
&= \sum_{k=m}^{n-1} b_{k-1}z_{k-1}y_k + a_k z_k y_k + b_k z_{k+1} y_k \\
&\quad - b_{k-1}y_{k-1}z_k - a_k y_k z_k - b_k y_{k+1} z_k \\
&= \sum_{k=m}^{n-1} b_k(z_{k+1}y_k - y_{k+1}z_k) - b_{k-1}(y_{k-1}z_k - z_{k-1}y_k) \\
&= b_m(z_{m+1}y_m - y_{m+1}z_m) - b_{m-1}(y_{m-1}z_m - z_{m-1}y_m) \\
&\quad + b_{m+1}(z_{m+2}y_{m+1} - y_{m+2}z_{m+1}) - b_m(y_m z_{m+1} - z_m y_{m+1}) \\
&\quad + \cdots + b_{n-1}(z_n y_{n-1} - y_n z_{n-1}) - b_{n-2}(y_{n-2}z_{n-1} - z_{n-2}y_{n-1}) \\
&= b_{n-1}(z_n y_{n-1} - y_n z_{n-1}) - b_{m-1}(y_{m-1}z_m - z_{m-1}y_m).
\end{aligned}$$

□

La siguiente propiedad es un caso particular de la proposición anterior.

Proposición 4.4. *(Fórmula de Christoffel-Darboux)*
Sea $\{P_k(\lambda)\}_{k=0}^{\infty}$ *la sucesión de polinomios dada por (4.2) y solución de las ecuaciones (4.1). Entonces se cumple que*

$$(\mu - \lambda) \sum_{k=0}^{n-1} P_k(\lambda) P_k(\mu) = b_{n-1}(P_{n-1}(\lambda)P_n(\mu) - P_n(\lambda)P_{n-1}(\mu)). \qquad (4.7)$$

Demostración. En (4.6) tomando $y_k = P_k(\lambda),\ z_k = P_k(\mu)$ y $m = 0$, dado que $b_{-1} = 0$, se obtiene el resultado deseado.

□

A continuación se define otro concepto el cual ayudará a mostrar más propiedades de los polinomios ortonormales.

Definición 4.5. *Sea* $\{P_k(\lambda)\}_{k=0}^{\infty}$ *la sucesión de polinomios ortonormales para una sucesión definida positiva* $\{s_k\}_{k=0}^{\infty}$. *Se llamará kernel polinomial de grado* n, *con* $n = 0, 1, 2, ...,$ *a la suma*

$$h_n(\lambda, \mu) = \sum_{k=0}^{n} P_k(\lambda) P_k(\mu). \qquad (4.8)$$

Proposición 4.6. *Sea* $\{P_k(\lambda)\}_{k=0}^{\infty}$ *la sucesión de polinomios ortonormales dada por (4.2) y solución de las ecuaciones (4.1). Entonces para* $n = 0, 1, 2, ...,$ *se tiene que*

$$\mathcal{S}_\lambda\{h_n(\lambda, \mu) P_k(\lambda)\} = P_k(\mu), \quad \text{donde} \quad k = 0, 1, ..., n.$$

Demostración. Usando la ortonormalidad de los polinomios $P_k(\lambda)$ y la linealidad del funcional $\mathcal{S}$ se tiene que

$$
\begin{aligned}
\mathcal{S}_\lambda\{h_n(\lambda,\mu)P_k(\lambda)\} &= \mathcal{S}_\lambda\left\{\left(\sum_{j=0}^{n} P_j(\lambda)P_j(\mu)\right) P_k(\lambda)\right\} \\
&= \sum_{j=0}^{n} \mathcal{S}_\lambda\{P_j(\lambda)P_j(\mu)P_k(\lambda)\} \\
&= \sum_{j=0}^{n} P_j(\mu)\, \mathcal{S}_\lambda\{P_j(\lambda)P_k(\lambda)\} \\
&= P_k(\mu).
\end{aligned}
$$

□

Proposición 4.7. *Sea $R(\lambda)$ un polinomio de grado menor o igual a n, para $n = 0, 1, 2, ...$ Entonces se cumple que*

$$\mathcal{S}_\lambda\{h_n(\lambda,\mu)R(\lambda)\} = R(\mu).$$

Demostración. Sea $R(\lambda)$ un polinomio de grado menor o igual a n, para alguna $n \geq 0$. Entonces se puede escribir a $R(\lambda)$ como $R(\lambda) = \sum_{k=0}^{n} a_k P_k(\lambda)$. Además, por la propiedad anterior se tiene que

$$
\begin{aligned}
\mathcal{S}_\lambda\{h_n(\lambda,\mu)R(\lambda)\} &= \mathcal{S}_\lambda\left\{h_n(\lambda,\mu)\sum_{k=0}^{n} a_k P_k(\lambda)\right\} \\
&= \sum_{k=0}^{n} a_k \mathcal{S}_\lambda\{h_n(\lambda,\mu)P_k(\lambda)\} \\
&= \sum_{k=0}^{n} a_k P_k(\mu) \\
&= R(\mu).
\end{aligned}
$$

□

El siguiente concepto generaliza a los polinomios ortogonales $P_k(\lambda)$, además será de mucha utilidad en la sección posterior.

Definición 4.8. *Sea $\{P_n(\lambda)\}_{n=0}^{\infty}$ la sucesión de polinomios ortonormales para una sucesión definida positiva $\{s_n\}_{n=0}^{\infty}$. Se llama polinomio cuasiortogonal de grado n, para cualquier real τ, al polinomio*

$$P_n(\lambda, \tau) = P_n(\lambda) - \tau P_{n-1}(\lambda), \quad \textit{para} \quad n = 1, 2, 3, ...$$

También para la sucesión de polinomios $\{Q_n(\lambda)\}_{n=0}^{\infty}$ se puede definir de manera análoga el polinomio cuasiortogonal de grado n, con $n = 1, 2, 3, ...$ y τ un número real, como

$$Q_n(\lambda, \tau) = Q_n(\lambda) - \tau Q_{n-1}(\lambda).$$

A continuación se verá una propiedad de ortogonalidad que cumplen los polinomios cuasiortogonales.

Proposición 4.9. *Sea $\{P_n(\lambda, \tau)\}_{n=1}^{\infty}$ la sucesión de polinomios cuasiortogonales asociada a la sucesión de polinomios $\{P_n(\lambda)\}_{n=0}^{\infty}$. Entonces para $k = 0, 1, ..., n-2$ se cumple la siguiente relación de ortogonalidad*

$$\mathcal{S}\{P_n(\lambda, \tau)\lambda^k\} = 0. \tag{4.9}$$

Además si $\tau = 0$ la propiedad se cumple también para $k = n - 1$.

Demostración. Por (3.4) se tiene que

$$\begin{aligned}
\mathcal{S}\{P_n(\lambda, \tau)\lambda^k\} &= \mathcal{S}\{(P_n(\lambda) - \tau P_{n-1}(\lambda))\lambda^k\} \\
&= \mathcal{S}\{P_n(\lambda)\lambda^k\} - \tau \mathcal{S}\{P_{n-1}(\lambda)\lambda^k\} \\
&= 0.
\end{aligned}$$

para $k = 0, 1, ..., n-2$.

Si $\tau = 0$, por las igualdades anteriores se tiene que

$$\mathcal{S}\{P_n(\lambda, \tau)\lambda^k\} = \mathcal{S}\{P_n(\lambda)\lambda^k\},$$

lo cual también se anula para $k = n - 1$.

$\square$

El siguiente teorema muestra como son las raíces de los polinomios cuasiortogonales y de los polinomios $Q_n(\lambda)$. Además muestra relaciones para las raíces de los polinomios ortonormales y las raíces de los polinomios $Q_n(\lambda)$.

Teorema 4.10. *Sean $\{P_n(\lambda)\}_{n=0}^{\infty}$ y $\{Q_n(\lambda)\}_{n=0}^{\infty}$ las sucesiones de polinomios dadas por (4.2), (4.3) y soluciones de las ecuaciones (4.1). Sea $\{P_n(\lambda, \tau)\}_{n=0}^{\infty}$ la sucesión de polinomios cuasiortogonales asociada a los polinomios $\{P_n(\lambda)\}_{n=0}^{\infty}$. Entonces*

(1) Todas las raíces de un polinomio cuasiortogonal son reales y simples.

(2) Cualesquiera dos raíces del polinomio $P_n(\lambda)$ están separadas por una raíz del polinomio $P_{n-1}(\lambda)$ y viceversa.

(3) Las raíces del polinomio $Q_n(\lambda)$ son reales y simples y cualesquiera dos de ellas están separadas por una raíz del polinomio $P_n(\lambda)$ y viceversa.

Demostración.

(1) Supóngase que el polinomio cuasiortogonal $P_n(\lambda, \tau)$ cambia de signo en los siguientes puntos del eje real

$$\lambda_1 < \lambda_2 < \cdots < \lambda_m,$$

donde $1 \leq m \leq n$. Entonces cada λ_k para $k = 1, 2, ..., m$ es una raíz de multiplicidad impar del polinomio cuasiortogonal, y en caso de que tuviera más raíces reales o complejas serían de multiplicidad par. Por lo anterior el polinomio

$$R(\lambda) = P_n(\lambda, \tau)(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_m),$$

cumple que todas sus raíces tienen multiplicidad par, por lo que es no negativo en el eje real, además es no idénticamente cero. Entonces para una sucesión definida positiva su funcional correspondiente cumple que $\mathcal{S}\{R(\lambda)\} > 0$.

Pero esto contradice la relación de ortogonalidad (4.9) si $m \leq n-2$, pues para esos casos se tiene por (3.4) que

$$\begin{aligned}
\mathcal{S}\{R(\lambda)\} &= \mathcal{S}\{P_n(\lambda,\tau)(\lambda^m + R_{m-1}(\lambda))\} \\
&= \mathcal{S}\{P_n(\lambda,\tau)\lambda^m\} + \mathcal{S}\{P_n(\lambda,\tau)R_{m-1}(\lambda)\} \\
&= 0,
\end{aligned}$$

donde $R_{m-1}(\lambda)$ es un polinomio de grado $m-1$. Por lo que $m \geq n-1$, es decir el polinomio $P_n(\lambda,\tau)$ tiene por lo menos $n-1$ raíces reales de grado impar que deben ser simples, pues por la cantidad de ellas no pueden tener multiplicidad mayor. Por lo tanto todas sus raíces son reales y simples, ya que si $m = n-1$ el número de raíces simples no permite que hubiera una de grado mayor o igual a dos, lo cual también impide que la raíz restante fuera complejo pues su conjugado también sería raíz, con lo cual queda demostrada la primera afirmación.

(2) De la fórmula (4.7) de Christoffel-Darboux se tiene que

$$\begin{aligned}
\sum_{k=0}^{n-1} P_k(\lambda)P_k(\mu) &= b_{n-1}\frac{P_{n-1}(\lambda)P_n(\mu) - P_n(\lambda)P_{n-1}(\mu)}{\mu-\lambda} \\
&= b_{n-1}\frac{P_{n-1}(\lambda)P_n(\mu) - P_{n-1}(\lambda)P_n(\lambda)}{\mu-\lambda} \\
&\quad + b_{n-1}\frac{P_{n-1}(\lambda)P_n(\lambda) - P_n(\lambda)P_{n-1}(\mu)}{\mu-\lambda} \\
&= b_{n-1}\left(P_{n-1}(\lambda)\frac{P_n(\mu) - P_n(\lambda)}{\mu-\lambda}\right) \\
&\quad - b_{n-1}\left(P_n(\lambda)\frac{P_{n-1}(\mu) - P_{n-1}(\lambda)}{\mu-\lambda}\right),
\end{aligned}$$

por lo que, cuando $\mu \longrightarrow \lambda$

$$\sum_{k=0}^{n-1}[P_k(\lambda)]^2 = b_{n-1}[P_{n-1}(\lambda)P_n'(\lambda) - P_n(\lambda)P_{n-1}'(\lambda)]. \tag{4.10}$$

Luego, si λ_1 y λ_2 son dos raíces consecutivas del polinomio $P_n(\lambda)$, entonces los números $P_n'(\lambda_1)$ y $P_n'(\lambda_2)$ son de signo diferente. Además por la igualdad anterior se tiene que

$$b_{n-1}P_{n-1}(\lambda_1)P_n'(\lambda_1) = \sum_{k=0}^{n-1}[P_k(\lambda_1)]^2 > 0,$$

$$b_{n-1}P_{n-1}(\lambda_2)P_n'(\lambda_2) = \sum_{k=0}^{n-1}[P_k(\lambda_2)]^2 > 0.$$

Por lo tanto, los números $P_{n-1}(\lambda_1)$ y $P_{n-1}(\lambda_2)$ son de signo diferente, así que entre λ_1 y λ_2 hay una raíz de $P_{n-1}(\lambda)$ con lo cual se sigue la afirmación (2).

(3) Esta afirmación se prueba de manera análoga al procedimiento empleado en las demostraciones de las propiedades (1) y (2), usando la ecuación (4.5).
$\square$

Anteriormente se definió el concepto de polinomios ortonormales, que en esta sección se extendió con los polinomios cuasiortogonales. Estos últimos polinomios servirán para la construcción de la función de distribución asociada al problema de los momentos, por lo que más adelante jugarán un papel importante.

5. Fórmula de cuadratura

Ahora se obtendrá una fórmula llamada de cuadratura. Esta fórmula tiene coeficientes particulares que ayudarán a obtener una solución del problema de los momentos y están construidos con los polinomios cuasiortogonales. Se muestran además dos representaciones adicionales de dichos coeficientes.

