

Yasmina Khlar Viana

# Condicionamiento de fórmulas de interpolación

Departamento

Instituto Universitario de Investigación en  
Matemáticas y sus Aplicaciones

Director/es

PEÑA FERRÁNDEZ, JUAN M.  
CARNICER ÁLVAREZ, JESÚS MIGUEL

<http://zaguan.unizar.es/collection/Tesis>



Reconocimiento – NoComercial – SinObraDerivada (by-nc-nd): No se permite un uso comercial de la obra original ni la generación de obras derivadas.

© Universidad de Zaragoza  
Servicio de Publicaciones

ISSN 2254-7606

Tesis Doctoral

# CONDICIONAMIENTO DE FÓRMULAS DE INTERPOLACIÓN

Autor

Yasmina Khier Viana

Director/es

PEÑA FERRÁNDEZ, JUAN M.  
CARNICER ÁLVAREZ, JESÚS MIGUEL

**UNIVERSIDAD DE ZARAGOZA**

Instituto Universitario de Investigación en Matemáticas y sus Aplicaciones

2018





**Universidad**  
Zaragoza

## Tesis Doctoral

# Condicionamiento de fórmulas de interpolación

Autora

**Yasmina Khiar Viana**

Directores

**Jesús Miguel Carnicer Álvarez**  
**Juan Manuel Peña Ferrández**

UNIVERSIDAD DE ZARAGOZA  
Departamento de Matemática Aplicada – IUMA  
2018



Instituto Universitario de Investigación  
**de Matemáticas**  
**y Aplicaciones**  
**Universidad Zaragoza**



# Agradecimientos

---

En primer lugar, quiero agradecer a Jesús Carnicer y Juan Manuel Peña por todo el trabajo y los consejos que me han dedicado y que me han guiado durante estos años. Su ayuda y experiencia han sido esenciales para la realización de esta tesis.

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad–Fondo Europeo de Desarrollo Regional (FEDER) MTM2012-31544 y MTM2015-65433-P, la Diputación General de Aragón–Fondo Social Europeo y la ayuda para contratos predoctorales para la formación de doctores del Ministerio de Economía y Competitividad BES-2013-065398.

Gracias a Tomas Sauer, responsable de una estancia de tres meses que realicé en el centro Chair of Mathematics/Image Processing and FORWISS en la Universidad de Passau, Alemania.

También quiero agradecer al Departamento de Matemática Aplicada de la Universidad de Zaragoza y al Instituto Universitario de Matemáticas y Aplicaciones (IUMA). En especial, quiero dar las gracias a Luis Rández por el apoyo y la confianza que siempre me ha demostrado.

A mis padres y a mi hermana, gracias por su apoyo incondicional. Sin ellos no habría llegado hasta aquí.

Gracias también a tantos amigos que han participado, directa o indirectamente, en esta tesis. Gracias a todos mis animalicos.

Finalmente, quiero dar las gracias de una manera muy especial a Eduardo. Sin su ayuda y su paciencia nada de esto sería posible.



# Índice general

---

<b>Introducción</b>	<b>v</b>
<b>Introduction</b>	<b>xi</b>
<b>1 Operadores de cuasi-interpolación</b>	<b>1</b>
1.1 Representaciones de operadores de cuasi-interpolación . . . . .	1
1.2 Condicionamientos de las representaciones . . . . .	4
1.3 Condicionamiento de Skeel y cotas de normas . . . . .	10
<b>2 Óptima estabilidad de la fórmula de Lagrange y estabilidad de la fórmula de Newton</b>	<b>17</b>
2.1 El operador de interpolación de Lagrange . . . . .	17
2.2 Estabilidad de la fórmula de Newton en interpolación polinómica . . . . .	23
2.3 Alta precisión relativa de las correspondientes matrices triangulares inferiores . . . . .	27
2.4 Diferentes normalizaciones de la fórmula de Newton . . . . .	28
2.5 Matrices totalmente positivas . . . . .	32
2.6 Matrices de Vandermonde, factorizaciones triangulares y alta precisión relativa . . . . .	36
<b>3 Estabilidad de la fórmula de Newton para diferentes elecciones de nodos con orden creciente</b>	<b>41</b>
3.1 Estabilidad de la fórmula de Newton para nodos equidistantes ordenados	41
3.2 Comparación de las normalizaciones de la fórmula de Newton en nodos equidistantes ordenados . . . . .	50
3.3 Mallas generales con nodos ordenados . . . . .	62
3.4 Nodos en progresión geométrica . . . . .	66

<b>4</b>	<b>Estabilidad de la fórmula de Newton para otras ordenaciones de nodos</b>	<b>83</b>
4.1	Orden de Leja . . . . .	83
4.2	Orden central . . . . .	89
4.3	Orden central respecto al punto de evaluación . . . . .	95
4.4	Orden central inverso . . . . .	100
4.5	Matriz de Vandermonde, eliminación y pivotaje . . . . .	118
4.6	Experimentos numéricos con el condicionamiento de las factorizaciones triangulares de una matriz de Vandermonde . . . . .	132
<b>5</b>	<b>Representaciones con bases ortogonales</b>	<b>141</b>
5.1	Aproximación por mínimos cuadrados . . . . .	141
5.2	Representaciones con bases ortogonales y normas . . . . .	143
5.3	Mínimos cuadrados en el espacio de funciones continuas . . . . .	144
5.4	Condicionamiento con polinomios ortogonales clásicos . . . . .	147
5.5	Condicionamiento de problemas de mínimos cuadrados discretos . . . . .	155
5.6	Condicionamiento del problema de interpolación de Lagrange en los ceros de los polinomios ortogonales . . . . .	162
<b>6</b>	<b>Bases y funcionales positivos</b>	<b>165</b>
6.1	Caracterización de funcionales no negativos . . . . .	165
6.2	Comparación entre sistemas de funcionales no negativos . . . . .	167
6.3	Comparación entre bases no negativas . . . . .	170
6.4	Bases totalmente positivas . . . . .	172
6.5	Base de Bernstein . . . . .	173
	<b>Conclusiones</b>	<b>179</b>
	<b>Conclusions</b>	<b>183</b>
	<b>Bibliografía</b>	<b>187</b>

# Introducción

---

Uno de los problemas que más influencia ha tenido en el desarrollo del Análisis Matemático moderno es el de la aproximación de funciones por polinomios trigonométricos. Fourier propuso en [32], para la resolución de la ecuación del calor, encontrar representaciones de las condiciones de contorno como series trigonométricas, llamadas a partir de entonces series de Fourier. Diferentes estudios mostraron que la serie de Fourier de una función continua podía ser divergente y se propusieron algunas condiciones para garantizar la convergencia de dicha serie. Lebesgue en 1909 [40] observó que el estudio del crecimiento de la norma del operador que a cada función asocia su  $n$ -ésima suma de Fourier permitía deducir resultados de convergencia y divergencia. La importancia de la divergencia de la norma de ciertas sucesiones de operadores dio lugar al principio de acotación uniforme, que Banach y Steinhaus [4] demostraron en un contexto muy general. Faber en 1914 [29] aplicó los resultados de divergencia de la constante de Lebesgue de los problemas de series de Fourier trigonométricas para demostrar que la norma de los operadores de interpolación de Lagrange en los espacios de polinomios trigonométricos y, también de polinomios algebraicos, diverge cuando el grado tiende a infinito. De aquí, Faber dedujo que para todo proceso de interpolación siempre existe una función tal que la sucesión de interpolantes diverge. Un primer ejemplo que ilustra esta divergencia de los interpolantes polinómicos en nodos equidistantes ya fue descrito por Runge en 1901 [55]. Estas propiedades de divergencia se observaron en problemas más generales. Cabe destacar que Nikolaev [45] dedujo la divergencia de los procesos de aproximación en problemas generales de mínimos cuadrados. En todos los casos, la demostración de la divergencia de la norma del operador juega un papel fundamental. Por ello, en honor a Lebesgue, la norma de los operadores asociados a problemas de aproximación reciben el nombre de constante de Lebesgue.

La constante de Lebesgue no solamente permite demostrar resultados de convergencia o divergencia de los procesos de aproximación. También es una medida de la estabilidad del proceso de aproximación. Por ello, la búsqueda de nodos de interpolación tales que

la constante de Lebesgue  $\Lambda_n$  del problema de interpolación de Lagrange de grado  $n$  sea pequeña ha sido objeto de constante estudio. En 1931 Bernstein [6] demostró que la constante de Lebesgue del problema de interpolación de grado  $n$  en los nodos de Chebyshev tiene un crecimiento asintótico del tipo  $2 \log(n+1)/\pi$ . Rivlin demostró en el Teorema 1.2 de [52] que la constante de Lebesgue para nodos de Chebyshev verifica la cota  $\Lambda_n \leq 1 + 2 \log(n+1)/\pi$ . Erdős [28] demostró que la constante de Lebesgue para nodos cualesquiera siempre era mayor que  $2 \log(n+1)/\pi - c$ , donde  $c$  es una constante positiva. Brutman [10] mejoró la cota obtenida por Erdős y demostró  $\Lambda_n \geq 2 \log(n+1)/\pi + 2(\gamma + \log 4/\pi)/\pi$ , donde  $\gamma \approx 0.5772156649$  es la constante de Euler–Mascheroni. Aunque existe una distribución de nodos tal que la constante de Lebesgue es mínima, estos nodos no han sido descritos de manera práctica. Sin embargo, la distribución de nodos de Chebyshev y otras relacionadas proporcionan constantes de Lebesgue próximas a la óptima.

El problema de interpolación de Lagrange consiste en encontrar un polinomio que coincida con una función dada en puntos distintos llamados nodos. El interpolante puede expresarse mediante fórmulas distintas. La fórmula de Lagrange expresa directamente el interpolante en términos de los valores de la función a través de los polinomios fundamentales de Lagrange  $l_0, \dots, l_n$ . Debido a que la constante de Lebesgue verifica  $\Lambda_n = \max_{x \in [a,b]} \sum_{i=0}^n |l_i(x)|$ , la propagación del error en la fórmula de Lagrange está relacionada con la constante de Lebesgue. Otra forma de representar el interpolante de Lagrange es la fórmula de Newton asociada a una sucesión de nodos. Esta fórmula tiene diferentes propiedades de estabilidad dependiendo del orden de los nodos. A pesar de que las fórmulas en las que los nodos están ordenados de manera creciente (fórmula progresiva) o decreciente (fórmula regresiva) parecen las más naturales, existe un amplio catálogo de variantes de la fórmula de Newton en las que los nodos se presentan en otros órdenes. El orden central, del que se hablará en la Sección 4.2, se encuentra implícito en las fórmulas basadas en diferencias centrales, entre las que destacan las fórmulas Gaussianas, de Stirling, fórmulas de Everett y de Bessel (véase el Capítulo 4 de [59]). Esta gran variedad de fórmulas está motivada no sólo para facilitar los cálculos en las tablas, sino también para controlar la propagación de los errores. De hecho, cuando los nodos no tienen una estructura especial se recomienda utilizar el orden central respecto al punto de evaluación (véase el párrafo 42 del Capítulo 5 de [59]), que estudiaremos en la Sección 4.3. En esta memoria introduciremos un condicionamiento de la fórmula de Newton que nos permitirá realizar un análisis cuantitativo de la estabilidad de dicha fórmula con distintas ordenaciones de los nodos.

La búsqueda de buenos conjuntos de nodos y cómo ordenarlos para aplicar eficiente-

mente la fórmula de Newton ha sido también objeto de estudio. Debido a la optimalidad de la norma de los polinomios de Chebyshev, la elección de los nodos Chebyshev proporciona cotas óptimas del error de interpolación, además del lento crecimiento de la constante de Lebesgue. Los polinomios de Fekete, cuyos ceros son los nodos de Fekete, están muy relacionados con los de Chebyshev. En Teoría del Potencial, los nodos de Fekete representan una distribución de cargas que minimizan la energía electrostática. La distribución límite de los nodos de Chebyshev y de Fekete corresponden a la medida de equilibrio, permitiendo aproximar la capacidad logarítmica del dominio (véase [56] y [57]). Los nodos propuestos por Leja [41], mediante un criterio simplificado, también proporcionar buenas aproximaciones de la capacidad logarítmica. Reichel [51] definió un condicionamiento asociado a la fórmula de Newton y mostró que este condicionamiento con nodos de Leja tiene un crecimiento subexponencial. Taylor [61] proporciona condiciones suficientes para el crecimiento subexponencial de las constantes de Lebesgue de un proceso de interpolación y demostró que la constante de Lebesgue en nodos de Leja tiene un crecimiento subexponencial. El criterio de selección de los nodos de Leja puede utilizarse para reordenar los puntos de un conjunto cualquiera. En la práctica, la ordenación de Leja presenta buenas propiedades de estabilidad.

Otras representaciones muy utilizadas en problemas de mínimos cuadrados son las representaciones de las aproximaciones a través de polinomios ortogonales. En esta memoria definiremos un condicionamiento general para las representaciones de problemas de aproximación. El resultado negativo de Nikolaev predice un crecimiento al menos logarítmico para cualquier representación. Veremos que el condicionamiento de las representaciones en términos de polinomios ortogonales más usuales tiene un crecimiento del tipo  $n^\alpha$  y, por tanto, subexponencial. En el importante caso de los polinomios de Chebyshev, el crecimiento del condicionamiento es lineal.

Las funciones polinómicas a trozos, también llamadas funciones spline, han recibido la atención de los investigadores en aproximación desde mitad del siglo XX. Una de las motivaciones para su uso es la estabilidad de las aproximaciones. Esta estabilidad se revela de varias maneras. Por un lado, las matrices que aparecen en la resolución de interpolación y mínimos cuadrados son matrices que están bien condicionadas. Por otro lado, existe una base, llamada base de B-splines (véase el Capítulo IX de [7]), que tiene importantes propiedades de estabilidad. Para demostrar el buen condicionamiento de la base de B-splines (véase el Capítulo XI de [7]) se introdujeron los funcionales de de Boor–Fix [8]. Estos funcionales son funcionales duales de la base de B-splines y permiten definir cuasi-interpolantes que dan lugar a representaciones cuyo condicionamiento puede acotarse en términos del grado, independientemente del número y posición de los nudos.

Precisamente, el nombre de cuasi-interpolante proviene de los operadores que se definen a través de funcionales duales en el contexto de las funciones spline. Para demostrar la existencia de cuasi-interpolantes con condicionamiento acotado en determinados espacios de funciones es preciso utilizar el Teorema de Hahn–Banach. En limitados casos pueden ofrecerse construcciones explícitas de estos cuasi-interpolantes. Para poder demostrar el poder de aproximación de los espacios de funciones spline también se utilizan los cuasi-interpolantes (véase el Capítulo XII de [7]). Entre los operadores de cuasi-interpolación definidos en espacios de funciones spline, podemos destacar el operador de Schoenberg.

El Teorema de Schoenberg–Whitney presenta condiciones necesarias y suficientes para la existencia y unicidad de solución del problema de interpolación en espacios de funciones spline. Una demostración de este teorema basada en la total positividad de la matriz de colocación de la base de los B-splines fue obtenida por de Boor (véase el Capítulo XIII de [7]). La total positividad de la base de los B-splines está relacionada con sus propiedades de estabilidad para la resolución de sistemas de ecuaciones lineales basados en la matriz de colocación (véase [9], [46]) o para algoritmos de evaluación (véase el Capítulo X de [7]). La total positividad proporciona herramientas para detectar bases con condicionamiento óptimo, como se muestra en [18], [48].

A pesar de los múltiples resultados sobre el condicionamiento con grado acotado, las técnicas de cuasi-interpolantes apenas se han utilizado en los clásicos problemas de interpolación y aproximación polinómica con grado  $n$ , ya que cuando  $n \rightarrow \infty$  el condicionamiento no está acotado, como indican los resultados de Faber y Nikolaev. En esta memoria ofrecemos medidas de la estabilidad en los problemas de aproximación por polinomios utilizando funcionales duales y cuasi-interpolantes.

En el Capítulo 1 estudiamos los operadores de cuasi-interpolación y sus representaciones. Definimos un condicionamiento asociado a cada representación. Recordamos el condicionamiento de Skeel de una matriz y lo aplicamos a la acotación de condicionamientos. Los capítulos 2, 3 y 4 están dedicados al estudio de operador de interpolación de Lagrange y sus representaciones.

En el Capítulo 2 demostramos que la representación de Lagrange es óptima y analizamos la estabilidad de la fórmula de Newton, que depende del orden de los nodos. Recordamos el concepto de matriz totalmente positiva y algunos tipos de matrices importantes en interpolación, como las matrices de Vandermonde. Demostramos que la inversa de una matriz de Vandermonde puede calcularse con alta precisión relativa (HRA) bajo ciertas condiciones. Consideramos las factorizaciones triangulares de Crout y de Doolittle de la matriz de Vandermonde y proponemos algoritmos para el cálculo de las correspondientes matrices triangulares y sus inversas con HRA. Dichas matrices triangulares inferiores

están relacionadas con diferentes normalizaciones de la fórmula de Newton y se pueden utilizar para la acotación del condicionamiento de la representación de Newton. Parte del material de este capítulo ha sido publicado en [14], [15].

En el Capítulo 3 se recuerda que la constante de Lebesgue con  $n + 1$  nodos equidistantes tiene un crecimiento del tipo  $2^n$ . Se estudia la estabilidad de la fórmula de Newton con nodos equidistantes en orden creciente y se generalizan estas técnicas para abordar los casos de nodos casi equidistantes y nodos en progresión geométrica. En el caso de nodos equidistantes demostramos que el condicionamiento está acotado por  $3^n$  y que dicho valor se alcanza en el último nodo  $x_n$ . Deducimos que el condicionamiento de Skeel de las inversas de las matrices triangulares inferiores asociadas a diferentes representaciones de la fórmula de Newton es  $3^n$ . Como caso particular importante obtenemos el condicionamiento de Skeel y el condicionamiento en norma  $\infty$  de la matriz de Pascal triangular inferior. Realizamos un estudio del condicionamiento en norma  $\infty$  de la matriz  $L$  que permite obtener las diferencias divididas en términos de los valores de la función. Este condicionamiento depende de la longitud del intervalo. Demostramos que el condicionamiento tiene un crecimiento exponencial y que los intervalos de longitud 3 corresponden a un crecimiento asintótico óptimo.

En el Capítulo 4 consideramos diferentes ordenaciones de nodos para la fórmula de Newton. Comenzamos con el orden de Leja y demostramos que el máximo condicionamiento en los nodos está acotado por  $2^{n+1} - 1$ . Sin embargo, los experimentos numéricos en nodos equidistantes muestran que esta cota es muy pesimista. La dificultad del análisis del orden de Leja no nos ha permitido mejorar la cota. Sin embargo, hemos estudiado en las restantes secciones otras ordenaciones con buenas propiedades y cuyo análisis es más simple. En la Sección 4.2 hemos considerado el orden central respecto a un centro y hemos realizado un análisis cuantitativo que permite justificar el uso de métodos relacionados con diferencias centrales. Demostramos que el condicionamiento de la fórmula de Newton está acotado por  $(1 + \sqrt{2})^{n+1}$ . Dedicamos especial atención al caso particular de orden central respecto al punto de evaluación que proporciona cotas muy ajustadas próximas a la constante de Lebesgue. También se obtienen cotas próximas a la constante de Lebesgue con el orden central inverso, estudiado en la Sección 4.4. En la Sección 4.5 interpretamos el orden de Leja y el central inverso en términos de eliminaciones matriciales con estrategias de pivotaje. Concretamente, recordamos la conexión entre el orden de Leja y la eliminación gaussiana con pivotaje parcial y demostramos una conexión similar entre el orden central inverso para nodos equidistantes y la eliminación de Neville con pivotaje parcial. Recordemos que la eliminación de Neville es un procedimiento de eliminación alternativo a la eliminación gaussiana en la que a cada fila se le resta

un múltiplo de la fila anterior (véase [34]). En la Sección 4.6 realizamos experimentos numéricos con el condicionamiento de las factorizaciones triangulares de una matriz de Vandermonde. Parte de dicho material ha sido publicado en [12], [13].

El Capítulo 5 está dedicado al importante caso de representaciones con bases ortogonales. Introducimos el problema de aproximación por mínimos cuadrados y vemos las buenas propiedades de los operadores en norma media cuadrática. En la Sección 5.3 analizamos el comportamiento en la norma  $\infty$ , que proporciona un condicionamiento más útil desde el punto de vista de propagación del error. Además, obtenemos fórmulas y cotas para el condicionamiento. En las secciones 5.4 y 5.5 estudiamos las representaciones en términos de polinomios ortogonales clásicos de los operadores asociados a problemas de mínimos cuadrados continuos y discretos. Estudiamos los casos de polinomios de Legendre y Chebyshev de primera especie y segunda especie. En el caso de Legendre el condicionamiento crece como  $n^{3/2}$ , para polinomios de Chebyshev de primera especie el crecimiento es lineal y para los de segunda especie es cuadrático. En el caso en que el número de nodos coincide con  $n + 1$ , la dimensión del espacio de los polinomios, tenemos problemas de interpolación en los ceros de polinomios ortogonales y el estudio anterior nos permite calcular el condicionamiento de representaciones en términos de polinomios ortogonales clásicos. A este respecto, cabe recordar las propiedades casi óptimas que tiene la constante de Lebesgue en los ceros de los polinomios de Chebyshev (véase [10], [52]).

En el Capítulo 6 estudiamos las consecuencias que tiene la positividad de las bases y de los funcionales en la búsqueda de bases que tengan un buen condicionamiento. Como consecuencia de los resultados, podemos destacar la optimalidad de la representación de Lagrange respecto a otras con funcionales positivos y la optimalidad de la base de Bernstein respecto a otras bases totalmente positivas, como la base de monomios.

# Introduction

---

One of the most influential problems in the development of modern mathematical analysis is the approximation of functions by trigonometric polynomials. For the solution of the heat equation, Fourier in [32] used representations of boundary conditions as trigonometric series, thereafter called Fourier series. Several studies showed that the Fourier series of a continuous function could be divergent and some conditions were proposed to ensure the convergence of this series. Lebesgue in 1909 [40] observed that the study of the growth of the norm of the operator that associates to each function its  $n$ -th Fourier sum can be used to deduce results of convergence and divergence. The importance of the divergence of the norm of certain sequences of operators gave rise to the uniform boundedness principle, proved by Banach and Steinhaus [4] in a very general context. Faber in 1914 [29] applied the divergence results of the Lebesgue constant of the trigonometric Fourier series problems to show that the norm of the Lagrange interpolation operators in trigonometric and algebraic polynomial spaces diverges when the degree tends to infinity. Hence, Faber deduced that, for every interpolation process, there is always a function such that the sequence of interpolating polynomials diverges. A first example illustrating this divergence of interpolating polynomial in equidistant nodes was already described by Runge in 1901 [55]. These divergence properties were observed in more general problems. It should be noted that Nikolaev [45] deduced the divergence of approximation processes of general least squares problems. In all cases, the proof of the divergence of the norm of the operator plays a fundamental role. Therefore, in honor of Lebesgue, the norm of operators associated to approximation problems is called the Lebesgue constant.

The Lebesgue constant can be used not only to show results of convergence or divergence of the approximation processes. It is also a measure of the stability of the approximation process. Therefore, the search of interpolation nodes such that the Lebesgue constant  $\Lambda_n$  of the Lagrange interpolation problem of degree  $n$  is small has been often considered. In 1931 Bernstein [6] showed that the Lebesgue constant of the problem of in-

terpolation of degree  $n$  in Chebyshev nodes has an asymptotic growth rate  $2 \log(n+1)/\pi$ . Rivlin showed in Theorem 1.2 of [52] that the Lebesgue constant for Chebyshev nodes satisfies  $\Lambda_n \leq 1 + 2 \log(n+1)/\pi$ . Erdős [28] proved that the Lebesgue constant of an arbitrary set of nodes is always greater than  $2 \log(n+1)/\pi - c$ , where  $c$  is a positive constant. Brutman [10] improved the bound obtained by Erdős and showed that  $\Lambda_n \geq 2 \log(n+1)/\pi + 2(\gamma + \log 4/\pi)/\pi$ , where  $\gamma \approx 0.5772156649$  is the constant of Euler–Mascheroni. Although there is a distribution of nodes such that the Lebesgue constant is minimal, these nodes have not been described in a practical way. However, the distribution of Chebyshev nodes and related nodes provide Lebesgue constants close to the optimal one.

The Lagrange interpolation problem consists in finding a polynomial whose values at different points, called nodes, are given. The interpolating polynomial can be expressed by different formulas. The Lagrange formula expresses the interpolating polynomial in terms of the values of the function and the Lagrange fundamental polynomials  $l_0, \dots, l_n$ . Since the Lebesgue constant  $\Lambda_n = \max_{x \in [a,b]} \sum_{i=0}^n |l_i(x)|$ , the propagation of the error in the Lagrange formula is related to the Lebesgue constant. Another representation of the Lagrange interpolating polynomial is the Newton formula associated with a sequence of nodes. This formula has different stability properties depending on the order of the nodes. Although the formulas with nodes arranged in an increasing (forward formula) or decreasing (backward formula) order seem the most natural, there is a wide range of variants of the Newton formula in which the nodes are presented in other orders. The central order, which will be discussed in Section 4.2, is implicit in the formulas based on central differences, including for example the Gaussian, Stirling, Everett and Bessel formulas (see Chapter 4 of [59]). This wide range of formulas is motivated not only to help in the computations in tables, but also to control the propagation of errors. In fact, when the nodes do not have a special structure, it is recommended to use the central order with respect to the evaluation point (see paragraph 42 of Chapter 5 of [59]), which we will study in Section 4.3. In this work we introduce a conditioning of the Newton formula which allows us to perform a quantitative analysis of the stability of this formula with different orderings of the nodes.

The search for good sets of nodes and how to order them to efficiently apply the Newton formula has also been object of study. Due to the optimality of the norm of Chebyshev polynomials, the choice of Chebyshev nodes provides optimal bounds of the interpolation error, in addition to the slow growth of the Lebesgue constant. The Fekete polynomials, whose zeros are the Fekete nodes, are closely related to Chebyshev ones. In Potential Theory, Fekete nodes represent a distribution of charges minimizing the elec-

trostatic energy. The limit distribution of the Chebyshev and Fekete nodes corresponds to the equilibrium measure, providing approximations to the logarithmic capacity of the domain (see [56] and [57]). The nodes proposed by Leja [41], using a simplified criterion, also provide good approximations to the logarithmic capacity. Reichel [51] defined a conditioning associated with the Newton formula and showed that this conditioning has subexponential growth for Leja nodes. Taylor [61] provided sufficient conditions for the subexponential growth of the Lebesgue constants of an interpolation process and showed that the Lebesgue constant in Leja nodes has a subexponential growth. The selection criteria for the Leja nodes can be used to reorder the points of any set. In practice, the Leja ordering presents good stability properties.

Other representations often used in least squares problems are the representations in terms of orthogonal polynomials. In this work we will define a general conditioning for the representations of approximation problems. Nikolaev's negative result predicts at least logarithmic growth for any representation. We will see that the conditioning of representations in terms of classical orthogonal polynomials has a growth of the type  $n^\alpha$  and, therefore, subexponential. In the important case of the Chebyshev polynomials, the conditioning growth is linear.