Teorema 5.1. *(Fórmula de cuadratura)*
Sea $P_n(\lambda, \tau)$ un polinomio cuasiortogonal de grado n, con $n \in \mathbb{N} \setminus \{0\}$, para algún τ real y sean $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ las raíces de ese polinomio. Entonces para cualquier polinomio $R(\lambda)$ de grado menor o igual a $2n-2$, se tiene que

$$\mathcal{S}\{R(\lambda)\} = \sum_{k=1}^{n} \mu_k R(\lambda_k), \tag{5.1}$$

donde

$$\mu_k = \mu_k^{(n)}(\tau) = \frac{Q_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)}. \tag{5.2}$$

Demostración. Sea $P_n(\lambda, \tau)$ un polinomio cuasiortogonal de grado n, es decir,

$$P_n(\lambda, \tau) = P_n(\lambda) - \tau P_{n-1}(\lambda),$$

donde τ es un número real. Sean $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ las raíces del polinomio. Así que

$$P_n(\lambda, \tau) = c \prod_{j=1}^{n} (\lambda - \lambda_j),$$

con c el coeficiente del término principal.

Sea $R_{2n-2}(\lambda)$ un polinomio de grado $2n-2$. Entonces

$$R_{2n-2}(\lambda) = P_n(\lambda, \tau) R_{n-2}(\lambda) + R_{n-1}(\lambda), \tag{5.3}$$

donde $R_{n-2}(\lambda)$ y $R_{n-1}(\lambda)$ son algunos polinomios de grado menor o igual a $n-2$ y menor o igual a $n-1$, respectivamente.

Se puede ver que para λ_k con $k = 1, 2, ..., n$ se cumple la igualdad de los polinomios $R_{2n-2}(\lambda_k) = R_{n-1}(\lambda_k)$.

Por la fórmula de interpolación de Lagrange, que se muestra en la segunda sección del apéndice de este trabajo, se tiene para $(\lambda_1, R_{n-1}(\lambda_1)), (\lambda_2, R_{n-1}(\lambda_2))$, ..., $(\lambda_n, R_{n-1}(\lambda_n))$ que

$$\begin{aligned} R_{n-1}(\lambda) &= \sum_{k=1}^{n} \frac{\prod_{j=1}^{n}(\lambda-\lambda_j)}{(\lambda-\lambda_k)\prod_{\substack{j=1\\ j\neq k}}^{n}(\lambda_k-\lambda_j)}\, R_{n-1}(\lambda_k) \\ &= \sum_{k=1}^{n} \frac{P_n(\lambda,\tau)}{(\lambda-\lambda_k)P_n'(\lambda_k,\tau)}\, R_{n-1}(\lambda_k) \\ &= \sum_{k=1}^{n} \frac{P_n(\lambda,\tau)R_{2n-2}(\lambda_k)}{(\lambda-\lambda_k)P_n'(\lambda_k,\tau)}. \end{aligned}$$

Por otro lado, utilizando la propiedad de ortogonalidad (4.9) de los polinomios cuasiortogonales en (5.3) se tiene que

$$\mathcal{S}_\lambda\{R_{2n-2}(\lambda)\} = \mathcal{S}_\lambda\{R_{n-1}(\lambda)\}, \tag{5.4}$$

por lo tanto,

$$\mathcal{S}_\lambda\{R_{2n-2}(\lambda)\} = \sum_{k=1}^{n} \frac{R_{2n-2}(\lambda_k)}{P_n'(\lambda_k,\tau)}\, \mathcal{S}_\lambda\left\{\frac{P_n(\lambda,\tau)}{\lambda-\lambda_k}\right\},$$

y ya que λ_k es una raíz del polinomio $P_n(\lambda,\tau)$ se tiene por (4.4) que

$$\begin{aligned} \mathcal{S}_\lambda\left\{\frac{P_n(\lambda,\tau)}{\lambda-\lambda_k}\right\} &= \mathcal{S}_\lambda\left\{\frac{P_n(\lambda,\tau)-P_n(\lambda_k,\tau)}{\lambda-\lambda_k}\right\} \\ &= \mathcal{S}_\lambda\left\{\frac{P_n(\lambda)-\tau P_{n-1}(\lambda)-P_n(\lambda_k)+\tau P_{n-1}(\lambda_k)}{\lambda-\lambda_k}\right\} \\ &= Q_n(\lambda_k)-\tau Q_{n-1}(\lambda_k) = Q_n(\lambda_k,\tau). \end{aligned}$$

Entonces

$$\mathcal{S}_\lambda\{R_{2n-2}(\lambda)\} = \sum_{k=1}^{n} \frac{Q_n(\lambda_k,\tau)}{P_n'(\lambda_k,\tau)}\ R_{2n-2}(\lambda_k).$$

Por lo tanto, lo anterior se cumple para cualquier polinomio $R(\lambda)$ de grado menor o igual a $2n-2$.

□

Para $\tau = 0$, es decir para $P_n(\lambda,\tau) = P_n(\lambda)$, la fórmula de cuadratura también es válida para cualquier polinomio $R(\lambda)$ de grado menor o igual a $2n-1$. Para esto basta utilizar la representación

$$R_{2n-1}(\lambda) = P_n(\lambda,\tau)R_{n-1}(\lambda) + R_{n-1}(\lambda),$$

y notar que (5.4) también se cumple para $\tau = 0$ por una observación de la propiedad de ortogonalidad (4.9).

Además de la propiedad establecida en (5.2), los coeficientes μ_k tienen otras dos representaciones importantes. La primera representación se verá en la siguiente proposición.

Proposición 5.2. *Sea $P_n(\lambda,\tau)$ un polinomio cuasiortogonal como en el Teorema 5.1. Entonces los coeficientes (5.2) de la fórmula de cuadratura también se pueden expresar de la siguiente forma*

$$\mu_k = \frac{1}{\sum\limits_{i=0}^{n-1} [P_i(\lambda_k)]^2}.$$

Demostración. Para obtener esta representación se usa el hecho de que

$$P_n(\lambda_k,\tau) = P_n(\lambda_k) - \tau P_{n-1}(\lambda_k) = 0,$$

por lo que

$$\tau = \frac{P_n(\lambda_k)}{P_{n-1}(\lambda_k)},$$

si $P_{n-1}(\lambda_k) \neq 0$, y usando lo anterior en (5.2) se obtiene que

$$\mu_k = \frac{Q_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)} = \frac{Q_n(\lambda_k) - \tau Q_{n-1}(\lambda_k)}{P_n'(\lambda_k) - \tau P_{n-1}'(\lambda_k)}$$

$$= \frac{Q_n(\lambda_k) - \dfrac{P_n(\lambda_k)}{P_{n-1}(\lambda_k)} Q_{n-1}(\lambda_k)}{P_n'(\lambda_k) - \dfrac{P_n(\lambda_k)}{P_{n-1}(\lambda_k)} P_{n-1}'(\lambda_k)}$$

$$= \frac{P_{n-1}(\lambda_k) Q_n(\lambda_k) - P_n(\lambda_k) Q_{n-1}(\lambda_k)}{P_{n-1}(\lambda_k) P_n'(\lambda_k) - P_n(\lambda_k) P_{n-1}'(\lambda_k)}.$$

El numerador de la última fórmula es igual a $1/b_{n-1}$ por (4.5) y usando el resultado obtenido en (4.10) se obtiene que

$$\begin{aligned} \mu_k &= \frac{P_{n-1}(\lambda_k) Q_n(\lambda_k) - P_n(\lambda_k) Q_{n-1}(\lambda_k)}{P_{n-1}(\lambda_k) P_n'(\lambda_k) - P_n(\lambda_k) P_{n-1}'(\lambda_k)} \\ &= \frac{\dfrac{1}{b_{n-1}}}{\dfrac{1}{b_{n-1}} \sum\limits_{i=0}^{n-1} [P_i(\lambda_k)]^2} = \frac{1}{\sum\limits_{i=0}^{n-1} |P_i(\lambda_k)|^2}, \end{aligned} \tag{5.5}$$

donde se ve que todos los coeficientes μ_k son positivos.

□

La segunda representación de los coeficientes de la fórmula de cuadratura se muestra en la siguiente proposición.

Proposición 5.3. *Sea $P_n(\lambda, \tau)$ un polinomio cuasiortogonal como en el Teorema 5.1. Entonces los coeficientes (5.2) de la fórmula de cuadratura también tienen la siguiente representación*

$$\mu_k = \mathcal{S}_\lambda \left\{ \left[\frac{P_n(\lambda, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)} \right]^2 \right\}. \tag{5.6}$$

Demostración. Para conseguir esta representación se considera para cada $k = 1, 2, ..., n$, el polinomio

$$R(\lambda) = \left[\frac{P_n(\lambda, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)}\right]^2,$$

el cual es de grado $2n - 2$, además es igual a 1 para $\lambda = \lambda_k$ pues

$$\begin{aligned} R(\lambda) &= \left[\frac{P_n(\lambda, \tau) - P_n(\lambda_k, \tau) + P_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)}\right]^2 \\ &= \left[\frac{P_n(\lambda, \tau) - P_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)} + \frac{P_n(\lambda_k, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)}\right]^2 \\ &= \left[\frac{1}{P_n'(\lambda_k, \tau)}\right]^2 \left[\frac{P_n(\lambda, \tau) - P_n(\lambda_k, \tau)}{(\lambda - \lambda_k)}\right]^2, \end{aligned}$$

ya que en la segunda igualdad el segundo sumando es cero. Entonces tomando el límite cuando $\lambda \longrightarrow \lambda_k$ de ambos lados se tiene que

$$R(\lambda_k) = \left[\frac{1}{P_n'(\lambda_k, \tau)}\right]^2 [P_n'(\lambda_k, \tau)]^2 = 1,$$

y claramente $R(\lambda)$ es cero en las otras raíces de $P_n(\lambda, \tau)$. Usando la fórmula de cuadratura (5.1) para $R(\lambda)$ se tiene que

$$\begin{aligned} \mathcal{S}_\lambda\{R(\lambda)\} &= \mathcal{S}_\lambda\left\{\left[\frac{P_n(\lambda, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)}\right]^2\right\} \\ &= \sum_{i=0}^{n} \mu_i R(\lambda_i) = \mu_k, \end{aligned}$$

por lo que se obtiene la representación

$$\mu_k = \mathcal{S}_\lambda\left\{\left[\frac{P_n(\lambda, \tau)}{P_n'(\lambda_k, \tau)(\lambda - \lambda_k)}\right]^2\right\}, \tag{5.7}$$

donde se observa nuevamente el signo positivo de μ_k debido a que el funcional $\mathcal{S}$ es positivo.

$\square$

Proposición 5.4. *Sea $P_n(\lambda,\tau)$ un polinomio cuasiortogonal como en el Teorema 5.1. Entonces para cualquier real τ se cumple la siguiente expansión*

$$\frac{Q_n(\lambda,\tau)}{P_n(\lambda,\tau)} = \frac{s_0}{\lambda} + \frac{s_1}{\lambda^2} + \cdots + \frac{s_{2n-2}}{\lambda^{2n-1}} + O\left(\frac{1}{\lambda^{2n}}\right). \tag{5.8}$$

Además para $\tau = 0$ esta expansión puede ser reemplazada por

$$\frac{Q_n(\lambda)}{P_n(\lambda)} = \frac{s_0}{\lambda} + \frac{s_1}{\lambda^2} + \cdots + \frac{s_{2n-1}}{\lambda^{2n}} + O\left(\frac{1}{\lambda^{2n+1}}\right). \tag{5.9}$$

Demostración. Para ver la primera expansión se debe notar que por la fórmula de interpolación de Lagrange, mostrada en la segunda sección del apéndice de este trabajo, se tiene para $(\lambda_1, Q_n(\lambda_1,\tau)), (\lambda_2, Q_n(\lambda_2,\tau)), ..., (\lambda_n, Q_n(\lambda_n,\tau))$

$$Q_n(\lambda,\tau) = \sum_{k=1}^{n} \frac{P_n(\lambda,\tau)}{P_n'(\lambda_k,\tau)(\lambda-\lambda_k)} Q_n(\lambda_k,\tau).$$

Entonces

$$\frac{Q_n(\lambda,\tau)}{P_n(\lambda,\tau)} = \sum_{k=1}^{n} \frac{Q_n(\lambda_k,\tau)}{P_n'(\lambda_k,\tau)(\lambda-\lambda_k)}.$$

Por la expresión de los coeficientes μ_k de la ecuación (5.2)

$$\frac{Q_n(\lambda,\tau)}{P_n(\lambda,\tau)} = \sum_{k=1}^{n} \frac{\mu_k}{\lambda-\lambda_k}, \tag{5.10}$$

donde cada sumando se puede expresar como

$$\frac{\mu_k}{\lambda-\lambda_k}=\frac{\mu_k}{\lambda}+\frac{\lambda_k\mu_k}{\lambda(\lambda-\lambda_k)}.$$

Por lo que

$$\begin{aligned}
\sum_{k=1}^{n}\frac{\mu_k}{\lambda-\lambda_k} &= \sum_{k=1}^{n}\frac{\mu_k}{\lambda}+\frac{\lambda_k\mu_k}{\lambda(\lambda-\lambda_k)}\\
&= \frac{1}{\lambda}\sum_{k=1}^{n}\mu_k+\sum_{k=1}^{n}\left(\frac{\lambda_k}{\lambda}\right)\left(\frac{\mu_k}{\lambda-\lambda_k}\right)\\
&= \frac{1}{\lambda}\sum_{k=1}^{n}\mu_k+\sum_{k=1}^{n}\left(\frac{\lambda_k}{\lambda}\right)\left[\frac{\mu_k}{\lambda}+\frac{\lambda_k\mu_k}{\lambda(\lambda-\lambda_k)}\right]\\
&= \frac{1}{\lambda}\sum_{k=1}^{n}\mu_k+\sum_{k=1}^{n}\frac{\lambda_k\mu_k}{\lambda^2}+\frac{\lambda_k^2\mu_k}{\lambda^2(\lambda-\lambda_k)}\\
&= \frac{1}{\lambda}\sum_{k=1}^{n}\mu_k+\frac{1}{\lambda^2}\sum_{k=1}^{n}\lambda_k\mu_k+\sum_{k=1}^{n}\left(\frac{\lambda_k^2}{\lambda^2}\right)\left(\frac{\mu_k}{\lambda-\lambda_k}\right),
\end{aligned}$$

y continuando sucesivamente se obtiene que

$$\sum_{k=1}^{n}\frac{\mu_k}{\lambda-\lambda_k}=\frac{1}{\lambda}\sum_{k=1}^{n}\mu_k+\frac{1}{\lambda^2}\sum_{k=1}^{n}\mu_k\lambda_k+\cdots+\frac{1}{\lambda^{m+1}}\sum_{k=1}^{n}\mu_k\lambda_k^m+\cdots \tag{5.11}$$

Por último falta notar que, por (5.1),

$$\sum_{k=1}^{n}\mu_k\lambda_k^m=\mathcal{S}\{\lambda^m\}=s_m, \tag{5.12}$$

para $m=0,1,...,2n-2$, por lo que sustituyendo en (5.11) se obtiene el resultado buscado. Si $\tau=0$ la última relación es válida también para $m=2n-1$, así que (5.9) también está demostrada.

$\square$

Con los conceptos vistos hasta ahora, finalmente será posible mostrar una solución al problema de los momentos. Como ya se ha mencionado, los coeficientes de la fórmula de cuadratura serán indispensables para construir dicha solución, lo cual se verá en la siguiente sección.

6. Un criterio de solubilidad del problema de los momentos

A continuación se muestra un criterio de solución al problema de los momentos. Para ello se agrega un requerimiento extra a las condiciones iniciales. Pero primero se recordará el enunciado del problema principal estudiado en este trabajo.

Dada una sucesión infinita de números reales s_k, con $k = 0, 1, 2, ...$, se busca una función no decreciente $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ que satisfaga las ecuaciones

$$s_k = \int_{-\infty}^{\infty} u^k d\alpha(u). \tag{6.1}$$

Y ahora, antes de continuar con el criterio de solución, se mencionará una característica importante del problema.

Definición 6.1. *Dos soluciones del problema de los momentos se consideran iguales si su diferencia es una constante en todos los puntos donde esa diferencia sea continua. El problema de los momentos es determinado si, dado el criterio anterior, tiene una única solución, e indeterminado de otra forma.*

Existen criterios con los cuales se puede establecer si el problema de los momentos es determinado o indeterminado. La exposición en el trabajo [9] profundiza en estos detalles y además presenta ejemplos de los dos casos. Uno de esos ejemplos está dado por la función

$$f(x) = \frac{\beta}{2\,\Gamma(1/\beta)}\, e^{-|x|^{\beta}},$$

para $x \in \mathbb{R}$ y $\beta > 0$, el cual es indeterminado si $0 < \beta < 1$ y determinado si $\beta \geq 1$.

En las ecuaciones (6.1), si la función α incrementa en un número finito de puntos, las integrales que definen la sucesión de momentos se convierten en sumas finitas. Como ejemplo de lo anterior se puede considerar una variable aleatoria X con distribución Ber(1/2) para la cual

$$s_k = \sum_{x=0}^{1} x^k (1/2)^x (1/2)^{1-x} = 1/2,$$

para cualquier entero positivo k.

En el presente trabajo se discutirá el caso en donde las ecuaciones (6.1) no se convierten en sumas finitas por lo que, como se mencionó anteriormente, se agrega un requerimiento adicional al problema. Dicho requerimiento es que la función α tenga un número infinito de puntos de incremento. De esta forma, utilizando el concepto de límite lateral, visto en el primer apéndice de este trabajo, se define un punto de incremento de una función de la siguiente manera.

Definición 6.2. *Sea $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ una función. Se dice que x_0 es un punto de incremento de α si cumple alguna de las siguientes condiciones*

a) $\alpha(x_0+) - \alpha(x_0-) > 0$.

b) $\alpha(x_0 + h) - \alpha(x_0) > 0$, *para cualquier número real $h > 0$.*

Se puede verificar fácilmente que, para la función dada por

$$\alpha(x) = \begin{cases} 1/3 & \text{si} \quad x \in [0, 1/2), \\ 2/3 & \text{si} \quad x \in [1/2, 1), \\ x^2 & \text{si} \quad x \in [1, 2], \end{cases}$$

$x_0 = 1/2$ sólo cumple la condición (a), $x_0 = 3/2$ sólo cumple la condición (b) y $x_0 = 1$ cumple ambas condiciones, por lo que los tres son puntos de incremento de α, además $x_0 = 1/4$ no es punto de incremento.

Así, el problema principal es resuelto mediante el siguiente resultado, en cuya demostración se utiliza teoría de la integral de Riemann-Stieltjes y los teoremas de Helly, lo cual se menciona en el apéndice de este trabajo.

Teorema 6.3. *(Hamburger)*

a) Sea $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ una función no decreciente con un número infinito de puntos de incremento. Entonces la sucesión de números reales dada por

$$s_k = \int_{-\infty}^{\infty} u^k d\alpha(u), \quad \text{para} \quad k = 0, 1, 2, ... \tag{6.2}$$

es definida positiva.

b) Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Entonces existe una función $\alpha : \mathbb{R} \longrightarrow \mathbb{R}$ no decreciente con un número infinito de puntos de incremento tal que

$$\int_{-\infty}^{\infty} u^k d\alpha(u) = s_k, \quad \text{para} \quad k = 0, 1, 2, ... \tag{6.3}$$

Demostración.

(a) Puesto que las ecuaciones (6.2) se cumplen, entonces para cualquier entero $m \geq 0$ y cualquier $\bar{x} = (x_0, x_1, ..., x_m) \in \mathbb{R}^{m+1}$ se tiene que

$$\begin{aligned}
\sum_{i=0}^{m}\sum_{k=0}^{m} s_{i+k}x_i x_k &= \sum_{i=0}^{m}\sum_{k=0}^{m}\left(\int_{-\infty}^{\infty} u^{i+k} d\alpha(u)\right) x_i x_k \\
&= \int_{-\infty}^{\infty}\left(\sum_{i=0}^{m} x_i u^i\right)\left(\sum_{k=0}^{m} x_k u^k\right) d\alpha(u) \\
&= \int_{-\infty}^{\infty}\left[\sum_{k=0}^{m} x_k u^k\right]^2 d\alpha(u) \geq 0,
\end{aligned}$$

por lo que

$$\sum_{i,k=0}^{m} s_{i+k}x_ix_k = \int_{-\infty}^{\infty}\left[\sum_{k=0}^{m} x_ku^k\right]^2 d\alpha(u) \geq 0. \tag{6.4}$$

Si la igualdad se cumple en esa relación, el polinomio

$$A(u) = \sum_{k=0}^{m} x_ku^k,$$

debe ser cero en todos los puntos en los cuales la función α se incrementa. Pero de acuerdo a la condición inicial, existe una infinidad de puntos de incremento, por lo tanto, el polinomio $A(u)$ debe ser idénticamente cero.

Así, la igualdad en (6.4) se da sólo si todos los $x_0, x_1, ..., x_m$ son cero, es decir, si $\bar{x} = \bar{0}$. Por lo tanto, la sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva y de esta manera se concluye la prueba de que la condición es necesaria.

La prueba del inciso (b) es más complicada por lo cual se dividirá en dos partes.

(b) i) Sea $\{s_k\}_{k=0}^{\infty}$ una sucesión definida positiva. Para cualquier entero positivo n se considera el llamado problema truncado de momentos de orden $2n-1$, a saber

$$s_k = \int_{-\infty}^{\infty} u^k d\alpha(u), \quad k = 0, 1, ..., 2n-1, \tag{6.5}$$

donde α se busca en la clase de las funciones no decrecientes. Se puede construir un conjunto de soluciones continuas para este problema de la siguiente manera, tomando un número real arbitrario τ y escribiendo la ecuación probada en (5.10) para $n+1$

$$\frac{Q_{n+1}(\lambda,\tau)}{P_{n+1}(\lambda,\tau)} = \frac{Q_{n+1}(\lambda) - \tau Q_n(\lambda)}{P_{n+1}(\lambda) - \tau P_n(\lambda)} = \sum_{k=1}^{n+1} \frac{\mu_k}{\lambda - \lambda_k}, \tag{6.6}$$

donde

$$\lambda_1 < \lambda_2 < \cdots < \lambda_{n+1} \quad \text{con} \quad \lambda_k = \lambda_k^{(n+1)}(\tau),$$

son la raíces del polinomio cuasiortogonal $P_{n+1}(\lambda, \tau)$ y

$$\mu_k = \mu_k^{(n+1)}(\tau) > 0,$$

son los coeficientes de la fórmula de cuadratura. Como se estableció en (5.12)

$$s_k = \sum_{i=1}^{n+1} \mu_i \lambda_i^k, \quad \text{con} \quad k = 0, 1, ..., 2n-1. \tag{6.7}$$

Ahora se introducen las funciones constantes a pedazos α_n, para $n = 1, 2, 3, ...,$ las cuales tienen como sus únicos puntos de incremento los λ_i, y en las cuales sus discontinuidades son los μ_i, es decir

$$\mu_i = \alpha_n(\lambda_i+) - \alpha_n(\lambda_i-) \quad \text{donde} \quad i = 1, 2, ..., n+1,$$

que también se pueden definir como

$$\alpha_n(\lambda) = \begin{cases} 0 & \text{si} \quad \lambda < \lambda_1, \\ \sum\limits_{k=1}^{i} \mu_k & \text{si} \quad \lambda \in [\lambda_i, \lambda_{i+1}), \quad \text{para} \quad i = 1, 2, ..., n, \\ \sum\limits_{k=1}^{n+1} \mu_k & \text{si} \quad \lambda \geq \lambda_{n+1}. \end{cases}$$

Entonces las ecuaciones (6.7) toman la forma

$$s_k = \sum_{i=1}^{n+1} \lambda_i^k \mu_i = \int_{-\infty}^{\infty} u^k d\alpha_n(u), \quad \text{con} \quad k = 0, 1, ..., 2n-1. \tag{6.8}$$

Así, la función α_n es una solución para el problema truncado de momentos de orden n.

ii) Ahora se toma cualquier sucesión de soluciones no decrecientes del problema de los momentos de orden $2n-1$, con $n = 1, 2, 3, ...,$ por ejemplo la sucesión que fue construida en el inciso (i) de esta demostración. Esta sucesión $\{\alpha_n\}_{n=1}^{\infty}$ satisface las condiciones del primer teorema de Helly, mostrado en la tercera sección del apéndice de este trabajo, puesto que para cualquier n la variación total de la función α_n es

$$\sum_{k=1}^{n+1} \alpha_{n_k} = \int_{-\infty}^{\infty} d\alpha_n(u) = s_0.$$

Por lo tanto, existe una función no decreciente y acotada α y una subsucesión $\{\alpha_{n_i}\}_{i=1}^{\infty}$ que en todos los puntos de continuidad de α la ecuación

$$\alpha = \lim_{i \longrightarrow \infty} \alpha_{n_i}$$

se mantiene.

Se probará que α es una solución del problema completo de momentos que se está discutiendo, en otras palabras que

$$s_k = \int_{-\infty}^{\infty} u^k d\alpha(u), \quad \text{con} \quad k = 0, 1, 2, ... \tag{6.9}$$

Sea $\epsilon > 0$ y sean $-A < 0$ y $B > 0$ dos puntos de continuidad de α, tales que para k fijo y r se tiene que $s_{2r} < \epsilon\, C^{2r-k}$, con $2r > k$ y $C = \text{mín}\{A, B\}$. En este caso se tiene por el segundo teorema de Helly, que se muestra en la tercera sección del apéndice de este trabajo, que

$$\int_{-A}^{B} u^k d\alpha(u) = \lim_{i \longrightarrow \infty} \int_{-A}^{B} u^k d\alpha_{n_i}(u). \tag{6.10}$$

Por propiedades de la integral de Riemann-Stieltjes se tiene que

$$\int_{-\infty}^{\infty} u^k d\alpha_{n_i}(u) = \int_{-A}^{B} u^k d\alpha_{n_i}(u) + \int_{-\infty}^{-A} u^k d\alpha_{n_i}(u) + \int_{B}^{\infty} u^k d\alpha_{n_i}(u),$$

por lo que

$$\left|\int_{-A}^{B} u^k d\alpha_{n_i}(u) - \int_{-\infty}^{\infty} u^k d\alpha_{n_i}(u)\right| = \left|\int_{-\infty}^{-A} u^k d\alpha_{n_i}(u) + \int_{B}^{\infty} u^k d\alpha_{n_i}(u)\right|,$$

entonces por (6.8)

$$\left|\int_{-A}^{B} u^k d\alpha_{n_i}(u) - s_k\right| = \left|\int_{-\infty}^{-A} u^k d\alpha_{n_i}(u) + \int_{B}^{\infty} u^k d\alpha_{n_i}(u)\right|,$$

luego, haciendo tender i a infinito y usando (6.10)

$$\left|\int_{-A}^{B} u^k d\alpha(u) - s_k\right| = \left|\lim_{i\longrightarrow\infty}\left(\int_{-\infty}^{-A} u^k d\alpha_{n_i}(u) + \int_{B}^{\infty} u^k d\alpha_{n_i}(u)\right)\right|,$$

Además como $2r > k$ y si $n_i > r$, se cumple que $2n_i - 1 \geq 2r$, que junto con (6.8) para $C = \text{mín}\{A, B\}$ se tiene que

$$\begin{aligned}
\left|\int_{-\infty}^{-A} u^k d\alpha_{n_i}(u) + \int_{B}^{\infty} u^k d\alpha_{n_i}(u)\right| &= \left|\int_{-\infty}^{-A} \frac{u^{2r}}{u^{2r-k}} d\alpha_{n_i}(u) + \int_{B}^{\infty} \frac{u^{2r}}{u^{2r-k}} d\alpha_{n_i}(u)\right| \\
&\leq \frac{1}{C^{2r-k}}\left(\int_{-\infty}^{-A} u^{2r} d\alpha_{n_i}(u) + \int_{B}^{\infty} u^{2r} d\alpha_{n_i}(u)\right) \\
&\leq \frac{1}{C^{2r-k}} \int_{-\infty}^{\infty} u^{2r} d\alpha_{n_i}(u) \\
&= \frac{s_{2r}}{C^{2r-k}},
\end{aligned}$$

ya que $1/u^{2r-k} \leq 1/C^{2r-k}$, para $u \leq -A$ y $u \geq B$.