The piecewise polynomial functions, also called spline functions, have been studied since the mid-twentieth century. One of the motivations for its use is the stability of the approximations. This stability is revealed in several ways. On the one hand, the matrices arising in the solution of interpolation and least squares problems are well conditioned. On the other hand, there is a basis, called the basis of B-splines (see Chapter IX of [7]), which has important stability properties. In order to prove the good conditioning of the basis of B-splines (see Chapter XI of [7]), the de Boor–Fix functionals [8] were introduced. These functionals are dual to the basis of B-splines and can be used to define quasi-interpolants whose conditioning can be bounded in terms of the degree, independently of the number and position of the nodes. In fact, the name quasi-interpolant comes from the operators that are defined in terms of dual functionals in the context of spline functions. In order to show the existence of quasi-interpolants with bounded conditioning in certain spaces of functions, it is necessary to use the Hahn–Banach Theorem. In certain cases, explicit constructions of these quasi-interpolants can be offered. In order to prove the approximation power of spline spaces, quasi-interpolators are also used (see Chapter XII of [7]). Among the quasi-interpolation operators defined in spline spaces, we can mention the Schoenberg operator.

The Schoenberg–Whitney Theorem presents necessary and sufficient conditions for the existence and uniqueness of solution of the interpolation problem in spaces of spline

functions. A proof of this theorem based on the total positivity of the collocation matrix of the B-spline basis was obtained by de Boor (see Chapter XIII of [7]). The total positivity of the B-spline basis is related to its stability properties for the solution of systems of linear equations based on the collocation matrix ( see [9], [46]) or for evaluation algorithms (see Chapter X of [7]). Total positivity provides tools to find bases with optimal conditioning, as shown in [18], [48].

Despite the multiple results on conditioning with bounded degree, the techniques of quasi-interpolants have barely been used in the classical interpolation and polynomial approximation problems with degree  $n$ , because the conditioning is not bounded when  $n \rightarrow \infty$ , as the results of Faber and Nikolaev indicate. In this work we offer measures of the stability in polynomial approximation using dual functionals and quasi-interpolants.

In Chapter 1 we study quasi-interpolation operators and their representations. We define a conditioning associated with each presentation. We recall the Skeel condition number of a matrix and apply it to bound conditionings. Chapters 2, 3 and 4 are devoted to the study of the Lagrange interpolation operator and its representations.

In Chapter 2 we prove that the Lagrange representation is optimal and we analyze the stability of the Newton formula, which depends on the order of the nodes. We recall the concept of totally positive matrix and some types of important matrices in interpolation, such as the Vandermonde matrices. We show that the inverse of a Vandermonde matrix can be calculated with high relative accuracy (HRA) under certain conditions. We consider the triangular factorizations of Crout and Doolittle of the Vandermonde matrix and we propose algorithms for the computation of the corresponding triangular matrices and their inverses with HRA. These lower triangular matrices are related to different normalizations of the Newton formula and can be used for the conditioning of the Newton representation. Some results of this chapter have been published in [14], [15].

In Chapter 3 it is recalled that the Lebesgue constant with  $n + 1$  equidistant nodes has growth  $2^n$ . The stability of the Newton formula with equidistant nodes in increasing order is studied and these techniques are generalized to approach the cases of almost equidistant nodes and nodes in geometric progression. In the case of equidistant nodes we show that the conditioning is bounded above by  $3^n$  and that this value is attained in the last node  $x_n$ . We deduce that the Skeel conditioning of the inverses of the lower triangular matrices associated with different representations of the Newton formula is  $3^n$ . As an important particular case we obtain the conditioning of Skeel and the conditioning in  $\infty$  norm of the lower triangular Pascal matrix. We carried out an analysis of the conditioning in  $\infty$ -norm of the matrix  $L$  relating the divided differences with the values of the function. This conditioning depends on the length of the interval. We show that

the conditioning has an exponential growth and that the intervals of length 3 correspond to an optimal asymptotic growth.

In Chapter 4 we consider different orderings of nodes for the Newton formula. We start with the Leja order and show that the maximum conditioning in the nodes is bounded above by  $2^{n+1} - 1$ . However, the numerical experiments in equidistant nodes show that this bound is very pessimistic. The difficulty of analyzing the Leja order has not allowed us to improve the bound. However, we have studied in the remaining sections other node arrangements with good properties and whose analysis is simpler. In Section 4.2 we have considered the central order and we have performed a quantitative analysis that allows us to justify the use of methods related to central differences. We show that the conditioning of the Newton formula is bounded above by  $(1 + \sqrt{2})^{n+1}$ . We put attention to the particular case of central order with respect to the evaluation point, providing very tight bounds close to the Lebesgue constant. We also obtain close bounds to the Lebesgue constant with the inverse central order, studied in Section 4.4. In Section 4.5 we interpret the Leja and the inverse central orderings in terms of matrix eliminations with pivoting strategies. Specifically, we recall the connection between the Leja order and the Gaussian elimination with partial pivoting and we show a similar connection between the inverse central order for equidistant nodes and the Neville elimination with partial pivoting. Recall that the Neville elimination is an elimination algorithm alternative to Gaussian elimination where we subtract to each row a multiple of the previous one (see [34]). In Section 4.6 we perform numerical tests for the conditioning of the triangular factorizations of a Vandermonde matrix. Some results of this section have been published in [12], [13].

Chapter 5 is devoted to the important case of representations with orthogonal bases. We introduce the least squares approximation problem and we confirm the good properties of the operators in 2-norm. In Section 5.3 we analyze the behavior in  $\infty$ -norm, providing a more useful conditioning from the point of view of error propagation. In addition, we obtain formulas and bounds for conditioning. In sections 5.4 and 5.5 we study representations in terms of classical orthogonal polynomials of operators associated with continuous and discrete least squares problems. We studied the cases of Legendre and Chebyshev polynomials of the first and second kinds. In the case of Legendre polynomials, the conditioning grows as  $n^{3/2}$ , for Chebyshev polynomials of the first kind the growth is linear and for Chebyshev polynomials of the second kind it is quadratic. If the number of nodes coincides with  $n + 1$ , the dimension of the polynomial space, we have interpolation problems at the zeros of orthogonal polynomials and the previous analysis allows us to compute the conditioning of representations in terms of classical orthogonal

polynomials. In this regard, it is worth mentioning the almost optimal properties of the Lebesgue constant at the zeros of the Chebyshev polynomials (see [10], [52]).

In Chapter 6 we study the consequences of the positivity of the bases and the functionals in the search for bases that have a good conditioning. As a consequence of the results, we can deduce the optimality of the Lagrange representation with respect to other ones with positive functionals and the optimality of the Bernstein basis with respect to other totally positive bases, such as the monomial basis.

# 1 Operadores de cuasi-interpolación

---

## 1.1 Representaciones de operadores de cuasi-interpolación

Un operador de interpolación asigna a cada función su interpolante. Introducimos los operadores de cuasi-interpolación como una generalización de los operadores de interpolación.

**Definición 1.1.1.** *Un operador de cuasi-interpolación es una aplicación lineal y continua  $Q : F \rightarrow U$  de un espacio normado de funciones  $F$  en un subespacio de funciones  $U \subset F$ ,  $U \neq 0$ , tal que*

$$Q[u] = u, \quad u \in U,$$

*es decir, es una proyección lineal sobre  $U$ .*

Frecuentemente los operadores de cuasi-interpolación se utilizan en espacios  $U$  que contienen al espacio de polinomios  $P_k$  de grado menor o igual que  $k$ . En el caso de cuasi-interpolantes en el espacio de splines polinómicos, el mayor entero  $k$  tal que  $P_k$  está contenido en  $U$  corresponde al grado de aproximación del cuasi-interpolante.

El operador que asocia a cada función en  $C[a, b]$  su interpolante polinómico de Lagrange en nodos  $x_0, \dots, x_n$  del intervalo  $[a, b]$  es un caso particular importante de un operador de cuasi-interpolación

$$L[f](x) := \sum_{i=0}^n f(x_i)l_i(x), \quad l_i(x) = \prod_{j \in \{0, \dots, n\} \setminus \{i\}} \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n.$$

En este ejemplo,  $F = C[a, b]$  y  $U = P_n$ .

**Definición 1.1.2.** *Una representación de un operador de cuasi-interpolación  $Q : F \rightarrow U$  es una expresión de la forma  $\sum_{i=0}^n \alpha_i f u_i$ , de modo que*

$$Q[f] = \sum_{i=0}^n \alpha_i f u_i, \tag{1.1.1}$$

donde  $\alpha_0, \dots, \alpha_n \in F'$ , el espacio de los funcionales lineales y continuos definidos en el espacio  $F$ , y  $(u_0, \dots, u_n)$  es una base de  $U$ .

Veamos que todo operador de cuasi-interpolación admite una única representación respecto a una base cualquiera.

**Teorema 1.1.3.** *Sea  $Q : F \rightarrow U$  un operador de cuasi-interpolación. Para cada base ordenada  $(u_0, \dots, u_n)$  de  $U$ , existe una única sucesión  $(\alpha_0, \dots, \alpha_n)$  de funcionales lineales y continuos  $\alpha_i \in F'$ ,  $i = 0, \dots, n$ , tales que  $\sum_{i=0}^n \alpha_i f u_i$  es una representación de  $Q$ .*

*Demostración.* Definamos  $\alpha_i f$  como el coeficiente en  $u_i$  al expresar  $Q[f]$  como combinación lineal de la base  $(u_0, \dots, u_n)$ . Como  $Q$  es un operador lineal y continuo, se deduce que los  $\alpha_i$  son funcionales lineales y continuos. Como  $(u_0, \dots, u_n)$  es una base de  $U$ , los coeficientes  $\alpha_i f$  de  $Q[f] = \sum_{i=0}^n \alpha_i f u_i$  están determinados únicamente, lo que implica que los funcionales  $\alpha_0, \dots, \alpha_n$  son únicos.  $\square$

Teniendo en cuenta que  $u_j = Q[u_j] = \sum_{i=0}^n \alpha_i u_j u_i$ , para todo  $j = 0, \dots, n$ , se deduce el siguiente resultado.

**Proposición 1.1.4.** *Sea  $Q[f] = \sum_{i=0}^n \alpha_i f u_i$  una representación de un operador de cuasi-interpolación respecto a una base  $(u_0, \dots, u_n)$  de  $U$ . Entonces*

$$\alpha_i u_j = \delta_{ij}, \quad i, j = 0, \dots, n,$$

es decir,  $(u_0, \dots, u_n)$  y  $(\alpha_0, \dots, \alpha_n)$  forman un par dual.

De la Proposición anterior se deduce que

$$\alpha_i Q[f] = \alpha_i f, \quad i = 0, \dots, n,$$

es decir,  $Q[f]$  es la función de  $U$  que mejor se ajusta a  $f$  en el sentido de que los funcionales  $\alpha_0, \dots, \alpha_n$  coinciden sobre ambas. Por lo tanto,  $Q[f]$  es la solución de un problema de interpolación lineal general en el sentido descrito en la Sección 2.2 de [24].

**Teorema 1.1.5.** *Sean  $(\alpha_0, \dots, \alpha_n) \in F'$  funcionales linealmente independientes sobre  $U$ . Entonces existe una única base  $(u_0, \dots, u_n)$  de  $U$  tal que  $(u_0, \dots, u_n)$  y  $(\alpha_0, \dots, \alpha_n)$  forman un par dual y queda determinado un único operador de cuasi-interpolación  $Q$  en  $U$  correspondiente a los funcionales  $(\alpha_0, \dots, \alpha_n)$ .*

*Demostración.* Sea  $(v_0, \dots, v_n)$  una base cualquiera de  $U$ . Entonces, la matriz  $M = (\alpha_i v_j)_{i,j=0,\dots,n}$  es no singular. Definiendo la base  $(u_0, \dots, u_n) := (v_0, \dots, v_n)M^{-1}$ , obtenemos la base dual de  $(\alpha_0, \dots, \alpha_n)$  y se tiene el resultado.  $\square$

*Nota 1.1.6.* Por el Teorema 1.1.5, la sucesión de funcionales independientes  $(\alpha_0, \dots, \alpha_n)$  determina unívocamente la base dual  $(u_0, \dots, u_n)$  de  $U$  y el operador de cuasi-interpolación  $Q$ . Sin embargo, una base  $(u_0, \dots, u_n)$  de  $U$  no permite determinar los funcionales  $\alpha_0, \dots, \alpha_n$  de  $F'$  ni el cuasi-interpolante  $Q$ . Sus valores quedan determinados en el espacio  $U$ , pero pueden admitir distintas extensiones al espacio total  $F$ .

Un cuasi-interpolante admite varias representaciones dependiendo de la base de  $U$  elegida. Sea  $(v_0, \dots, v_n)$  otra base del espacio  $U$  y  $M$  la matriz de cambio de base. En el siguiente resultado expresamos la matriz de cambio de base en términos de las bases y de los funcionales y obtenemos la relación entre funcionales correspondientes a bases distintas.

**Teorema 1.1.7.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f v_i$  dos representaciones de un cuasi-interpolante en  $U$  y sea  $M$  la matriz de cambio de base tal que

$$(v_0, \dots, v_n) = (u_0, \dots, u_n)M. \quad (1.1.2)$$

Entonces

$$M = (\alpha_i v_j)_{0 \leq i, j \leq n}, \quad (1.1.3)$$

y

$$M^{-1} = (\beta_i u_j)_{0 \leq i, j \leq n}.$$

Además, los funcionales se relacionan mediante

$$\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = M^{-1} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}. \quad (1.1.4)$$

*Demostración.* Sea  $Q$  el operador de cuasi-interpolación dado por (1.1.1). Aplicando  $Q$  a los elementos  $v_j$ ,  $j = 0, \dots, n$ , obtenemos

$$v_j(x) = Q[v_j](x) = \sum_{i=0}^n \alpha_i v_j Q[u_i](x) = \sum_{i=0}^n \alpha_i v_j u_i(x),$$

de donde se deduce (1.1.3).