Por lo tanto se sigue que

$$\left|\int_{-A}^{B} u^k d\alpha(u) - s_k\right| \leq \frac{s_{2r}}{C^{2r-k}},$$

y haciendo tender A y B a infinito, de tal forma que $-A$ y B siempre sigan siendo puntos de continuidad de α, se obtiene

$$\left|\int_{-\infty}^{\infty} u^k d\alpha(u) - s_k\right| \leq \frac{s_{2r}}{C^{2r-k}} < \epsilon.$$

Esto demuestra las ecuaciones (6.9).

Se puede ver que si se cumplen las ecuaciones (6.9), entones para cualquier polinomio $P(u) = a_0 + a_1 u + \cdots + a_n u^n$ se tiene que

$$\mathcal{S}\{P(u)\} = \int_{-\infty}^{\infty} P(u) d\alpha(u), \tag{6.11}$$

por linealidad de la integral, pues

$$\begin{aligned}
\mathcal{S}\{P(u)\} &= a_0 s_0 + a_1 s_1 + \cdots + a_n s_n \\
&= a_0 \int_{-\infty}^{\infty} u^0 d\alpha(u) + a_1 \int_{-\infty}^{\infty} u^1 d\alpha(u) + \cdots + a_n \int_{-\infty}^{\infty} u^n d\alpha(u) \\
&= \int_{-\infty}^{\infty} a_0 u^0 + a_1 u^1 + \cdots + a_n u^n d\alpha(u) \\
&= \int_{-\infty}^{\infty} P(u) d\alpha(u).
\end{aligned}$$

Por último, se demostrará que la función α que se ha encontrado tiene un número infinito de puntos de incremento. Suponiendo lo contrario, es decir, que α tiene sólo un número finito de puntos de incremento, $u_1, u_2, ..., u_N$, se tendría por (6.11), para el polinomio

$$P(u) = (u - u_1)^2 (u - u_2)^2 \cdots (u - u_N)^2,$$

la igualdad

$$
\begin{aligned}
\mathcal{S}\{P(u)\} &= \int_{-\infty}^{\infty} P(u)d\alpha(u) \\
&= \int_{-\infty}^{\infty} (u-u_1)^2(u-u_2)^2 \cdots (u-u_N)^2 d\alpha(u) \\
&= \sum_{k=1}^{N} P(u_k)\alpha_k = 0,
\end{aligned}
$$

lo cual es contradictorio pues la sucesión $\{s_k\}_{k=0}^{\infty}$ es definida positiva, por lo que su funcional correspondiente debe ser positivo.

Así la suficiencia de la condición del teorema también está probada.

□

Considérese nuevamente una variable aleatoria X con distribución Ber(1/2), la cual tiene una función de distribución con un número finito de puntos de incremento, además $s_0 = 1$ y $s_k = 1/2$ para cualquier entero $k > 0$. Entonces se tiene que esta sucesión de momentos no es definida positiva pues

$$
D_2 = \begin{vmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \end{vmatrix} = 0.
$$

Por lo tanto se tiene que el Teorema 6.3 no necesariamente se cumple si la función α tiene un número finito de puntos de incremento.

Para concluir, recuérdese el ejemplo dado en la primera sección que considera la sucesión de números reales $0!, 1!, 2!, \ldots$ Para ver que esta sucesión es definida positiva, también hay que recordar que para una variable aleatoria X con distribución exp(1), su función de distribución está dada por

$$
F(x) = \begin{cases} 1 - e^{-x} & \text{si} \quad x > 0, \\ 0 & \text{en otro caso.} \end{cases}
$$

Esta función es no decreciente, tiene un número infinito de puntos de incremento y cumple que

$$s_k = E(X^k) = \int_{-\infty}^{\infty} x^k dF(x) = \frac{k!}{1^k}, \quad \text{para} \quad k = 0, 1, 2, ...,$$

por lo tanto, por el Teorema 6.3, se puede decir que la sucesión $\{k!\}_{k=0}^{\infty}$ es definida positiva.

Más aún, se puede afirmar que toda sucesión de momentos de una función de distribución con un número infinito de puntos de incremento, es una sucesión definida positiva que también es normalizada.

Conclusiones

En esta tesis se expuso acerca de una solución al problema de los momentos de Hamburger, es decir se discutió acerca de una condición para que una sucesión de números reales $\{s_k\}_{k=0}^{\infty}$ cumpliera ser la sucesión de momentos de una función de distribución α y lo que esta función debe de satisfacer para que tenga una sucesión de momentos con ciertas características, para lo cual se agregó el requerimiento adicional a la función α de tener un número infinito de puntos de incremento.

El resultado principal estudiado fue mostrar que una sucesión de números reales definida positiva es la sucesión de momentos de una función no decreciente con un número infinito de puntos de incremento.

De manera sencilla se mostró que una sucesión de momentos debe de ser definida positiva si es la sucesión de momentos de una función no decreciente con un número infinito de puntos.

Para mostrar que toda sucesión definida positiva es la sucesión de momentos de una función no decreciente con un número infinito de puntos de incremento, primero se construyeron soluciones para el problema truncado de los momentos, o con una sucesión de momentos finita, con las cuales, por medio del primer teorema de Helly se construye la solución al problema infinito y se demuestra que lo es apoyándose del segundo teorema de Helly.

Gracias al resultado principal estudiado en esta tesis se puede decir que la sucesión de momentos de una función de distribución continua es una sucesión definida positiva. De la misma forma, la sucesión de momentos de una función de distribución discreta con un número infinito de puntos de incremento es una sucesión definida positiva. Sin embargo se sigue teniendo la dificultad para determinar de manera general cuándo una sucesión de números reales es definida positiva, por lo que encontrar criterios que lo faciliten es un buen reto. Además, en la demostración del Teorema 6.3, la función que se obtiene como solución al problema de los momentos puede ser continua o discreta. Otro problema interesante sería proporcionar criterios para determinar cuándo se tiene una función continua o cuándo se tiene una función discreta.

La teoría presentada en las seis secciones de este trabajo está basada principalmente en [1], pero todas las demostraciones se trabajaron para presentarse con un desarrollo que pudiera dar una comprensión más completa y detallada para el lector. Algunas de estas demostraciones se desarrollaron casi completamente ya que en los libros citados sólo se mencionaba, de manera breve o resumida, la idea principal, como es el caso del Lema 2.3 y del Teorema 2.4, por mencionar algunas.

Apéndices

A. La integral de Riemann-Stieltjes

Con el propósito de comprender mejor la definición del problema central discutido en este trabajo, en este apéndice se recordará teoría de la integral de Riemann-Stieltjes. Dicha teoría se tomó principalmente de [3] y es de gran utilidad para la demostración del resultado principal estudiado en esta tesis en la sexta sección.

Definición A.1. *Sea $[a,b]$ un intervalo en $\mathbb{R}$. Una partición P de $[a,b]$ es un conjunto finito de puntos*

$$P = \{x_0, x_1, ..., x_n\},$$

tal que $a = x_0 < x_1 < \cdots < x_n = b$.

Se denota la longitud del k-ésimo subintervalo $[x_{k-1}, x_k]$ de una partición P de $[a,b]$ por $\triangle x_k = x_k - x_{k-1}$, por lo que $\sum_{k=1}^{n} \triangle x_k = b - a$.

Para una función f en $[a,b]$ se denota el incremento de f en el intervalo $[x_{k-1}, x_k]$ por $\triangle f_k = f(x_k) - f(x_{k-1})$, entonces $\sum_{k=1}^{n} \triangle f_k = f(b) - f(a)$.

Definición A.2. *Sea $P = \{x_0, x_1, ..., x_n\}$ una partición de $[a,b]$. Un refinamiento P' de P es una partición de $[a,b]$ tal que $P' \supseteq P$.*

Definición A.3. *Dada una partición P, se define su norma como*

$$||P|| = \text{máx}\{\triangle x_k;\ k = 1, 2, ..., n\}.$$

Se observa que si $P' \supseteq P$ entonces $||P'|| \leq ||P||$. La colección de todas las particiones del intervalo $[a,b]$ se denota por $\mathcal{P}[a,b]$.

Definición A.4. *Sea f una función definida en $[a,b]$. Si existe un número $M > 0$ tal que*

$$\sum_{k=1}^{n} |\triangle f_k| \leq M,$$

para toda partición de $[a, b]$, *se dice que* f *es de variación acotada en* $[a, b]$.

Dada una función f su variación con respecto de una partición P, calculada como $\sum_{k=1}^{n} |\triangle f_k|$, se denotará por $\sum(P)$.

Definición A.5. *Sea* f *una función de variación acotada en* $[a, b]$. *Se llamará al número*

$$V_f(a, b) = \sup \left\{ \sum(P) : P \in \mathcal{P}[a, b] \right\},$$

variación total de f *en el intervalo* $[a, b]$.

Definición A.6. *Sean* f *y* α *funciones definidas en* $[a, b]$, $P = \{x_0, x_1, ..., x_n\}$ *una partición de* $[a, b]$ *y sea* t_k *un punto en el intervalo* $[x_{k-1}, x_k]$. *Se llamará suma de Riemann-Stieltjes de* f *respecto de* α *a una suma de la forma*

$$S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k) \triangle \alpha_k,$$

donde $\triangle \alpha_k = \alpha(x_k) - \alpha(x_{k-1})$.

Con los conceptos anteriores ahora se puede definir la integral de Riemann-Stieltjes de la siguiente manera.

Definición A.7. *Sean* f *y* α *funciones definidas en* $[a, b]$. *Se dice que* f *es Riemann integrable respecto de* α *en* $[a, b]$, *denotado por* $f \in R(\alpha)$ *en* $[a, b]$, *si existe un número* A *que cumple que para cada* $\epsilon > 0$ *existe una partición* P_ϵ *de* $[a, b]$ *tal que para cada refinamiento* P *de* P_ϵ *se tiene que*

$$|S(P, f, \alpha) - A| < \epsilon.$$

Cuando A *existe, es único y se representa por medio de*

$$\int_a^b f(x)d\alpha(x) \quad o \quad \int_a^b f d\alpha.$$

Se dice también que existe la integral de Riemann-Stieltjes $\int_a^b f d\alpha$. Las funciones f y α son llamadas, respectivamente, integrando e integrador.

Cabe mencionar que en el caso particular cuando $\alpha(x) = x$, la integral se reduce a la integral de Riemann que se denota por $\int_a^b f(x)dx$ o por $\int_a^b f dx$.

Además se puede verificar fácilmente que para una función definida como $f(x) = c$ para todo $x \in [a, b]$ y c constante se tiene que

$$\int_a^b f(x)d\alpha(x) = c[\alpha(b) - \alpha(a)].$$

La integral de Riemann-Stieltjes es lineal tanto en el integrando como en el integrador. Esto se mostrará a continuación.

Teorema A.8. *Sean $f, g \in R(\alpha)$ en $[a, b]$ y c constante. Entonces se tiene que $cf + g \in R(\alpha)$ en $[a, b]$ y además*

$$\int_a^b (cf + g)d\alpha = c\int_a^b f d\alpha + \int_a^b g d\alpha.$$

Demostración. Sea $h = cf + g$ y $\epsilon > 0$. Como $f \in R(\alpha)$ en $[a, b]$, entonces para $\epsilon' = \epsilon/2|c|$ existe una partición $P'_{\epsilon'}$ de $[a, b]$, tal que para cualquier refinamiento $P \supseteq P'_{\epsilon'}$,

$$\left| S(P, f, \alpha) - \int_a^b f d\alpha \right| < \epsilon',$$

y como $g \in R(\alpha)$ para $\epsilon'' = \epsilon/2$ existe una partición $P''_{\epsilon''}$ de $[a, b]$, tal que para cualquier refinamiento $P \supseteq P''_{\epsilon''}$,

$$\left| S(P, g, \alpha) - \int_a^b g d\alpha \right| < \epsilon''.$$

Como

$$
\begin{aligned}
S(P,h,\alpha) &= \sum_{k=1}^{n} h(t_k)\,\triangle\alpha_k = c\sum_{k=1}^{n} f(t_k)\,\triangle\alpha_k + \sum_{k=1}^{n} g(t_k)\,\triangle\alpha_k \\
&= cS(P,f,\alpha) + S(P,g,\alpha),
\end{aligned}
$$

entonces si se toma a $P_\epsilon = P'_{\epsilon'} \cup P''_{\epsilon''}$, se tiene que para P refinamiento de P_ϵ

$$
\begin{aligned}
&\left|S(P,h,\alpha) - c\int_a^b fd\alpha - \int_a^b gd\alpha\right| \\
=&\left|cS(P,f,\alpha) + S(P,g,\alpha) - c\int_a^b fd\alpha - \int_a^b gd\alpha\right| \\
\leq&\,|c|\left|S(P,f,\alpha) - \int_a^b fd\alpha\right| + \left|S(P,g,\alpha) - \int_a^b gd\alpha\right| \\
<&\,|c|\left(\frac{\epsilon}{2|c|}\right) + \frac{\epsilon}{2} \\
=&\epsilon,
\end{aligned}
$$

por lo que $h \in R(\alpha)$. Finalmente por la unicidad de la integral, se tiene que $\int_b^a (cf+g)d\alpha = c\int_b^a fd\alpha + \int_b^a gd\alpha$.

$\square$

Teorema A.9. *Sean $f \in R(\alpha) \cap R(\beta)$ en $[a,b]$ y c constante. Entonces $f \in R(c\alpha+\beta)$ y se tiene que*

$$
\int_a^b fd(c\alpha+\beta) = c\int_a^b fd\alpha + \int_a^b fd\beta.
$$

Demostración. Sea $\gamma = c\alpha + \beta$ y $\epsilon > 0$. Como $f \in R(\alpha)$ para $\epsilon' = \epsilon/2|c|$ existe una partición $P'_{\epsilon'}$ de $[a,b]$, tal que para cualquier refinamiento $P \supseteq P'_{\epsilon'}$,

$$\left|S(P,f,\alpha)-\int_a^b f d\alpha\right|<\epsilon',$$

y como $f\in R(\beta)$ para $\epsilon''=\epsilon/2$ existe una partición $P''_{\epsilon''}$ de $[a,b]$, tal que para cualquier refinamiento $P\supseteq P''_{\epsilon''}$, se tiene que

$$\left|S(P,f,\beta)-\int_a^b f d\beta\right|<\epsilon''.$$

Como

$$\begin{aligned}S(P,f,\gamma)&=\sum_{k=1}^{n} f(t_k)\triangle\gamma=c\sum_{k=1}^{n}f(t_k)\triangle\alpha_k+\sum_{k=1}^{n} f(t_k)\triangle\beta_k\\&=cS(P,f,\alpha)+S(P,f,\beta),\end{aligned}$$

tomando a $P_\epsilon=P'_{\epsilon'}\cup P''_{\epsilon''}$, se tiene que para $P\supseteq P_\epsilon$

$$\begin{aligned}&\left|S(P,f,\gamma)-c\int_a^b f d\alpha-\int_a^b f d\beta\right|\\&=\left|cS(P,f,\alpha)+S(P,f,\beta)-c\int_a^b f d\alpha-\int_a^b f d\beta\right|\\&\le|c|\left|S(P,f,\alpha)-\int_a^b f d\alpha\right|+\left|S(P,f,\beta)-\int_a^b f d\beta\right|\\&<|c|\left(\frac{\epsilon}{2|c|}\right)+\frac{\epsilon}{2}\\&=\epsilon.\end{aligned}$$

□

En el siguiente resultado se muestra que, bajo ciertas condiciones, se puede dividir el intervalo de integración.

Teorema A.10. *Sea* $c \in (a,b)$*. Si dos de las integrales* $\int_a^b f d\alpha$, $\int_a^c f d\alpha$, $\int_c^b f d\alpha$ *existen, entonces la tercera también existe y*

$$\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha.$$

Demostración. Sea $\epsilon > 0$. Si se supone que $\int_a^c f d\alpha$ y $\int_c^b$ existen, entonces para $\epsilon' = \epsilon/2$ existe una partición $P'_{\epsilon'}$ de $[a,c]$ tal que para cualquier refinamiento $P \supseteq P'_{\epsilon'}$ se tiene que

$$\left| S(P,f,\alpha) - \int_a^c f d\alpha \right| < \epsilon',$$

y existe una partición $P''_{\epsilon'}$ de $[c,b]$ tal que para cualquier refinamiento $P \supseteq P''_{\epsilon'}$ se tiene que

$$\left| S(P,f,\alpha) - \int_c^b f d\alpha \right| < \epsilon'.$$

Entonces tomando a $P_\epsilon = P'_{\epsilon'} \cup P''_{\epsilon'}$ partición de $[a,b]$ y un refinamiento $P \supseteq P_\epsilon$, se tiene para $Q' = P \cap [a,c] \supseteq P'_{\epsilon'}$ y $Q'' = P \cap [c,b] \supseteq P''_{\epsilon'}$ que

$$\begin{aligned}
&\left| S(P,f,\alpha) - \left(\int_a^c f d\alpha + \int_c^b f d\alpha \right) \right| \\
&= \left| S(Q',f,\alpha) + S(Q'',f,\alpha) - \int_a^c f d\alpha - \int_c^b f d\alpha \right| \\
&\leq \left| S(Q',f,\alpha) - \int_a^c f d\alpha \right| + \left| S(Q'',f,\alpha) - \int_c^b f d\alpha \right| \\
&< \epsilon.
\end{aligned}$$

Por lo tanto la integral $\int_a^b f d\alpha$ existe y es igual a $\int_a^c f d\alpha + \int_c^b f d\alpha$.

Ahora supóngase que $\int_a^c f d\alpha$ y $\int_a^b f d\alpha$ existen, entonces para $\epsilon' = \epsilon/2$ existe una partición $P'_{\epsilon'}$ de $[a,c]$ tal que para cualquier refinamiento $P \supseteq P'_{\epsilon'}$ se tiene que

$$\left|S(P,f,\alpha)-\int_a^c fd\alpha\right|<\epsilon',$$

y existe una partición $P''_{\epsilon'}$ de $[a,b]$ tal que para cualquier refinamiento $P\supseteq P''_{\epsilon'}$ se tiene que

$$\left|S(P,f,\alpha)-\int_a^b fd\alpha\right|<\epsilon'.$$

Sin pérdida de generalidad, si se supone que $P''_{\epsilon'}\supseteq P'_{\epsilon'}$, entonces existe una partición $P'''_{\epsilon'}$ de $[c,b]$ tal que $P''_{\epsilon'}=P'_{\epsilon'}\cup P'''_{\epsilon'}$. Si P es una partición de $[c,b]$ tal que $P\supseteq P'''_{\epsilon'}$, se tiene que

$$\begin{aligned}
&\left|S(P,f,\alpha)-\left(\int_a^b fd\alpha-\int_a^c fd\alpha\right)\right|\\
=&\left|S(P'_{\epsilon'}\cup P,f,\alpha)-S(P'_{\epsilon'},f,\alpha)-\int_a^b fd\alpha+\int_a^c fd\alpha\right|\\
\leq&\left|S(P'_{\epsilon'}\cup P,f,\alpha)-\int_a^b fd\alpha\right|+\left|S(P'_{\epsilon'},f,\alpha)-\int_a^c fd\alpha\right|\\
<&\ \epsilon.
\end{aligned}$$

Por lo que $\int_c^b fd\alpha$ existe y es igual a $\int_a^b fd\alpha-\int_a^c fd\alpha$.

□

Se recordará el siguiente resultado el cual será de utilidad más adelante pero cuya prueba se omite para no desviarse del objetivo principal de la presente sección.

Teorema A.11. *(Teorema del valor medio)*
Sea $f:[a,b]\longrightarrow\mathbb{R}$ continua en $[a,b]$ y derivable en (a,b). Entonces existe un número $x\in(a,b)$ tal que

$$f'(x)=\frac{f(b)-f(a)}{b-a}.$$

El siguiente resultado muestra condiciones para que una integral de Riemann-Stieltjes se pueda calcular como una integral de Riemann.

Teorema A.12. *Sean $f \in R(\alpha)$ en $[a,b]$ acotada y α' continua en $[a,b]$. Entonces existe la integral de Riemann $\int_a^b f(x)\alpha'(x)dx$ y además*

$$\int_a^b f(x)d\alpha(x) = \int_a^b f(x)\alpha'(x)dx.$$

Demostración. Sean $\epsilon > 0$, $g = f\alpha'$ y P una partición de $[a,b]$. Entonces

$$S(P,g) = \sum_{k=1}^{n} g(t_k) \bigtriangleup x_k = \sum_{k=1}^{n} f(t_k)\alpha'(t_k) \bigtriangleup x_k.$$

Por otro lado,

$$S(P,f,\alpha) = \sum_{k=1}^{n} f(t_k) \bigtriangleup \alpha_k.$$

Aplicando el teorema del valor medio a la función α en cada subintervalo $[x_{k-1}, x_k]$ de P, se encuentra un punto $v_k \in (x_{k-1}, x_k)$, para $k = 1, 2, ..., n$, tal que

$$\bigtriangleup \alpha_k = \alpha'(v_k) \bigtriangleup x_k,$$

por lo que

$$S(P,f,\alpha) - S(P,g) = \sum_{k=1}^{n} f(t_k)[\alpha'(v_k) - \alpha'(t_k)] \bigtriangleup x_k.$$

Como f es acotada, existe $M > 0$ tal que $|f(x)| < M$ para toda $x \in [a,b]$.

Como α' es continua en $[a,b]$ entonces es uniformemente continua en $[a,b]$, por lo que existe $\delta > 0$, que depende de ϵ, tal que para cada $x, y \in [a,b]$ si

$$0 \leq |x-y| < \delta \Longrightarrow |\alpha'(x) - \alpha'(y)| < \frac{\epsilon}{2M(b-a)}.$$

Entonces, si la partición P tiene norma menor que δ, se tiene que

$$|S(P,f,\alpha)-S(P,g)| \leq \sum_{k=1}^{n}|f(t_k)||\alpha'(v_k)-\alpha'(t_k)|\triangle x_k$$

$$< M\frac{\epsilon}{2M(b-a)}\sum_{k=1}^{n}\triangle x_k=\frac{\epsilon}{2}.$$

Por otro lado, para $\epsilon'=\epsilon/2$ existe una partición $P_{\epsilon'}$ de $[a,b]$ tal que para cualquier refinamiento $P\supseteq P_{\epsilon'}$ se tiene que

$$\left|S(P,f,\alpha)-\int_a^b fd\alpha\right|<\epsilon'.$$

Por lo tanto, si se toma una partición $P\supseteq P_{\epsilon'}$ tal que $||P||<\delta$, se tiene

$$\left|S(P,g)-\int_a^b fd\alpha\right| \leq \left|S(P,g)-S(P,f,\alpha)\right|+\left|S(P,f,\alpha)-\int_a^b fd\alpha\right|$$

$$<\frac{\epsilon}{2}+\epsilon'=\epsilon.$$

□

En lo siguiente se utilizarán los límites laterales de una función, por lo que a continuación se recuerda su definición.

Definición A.13. *Sea $f:\mathbb{R}\longrightarrow\mathbb{R}$ una función y sean $t,l\in\mathbb{R}$. Entonces*

1) *Se dice que el límite de f por la izquierda de t es l y se denota como $\lim\limits_{x\longrightarrow t-} f(x)=l$ si y sólo si para cualquier $\epsilon>0$, existe $\delta>0$ tal que si*

$$t-\delta<x<t \implies |f(x)-l|<\epsilon.$$

En tal caso se dice que $f(x)$ tiende a l cuando x tiende a t por la izquierda y se denota como $f(x-)$.

2) Se dice que el límite de f por la derecha de t es l y se denota como $\lim_{x \longrightarrow t+} f(x) = l$ si y sólo si para cualquier $\epsilon > 0$, existe $\delta > 0$ tal que si

$$t < x < t + \delta \implies |f(x) - l| < \epsilon.$$

En este caso se dice que $f(x)$ tiende a l cuando x tiende a t por la derecha y se denota como $f(x+)$.

Dada la definición de límite lateral ahora es posible establecer el concepto de función escalonada, lo cual se realiza a continuación.

Definición A.14. *Una función α en $[a, b]$ se llama función escalonada si existen una partición $P = \{x_0, x_1, ..., x_n\}$ de $[a, b]$ y constantes $\lambda_1, \lambda_2, ..., \lambda_n$ tales que*

$$\alpha(x) = \lambda_{k+1} \quad \text{si} \quad x \in (x_k, x_{k+1}), \quad \text{para} \quad k = 0, 1, ..., n-1.$$

Se llama el salto de α en x_k al número $\alpha_k = \alpha(x_k+) - \alpha(x_k-)$, para cada $k = 1, 2, ..., n-1$, el salto de α en x_0 es $\alpha(x_0+) - \alpha(x_0) = \lambda_1 - \alpha(x_0)$ y en x_n es $\alpha(x_n) - \alpha(x_n-) = \alpha(x_n) - \lambda_n$.

Cabe notar que en la definición anterior no se establecen los valores de la función α en $x_0, x_1, ..., x_n$. Esto es debido a que para la teoría posterior no afectan dichos valores, por lo que se dejan sin especificar.

Ahora se muestra cómo se simplifica el cálculo de una integral de Riemann-Stieltjes si el integrador es una función escalonada.

Teorema A.15. *Sea α una función escalonada definida en $[a, b]$ como en la Definición A.14, con salto α_k en x_k, para $k = 0, 1, ..., n$. Sea f definida en $[a, b]$ tal que f y α no son ambas discontinuas por la derecha o por la izquierda de cada x_k. Entonces $f \in R(\alpha)$ en $[a, b]$ y*

$$\int_a^b f(x) d\alpha(x) = \sum_{k=0}^{n} f(x_k) \alpha_k.$$

Demostración. Sean $c_1, c_2, ..., c_n$ los puntos medios de cada uno de los los intervalos $(a, x_1), (x_1, x_2), ..., (x_{n-1}, b)$, en donde α toma los valores $\lambda_1, \lambda_2, ..., \lambda_n$, respectivamente. Se probará que $f \in R(\alpha)$ en cada uno de los siguientes intervalos cerrados $[a, c_1], [c_1, x_1], [x_1, c_2], ..., [x_{n-1}, c_n], [c_n, b]$. En cada uno de esos intervalos α es constante excepto tal vez, en uno de los extremos. Se verá lo que sucede en el primer intervalo de la siguiente manera.

Sea $\epsilon > 0$. Para cualquier partición P del intervalo $[a, c_1]$ se tiene que

$$\begin{aligned} S(P, f, \alpha) = \sum_{k=1}^{n} f(t_k) \triangle \alpha_k &= f(t_1)[\lambda_1 - \alpha(a)] + \sum_{k=2}^{n} f(t_k)[\lambda_1 - \lambda_1] \\ &= f(t_1)[\lambda_1 - \alpha(a)]. \end{aligned}$$

Por lo que

$$|S(P, f, \alpha) - f(a)[\lambda_1 - \alpha(a)]| = |f(t_1) - f(a)||\lambda_1 - \alpha(a)|.$$

Si f es continua por la derecha en a, se tiene que para $\epsilon' = \epsilon/|\lambda_1 - \alpha(a)| > 0$ existe $\delta > 0$ tal que si $|x - a| < \delta$, entonces $|f(x) - f(a)| < \epsilon'$, donde $x > a$.