Intercambiando los papeles de  $(\alpha_0, \dots, \alpha_n)$ ,  $(u_0, \dots, u_n)$  con  $(\beta_0, \dots, \beta_n)$ ,  $(v_0, \dots, v_n)$  en la fórmula (1.1.3), obtenemos

$$M^{-1} = (\beta_i u_j)_{0 \leq i, j \leq n}.$$

De la fórmula

$$\sum_{j=0}^n m_{ij} \beta_j f = \sum_{j=0}^n \alpha_i v_j \beta_j f = \alpha_i \left( \sum_{j=0}^n \beta_j f v_j \right) = \alpha_i Q[f] = \alpha_i f,$$

se deduce que

$$\begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = M \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Por lo tanto, la base de funcionales  $(\beta_0, \dots, \beta_n)$  se expresa en términos de los funcionales  $(\alpha_0, \dots, \alpha_n)$  a través de la matriz inversa del cambio de base, obteniéndose (1.1.4).  $\square$

## 1.2 Condicionamientos de las representaciones

Sea  $Q : C[a, b] \rightarrow U$ ,  $U \subset C[a, b]$ , un operador de cuasi-interpolación definido en el espacio normado  $C[a, b]$  dotado de la norma del supremo

$$\|f\|_\infty := \sup_{x \in [a, b]} |f(x)|.$$

Se define la norma del cuasi-interpolante  $Q$

$$\|Q\|_\infty := \sup_{\|f\|_\infty=1} \|Q[f]\|_\infty.$$

Observemos que, como  $Q$  fija las funciones del espacio  $U$ , tenemos que  $\|Q\|_\infty \geq 1$ . La norma del cuasi-interpolante recibe el nombre de *constante de Lebesgue*.

**Proposición 1.2.1.** *La constante de Lebesgue de un cuasi-interpolante verifica*

$$\|Q\|_\infty = \sup_{x \in [a, b]} \lambda_Q(x),$$

donde

$$\lambda_Q(x) := \sup_{\|f\|_\infty=1} |Q[f](x)|,$$

es la llamada *función de Lebesgue*.

*Demostración.* Para cada  $\varepsilon > 0$  elijamos una función  $f \in C[a, b]$  con  $\|f\|_\infty = 1$  tal que

$$\|Q[f]\|_\infty \geq \|Q\|_\infty - \varepsilon.$$

Entonces, existe  $x \in [a, b]$  tal que

$$|Q[f](x)| = \|Q[f]\|_\infty.$$

Se deduce que,

$$\lambda_Q(x) \geq |Q[f](x)| \geq \|Q\|_\infty - \varepsilon.$$

Como  $\varepsilon$  es arbitrario obtenemos la desigualdad  $\sup_{x \in [a, b]} \lambda_Q(x) \geq \|Q\|_\infty$ .

Recíprocamente, para todo  $x \in [a, b]$  tenemos

$$\lambda_Q(x) = \sup_{\|f\|_\infty=1} |Q[f](x)| \leq \sup_{\|f\|_\infty=1} \|Q[f]\|_\infty = \|Q\|_\infty,$$

de donde  $\sup_{x \in [a, b]} \lambda_Q(x) \leq \|Q\|_\infty$ .  $\square$

La función de Lebesgue es una medida puntual de la estabilidad del cuasi- interpolante. Si  $\hat{f}$  es una aproximación de  $f$  con error  $\|\hat{f} - f\|_\infty \leq \varepsilon$  entonces

$$|Q[f](x) - Q[\hat{f}](x)| \leq \varepsilon \lambda_Q(x). \quad (1.2.1)$$

La norma de un funcional lineal  $\alpha : C[a, b] \rightarrow \mathbb{R}$  se define como

$$\|\alpha\|_\infty := \sup_{\|f\|_\infty=1} |\alpha f|.$$

Si este supremo es finito se dice que el funcional está acotado. Observemos que un funcional es continuo si y sólo si está acotado. Recordemos ahora el Teorema de Riesz en el espacio de las funciones continuas, que caracteriza a los funcionales lineales acotados (véase el Teorema 6.19 del Capítulo 6 de [54]).

**Teorema 1.2.2** (Teorema de representación de Riesz para funcionales acotados). *Todo funcional lineal acotado  $\alpha : C[a, b] \rightarrow \mathbb{R}$  es de la forma*

$$\alpha f = \int_a^b f(x) d\mu(x),$$

donde  $\mu$  es una medida de Borel regular finita. Además

$$\|\alpha\|_\infty = |\mu|[a, b].$$

**Definición 1.2.3.** Sean  $u = (u_0, \dots, u_n)$  una base de un espacio  $U \subset C[a, b]$  y  $\alpha = (\alpha_0, \dots, \alpha_n)$  una base de funcionales lineales y acotados de  $C[a, b]$  de modo que  $u, \alpha$  forman un par de bases duales. El condicionamiento de la representación (1.1.1) en  $x$  viene dado por

$$\text{cond}(x; \alpha, u) := \sum_{i=0}^n \|\alpha_i\|_\infty |u_i(x)|, \quad x \in [a, b]. \quad (1.2.2)$$

Si el espacio  $U$  está claro por el contexto, utilizaremos la notación simplificada  $\text{cond}(x; \alpha)$  ya que la base dual  $u$  está determinada por  $\alpha$ , como se indica en la Nota 1.1.6. Sin embargo, dada una base  $u$  de  $U$ , la relación de dualidad determina los valores de la base  $\alpha$  sobre los elementos de  $U$ . Teniendo en cuenta la Nota 1.1.6,  $\text{cond}(x; \alpha, u)$  dependerá de cómo la base  $\alpha$  se extiende a todo el espacio  $C[a, b]$ .

*Nota 1.2.4.* Dada una base  $u$  de  $U$  existen funcionales duales definidos en el espacio  $U$  por la relación de dualidad. El Teorema de Hahn-Banach (véase la Sección 1.6 de [21]) permite garantizar que existen extensiones  $\gamma_i$  definidas en todo el espacio  $C[a, b]$  que preservan la norma, es decir,

$$\|\gamma_i\|_\infty = \sup_{f \in U; \|f\|_\infty=1} |\gamma_i f|,$$

definiendo un operador de cuasi-interpolación  $Q[f] = \sum_{i=0}^n \gamma_i f u_i$  y una representación asociada tal que

$$\text{cond}(x; \gamma, u) = \sum_{i=0}^n \sup_{f \in U; \|f\|_\infty=1} |\gamma_i f| |u_i(x)|, \quad x \in [a, b].$$

Observemos que este operador de cuasi-interpolación y esta representación son óptimas respecto a la base dada ya que otras representaciones  $\sum_{i=0}^n \alpha_i f u_i$  verificarán  $\|\alpha_i\|_\infty \geq \|\gamma_i\|_\infty$  y darán lugar a otro operador de cuasi-interpolación que verificará  $\text{cond}(x; \alpha, u) \geq \text{cond}(x; \gamma, u)$ . A pesar del interés teórico de la existencia de estos operadores de cuasi-interpolación, la dificultad de encontrar extensiones en la práctica conlleva a que se utilicen otros basados en funcionales cuya definición es más simple.

El siguiente resultado relaciona el condicionamiento con la estabilidad de la representación.

**Proposición 1.2.5.** *Sea  $Q[f] = \sum_{i=0}^n \alpha_i f u_i$  una representación del operador de cuasi-interpolación  $Q : C[a, b] \rightarrow U$ . Sea  $\hat{f}$  una aproximación de  $f$ . Entonces,*

$$|Q[f](x) - Q[\hat{f}](x)| \leq \|\hat{f} - f\|_\infty \text{cond}(x; \alpha, u).$$

*Demostración.* De la definición de norma de funcional, se tiene

$$|Q[f](x) - Q[\hat{f}](x)| \leq \sum_{i=0}^n |\alpha_i [f - \hat{f}]| |u_i(x)| \leq \sum_{i=0}^n \|\alpha_i\|_\infty \|\hat{f} - f\|_\infty |u_i(x)|.$$

Usando la definición (1.2.2), obtenemos el resultado

$$|Q[f](x) - Q[\hat{f}](x)| \leq \|\hat{f} - f\|_\infty \text{cond}(x; \alpha, u).$$

□

Si calculamos el cuasi-interpolante evaluando cada término de la suma y sumando los términos resultantes, puede suceder que los valores de  $\alpha_i[f - \hat{f}]$  alcancen su valor máximo  $\|\alpha_i\|_\infty \|f - \hat{f}\|_\infty$  con el mismo signo que  $u_i(x)$ . Así que la función  $\text{cond}(x; \alpha, u)$  puede interpretarse como una medida puntual de la estabilidad de la representación del cuasi-interpolante en el sentido de la propagación del error.

Para funciones  $f$  pertenecientes al espacio  $U$ ,  $\text{cond}(x; \alpha, u)$  está relacionado con el número de condición para la evaluación introducido por Farouki-Rajan en [31] y con las generalizaciones de condición de bases en [42]. En particular, en la fórmula (2.2) de [42] se definió para una base  $u = (u_0, \dots, u_n)$  y funciones  $f(x) = \sum_{i=0}^n c_i u_i(x)$  el siguiente número de condición

$$\text{cond}(u; f, x) := \frac{\sum_{i=0}^n |c_i u_i(x)|}{\|\sum_{i=0}^n c_i u_i\|_\infty}. \quad (1.2.3)$$

Se tiene

$$\sup_{f \in U \setminus \{0\}} \text{cond}(u; f, x) = \sup_{f \in U \setminus \{0\}; \|f\|_\infty = 1} \sum_{i=0}^n |\alpha_i f| |u_i(x)| \leq \text{cond}(x; \alpha, u), \quad (1.2.4)$$

donde  $u$  y  $\alpha$  son bases duales.

Observemos que el número de condición (1.2.3) sólo puede aplicarse a funciones  $f \in U$ , mientras que (1.2.2) establece una relación entre las funciones y sus cuasi-interpolantes. Como el error cometido no tiene por qué describirse como una función del espacio  $U$ , el condicionamiento (1.2.3) no resulta tan útil para estudiar la propagación del error como (1.2.2).

**Definición 1.2.6.** *Una matriz generalizada de permutación  $K$  es una matriz tal que en cada fila y en cada columna hay exactamente un elemento no nulo. Si el elemento no nulo de cada fila y columna es 1, la matriz se llama matriz de permutación.*

Podemos considerar que una matriz de permutación puede obtenerse al aplicar una permutación  $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  a las filas de la matriz identidad, de modo que su elemento  $(i, j)$  es  $\delta_{\sigma(i), j}$ , donde  $\delta_{ij}$  es la delta de Kroenecker. Extrayendo de cada fila el factor no nulo, se deduce que toda matriz generalizada de permutación,  $K$ , es de la forma  $K = DP$ , donde  $D$  es una matriz diagonal y  $P$  una matriz de permutación.

Dada una base ordenada  $(u_0, \dots, u_n)$  y  $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  una permutación, podemos obtener otra base  $(u_{\sigma(0)}, \dots, u_{\sigma(n)})$  reordenando los elementos de la base original. La matriz de cambio de base entre ambas es una matriz de permutación.

De una base dada  $(u_0, \dots, u_n)$ , podemos obtener otras bases a partir de la anterior mediante reescalado  $(r_0 u_0, \dots, r_n u_n)$ , donde  $r_0, \dots, r_n$  son constantes no nulas. La matriz de cambio de base es la matriz diagonal  $\text{diag}(r_0, \dots, r_n)$ .

Como consecuencia de los dos comentarios anteriores, si una base coincide con otra salvo permutación y reescalado, la matriz de cambio de base es una matriz generalizada de permutación. En el siguiente resultado veremos que el condicionamiento es invariante bajo permutación y reescalado de las bases.

**Proposición 1.2.7.** *Sea  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f v_i$  dos representaciones de un operador de cuasi-interpolación  $Q : C[a, b] \rightarrow U$ . Si  $v$  se obtiene de  $u$  mediante reescalado y permutación, es decir, la matriz de cambio de base  $K$  tal que  $v^T = u^T K$  es una matriz generalizada de permutación, entonces*

$$\text{cond}(x; \alpha, u) = \text{cond}(x; \beta, v).$$

*Demostración.* Por hipótesis existe una permutación  $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  tal que

$$v_i = r_i u_{\sigma(i)}, \quad i = 0, \dots, n,$$

con  $r_i$  distinto de cero,  $i = 0, \dots, n$ . De lo anterior, se deduce la siguiente relación entre los funcionales

$$\beta_i = \frac{1}{r_i} \alpha_{\sigma(i)}, \quad i = 0, \dots, n.$$

Teniendo en cuenta que los sumandos que contribuyen al condicionamiento son los mismos, pero presentados en orden diferente, se tiene que el condicionamiento no varía

$$\text{cond}(x; \beta, v) = \sum_{i=0}^n \|\beta_i\|_{\infty} |v_i(x)| = \sum_{i=0}^n \|\alpha_{\sigma(i)}\|_{\infty} |u_{\sigma(i)}(x)| = \text{cond}(x; \alpha, u).$$

□

*Nota 1.2.8.* Podemos utilizar la propiedad de que el condicionamiento es invariante bajo el reescalado de las bases para normalizar las representaciones sin modificar el condicionamiento. Si normalizamos los funcionales para que tengan norma 1

$$\hat{u}_i = \|\alpha_i\|_{\infty} u_i, \quad \hat{\alpha}_i = \frac{\alpha_i}{\|\alpha_i\|_{\infty}}, \quad i = 0, \dots, n, \quad (1.2.5)$$

tenemos la representación normalizada  $Q[f] = \sum_{i=0}^n \hat{\alpha}_i f \hat{u}_i$ . Matricialmente esta normalización puede expresarse de la siguiente manera

$$\begin{pmatrix} \hat{\alpha}_0 \\ \vdots \\ \hat{\alpha}_n \end{pmatrix} = D_{\alpha}^{-1} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad (\hat{u}_0, \dots, \hat{u}_n) = (u_0, \dots, u_n) D_{\alpha},$$

donde  $D_{\alpha} := \text{diag}(\|\alpha_0\|_{\infty}, \dots, \|\alpha_n\|_{\infty})$ .

Veamos que  $\text{cond}(x; \alpha)$  permite acotar la norma del cuasi-interpolante.

**Proposición 1.2.9.** *Sea  $Q$  un cuasi-interpolante de un espacio  $U$  que admite la representación  $\sum_{i=0}^n \alpha_i f u_i(x)$ . Entonces*

$$\begin{aligned} \|Q\|_\infty &\leq \max_{x \in [a, b]} \text{cond}(x; \alpha), \\ \lambda_Q(x) &\leq \text{cond}(x; \alpha). \end{aligned}$$

*Demostración.* Dada la representación (1.1.1) tenemos

$$|Q[f](x)| \leq \sum_{i=0}^n |\alpha_i f| |u_i(x)|,$$

de donde se deduce

$$|Q[f](x)| \leq \|f\|_\infty \sum_{i=0}^n \|\alpha_i\|_\infty |u_i(x)| = \|f\|_\infty \text{cond}(x; \alpha).$$

Por tanto,

$$\lambda_Q(x) = \sup_{\|f\|_\infty=1} |Q[f](x)| \leq \text{cond}(x; \alpha)$$

y por la Proposición 1.2.1

$$\|Q\|_\infty \leq \max_{x \in [a, b]} \text{cond}(x; \alpha).$$

□

*Nota 1.2.10.* Sea  $M$  la matriz de cambio de base  $v^T = u^T M$  de dos representaciones  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f v_i$  del operador de cuasi-interpolación  $Q : C[a, b] \rightarrow U$ . Del Teorema 1.1.7 y de la Nota 1.2.8, se deduce que la matriz de cambio de base entre las representaciones normalizadas (1.2.5)  $\hat{M}$ ,  $\hat{v}^T = \hat{u}^T \hat{M}$ , es de la forma

$$\hat{M} = D_\alpha^{-1} M D_\beta, \quad D_\alpha := \text{diag}(\|\alpha_0\|_\infty, \dots, \|\alpha_n\|_\infty), \quad D_\beta := \text{diag}(\|\beta_0\|_\infty, \dots, \|\beta_n\|_\infty). \quad (1.2.6)$$

Recordemos que  $\|M\|_\infty := \max_{i=0, \dots, n} \sum_{j=0}^n |m_{ij}|$ . Comparemos ahora condicionamientos correspondientes a diferentes representaciones del cuasi-interpolante.

**Teorema 1.2.11.** *Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f v_i$  dos representaciones de un operador de cuasi-interpolación en  $C[a, b]$ . Sea  $M$  la matriz  $M = (\alpha_i v_j)_{0 \leq i, j \leq n}$ . Entonces*

$$\text{cond}(x; \beta, v) \leq \|\hat{M}\|_\infty \text{cond}(x; \alpha, u), \quad (1.2.7)$$

donde  $\hat{M}$  es la matriz de cambio de las bases normalizadas dada por (1.2.6).

*Demostración.* De (1.1.2) obtenemos

$$|v_j(x)| \leq \sum_{i=0}^n |m_{ij}| |u_i(x)|, \quad x \in [a, b],$$

Por tanto, tenemos

$$\begin{aligned} \text{cond}(x; \beta, v) &= \sum_{j=0}^n \|\beta_j\|_\infty |v_j(x)| \leq \sum_{j=0}^n \sum_{i=0}^n \|\beta_j\|_\infty |m_{ij}| |u_i(x)| \\ &= \sum_{i=0}^n \sum_{j=0}^n \frac{\|\beta_j\|_\infty |m_{ij}|}{\|\alpha_i\|_\infty} \|\alpha_i\|_\infty |u_i(x)| \leq \max_{i=0, \dots, n} \sum_{j=0}^n \frac{\|\beta_j\|_\infty |m_{ij}|}{\|\alpha_i\|_\infty} \text{cond}(x; \alpha, u). \end{aligned}$$

Teniendo en cuenta que

$$|\hat{m}_{ij}| = \frac{\|\beta_j\|_\infty |m_{ij}|}{\|\alpha_i\|_\infty},$$

tenemos que la norma de la matriz  $\hat{M}$  es

$$\|\hat{M}\|_\infty = \max_{i=0, \dots, n} \sum_{j=0}^n \frac{\|\beta_j\|_\infty |m_{ij}|}{\|\alpha_i\|_\infty}$$

y se deduce el resultado. □

Observemos que la fórmula (1.2.7) presenta una relación entre condicionamientos invariante por reescalado de las bases. Por la Proposición 1.2.7,  $\text{cond}(x; \alpha, u)$  y  $\text{cond}(x; \beta, v)$  son invariantes por reescalado y por la Nota 1.2.10, los elementos de  $\hat{M}$  sólo dependen de las representaciones normalizadas y también son invariantes por reescalado de la base.

### 1.3 Condicionamiento de Skeel y cotas de normas

Introduzcamos algunas definiciones y notaciones matriciales.

**Definición 1.3.1.** Una matriz  $M = (m_{ij})_{0 \leq i, j \leq n}$  se dice que es no negativa si  $m_{ij} \geq 0$ ,  $i, j = 0, \dots, n$ , y se denota por  $M \geq 0$ . Dadas dos matrices  $M_1, M_2$  decimos que  $M_1 \geq M_2$  si  $M_1 - M_2 \geq 0$ .

Estas definiciones anteriores también se pueden aplicar a vectores, considerados como matrices filas o columnas, para expresar desigualdades vectoriales componente a componente.

Sea la matriz  $M = (m_{ij})_{0 \leq i, j \leq n}$ . Denotamos por

$$|M| := (|m_{ij}|)_{0 \leq i, j \leq n}.$$

a la matriz cuyos elementos son los valores absolutos de los correspondientes elementos de  $M$ . Utilizando esta notación, la norma  $\infty$  de una matriz puede expresarse de la forma

$$\|M\|_{\infty} := \max_{i=0,\dots,n} \sum_{j=0}^n |m_{ij}| = \||M|e\|_{\infty},$$

donde  $e := (1, \dots, 1)^T$ .

**Definición 1.3.2.** Sea  $M$  una matriz no singular. El condicionamiento en norma  $\infty$  se define como

$$\kappa_{\infty}(M) := \|M\|_{\infty} \|M^{-1}\|_{\infty}.$$

La acotación de errores en el cálculo de un operador de cuasi-interpolación a través de una representación dada puede analizarse desde el punto de vista matricial. De la fórmula (1.1.4) tenemos

$$\left\| \begin{pmatrix} \beta_0 f \\ \vdots \\ \beta_n f \end{pmatrix} \right\|_{\infty} \leq \|M^{-1}\|_{\infty} \left\| \begin{pmatrix} \alpha_0 f \\ \vdots \\ \alpha_n f \end{pmatrix} \right\|_{\infty}.$$

Por otro lado, como  $v(x)^T = u(x)^T M$ ,

$$\sum_{i=0}^n |v_i(x)| \leq \sum_{i=0}^n |u_i(x)| \|M\|_{\infty}.$$

De las desigualdades anteriores se deduce que

$$\text{cond}(x; \beta, v) \leq \kappa_{\infty}(M) \text{cond}(x; \alpha, u). \quad (1.3.1)$$

La cota (1.3.1) es más pesimista que la del Teorema 1.2.11, ya que no tiene en cuenta los diferentes posibles escalados. Un gran error en el cálculo de  $\beta_i f$  al resolver el sistema

$$M \begin{pmatrix} \beta_0 f \\ \vdots \\ \beta_n f \end{pmatrix} = \begin{pmatrix} \alpha_0 f \\ \vdots \\ \alpha_n f \end{pmatrix}$$

quedará compensado si la función  $v_i(x)$  toma valores pequeños.

A continuación, introducimos un condicionamiento matricial que es invariante por escalado de filas y que proporcionará cotas más ajustadas que el condicionamiento en norma  $\infty$ .

**Definición 1.3.3.** *El condicionamiento de Skeel de una matriz  $M$  no singular es*

$$\text{Cond}(M) := \| |M^{-1}| |M| \|_{\infty}.$$

El número de condición de Skeel de una matriz permite analizar las perturbaciones, componente a componente, que se producen al multiplicar la matriz por un vector, en contraste con el número de condición en norma que mide la perturbación de la norma del vector y no de sus componentes. Si consideramos que las evaluaciones de una función en los puntos  $x \in [a, b]$  corresponden a las componentes de la función, puede mostrarse que el número de condición (1.2.3) tiene una naturaleza similar al número de condición de Skeel (véase también [42]). En este sentido,  $\text{cond}(x; \alpha)$  es un número de condición que permite realizar un análisis por componentes.

Se deduce de la propiedad submultiplicativa de la norma matricial  $\infty$  que  $\text{Cond}(M) \leq \kappa_{\infty}(M)$ , por lo que el condicionamiento de Skeel proporciona cotas más ajustadas que el condicionamiento en norma  $\infty$ . Otra de las ventajas de usar el condicionamiento de Skeel frente al condicionamiento en norma  $\infty$  es que el primero es invariante bajo escalados de fila (véase la Sección 7.2 de [37]).

**Proposición 1.3.4.** *Sea  $M$  una matriz no singular. Entonces*

$$\text{Cond}(M) = \text{Cond}(DM) \leq \kappa_{\infty}(DM), \quad (1.3.2)$$

para toda matriz diagonal no singular  $D$ .

*Demostración.* Se tiene que

$$\text{Cond}(DM) = \| |M^{-1}D^{-1}| |DM| \|_{\infty} = \| |M^{-1}| |D^{-1}| |D| |M| \|_{\infty} = \text{Cond}(M).$$

Como el condicionamiento de Skeel es menor o igual que el condicionamiento en norma  $\infty$ , se tiene el resultado.  $\square$

El condicionamiento de Skeel puede ser mucho menor que  $\kappa_{\infty}$ . Aplicando (1.3.2) a la matriz identidad  $I$ , tenemos

$$1 = \text{Cond}(I) = \text{Cond}(D) \leq \kappa_{\infty}(D) = \frac{\max_{i=0, \dots, n} |d_{ii}|}{\min_{i=0, \dots, n} |d_{ii}|},$$

pudiendo ser  $\kappa_{\infty}(D)$  arbitrariamente grande.

Podemos caracterizar las matrices con condicionamiento de Skeel óptimo (véase también [47]).

**Proposición 1.3.5.** *Sea  $M$  una matriz no singular. Entonces  $\text{Cond}(M) = 1$  si y sólo si  $M$  es una matriz generalizada de permutación.*

*Demostración.* Por la Definición 1.2.6, una matriz generalizada de permutación es de la forma  $M = DP$ , donde  $M$  es una matriz diagonal y  $P$  una matriz de permutación. Así,

$$|M^{-1}| = |P^{-1}||D^{-1}| = |P|^{-1}|D|^{-1} = |M|^{-1}.$$

Por tanto,  $|M^{-1}||M| = I$ , y el condicionamiento de Skeel de  $M$  es 1.

Si  $\text{Cond}(M) = 1$ ,  $\||M^{-1}||M|\|_{\infty} = 1$ . Como  $|M^{-1}||M| \geq I$ , se deduce que  $|M|^{-1} = |M^{-1}|$ . Por tanto,  $|M|$  es una matriz no negativa con inversa no negativa. Es bien conocido que una matriz no negativa con inversa no negativa es una matriz generalizada de permutación (cf. Lema 1.1 de [44]). En conclusión,  $|M|$  y  $M$  son matrices generalizadas de permutación. □

Veamos en la siguiente proposición, que es bien conocida, la matriz de escalado  $D_M$  óptima que consigue la igualdad en (1.3.2).

**Proposición 1.3.6.** *Sea  $M$  una matriz no singular y  $D_M := (\text{diag}(|M|e))^{-1}$ , donde  $e := (1, \dots, 1)^T$ . Entonces*

$$\text{Cond}(M) = \kappa_{\infty}(D_M M) = \min\{\kappa_{\infty}(DM) : D \text{ diagonal no singular}\},$$

*es decir,  $D_M$  es la matriz que minimiza  $\kappa_{\infty}(DM)$  para cualquier  $D$  matriz diagonal no singular.*

*Demostración.* Como  $\|D_M M\|_{\infty} = \|D_M |M|e\|_{\infty} = \|e\|_{\infty} = 1$ , se deduce que

$$\begin{aligned} \kappa_{\infty}(D_M M) &= \|D_M M\|_{\infty} \|M^{-1} D_M^{-1}\|_{\infty} = \|M^{-1} D_M^{-1}\|_{\infty} = \||M^{-1}| D_M^{-1} e\|_{\infty} \\ &= \||M^{-1}||M|e\|_{\infty} = \text{Cond}(M). \end{aligned}$$

Por (1.3.2),  $\text{Cond}(M) \leq \kappa_{\infty}(DM)$  para cualquier  $D$  matriz diagonal no singular y el resultado queda probado. □

*Nota 1.3.7.* A diferencia del condicionamiento en norma  $\infty$ , el condicionamiento de Skeel de la inversa de una matriz no tiene por qué coincidir con el de dicha matriz. El número de condición de Skeel de la inversa de una matriz

$$\text{Cond}(M^{-1}) = \||M||M^{-1}|\|_{\infty} = \||M||M^{-1}|e\|_{\infty}, \quad e = (1, \dots, 1)^T,$$

puede considerarse como un número de condición invariante por escalado de columnas. De hecho, por la Proposición 1.3.4,  $\text{Cond}((MD)^{-1}) = \text{Cond}(D^{-1}M^{-1}) = \text{Cond}(M^{-1})$ .

Del Teorema 1.2.11 y de la Nota 1.3.7 se deduce el siguiente resultado.