Entonces, para una partición P_ϵ de $[a, c_1]$ tal que $||P_\epsilon|| < \delta$, se tiene que para cualquier refinamiento $P \supseteq P_\epsilon$

$$\begin{aligned} |S(P, f, \alpha) - f(a)[\lambda_1 - \alpha(a)]| &< \epsilon' |\lambda_1 - \alpha(a)| \\ &= \frac{\epsilon}{|\lambda_1 - \alpha(a)|} |\lambda_1 - \alpha(a)| \\ &= \epsilon. \end{aligned}$$

Si se tiene el caso en que α es continua por la derecha en a, entonces tomando a $\epsilon' = \epsilon/|f(t_1) - f(a)|$ se obtiene de manera análoga que

$$
\begin{aligned}
|S(P,f,\alpha) - f(a)[\lambda_1 - \alpha(a)]| &< |f(t_1) - f(a)|\epsilon' \\
&= |f(t_1) - f(a)|\frac{\epsilon}{|f(t_1) - f(a)|} \\
&= \epsilon.
\end{aligned}
$$

Así, queda demostrado que

$$
\int_a^{c_1} f(x)d\alpha(x) = f(a)[\lambda_1 - \alpha(a)].
$$

De manera similar se puede ver que

$$
\begin{aligned}
\int_{c_1}^{x_1} f(x)d\alpha(x) &= f(x_1)[\alpha(x_1) - \lambda_1], \\
\int_{x_1}^{c_2} f(x)d\alpha(x) &= f(x_1)[\lambda_2 - \alpha(x_1)], \\
&\vdots \\
\int_{x_{n-1}}^{c_n} f(x)d\alpha(x) &= f(x_{n-1})[\lambda_n - \alpha(x_{n-1})], \\
\int_{c_n}^{b} f(x)d\alpha(x) &= f(b)[\alpha(b) - \lambda_n].
\end{aligned}
$$

Por lo que, utilizando el Teorema A.10, se concluye que

$$
\begin{aligned}
\int_a^b f(x)d\alpha(x) &= \int_a^{c_1} f(x)d\alpha(x) + \int_{c_1}^{x_1} f(x)d\alpha(x) + \cdots + \int_{c_n}^{b} f(x)d\alpha(x) \\
&= f(x_0)[\lambda_1 - \alpha(x_0)] + f(x_1)[\alpha(x_1) - \lambda_1] \\
&\quad + f(x_1)[\lambda_2 - \alpha(x_1)] + \cdots + f(x_{n-1})[\alpha(x_{n-1}) - \lambda_{n-1}] \\
&\quad + f(x_{n-1})[\lambda_n - \alpha(x_{n-1})] + f(x_n)[\alpha(x_n) - \lambda_n] \\
&= f(x_0)[\lambda_1 - \alpha(x_0)] + f(x_1)[\lambda_2 - \lambda_1] \\
&\quad + \cdots + f(x_{n-1})[\lambda_n - \lambda_{n-1}] + f(x_n)[\alpha(x_n) - \lambda_n] \\
&= \sum_{k=0}^{n} f(x_k)\alpha_k.
\end{aligned}
$$

□

Para el desarrollo de lo siguiente se trabaja con una función no decreciente como integrador, por lo cual se hace mención de la siguiente definición.

Definición A.16. *Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ una función y sean $x_1, x_2 \in \mathbb{R}$ tales que $x_1 < x_2$. Se dice que f es no decreciente si $f(x_1) \leq f(x_2)$ o no creciente si $f(x_1) \geq f(x_2)$.*

Definición A.17. *Sea P una partición de $[a, b]$ y sean*

$$
M_k(f) = \sup\{f(x) : x \in [x_{k-1}, x_k]\},
$$

$$
m_k(f) = \text{ínf}\{f(x) : x \in [x_{k-1}, x_k]\}.
$$

Los siguientes números

$$
U(P, f, \alpha) = \sum_{k=1}^{n} M_k(f) \triangle\alpha_k,
$$

$$
L(P, f, \alpha) = \sum_{k=1}^{n} m_k(f) \triangle\alpha_k,
$$

se llamarán suma superior e inferior de Stieltjes, respectivamente, de f con respecto de α para la partición P.

Se puede ver que siempre $m_k(f) \leq M_k(f)$. Además, si α es no decreciente en $[a, b]$, entonces $\triangle\alpha_k \geq 0$ por lo que $m_k(f) \triangle\alpha_k \leq M_k(f) \triangle\alpha_k$, y entonces

$$L(P, f, \alpha) \leq U(P, f, \alpha).$$

También se tiene que si $t_k \in [x_{k-1}, x_k]$, entonces $m_k(f) \leq f(t_k) \leq M_k(f)$. Por lo tanto, para α no decreciente

$$L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(P, f, \alpha).$$

Teorema A.18. *Sea α una función no decreciente en $[a, b]$. Entonces*

i) Si P' es un refinamiento de P, se tiene que

$$U(P', f, \alpha) \leq U(P, f, \alpha) \quad y \quad L(P', f, \alpha) \geq L(P, f, \alpha).$$

ii) Para cualesquiera dos particiones P_1 y P_2, se tiene que

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha).$$

Demostración.

i) Se observa que es suficiente demostrar el resultado cuando un refinamiento $P' \supseteq P$ tiene exactamente un punto más que P, por ejemplo el punto c del intervalo $I_j = [x_{j-1}, x_j]$, para algún $j = 1, 2, ..., n$. Entonces

$$U(P', f, \alpha) = \sum_{\substack{k=1 \\ k \neq j}}^{n} M_k(f) \triangle\alpha_k + M'[\alpha(c) - \alpha(x_{j-1})] + M''[\alpha(x_j) - \alpha(c)],$$

donde

$$M' = \sup\{f(x); x \in [x_{j-1}, c]\} \quad \text{y} \quad M'' = \sup\{f(x); x \in [c, x_j]\},$$

pero como

$$M' \leq M_j \quad \text{y} \quad M'' \leq M_j,$$

con M_j el supremo en el intervalo $[x_{j-1}, x_j]$, entonces

$$\begin{aligned} M'[\alpha(c) - \alpha(x_{j-1})] + M''[\alpha(x_j) - \alpha(c)] &\leq M_j[\alpha(c) - \alpha(x_{j-1})] \\ &\quad + M_j[\alpha(x_j) - \alpha(c)] \\ &= M_j[\alpha(x_j) - \alpha(x_{j-1})]. \end{aligned}$$

Por lo tanto, $U(P', f, \alpha) \leq U(P, f, \alpha)$.

La desigualdad para la suma inferior se prueba de manera similar.

ii) Sea $P = P_1 \cup P_2$, refinamiento de ambas particiones. Entonces por (i) y una observación previa al teorema se tiene que

$$L(P_1, f, \alpha) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq U(P_2, f, \alpha),$$

con lo cual se obtiene el resultado deseado.

□

Dadas las propiedades de las sumas superior e inferior de Stieltjes, que se obtienen debido a que el integrador es una función no decreciente, a continuación se pueden definir las integrales superior e inferior de Stieltjes.

Definición A.19. *Sea α una función no decreciente en $[a, b]$. Se define la integral superior de Stieltjes de f con respecto de α como*

$$\overline{\int_a^b} f d\alpha = \text{ínf}\{U(P, f, \alpha) : P \in \mathcal{P}[a, b]\},$$

y la integral inferior de Stieltjes como

$$\underline{\int_a^b} f d\alpha = \sup\{L(P, f, \alpha) : P \in \mathcal{P}[a, b]\}.$$

Se denotan también como $\bar{I}(f, \alpha)$ y $\underline{I}(f, \alpha)$ a la integral superior e inferior respectivamente.

A continuación se muestra una comparación entre las integrales superior e inferior de Stieltjes.

Teorema A.20. *Sea α una función no decreciente en $[a, b]$. Entonces se cumple que $\underline{I}(f, \alpha) \leq \bar{I}(f, \alpha)$.*

Demostración. Como $\bar{I}(f, \alpha) = \text{ínf}\{U(P, f, \alpha) : P \in \mathcal{P}[a, b]\}$, entonces para $\epsilon > 0$, existe una partición P' tal que

$$U(P', f, \alpha) < \bar{I}(f, \alpha) + \epsilon,$$

y por el Teorema A.18 (ii) se tiene que $\bar{I}(f, \alpha) + \epsilon$ es cota superior de las sumas inferiores, es decir $L(P, f, \alpha) \leq \bar{I}(f, \alpha) + \epsilon$, para toda partición de $[a, b]$.

Luego como $\underline{I}(f, \alpha) = \sup\{L(P, f, \alpha) : P \in \mathcal{P}[a, b]\}$, entonces se tiene que $\underline{I}(f, \alpha) \leq \bar{I}(f, \alpha) + \epsilon$. Lo anterior se tiene para todo $\epsilon > 0$, por lo tanto, $\underline{I}(f, \alpha) \leq \bar{I}(f, \alpha)$.

$\square$

El siguiente teorema muestra equivalencias a la definición de la integral de Riemann-Stieltjes para integradores no decrecientes.

Teorema A.21. *Sea α una función no decreciente en $[a, b]$. Las siguientes tres afirmaciones son equivalentes.*

i) $f \in R(\alpha)$ *en* $[a, b]$.

ii) *Para cada $\epsilon > 0$, existe una partición P_ϵ, tal que para cualquier refinamiento $P \supseteq P_\epsilon$ se tiene que*

$$0 \leq U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

Si f cumple lo anterior se dice que f satisface la condición de Riemann con respecto de α en $[a, b]$.

iii) $\bar{I}(f, \alpha) = \underline{I}(f, \alpha)$.

Demostración.

i) $\Rightarrow$ *ii)* Si $\alpha(b) = \alpha(a)$ se tiene que α es constante pues es no decreciente, por lo que $\triangle\alpha_k = 0$ para cada $k = 1, 2, ..., n$, y entonces las sumas inferiores y superiores siempre son cero.

Supóngase entonces que $\alpha(a) < \alpha(b)$. Dado $\epsilon > 0$, se elige una partición P_ϵ de $[a, b]$ tal que para todo refinamiento P de P_ϵ y cualquier elección de t_k y t'_k en $[x_{k-1}, x_k]$, se tenga que

$$\left| \sum_{k=1}^{n} f(t_k) \triangle\alpha_k - \int_a^b f d\alpha \right| < \frac{\epsilon}{3} \quad \text{y} \quad \left| \sum_{k=1}^{n} f(t'_k) \triangle\alpha_k - \int_a^b f d\alpha \right| < \frac{\epsilon}{3}.$$

De estas ecuaciones se obtiene que

$$\begin{aligned} \left| \sum_{k=1}^{n} [f(t_k) - f(t'_k)] \triangle\alpha_k \right| &\leq \left| \sum_{k=1}^{n} f(t_k) \triangle\alpha_k - \int_a^b f d\alpha \right| \\ &\quad + \left| \int_a^b f d\alpha - \sum_{k=1}^{n} f(t'_k) \triangle\alpha_k \right| \\ &< \frac{2}{3}\, \epsilon. \end{aligned}$$

Como $M_k(f) - m_k(f) = \sup\{f(x) - f(x') : x, x' \in [x_{k-1}, x_k]\}$, entonces para $\delta > 0$ se pueden elegir t_k y t'_k en $[x_{k-1}, x_k]$ tales que

$$M_k(f) - m_k(f) < f(t_k) - f(t'_k) + \delta,$$

por lo que para $\delta = \frac{1}{3}\,\epsilon/[\alpha(b) - \alpha(a)]$ se tiene que

$$\begin{aligned}
U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{k=1}^{n} [M_k(f) - m_k(f)] \triangle\alpha_k \\
&< \sum_{k=1}^{n} [f(t_k) - f(t'_k) + \delta] \triangle\alpha_k \\
&= \sum_{k=1}^{n} [f(t_k) - f(t'_k)] \triangle\alpha_k + \delta \sum_{k=1}^{n} \triangle\alpha_k \\
&< \frac{2}{3}\,\epsilon + \frac{\epsilon}{3[\alpha(b) - \alpha(a)]}\,[\alpha(b) - \alpha(a)] \\
&= \epsilon.
\end{aligned}$$

ii) $\Rightarrow$ *iii)* Dada $\epsilon > 0$, por (ii) existe una partición P_ϵ tal que para un refinamiento $P \supseteq P_\epsilon$ se tiene que

$$U(P, f, \alpha) < L(P, f, \alpha) + \epsilon.$$

Por lo tanto, para esa partición se cumple que

$$\bar{I}(f, \alpha) \le U(P, f, \alpha) < L(P, f, \alpha) + \epsilon \le \underline{I}(f, \alpha) + \epsilon.$$

Lo cual se tiene para todo $\epsilon > 0$, por lo que $\bar{I}(f, \alpha) \le \underline{I}(f, \alpha)$. Además, por el Teorema A.20 $\bar{I}(f, \alpha) \ge \underline{I}(f, \alpha)$, por lo tanto se tiene la igualdad.

iii) $\Rightarrow$ *i)* Sea $A = \bar{I}(f, \alpha) = \underline{I}(f, \alpha)$ y $\epsilon > 0$. Se elige P'_ϵ tal que $U(P, f, \alpha) < \bar{I}(f, \alpha) + \epsilon$ para todo refinamiento $P \supseteq P'_\epsilon$. También se elige P''_ϵ tal que

$L(P, f, \alpha) > \underline{I}(f, \alpha) - \epsilon$ para todo refinamiento $P \supseteq P''_\epsilon$.