**Corolario 1.3.8.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f v_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$ . Sea  $M$  la matriz  $M = (\alpha_i v_j)_{0 \leq i, j \leq n}$  y sea  $\hat{M}$  la matriz de cambio de base de las representaciones normalizadas dada por (1.2.6). Entonces

$$\|\hat{M}\|_\infty \leq \text{Cond}(M^{-1}D_\alpha) = \text{Cond}(\hat{M}^{-1}) \quad (1.3.3)$$

y

$$\text{cond}(x; \beta, v) \leq \text{Cond}(\hat{M}^{-1}) \text{cond}(x; \alpha, u). \quad (1.3.4)$$

*Demostración.* Por (1.1.4) se tiene

$$\beta_i = \sum_{j=0}^n m_{ij}^{(-1)} \alpha_j, \quad i = 0, \dots, n,$$

donde  $M^{-1} = (m_{ij}^{(-1)})_{0 \leq i, j \leq n}$ . Por lo tanto

$$\begin{aligned} \|\beta_i\|_\infty &= \sup_{\|f\|_\infty=1} |\beta_i f| = \sup_{\|f\|_\infty=1} \left| \sum_{j=0}^n m_{ij}^{(-1)} \alpha_j f \right| \leq \sup_{\|f\|_\infty=1} \sum_{j=0}^n |m_{ij}^{(-1)} \alpha_j f| \\ &= \sum_{j=0}^n |m_{ij}^{(-1)}| \|\alpha_j\|_\infty, \end{aligned}$$

de lo que se deduce

$$\begin{pmatrix} \|\beta_0\|_\infty \\ \vdots \\ \|\beta_n\|_\infty \end{pmatrix} \leq |M^{-1}| \begin{pmatrix} \|\alpha_0\|_\infty \\ \vdots \\ \|\alpha_n\|_\infty \end{pmatrix},$$

es decir,

$$D_\beta e \leq |M^{-1}| D_\alpha e.$$

Así, por (1.2.6)  $\hat{M} = D_\alpha^{-1} M D_\beta$ , y tenemos

$$|\hat{M}|e = D_\alpha^{-1} |M| D_\beta e \leq D_\alpha^{-1} |M| |M^{-1}| D_\alpha e,$$

y

$$\hat{M}^{-1} = D_\beta^{-1} M^{-1} D_\alpha.$$

Por la Proposición 1.3.4, se obtiene la cota

$$\|\hat{M}\|_\infty = \|\hat{M}|e\|_\infty \leq \|D_\alpha^{-1} |M| |M^{-1}| D_\alpha\|_\infty = \text{Cond}(M^{-1}D_\alpha) = \text{Cond}(\hat{M}^{-1}).$$

La fórmula (1.3.4) se deduce del Teorema 1.2.11. □

De la Nota 1.2.10 se deduce que tanto  $\|\hat{M}\|_\infty$  como  $\text{Cond}(\hat{M}^{-1})$  son magnitudes invariantes por el escalado de ambas representaciones. La fórmula (1.3.3) expresa una desigualdad independiente del escalado de las representaciones.

En el caso importante de los operadores de interpolación, los funcionales evaluación tienen norma 1, lo que permite simplificar el condicionamiento de la representación de Lagrange. En general, si los funcionales  $\alpha$  están normalizados, entonces  $D_\alpha$  es la matriz identidad y, por tanto,  $\text{Cond}(\hat{M}^{-1}) = \text{Cond}(M^{-1})$ . La fórmula (1.3.3) y la fórmula (1.3.4) se reducen a

$$\|\hat{M}\|_\infty \leq \text{Cond}(M^{-1}), \quad (1.3.5)$$

$$\text{cond}(x; \beta, v) \leq \text{Cond}(M^{-1}) \text{cond}(x; \alpha, u). \quad (1.3.6)$$



# 2

## Óptima estabilidad de la fórmula de Lagrange y estabilidad de la fórmula de Newton

---

### 2.1 El operador de interpolación de Lagrange

En esta sección vamos a considerar el problema de interpolación de Lagrange en un subespacio  $U$  de  $C[a, b]$ .

**Problema de interpolación de Lagrange.** Sea  $U$  un espacio de funciones continuas en  $[a, b]$  de dimensión  $n + 1$  y  $x_0, \dots, x_n$  nodos distintos en  $[a, b]$ . Se trata de encontrar  $u \in U$  tal que  $u(x_i) = f(x_i)$ ,  $i = 0, \dots, n$ , para cierta función  $f \in C[a, b]$ .

Definimos los funcionales de evaluación

$$\lambda_i(f) := f(x_i), \quad i = 0, \dots, n.$$

Si el problema de interpolación de Lagrange admite una única solución para los nodos  $x_0, \dots, x_n$  entonces podemos definir las funciones fundamentales de Lagrange  $l_i \in U$  tales que

$$l_i(x_j) = \delta_{ij}, \quad i, j = 0, \dots, n,$$

de manera que la solución del problema de interpolación puede expresarse en la forma de Lagrange  $\sum_{i=0}^n f(x_i)l_i(x)$ . Observemos que el operador de interpolación

$$L[f](x) = \sum_{i=0}^n f(x_i)l_i(x)$$

es un caso particular de cuasi-interpolante representado a través de los funcionales evaluación.

Observemos que los funcionales evaluación están normalizados ya que

$$\|\lambda_i\|_\infty = \sup_{\|f\|_\infty=1} |f(x_i)| = 1$$

Para obtener una fórmula para la función de Lebesgue demostramos el siguiente lema.

**Lema 2.1.1.** *Dados  $x_0, \dots, x_n$  en  $[a, b]$  y  $s_0, \dots, s_n \in [-1, 1]$ , existe  $f \in C[a, b]$  con  $\|f\|_\infty = 1$  tal que  $f(x_i) = s_i$ ,  $i = 0, \dots, n$ .*

*Demostración.* Sea  $f \in C[a, b]$  la función lineal a trozos asociada a la partición  $a \leq x_0 < x_1 < \dots < x_n \leq b$  tal que

$$f(a) = s_0, \quad f(x_i) = s_i, \quad i = 0, \dots, n, \quad f(b) = s_n.$$

Entonces,  $\|f\|_\infty = 1$  y se verifican las condiciones del enunciado. □

**Teorema 2.1.2.** *La función de Lebesgue  $\lambda_L(x)$  del operador de interpolación de Lagrange coincide con el condicionamiento de la representación de Lagrange, es decir,*

$$\lambda_L(x) = \text{cond}(x; \lambda) = \sum_{i=0}^n |l_i(x)|.$$

Además,

$$\|L\|_\infty = \max_{x \in [a, b]} \sum_{i=0}^n |l_i(x)|.$$

*Demostración.* Como los funcionales están normalizados, tenemos que  $\text{cond}(x; \lambda) = \sum_{i=0}^n |l_i(x)|$ . También tenemos por la Proposición 1.2.9 que  $\lambda_L(x) \leq \text{cond}(x; \lambda)$ . Sin pérdida de generalidad podemos suponer que los puntos están ordenados,  $x_0 < \dots < x_n$ . Por el Lema 2.1.1 existe  $s \in C[a, b]$  con  $\|s\|_\infty = 1$  tal que

$$s(x_i) = \text{sign}(l_i(x)), \quad i = 0, \dots, n.$$

Entonces

$$\lambda_L(x) \geq |L[s](x)| = \left| \sum_{i=0}^n \text{sign}(l_i(x)) l_i(x) \right| = \sum_{i=0}^n |l_i(x)| = \text{cond}(x; \lambda).$$

La fórmula para la constante de Lebesgue se deduce de la Proposición 1.2.1 teniendo en cuenta que  $\lambda_L$  es una función continua. □

Consideraremos otras representaciones del operador de interpolación de Lagrange que utilizan otros funcionales diferentes de los funcionales evaluación. En el siguiente resultado calculamos su norma.

**Proposición 2.1.3.** *Sea  $\beta$  un funcional de la forma*

$$\beta f = \sum_{i=0}^n c_i f(x_i), \quad c = (c_0, \dots, c_n).$$

Entonces,

$$\|\beta\|_\infty = \|c\|_1 = \sum_{i=0}^n |c_i|.$$

*Demostración.*

$$\|\beta\|_\infty = \max_{\|f\|_\infty=1} \left| \sum_{i=0}^n c_i f(x_i) \right| \leq \max_{\|f\|_\infty=1} \sum_{i=0}^n |c_i| |f(x_i)| \leq \max_{\|f\|_\infty=1} \sum_{i=0}^n |c_i| = \|c\|_1.$$

Veamos la otra desigualdad. Por el Lema 2.1.1 existe  $f$  tal que  $f(x_i) = \text{sign}(c_i)$ ,  $i = 0, \dots, n$ , con  $\|f\|_\infty = 1$ . Entonces

$$\max_{\|f\|_\infty=1} \left| \sum_{i=0}^n c_i f(x_i) \right| \geq \left| \sum_{i=0}^n c_i \text{sign}(c_i) \right| = \sum_{i=0}^n |c_i| = \|c\|_1.$$

□

En problemas de interpolación de Lagrange, el concepto de matriz de colocación juega un papel relevante, al ser la matriz de coeficientes del problema que resulta de expresar el interpolante en términos de una base.

**Definición 2.1.4.** *Si  $u_0(t), \dots, u_n(t)$  son funciones reales definidas en un intervalo  $[a, b]$ , definimos la matriz de colocación del sistema de funciones  $(u_0(t), \dots, u_n(t))$  en  $t_0, t_1, \dots, t_n$  en  $[a, b]$  a la matriz*

$$M \begin{pmatrix} u_0, u_1, \dots, u_n \\ t_0, t_1, \dots, t_n \end{pmatrix} := (u_j(t_i))_{0 \leq i, j \leq n}.$$

Aplicando la Proposición 2.1.3 y el Teorema 1.1.7, tenemos el siguiente resultado.

**Teorema 2.1.5.** *Sean  $\sum_{i=0}^n \beta_i f v_i$  una representación del interpolador de Lagrange y sea  $M$  la matriz de cambio de base entre la base  $v = (v_0, \dots, v_n)^T$  y la base de Lagrange  $l$ ,  $v^T = l^T M$ . Entonces,  $M$  es la matriz de colocación de la base  $v$  en los nodos  $x_0, \dots, x_n$ , es decir,*

$$M = (v_j(x_i))_{0 \leq i, j \leq n}. \quad (2.1.1)$$

Además, los funcionales  $(\beta_0, \dots, \beta_n)$  pueden expresarse en términos de los funcionales de evaluación  $(\lambda_0, \dots, \lambda_n)$  de la forma

$$\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = M^{-1} \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_n \end{pmatrix} \quad (2.1.2)$$

y

$$\begin{pmatrix} \|\beta_0\|_\infty \\ \vdots \\ \|\beta_n\|_\infty \end{pmatrix} = |M^{-1}|e, \quad e = (1, \dots, 1)^T. \quad (2.1.3)$$

Veamos cómo podemos expresar el condicionamiento del interpolante respecto a una base cualquiera. Para ello, comparamos esta representación con la de Lagrange.

**Teorema 2.1.6.** *Sea  $\sum_{i=0}^n \beta_i f v_i$  una representación del operador de interpolación de Lagrange  $L[f] = \sum_{j=0}^n \lambda_j f l_j$  en nodos  $x_0, \dots, x_n$ , en el subespacio  $U$  de  $C[a, b]$  y  $M = (v_j(x_i))_{0 \leq i, j \leq n}$ . Entonces*

$$\text{cond}(x; \beta) = |v(x)|^T |M^{-1}| e,$$

donde  $|v(x)|^T = (|v_0(x)|, \dots, |v_n(x)|)$  y  $e = (1, \dots, 1)^T$ .

*Demostración.* Por la fórmula (2.1.3)

$$D_\beta e = |M^{-1}|e,$$

donde  $D_\beta = \text{diag}(\|\beta_0\|_\infty, \dots, \|\beta_n\|_\infty)$ . Así, obtenemos el resultado

$$\text{cond}(x; \beta) = \sum_{j=0}^n \|\beta_j\|_\infty |v_j(x)| = |v(x)|^T D_\beta e = |v(x)|^T |M^{-1}| e.$$

□

En el siguiente corolario demostramos que la representación de Lagrange es la más estable en términos de la propagación del error y que el peor comportamiento de cualquier otra representación puede detectarse analizando el condicionamiento de Skeel de la inversa de la matriz de cambio de base.

**Proposición 2.1.7.** Sea  $\sum_{i=0}^n \beta_i f v_i$  una representación del operador de interpolación de Lagrange  $L[f] = \sum_{j=0}^n \lambda_j f l_j$  en nodos  $x_0, \dots, x_n$ , en el subespacio  $U$  de  $C[a, b]$  y  $M = (v_j(x_i))_{0 \leq i, j \leq n}$ . Entonces, podemos establecer las siguientes desigualdades que relacionan los condicionamientos de la representación considerada con respecto a la representación de Lagrange

$$\text{cond}(x; \lambda) \leq \text{cond}(x; \beta) \leq \text{Cond}(M^{-1}) \text{cond}(x; \lambda).$$

*Demostración.* La primera desigualdad se tiene por el Teorema 2.1.2 y la Proposición 1.2.9, respectivamente. La segunda desigualdad se obtiene aplicando (1.3.6).  $\square$

La Proposición 2.1.7 demuestra que el cociente entre el condicionamiento de una representación cualquiera y el condicionamiento de la representación de Lagrange está acotado por  $\text{cond}(M^{-1})$ . Vamos a ver que esta cota se alcanza en uno de los nodos.

**Teorema 2.1.8.** Sea  $\sum_{i=0}^n \beta_i f v_i$  una representación del operador de interpolación de Lagrange  $L[f] = \sum_{j=0}^n f(x_j) l_j$  en nodos  $x_0, \dots, x_n$  en el subespacio  $U$  de  $C[a, b]$ . Entonces

$$\max_{\lambda_L(x) \neq 0} \frac{\text{cond}(x; \beta)}{\lambda_L(x)} = \max_{i=0, \dots, n} \text{cond}(x_i; \beta) = \text{Cond}(M^{-1}) = \|MD_\beta\|_\infty, \quad (2.1.4)$$

donde  $D_\beta = \text{diag}(\|\beta_0\|_\infty, \dots, \|\beta_n\|_\infty)$  y  $M = (v_j(x_i))_{0 \leq i, j \leq n}$ .

*Demostración.* Por la fórmula (2.1.3),  $D_\beta e = |M^{-1}|e$ , donde  $e = (1, \dots, 1)^T$ . Sea  $\hat{M}$  la matriz de cambio de base de las representaciones normalizadas dada por (1.2.6). Como los funcionales evaluación tienen norma unitaria,  $D_\lambda$  es la matriz identidad y  $\hat{M} = MD_\beta$ . Por (2.1.3) tenemos,

$$|\hat{M}|e = |M|D_\beta e = |M| |M^{-1}|e,$$

y, por tanto,

$$\|\hat{M}\|_\infty = \|MD_\beta\|_\infty = \||M| |M^{-1}|\|_\infty = \text{Cond}(M^{-1}). \quad (2.1.5)$$

Recordemos que

$$\text{cond}(x_i; \beta) = \sum_{j=0}^n \|\beta_j\|_\infty |v_j(x_i)| = \sum_{j=0}^n \|\beta_j\|_\infty |m_{ij}|, \quad i = 0, \dots, n,$$

de donde se tiene

$$\begin{pmatrix} \text{cond}(x_0; \beta) \\ \vdots \\ \text{cond}(x_n; \beta) \end{pmatrix} = |M|D_\beta e.$$

Tomando normas y usando (2.1.5),

$$\max_{i=0,\dots,n} \text{cond}(x_i, \beta) = \|MD_\beta\|_\infty = \text{Cond}(M^{-1}).$$

Teniendo en cuenta que  $\lambda_L(x_i) = 1$ ,  $i = 0, \dots, n$ , el Teorema 2.1.2 y la Proposición 2.1.7, obtenemos

$$\max_{i=0,\dots,n} \text{cond}(x_i, \beta) \leq \sup_{\lambda_L(x) \neq 0} \frac{\text{cond}(x; \beta)}{\lambda_L(x)} \leq \text{Cond}(M^{-1}) = \max_{i=0,\dots,n} \text{cond}(x_i, \beta).$$

Luego todas las desigualdades anteriores deben ser igualdades y el supremo debe ser un máximo porque se alcanza en el conjunto de nodos.  $\square$

El siguiente resultado muestra que la representación de Lagrange es la única óptima, salvo reescalado y permutación (véase la Proposición 1.2.7).

**Teorema 2.1.9.** *Sea  $\sum_{i=0}^n \beta_i f v_i$  una representación del operador de interpolación de Lagrange en el subespacio  $U$  de  $C[a, b]$ . Si*

$$\text{cond}(x; \beta) \leq \lambda_L(x), \quad x \in [a, b],$$

entonces  $\text{cond}(x; \beta) = \lambda_L(x)$  para todo  $x \in [a, b]$  y la matriz  $M = (v_j(x_i))_{0 \leq i, j \leq n}$  es una matriz generalizada de permutación. Por tanto, la representación de Lagrange es la única con condicionamiento mínimo, salvo reescalado y permutación de los funcionales.

*Demostración.* Usando la Proposición 2.1.7, obtenemos la igualdad  $\text{cond}(x; \beta) = \lambda_L(x)$ . Además, por Teorema 2.1.8,  $\text{Cond}(M^{-1}) = 1$ , de donde se deduce por la Proposición 1.3.5 que  $M^{-1}$  y  $M$  son matrices generalizadas de permutación.  $\square$

Teniendo en cuenta los teoremas 2.1.2 y 2.1.8 y la Proposición 2.1.7, que establece una relación entre  $\text{cond}(x; \beta)$  y  $\lambda_L(x) = \text{cond}(x; \lambda)$ , deducimos la siguiente fórmula

$$\lambda_L(x) \leq \text{cond}(x; \beta) \leq \lambda_L(x) \max_{i=0,\dots,n} \text{cond}(x_i; \beta). \quad (2.1.6)$$

Si los funcionales  $\beta_i$  tienen una descripción simple, tenemos

$$M^{-1} = (\beta_k l_j)_{k,j=0,\dots,n}$$

y deducimos

$$\beta_k = \sum_{j=0}^n \beta_k l_j \lambda_j, \quad \|\beta_k\|_\infty = \sum_{j=0}^n |\beta_k l_j|,$$

obteniendo la expresión

$$\text{Cond}(M^{-1}) = \max_{i=0,\dots,n} \text{cond}(x_i; \beta) = \max_{i=0,\dots,n} \sum_{j=0}^n \sum_{k=0}^n |\beta_k l_j| |v_k(x_i)|. \quad (2.1.7)$$

Hemos mencionado en la definición de cuasi-interpolante la importancia de que  $U$  contenga subespacios de polinomios. En el caso de que las funciones constantes estén en el espacio  $U$ , la función de Lebesgue,  $\lambda_L(x)$ , es positiva en todos los puntos del intervalo  $[a, b]$  y, por tanto, el cociente  $\text{cond}(x; \beta)/\lambda_L(x)$  está definido para todo  $x$  en  $[a, b]$ .

**Proposición 2.1.10.** *Sea  $\sum_{i=0}^n \beta_i f v_i$  una representación del operador de interpolación de Lagrange  $L[f] = \sum_{j=0}^n f(x_j) l_j$  en nodos  $x_0, \dots, x_n$  en el subespacio  $U$  de  $C[a, b]$ . Si  $1 \in U$ , entonces*

$$\sum_{i=0}^n l_i(x) = 1, \quad x \in [a, b] \quad (2.1.8)$$

y

$$\lambda_L(x) \geq 1, \quad x \in [a, b].$$

Además,

$$\max_{x \in [a, b]} \frac{\text{cond}(x; \beta)}{\lambda_L(x)} = \max_{i=0,\dots,n} \text{cond}(x_i; \beta).$$

*Demostración.* Como  $L[1] = 1$  se deduce (2.1.8). La segunda fórmula se deduce del Teorema 2.1.2 aplicando la desigualdad triangular. Y la última igualdad se deduce del Teorema 2.1.8 teniendo en cuenta que el conjunto  $\{x \in [a, b] \mid \lambda_L(x) > 0\}$  es  $[a, b]$ .  $\square$

## 2.2 Estabilidad de la fórmula de Newton en interpolación polinómica

Consideremos ahora el caso en el que  $U = P_n$ , el espacio de los polinomios de grado menor o igual que  $n$ . Existen diferentes representaciones del operador de interpolación de Lagrange que se manejan en la práctica. Las más importantes son las correspondientes a las fórmulas de Lagrange y de Newton.

La representación de Lagrange ofrece muchas ventajas. Como ya hemos visto en la sección anterior, el condicionamiento correspondiente a la representación de Lagrange es óptimo. Además, en este caso, los polinomios fundamentales de Lagrange, correspondientes a los nodos  $x_0, \dots, x_n$ , admiten la expresión explícita

$$l_i(x) = \prod_{j \in \{0, \dots, n\} \setminus \{i\}} \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n,$$

por lo que la representación de Lagrange del operador de interpolación admite una expresión fácilmente manejable. La función de Lebesgue en este caso tiene la expresión explícita

$$\lambda_L(x) = \sum_{i=0}^n \prod_{j \in \{0, \dots, n\} \setminus \{i\}} \frac{|x - x_j|}{|x_i - x_j|}. \quad (2.2.1)$$

Presentamos ahora la fórmula de Newton. Definamos los polinomios mónicos de grado  $i$  que se anulan en los puntos cuyos índices son menores que  $i$

$$\omega_0(x) := 1, \quad \omega_i(x) := (x - x_0) \cdots (x - x_{i-1}), \quad i = 1, \dots, n+1. \quad (2.2.2)$$

La fórmula de Newton es la representación del polinomio de interpolación de Lagrange en nodos  $x_0, \dots, x_n$ , dada por

$$L[f](x) = \sum_{i=0}^n d_i f \omega_i(x)$$

donde  $d_i$  son los funcionales diferencia dividida

$$d_i f := [x_0, \dots, x_i] f, \quad i = 0, \dots, n, \quad (2.2.3)$$

y  $(\omega_0, \dots, \omega_n)$  es la base de Newton.

Teniendo en cuenta que la diferencia dividida es el coeficiente en  $x^n$  de  $L[f]$  tenemos

$$d_n f = \sum_{i=0}^n \frac{f(x_i)}{\prod_{j \in \{0, \dots, n\} \setminus \{i\}} (x_i - x_j)}.$$

Como

$$\omega'_{n+1}(x_i) = \prod_{j \in \{0, \dots, n\} \setminus \{i\}} (x_i - x_j)$$

deducimos

$$d_n f = \sum_{i=0}^n \frac{f(x_i)}{\omega'_{n+1}(x_i)}. \quad (2.2.4)$$

*Nota 2.2.1.* De la fórmula (2.2.4) se deduce que la diferencia dividida en los nodos  $x_0, \dots, x_n$  no depende del orden en el que se presenten estos nodos.

De la Proposición 2.1.3 y la fórmula (2.2.4) se deduce el siguiente resultado.

**Proposición 2.2.2.** *La norma del funcional diferencia dividida  $d_k = [x_0, \dots, x_k]$  es*

$$\|d_k\|_\infty = \sum_{j=0}^k \frac{1}{|\omega'_{k+1}(x_j)|} = \sum_{j=0}^k \frac{1}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|}.$$

Notemos que el mismo interpolante en un conjunto de nodos admite distintas representaciones de Newton, dependiendo del orden de los nodos. Una ordenación diferente de los nodos da lugar a diferentes funcionales  $d_0, \dots, d_n$  y diferentes funciones básicas  $\omega_0, \dots, \omega_n$ . Dependiendo del orden de los nodos, la fórmula de Newton puede tener diferentes propiedades de estabilidad que pueden analizarse mediante la función

$$\text{cond}(x; d) = \sum_{k=0}^n \|d_k\|_{\infty} |\omega_k(x)|.$$

**Teorema 2.2.3.** Sean  $x_0, \dots, x_n$  una sucesión de nodos y  $L$  la matriz de cambio de base entre la base de Newton y la de Lagrange

$$(\omega_0, \dots, \omega_n) = (l_0, \dots, l_n)L. \quad (2.2.5)$$

Entonces  $L$  es una matriz triangular inferior y coincide la matriz de colocación

$$L = M \begin{pmatrix} \omega_0, \dots, \omega_n \\ x_0, \dots, x_n \end{pmatrix}, \quad (2.2.6)$$

cuyos elementos vienen dados por

$$l_{ik} = \omega_k(x_i) = (x_i - x_0) \cdots (x_i - x_{k-1}), \quad i, k = 0, \dots, n. \quad (2.2.7)$$

Además, los elementos de la matriz inversa  $L^{-1} = (l_{kj}^{(-1)})_{0 \leq k, j \leq n}$  son de la forma

$$l_{kj}^{(-1)} = \begin{cases} 0, & \text{si } j > k, \\ \frac{1}{\omega'_{k+1}(x_j)}, & \text{si } j \leq k. \end{cases} \quad (2.2.8)$$

*Demostración.* Por el Teorema 2.1.5, se cumplen (2.2.6) y (2.2.7). Además,  $L$  es triangular inferior ya que

$$\omega_k(x_i) = 0, \quad \text{si } k > i. \quad (2.2.9)$$

Aplicando la fórmula (2.2.4), tenemos

$$d_k = \sum_{j=0}^k \frac{1}{\omega'_{k+1}(x_j)} \lambda_j,$$

y, de la fórmula (2.1.2), se deduce (2.2.8).  $\square$

Para poder aplicar el Teorema 2.1.8 y dar una cota, calculamos

$$\max_{i=0, \dots, n} \text{cond}(x_i, d) = \max_{i=0, \dots, n} \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x_i)|}{|\omega'_{k+1}(x_j)|} = \max_{i=0, \dots, n} \sum_{k=0}^i \sum_{j=0}^k \frac{\prod_{l \in \{0, \dots, k-1\}} |x_i - x_l|}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|}, \quad (2.2.10)$$

y, usando la Proposición 2.1.10, tenemos el siguiente resultado.

**Teorema 2.2.4.** Sea  $\sum_{i=0}^n d_i f \omega_i$  la representación de Newton del polinomio de interpolación de Lagrange  $L[f] = \sum_{j=0}^n f(x_j) l_j$  en la sucesión nodos  $x_0, \dots, x_n$ . Entonces

$$\text{cond}(x; d) = \sum_{k=0}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} = \sum_{k=0}^n \sum_{j=0}^k \frac{\prod_{l \in \{0, \dots, k-1\}} |x - x_l|}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|}$$

y

$$\max_{x \in [a, b]} \frac{\text{cond}(x; d)}{\lambda_L(x)} = \max_{i=0, \dots, n} \text{cond}(x_i; d) = \max_{i=0, \dots, n} \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x_i)|}{|\omega'_{k+1}(x_j)|} = \text{Cond}(L^{-1}),$$

donde  $L$  es la matriz triangular inferior dada por (2.2.6).

*Demostración.* La expresión de  $\text{cond}(x; d)$  se deduce aplicando la Proposición 2.2.2. La fórmula final se deduce del Teorema 2.1.8 y la Proposición 2.1.10, teniendo en cuenta que  $\omega_k(x_i) = 0$  si  $k > i$ .  $\square$

Veamos que la función  $\text{cond}(x; d)$  es invariante por transformaciones afines.