Si $P_\epsilon = P'_\epsilon \cup P''_\epsilon$, se cumple

$$\underline{I}(f, \alpha) - \epsilon < L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(P, f, \alpha) < \bar{I}(f, \alpha) + \epsilon,$$

para cada refinamiento $P \supseteq P_\epsilon$. Pero como $\bar{I}(f, \alpha) = \underline{I}(f, \alpha) = A$, entonces

$$|S(P, f, \alpha) - A| < \epsilon,$$

para cualquier refinamiento P de P_ϵ. Por lo tanto $\int_a^b f d\alpha$ existe y además se tiene que

$$\int_a^b f d\alpha = \overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha,$$

con lo cual concluye la demostración.

$\square$

Ahora se muestra un criterio de comparación para la integral de Riemann-Stieltjes con integradores no decrecientes.

Teorema A.22. *Sean α una función no decreciente y $f, g \in R(\alpha)$ en $[a, b]$ tales que $f(x) \leq g(x)$ para todo $x \in [a, b]$. Entonces*

$$\int_a^b f(x) d\alpha(x) \leq \int_a^b g(x) d\alpha(x).$$

Demostración. Se tiene que para cualquier partición P de $[a, b]$

$$U(P, f, \alpha) = \sum_{k=1}^{n} M_k(f) \bigtriangleup \alpha_k \leq \sum_{k=1}^{n} M_k(g) \bigtriangleup \alpha_k = U(P, g, \alpha),$$

ya que $f(x) \leq g(x)$ y α es no decreciente en $[a, b]$. Por lo que

$$
\begin{aligned}
\int_a^b f d\alpha &= \bar{I}(f,\alpha) \\
&= \text{ínf}\{U(P,f,\alpha) : P \in \mathcal{P}[a,b]\} \\
&\leq \text{ínf}\{U(P,g,\alpha) : P \in \mathcal{P}[a,b]\} \\
&= \bar{I}(g,\alpha) \\
&= \int_a^b g d\alpha.
\end{aligned}
$$

□

A continuación se verá una propiedad que muestra una comparación del valor absoluto de una integral con la integral del valor absoluto de una función.

Teorema A.23. *Sea α una función no decreciente y $f \in R(\alpha)$ en $[a,b]$. Entonces*

$$\left| \int_a^b f(x) d\alpha(x) \right| \leq \int_a^b |f(x)| d\alpha(x).$$

Demostración. Sea $P_\epsilon = \{x_0, x_1, ..., x_n\}$ partición de $[a,b]$ tal que para cada refinamiento $P \supseteq P_\epsilon$ cumple que

$$U(P,f,\alpha) - L(P,f,\alpha) \leq \epsilon.$$

Por otro lado se tiene para cada $x, y \in [a,b]$ que

$$||f(x)| - |g(x)|| \leq |f(x) - g(x)|,$$

por lo que

$$\begin{aligned} M_k(|f|) - m_k(|f|) &= \sup\{|f(x)| - |g(x)| : x, y \in [x_{k-1}, x_k]\} \\ &\leq \sup\{f(x) - f(y) : x, y \in [x_{k-1}, x_k]\} \\ &= M_k(f) - m_k(f), \end{aligned}$$

entonces

$$U(P, |f|, \alpha) - L(P, |f|, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha) \leq \epsilon,$$

por lo tanto $|f| \in R(\alpha)$ en $[a, b]$. Luego, por el Teorema A.22 tomando $g = |f|$, se tiene que

$$\int_a^b f(x) d\alpha(x) \leq \int_a^b |f(x)| d\alpha(x).$$

Tomando valor absoluto de ambos lados se obtiene

$$\left| \int_a^b f(x) d\alpha(x) \right| \leq \int_a^b |f(x)| d\alpha(x),$$

pues la integral del lado derecho es positiva ya que α es no decreciente.

$\square$

Por último se define un tipo particular de integral, el cual se utiliza en el problema principal de esta tesis.

Definición A.24. *Sean f y α funciones en $\mathbb{R}$ tales que $f \in R(\alpha)$ en $[a, b]$ para cualesquiera $a, b \in \mathbb{R}$, con $a < b$. Si existe el límite*

$$\lim_{\substack{a \longrightarrow \infty \\ b \longrightarrow -\infty}} \int_a^b f(x) d\alpha(x),$$

se le llama integral impropia y se denota por

$$\int_{-\infty}^{\infty} f(x) d\alpha(x).$$

Además, de manera análoga a la definición anterior, también se pueden establecer las integrales impropias

$$\int_a^{\infty} f(x)d\alpha(x) \quad \text{y} \quad \int_{-\infty}^{b} f(x)d\alpha(x).$$

B. Fórmula de interpolación de Lagrange

El resultado que se muestra en este apéndice es utilizado para obtener la fórmula de cuadratura en la quinta sección de este trabajo, y se tomó del material proporcionado en [11].

Dada una función y $n+1$ puntos por los que ésta pasa, el objetivo de la fórmula de interpolación de Lagrange es encontrar un polinomio, a lo más de grado n, que pase a través de esos puntos. Ese resultado se muestra con el siguiente teorema.

Teorema B.1. *(Fórmula de interpolación de Lagrange)*
Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ una función que pasa por los puntos $(x_0, f(x_0)), (x_1, f(x_1))$, $..., (x_n, f(x_n))$, con $x_k \neq x_j$ para $k \neq j$. Entonces existe un único polinomio $P(x)$ de grado a lo más n tal que

$$P(x_k) = f(x_k),$$

para $i = 0, 1, ..., n$.

Demostración. Primero se verá la unicidad. Sean dos polinomios P_1 y P_2 de grado a lo más n, tales que $P_1(x_k) = P_2(x_k) = f(x_k)$, para $k = 0, 1, ..., n$. Entonces el polinomio $P = P_1 - P_2$ tiene a lo más grado n, pero también tiene $n+1$ raíces $x_0, x_1, ..., x_n$. Por lo tanto, P debe de ser igual a cero y la unicidad queda demostrada.

Ahora se demostrará la existencia. Considérense los siguientes polinomios

$$L_k(x) = \frac{(x-x_0)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)}$$
$$= \prod_{\substack{j\neq k\\ j=0}}^{n} \frac{x-x_j}{x_k-x_j},$$

llamados polinomios de Lagrange, los cuales cumplen

$$L_k(x_j) = \delta_{kj} = \begin{cases} 1 & \text{si} \quad j = k, \\ 0 & \text{si} \quad j \neq k. \end{cases}$$

El polinomio $P(x)$ buscado se obtiene con ayuda de los polinomios L_k de la siguiente forma

$$P(x) = \sum_{k=0}^{n} f(x_k) L_k(x) = \sum_{k=0}^{n} f(x_k) \prod_{\substack{j\neq k\\ j=0}}^{n} \frac{x-x_j}{x_k-x_j}, \tag{B.1}$$

lo que concluye la demostración.

□

A la representación (B.1) se le conoce como fórmula de interpolación de Lagrange, en donde se puede ver que el polinomio es a lo más de grado n. Además si se denota por

$$\pi(x) = \prod_{j=0}^{n} (x-x_j),$$

entonces

$$\pi'(x_k) = \left[\frac{d\pi}{dx}\right]_{x=x_k} = \prod_{\substack{j\neq k\\ j=0}}^{n} (x_k-x_j),$$

por lo que (B.1) también se puede expresar como

$$P(x) = \sum_{k=0}^{n} f(x_k)\, \frac{\pi(x)}{(x - x_k)\pi'(x_k)}.$$

C. Teoremas de Helly

A continuación se verán dos resultados que se utilizan en la sexta sección. Estos resultados son empleados para construir la función de distribución que da solución al problema de los momentos. La información presentada en esta sección se tomó principalmente de [7].

Primero se hará mención de algunas definiciones las cuales servirán para enunciar los teoremas de Helly.

Definición C.1. *Sean f y $\{f_n\}_{n=0}^{\infty}$ funciones no decrecientes y acotadas. Se dice que*

a) La sucesión $\{f_n\}_{n=0}^{\infty}$ converge débilmente a f si la sucesión converge a f para todo punto de continuidad de la función f.

b) La sucesión $\{f_n\}_{n=0}^{\infty}$ converge completamente a f si la sucesión converge débilmente a f y además $\{f_n(-\infty)\}_{n=0}^{\infty}$ converge a $f(-\infty)$ y $\{f_n(\infty)\}_{n=0}^{\infty}$ converge a $f(\infty)$.

A continuación se probará un lema que ayudará en la demostración del primer teorema de Helly.

Lema C.2. *Sea D un conjunto denso en $\mathbb{R}$ y $\{f_n\}_{n=0}^{\infty}$ una sucesión de funciones no decrecientes que converge a f para todo $x \in D$. Entonces $\{f_n\}_{n=0}^{\infty}$ converge débilmente a f.*

Demostración. Sea x un punto de continuidad de f. Como D es denso en $\mathbb{R}$ se pueden tomar dos sucesiones de números reales $\{x'_k\}_{k=0}^{\infty}$ y $\{x''_k\}_{k=0}^{\infty}$ tales que $x'_k < x < x''_k$ para todo k y

$$\lim_{x\longrightarrow\infty} x'_k = \lim_{x\longrightarrow\infty} x''_k = x.$$

Como f_n es no decreciente para cada n, entonces para cada k se tiene que

$$f_n(x'_k) \leq f_n(x) \leq f_n(x''_k),$$

y además como $\{f_n\}_{n=0}^{\infty}$ converge a f en D

$$\lim_{n\longrightarrow\infty} f_n(x'_k) = f(x'_k) \quad \text{y} \quad \lim_{n\longrightarrow\infty} f_n(x''_k) = f(x''_k),$$

por lo que se obtiene

$$f(x'_k) \leq \liminf_{n\longrightarrow\infty} f_n(x) \leq \limsup_{n\longrightarrow\infty} f_n(x) \leq f(x''_k). \tag{C.1}$$

Como x es punto de continuidad de f,

$$\lim_{k\longrightarrow\infty} f(x'_k) = \lim_{k\longrightarrow\infty} f(x''_k) = f(x),$$

por lo que haciendo tender k a infinito en (C.1) se tiene que

$$f(x) \leq \liminf_{n\longrightarrow\infty} f_n(x) \leq \limsup_{n\longrightarrow\infty} f_n(x) \leq f(x).$$

Entonces $\lim_{n\longrightarrow\infty} f_n(x)$ existe y es igual a $f(x)$.

$\square$

Teorema C.3. *(Primer teorema de Helly)*
Sea $\{f_n\}_{n=0}^{\infty}$ una sucesión de funciones no decrecientes y $A, B \in \mathbb{R}$ tales que $A \leq f_n(x) \leq B$ para todo x y toda n. Entonces la sucesión dada contiene una subsucesión $\{f_{n_k}\}_{k=0}^{\infty}$ que converge débilmente a una función f, no decreciente y acotada.

Demostración. Sea $D = \{x_0, x_1, ...\}$ un conjunto denso numerable de reales. La sucesión $\{f_n(x_0)\}_{n=0}^{\infty}$ es acotada por lo que contiene una subsucesión

$\{f_{0n}(x_0)\}_{n=0}^{\infty}$ que converge a un cierto límite L_0.

De la misma forma, la sucesión $\{f_{0n}(x_1)\}_{n=0}^{\infty}$ contiene una subsucesión $\{f_{1n}(x_1)\}_{n=0}^{\infty}$ convergente a cierto límite L_1, que también cumple

$$\lim_{n\longrightarrow\infty} f_{1n}(x_0) = \lim_{n\longrightarrow\infty} f_{0n}(x_0) = L_0.$$

Continuando este proceso se obtiene que para cualquier natural k, existen k sucesiones $\{f_{kn}(x_i)\}_{n=0}^{\infty},\ i = 0, 1, ..., k-1$, para las cuales

$$\lim_{n\longrightarrow\infty} f_{kn}(x_i) = L_i. \tag{C.2}$$

Considérese ahora la sucesión diagonal compuesta por las funciones $\{f_{nn}\}_{n=0}^{\infty}$.

Sea $x_k \in D$ para alguna $k \in \mathbb{N}$. Entonces

$$\lim_{n\longrightarrow\infty} f_{nn}(x_k) = L_k,$$

por (C.2) y dado que $\{f_{nn}\}_{n=0}^{\infty} \subseteq \{f_{kn}\}_{n=0}^{\infty}$, si $n \geq k$. De esta forma se define la función f como

$$f(x_k) = L_k = \lim_{n\longrightarrow\infty} f_{nn}(x_k),$$

para $k = 0, 1, 2, ...$, la cual está definida para todo elemento de D.

Si $x < y$ son dos elementos en D, como f_{nn} es no decreciente para cada $n = 0, 1, 2, ...$, entonces

$$f(x) = \lim_{n\longrightarrow\infty} f_{nn}(x) \leq \lim_{n\longrightarrow\infty} f_{nn}(y) = f(y),$$

por lo que f es no decreciente en D.

Además, dado que para toda $k = 0, 1, 2, ...,\ A \leq f(x_k) = L_k \leq B$ se tiene que f es acotada en D.

Estas propiedades permiten extender la función f a toda la recta real, conservando las propiedades antes mencionadas. Luego, por el Lema C.2 se tiene que $\{f_{nn}\}_{n=0}^{\infty}$ converge débilmente a f.

$\square$

Para el siguiente teorema se recordará un resultado sobre las funciones monótonas.

Teorema C.4. *Sea f una función monótona en $[a,b]$. Entonces el conjunto de discontinuidades de f es a lo más numerable.*

Demostración. Supóngase que f es no decreciente, si f es no creciente se puede considerar a la función $-f$ y el siguiente desarrollo se trabajaría de la misma forma. Sea S_m el conjunto de puntos de discontinuidad de f en los que el salto es mayor a $1/m$, con $m > 0$, es decir, si $x \in S_m$ se tiene que $f(x+) - f(x-) > 1/m$. Entonces

$$f(b) - f(a) \geq \sum_{x \in S_m} f(x+) - f(x-) > \frac{|S_m|}{m},$$

donde $|S_m|$ es el número de elementos del conjunto S_m que, por las desigualdades anteriores, debe ser finito. Finalmente, el conjunto de discontinuidades de f es un subconjunto de $\bigcup_{m=1}^{\infty} S_m$ por lo que es a lo más numerable.

$\square$

Teorema C.5. *(Segundo teorema de Helly)*
Sean f, $\{f_n\}_{n=0}^{\infty}$ funciones no decrecientes y acotadas tales que la sucesión converge completamente a f. Entonces para cualquier función g continua y acotada

$$\int_{-\infty}^{\infty} g(x)df(x) = \lim_{n \longrightarrow \infty} \int_{-\infty}^{\infty} g(x)df_n(x).$$

Demostración. Sea $\epsilon > 0$ y sean

$$B = \sup_{x \in \mathbb{R}} |g(x)| \quad \text{y} \quad C = f(\infty) - f(-\infty) \geq 0.$$

Si $f(x) = c$ para todo $x \in \mathbb{R}$, es claro que

$$\int_{-\infty}^{\infty} g(x) df(x) = 0,$$

y por la convergencia completa se tiene que

$$|f_n(x) - c| \leq \frac{\epsilon}{2B},$$

para n suficientemente grande, por lo que

$$c - \frac{\epsilon}{2B} \leq f_n(x) \leq c + \frac{\epsilon}{2B},$$

de donde se obtiene que

$$\int_{-\infty}^{\infty} g(x) df_n(x) \leq B(f_n(\infty) - f_n(-\infty)) \leq B\left(c + \frac{\epsilon}{2B} - \left(c - \frac{\epsilon}{2B}\right)\right) = \epsilon,$$

por lo tanto

$$\lim_{n \longrightarrow \infty} \int_{-\infty}^{\infty} g(x) df_n(x) = \int_{-\infty}^{\infty} g(x) df(x) = 0.$$

Suponiendo que f no es constante, por el Teorema C.4 y como

$$\lim_{x \longrightarrow \infty} f(x) = f(\infty) \quad \text{y} \quad \lim_{x \longrightarrow -\infty} f(x) = f(-\infty),$$

existen $a < b$, puntos de continuidad de $f(x)$, tales que $f(a) < f(b)$ y

$$f(\infty) - f(b) < \frac{\epsilon}{9B} \quad \text{y} \quad f(a) - f(-\infty) < \frac{\epsilon}{9B}. \tag{C.3}$$

Como $g(x)$ es una función continua existe $P = \{x_0, x_1, ..., x_N\}$ una partición del intervalo $[a, b]$, formada por puntos de continuidad de $f(x)$. Además por el Teorema C.4 los puntos de continuidad de P pueden tomarse de tal forma que si $x \in (x_{k-1}, x_k)$ entonces para $k = 1, 2, ...N$ se tiene que

$$|g(x) - g(x_k)| < \delta = \frac{\epsilon}{9(f(b) - f(a))}. \tag{C.4}$$

Por otro lado, considérese a la función auxiliar

$$g_0(x) = \begin{cases} g(x_k) & \text{si} \quad x \in (x_{k-1}, x_k], \\ 0 & \text{en otro caso,} \end{cases}$$

entonces por (C.4), $|g(x) - g_0(x)| < \delta$ en $[a, b]$. Luego

$$\begin{aligned} \int_{-\infty}^{\infty} g(x) df_n(x) - \int_{-\infty}^{\infty} g(x) df(x) &= \int_{-\infty}^{\infty} (g(x) - g_0(x)) df_n(x) \\ &\quad + \int_{-\infty}^{\infty} g_0(x) df_n(x) - \int_{-\infty}^{\infty} g_0(x) df(x) \\ &\quad + \int_{-\infty}^{\infty} (g_0(x) - g(x)) df(x) \\ &= S_1 + S_2 + S_3. \end{aligned}$$

Ahora se acotará a cada uno de los tres sumandos anteriores. Para S_3 se tiene que

$$\begin{aligned} |S_3| &\leq \int_{-\infty}^{a} |g(x)| df(x) + \int_{a}^{b} |g(x) - g_0(x)| df(x) + \int_{b}^{\infty} |g(x)| df(x) \\ &\leq B(f(a) - f(-\infty)) + \delta(f(b) - f(a)) + B(f(\infty) - f(b)) \\ &\leq \epsilon/9 + \epsilon/9 + \epsilon/9 \\ &= \epsilon/3, \end{aligned}$$

debido a (C.3) y (C.4).

Para S_1,

$$
\begin{aligned}
|S_1| &\leq \int_{-\infty}^{a} |g(x)| df_n(x) + \int_{a}^{b} |g(x) - g_0(x)| df_n(x) + \int_{b}^{\infty} |g(x)| df_n(x) \\
&\leq B(f_n(a) - f_n(-\infty)) + \delta(f_n(b) - f_n(a)) + B(f_n(\infty) - f_n(b)) \\
&\underset{n\longrightarrow\infty}{\longrightarrow} B(f(a) - f(-\infty)) + \delta(f(b) - f(a)) + B(f(\infty) - f(b)) \\
&\leq \epsilon/9 + \epsilon/9 + \epsilon/9 \\
&= \epsilon/3,
\end{aligned}
$$

debido a (C.3), (C.4) y a que $\{f_n\}_{n=0}^{\infty}$ converge completamente a f.

Finalmente, utilizando también que $\{f_n\}_{n=0}^{\infty}$ converge completamente a f, se tiene que

$$
\begin{aligned}
\int_{-\infty}^{\infty} g_0(x) df_n(x) &= \sum_{k=1}^{N} g(x_k)(f_n(x_k) - f_n(x_{k-1})) \\
&\qquad \underset{n\longrightarrow\infty}{\longrightarrow} \sum_{k=1}^{N} g(x_k)(f(x_k) - f(x_{k-1})) = \int_{-\infty}^{\infty} g_0(x) df(x),
\end{aligned}
$$

por lo que

$$
S_2 = \int_{-\infty}^{\infty} g_0(x) df_n(x) - \int_{-\infty}^{\infty} g_0(x) df(x) \underset{n\longrightarrow\infty}{\longrightarrow} 0. \tag{C.5}
$$

Por lo tanto, para n suficientemente grande, se tiene que

$$
\left| \int_{\infty}^{\infty} g(x) df_n(x) - \int_{\infty}^{\infty} g(x) df(x) \right| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.
$$

$\square$

En el presente trabajo se utiliza una variante del resultado anterior en la cual se cambia el dominio de las funciones f, f_n y g a un intervalo cerrado $[a, b]$ obteniendo lo siguiente.

Teorema C.6. *(Variante segundo teorema de Helly)*
Sean f, $\{f_n\}_{n=0}^{\infty}$ funciones no decrecientes y acotadas en $[a,b]$ y g continua y acotada en $[a,b]$, tales que la sucesión converge débilmente a f, donde a y b son puntos de continuidad de f. Entonces

$$\int_a^b g(x)df(x) = \lim_{n\longrightarrow\infty}\int_a^b g(x)df_n(x).$$

Para la prueba de este resultado simplemente se sustituye el hecho de que

$$\lim_{x\longrightarrow\infty} f(x) = f(\infty) \quad \text{y} \quad \lim_{x\longrightarrow-\infty} f(x) = f(-\infty),$$

por

$$\lim_{x\longrightarrow a} f(x) = f(a) \quad \text{y} \quad \lim_{x\longrightarrow b} f(x) = f(b),$$

en la demostración proporcionada para el segundo teorema de Helly.

D. Determinantes de sucesiones definidas positivas

A continuación se muestran ejemplos de los determinantes que se vieron en el Teorema 1.12 para sucesiones definidas positivas particulares con la intención de mostrar que su signo es positivo. La expresión general de dichos determinantes esta dada por

$$D_n = \begin{vmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{vmatrix}, \quad \text{donde} \quad n = 0, 1, 2, \dots$$

Primero considérese una variable aleatoria X con distribución $\mathrm{N}(0,1)$, para la cual se tiene que el n-ésimo momento está dado por

$$E(X^n) = \begin{cases} \dfrac{n!}{(n/2)!}\left(\dfrac{1}{2}\right)^{n/2} & \text{si } n \text{ es par,} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

Entonces, se puede verificar que los primeros trece momentos están dados por $1, 0, 1, 0, 3, 0, 15, 0, 105, 0, 945, 0, 10395$, y también que los determinantes para $n = 0, 1, ..., 6$ son

$$D_0 = |1|,$$

$$D_1 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1,$$

$$D_2 = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{vmatrix} = 2,$$

$$D_3 = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 3 \\ 1 & 0 & 3 & 0 \\ 0 & 3 & 0 & 15 \end{vmatrix} = 12,$$

$$D_4 = \begin{vmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & 0 & 3 & 0 \\ 1 & 0 & 3 & 0 & 15 \\ 0 & 3 & 0 & 15 & 0 \\ 3 & 0 & 15 & 0 & 105 \end{vmatrix} = 288,$$

$$D_5 = \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 & 0 & 15 \\ 1 & 0 & 3 & 0 & 15 & 0 \\ 0 & 3 & 0 & 15 & 0 & 105 \\ 3 & 0 & 15 & 0 & 105 & 0 \\ 0 & 15 & 0 & 105 & 0 & 945 \end{vmatrix} = 34560,$$

$$D_6 = \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 & 15 \\ 0 & 1 & 0 & 3 & 0 & 15 & 0 \\ 1 & 0 & 3 & 0 & 15 & 0 & 105 \\ 0 & 3 & 0 & 15 & 0 & 105 & 0 \\ 3 & 0 & 15 & 0 & 105 & 0 & 945 \\ 0 & 15 & 0 & 105 & 0 & 945 & 0 \\ 15 & 0 & 105 & 0 & 945 & 0 & 10395 \end{vmatrix} = 24883200.$$

Ahora considérese una variable aleatoria X con distribución Unif$(-1, 1)$, para la cual se tiene que el n-ésimo momento está dado por

$$E(X^n) = \frac{1^{n+1} - (-1)^{n+1}}{2(n+1)} = \begin{cases} \dfrac{1}{n+1} & \text{si } n \text{ es par,} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

Por lo tanto, es fácil verificar que los primeros trece momentos están dados por $1, 0, 1/3, 0, 1/5, 0, 1/7, 0, 1/9, 0, 1/11, 0, 1/13$, y que los determinantes para $n = 0, 1, ..., 6$ son

$$D_0 = |1|,$$

$$D_1 = \begin{vmatrix} 1 & 0 \\ 0 & 1/3 \end{vmatrix} \approx 0.333,$$

$$D_2 = \begin{vmatrix} 1 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1/3 & 0 & 1/5 \end{vmatrix} \approx 0.0296,$$

$$D_3 = \begin{vmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/5 \\ 1/3 & 0 & 1/5 & 0 \\ 0 & 1/5 & 0 & 1/7 \end{vmatrix} \approx 0.000677,$$

$$D_4 = \begin{vmatrix} 1 & 0 & 1/3 & 0 & 1/5 \\ 0 & 1/3 & 0 & 1/5 & 0 \\ 1/3 & 0 & 1/5 & 0 & 1/7 \\ 0 & 1/5 & 0 & 1/7 & 0 \\ 1/5 & 0 & 1/7 & 0 & 1/9 \end{vmatrix} \approx 0.00000393,$$

$$D_5 = \begin{vmatrix} 1 & 0 & 1/3 & 0 & 1/5 & 0 \\ 0 & 1/3 & 0 & 1/5 & 0 & 1/7 \\ 1/3 & 0 & 1/5 & 0 & 1/7 & 0 \\ 0 & 1/5 & 0 & 1/7 & 0 & 1/9 \\ 1/5 & 0 & 1/7 & 0 & 1/9 & 0 \\ 0 & 1/7 & 0 & 1/9 & 0 & 1/11 \end{vmatrix} \approx 0.00000000576,$$

$$D_6 = \begin{vmatrix} 1 & 0 & 1/3 & 0 & 1/5 & 0 & 1/7 \\ 0 & 1/3 & 0 & 1/5 & 0 & 1/7 & 0 \\ 1/3 & 0 & 1/5 & 0 & 1/7 & 0 & 1/9 \\ 0 & 1/5 & 0 & 1/7 & 0 & 1/9 & 0 \\ 1/5 & 0 & 1/7 & 0 & 1/9 & 0 & 1/11 \\ 0 & 1/7 & 0 & 1/9 & 0 & 1/11 & 0 \\ 1/7 & 0 & 1/9 & 0 & 1/11 & 0 & 1/13 \end{vmatrix} \approx 0.00000000000213.$$

En ambos ejemplos se observa, como se esperaba, que el signo de los determinantes es positivo. Además se puede apreciar que en el primer ejercicio el valor de los determinantes va incrementando rápidamente, mientras que en el segundo el valor de los determinantes va disminuyendo.

Bibliografía

[1] Akhiezer N. I. *The Classical Moment Problem and Some Related Questions in Analysis.* Oliver & Boyd, 1965.

[2] Apostol T. M. *Calculus, Volumen II.* Reverté, 2002.

[3] Apostol T. M. *Mathematical Analysis.* Addison-Wesley Publishing Company, 1974.

[4] Arnol'd V. I. *Ordinary Differential Equations.* Springer-Verlag, 1992.

[5] Howard A. *Introducción al Álgebra Lineal.* Limusa, 2001.

[6] Kurosch A.G. *Curso de Álgebra Superior.* Mir, 1968.

[7] Petrov V., Mordecki E. *Teoría de Probabilidades.* URSS, 2003.

[8] Rincón L. A. *Introducción a la Probabilidad.* Las Prensas de Ciencias, 2014.

[9] Rivera F. J. *Clases de Stieltjes.* Tesis de Licenciatura, Facultad de Ciencias UNAM, 2012.

[10] Spivak M. *Calculus: Cálculo Infinitesimal.* Reverté, 1992.

[11] Stoer J., Bulirsch R. *Introduction to Numerical Analysis.* Springer-Verlag, 1993.