**Proposición 2.2.5.** Sea  $T$  una transformación afín biyectiva. Entonces

$$\text{cond}(x; d) = \text{cond}(T(x), T^*(d)),$$

con

$$T^*(d_k) = [T(x_0), \dots, T(x_k)], \quad k = 0, \dots, n.$$

*Demostración.* Toda transformación afín biyectiva es de la forma  $T(x) = rx + s$ , con  $r \neq 0$ . Si consideramos el problema de interpolación con los nodos transformados  $T(x_i)$ ,  $i = 0, \dots, n$ , aparecerán nuevos funcionales diferencia dividida

$$T^*(d_k) = [T(x_0), \dots, T(x_k)]$$

cuyas normas verifican

$$\|T^*(d_k)\|_\infty = \|[T(x_0), \dots, T(x_k)]\|_\infty = |r|^{-k} \|[x_0, \dots, x_k]\|_\infty.$$

También tenemos que

$$\prod_{l=0}^{k-1} |T(x) - T(x_k)| = |r|^k \prod_{l=0}^{k-1} |x - x_k|.$$

Por tanto,

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \sum_{j=0}^k \frac{\prod_{l \in \{0, \dots, k-1\}} |x - x_l|}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|} \\ &= \sum_{k=0}^n \sum_{j=0}^k \frac{\prod_{l \in \{0, \dots, k-1\}} |T(x) - T(x_l)|}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |T(x_j) - T(x_l)|} = \text{cond}(T(x), T^*(d)). \end{aligned}$$

$\square$

### 2.3 Alta precisión relativa de las correspondientes matrices triangulares inferiores

Una expresión  $y$  puede obtenerse con *alta precisión relativa* (high relative accuracy o HRA) si el error relativo del valor calculado  $\hat{y}$  puede acotarse del siguiente modo ([19]):

$$\frac{\|y - \hat{y}\|}{\|y\|} \leq Cu,$$

donde  $C$  es una constante positiva independiente de la precisión aritmética y  $u$  es la unidad de redondeo.

Podemos asegurar que podemos calcular con alta precisión relativa los productos, cocientes y sumas verdaderas (sumas de números del mismo signo) de expresiones que se pueden calcular con alta precisión relativa. Sólo se permite efectuar operaciones de diferencia (sumas de números de signo opuesto) con los datos iniciales del problema (véase [26]). La importancia de poder asegurar alta precisión relativa proviene de que podemos asegurar que los errores relativos son del orden de la unidad de redondeo, independientemente del condicionamiento del problema. Muy pocos problemas se han podido resolver con métodos numéricos con HRA hasta ahora, aunque actualmente la actividad investigadora en este campo es muy intensa.

Vamos a ver que la matriz de colocación  $L$ , dada por (2.2.6), y su inversa pueden calcularse con HRA. Teniendo en cuenta (2.2.7) y (2.2.8),  $L$  y  $L^{-1}$  pueden calcularse con HRA porque sus elementos son productos o cocientes de diferencias de datos originales. Además, por (2.2.7) los elementos de la matriz  $L$  pueden calcularse por recurrencia, dando lugar al siguiente algoritmo

$$\begin{aligned} l_{i0} &= 1, \quad i = 0, \dots, n, \\ l_{ij} &= l_{i,j-1}(x_i - x_{j-1}), \quad i = j, \dots, n, \quad j = 1, \dots, n. \end{aligned} \quad (2.3.1)$$

Análogamente, por (2.2.8) los elementos de  $L^{-1}$  pueden calcularse siguiendo el siguiente algoritmo

$$l_{ij}^{(-1)} = -\frac{l_{i-1,j}^{(-1)}}{x_i - x_j}, \quad j = 0, \dots, i-1, \quad l_{ii}^{(-1)} = \frac{1}{\prod_{j=0}^{i-1} (x_i - x_j)}, \quad i = 0, \dots, n. \quad (2.3.2)$$

Con los algoritmos (2.3.1) y (2.3.2),  $L$  y  $L^{-1}$  pueden calcularse con HRA sin restricciones sobre los nodos, es decir, el cálculo con HRA puede asegurarse para todos los posibles signos y configuraciones de los nodos.

**Teorema 2.3.1.** *Sea  $L$  la matriz triangular inferior dada por (2.2.6). Entonces  $L$  y  $L^{-1}$  pueden calcularse con HRA para cualesquiera nodos usando los algoritmos (2.3.1) y (2.3.2), respectivamente.*

*Demostración.* Los algoritmos (2.3.1) y (2.3.2) solamente usan restas en los nodos (datos iniciales) y productos y cocientes de ellos.  $\square$

Del Teorema 2.3.1 se deduce que  $\max_{i=0,\dots,n} \text{cond}(x_i; d) = \text{Cond}(L^{-1})$  también se puede calcular con alta precisión relativa.

*Nota 2.3.2.* Recordemos que por el Teorema 2.2.4,

$$\text{cond}(x; d) = \sum_{k=0}^n \sum_{j=0}^k \frac{\prod_{l \in \{0, \dots, k-1\}} |x - x_l|}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|},$$

y la fórmula (2.2.1)

$$\text{cond}(x; \lambda) = \sum_{i=0}^n \prod_{j \in \{0, \dots, n\} \setminus \{i\}} \frac{|x - x_j|}{|x_i - x_j|}.$$

El punto de evaluación puede considerarse un dato inicial. Así, en ambos casos, los sumandos son productos o cocientes de diferencias de datos originales, y por tanto, tenemos que los condicionamientos de las fórmulas de Lagrange y de Newton pueden calcularse con HRA. Por tanto, a pesar de la inestabilidad de las fórmulas de Lagrange y de Newton para grado alto, podemos ofrecer resultados fiables para el cálculo del condicionamiento en nuestros experimentos numéricos.

## 2.4 Diferentes normalizaciones de la fórmula de Newton

Al igual que cualquier representación de un cuasi-interpolante, la representación de Newton es invariante por escalado de los funcionales. Una posibilidad consiste en utilizar funcionales normalizados de diferencias divididas

$$\hat{d}_i = \frac{d_i}{\|d_i\|_\infty}, \quad \hat{\omega}_i = \|d_i\|_\infty \omega_i, \quad i = 0, \dots, n. \quad (2.4.1)$$

La norma de los funcionales diferencia dividida puede calcularse explícitamente en términos de los nodos usando la Proposición 2.2.2.

Otra representación interesante en este contexto consiste en considerar la representación de Newton con los funcionales diferencia finita

$$\sum_{k=0}^n \Delta(x_0, \dots, x_k) f \tilde{\omega}_k(x),$$

donde  $\Delta(x_0, \dots, x_k)$  denota el funcional diferencia finita definido por

$$\Delta(x_0, \dots, x_k)f := \omega_k(x_k)[x_0, \dots, x_k]f = \sum_{j=0}^k \frac{\omega_k(x_k)}{\omega'_{k+1}(x_j)} f(x_j), \quad (2.4.2)$$

cuya norma es

$$\|\Delta(x_0, \dots, x_k)\|_\infty = \sum_{j=0}^k \frac{|\omega_k(x_k)|}{|\omega'_{k+1}(x_j)|}, \quad k = 0, \dots, n, \quad (2.4.3)$$

dando lugar a otra normalización

$$\tilde{d}_k = \omega_k(x_k)d_k, \quad \tilde{\omega}_k(x) = \frac{\omega_k(x)}{\omega_k(x_k)}, \quad k = 0, \dots, n.$$

Al contrario que los funcionales diferencia dividida  $d_k$ , que no dependen del orden en el que se presenten sus nodos  $x_0, \dots, x_k$ , el funcional diferencia finita  $\Delta(x_0, \dots, x_k)$  no es una función simétrica de los nodos, puesto que depende de manera especial del último nodo elegido, siendo simétrica respecto a los demás nodos.

Observemos que tanto  $\hat{d}_k$ ,  $\hat{\omega}_k$ , como  $\tilde{d}_k$ ,  $\tilde{\omega}_k$  son invariantes por transformaciones afines, mientras que  $d_k$  y  $\omega_k$  solamente son invariantes por traslaciones. Así, en el caso de los funcionales diferencias divididas, se tiene

$$[x_0 + h, \dots, x_n + h]f = [x_0, \dots, x_n]f(\cdot + h),$$

es decir,

$$(T^*d)(f) = d(f \circ T), \quad T(x) = x + h.$$

Análogamente, los valores en  $x+h$  de los elementos de la base de Newton correspondientes a los nodos  $x_0 + h, \dots, x_n + h$  coinciden con los valores de la base de Newton en  $x$  en los nodos  $x_0, \dots, x_n$ .

Las tres representaciones son útiles para estimar el condicionamiento. Debido a la invariancia de  $\text{cond}(x; d)$  por el reescalado de los funcionales  $d_0, \dots, d_n$  tenemos que

$$\text{cond}(x; d) = \text{cond}(x; \tilde{d}) = \text{cond}(x; \hat{d}).$$

*Nota 2.4.1.* Notemos que, por (2.2.7) y (2.2.8), la traslación de los nodos no tiene ninguna influencia sobre el comportamiento de  $L$  y  $L^{-1}$  ya que sus elementos son productos o cocientes de diferencias de nodos.

La representación de Newton  $L[f] = \sum_{i=0}^n c_i \omega_i(x)$ ,  $c_i = d_i f$ ,  $i = 0, \dots, n$ , puede interpretarse matricialmente. El vector de coeficientes  $c = (c_0, \dots, c_n)^T$ , se obtiene resolviendo el sistema

$$Lc = \begin{pmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{pmatrix}. \quad (2.4.4)$$

Para ciertas distribuciones de nodos, la matriz  $L$  puede estar mal condicionada dando lugar a coeficientes  $c$  grandes en comparación con los valores de  $f$  en los nodos. Esto implica también grandes errores en el cálculo de los coeficientes, que son diferencias divididas de la función en los nodos. El mal condicionamiento de  $L$  no siempre conduce a grandes errores de la representación. Frecuentemente los coeficientes con mayor error suelen estar multiplicados por las funciones con valores más bajos. A este respecto conviene tener en cuenta el comentario de Higham en la Sección 5.3 de [37], en el que se advierte que la representación de Newton puede reproducir los datos originales bien aunque las diferencias divididas se calculen con muy poca precisión.

En general tenemos

$$\frac{\max_{i=0, \dots, n} |f(x_i)|}{\|L\|_\infty} \leq \|c\|_\infty \leq \|L^{-1}\|_\infty \max_{i=0, \dots, n} |f(x_i)|,$$

de manera que la discrepancia entre los valores de la función en los nodos y sus coeficientes puede medirse utilizando  $\kappa_\infty(L)$ . Para la misma distribución relativa de los nodos,  $\kappa_\infty(L)$  puede cambiar dependiendo de la longitud del intervalo elegido, pudiendo ser arbitrariamente grande. Esta desventaja no se encuentra en la representación normalizada o en la representación de diferencias finitas.

Las diferentes normalizaciones de la representación de Newton corresponden a sistemas como (2.4.4), donde la matriz de coeficientes es de la forma  $LD$ , siendo  $D$  una matriz diagonal no singular. Observemos que por la Nota 1.3.7 el condicionamiento de Skeel de las inversas de todas estas matrices es el mismo.

Por la Nota 1.2.10,  $\hat{M}$  es la matriz de cambio de base entre dos representaciones con funcionales normalizados. En el caso de la representación de Newton,

$$\hat{L} = LD_d, \quad D_d = \text{diag}(\|d_0\|_\infty, \dots, \|d_n\|_\infty), \quad (2.4.5)$$

y la matriz asociada a la representación de Newton de los funcionales diferencia dividida normalizados (2.4.1). Por tanto,

$$\hat{L} = M \begin{pmatrix} \hat{\omega}_0, \dots, \hat{\omega}_n \\ x_0, \dots, x_n \end{pmatrix}. \quad (2.4.6)$$

Como

$$\hat{\omega}_k(x) = \|d_k\|_\infty \omega_k(x) = \sum_{j=0}^k \frac{(x - x_0) \cdots (x - x_{k-1})}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|},$$

tenemos

$$\hat{l}_{ik} = \hat{\omega}_k(x_i) = \sum_{j=0}^k \frac{(x_i - x_0) \cdots (x_i - x_{k-1})}{\prod_{l \in \{0, \dots, k\} \setminus \{j\}} |x_j - x_l|}, \quad i = 0, \dots, n, \quad k = i, \dots, n.$$

El siguiente resultado muestra que  $\text{Cond}(L^{-1}) = \kappa_\infty(\hat{L})$ .

**Proposición 2.4.2.** *Sea  $\hat{L}$  dada por (2.4.6),  $L$  dada por (2.2.6) y  $D_{L^{-1}} = \text{diag}(|L^{-1}|e)$ , con  $e = (1, \dots, 1)^T$ . Entonces*

$$\hat{L}^{-1} = D_{L^{-1}} L^{-1}$$

y se verifica la siguiente propiedad de optimalidad

$$\kappa_\infty(\hat{L}) \leq \kappa_\infty(LD),$$

para toda  $D$  matriz diagonal no singular. Además,

$$\text{Cond}(L^{-1}) = \kappa_\infty(\hat{L}).$$

*Demostración.* Por la Proposición 2.2.2 y la fórmula (2.2.8)

$$|L^{-1}|e = \begin{pmatrix} \|d_0\|_\infty \\ \vdots \\ \|d_n\|_\infty \end{pmatrix}.$$

Por tanto,

$$D_{L^{-1}} = (\text{diag}(\|d_0\|_\infty, \dots, \|d_n\|_\infty))^{-1} = D_d^{-1}$$

y

$$\hat{L}^{-1} = D_{L^{-1}} L^{-1}$$

Por la Proposición 1.3.6,

$$\text{Cond}(L^{-1}) = \kappa_\infty(D_{L^{-1}} L^{-1}) = \kappa_\infty(\hat{L}^{-1}) \leq \kappa_\infty(D^{-1} L^{-1}),$$

para toda matriz  $D$  diagonal no singular.

Por tanto

$$\text{Cond}(L^{-1}) = \kappa_\infty(\hat{L}) \leq \kappa_\infty(LD)$$

para toda  $D$  matriz diagonal no singular. □

Finalmente, aplicando el Teorema 2.1.8 se deduce el siguiente resultado.

**Proposición 2.4.3.** *Sea  $\hat{L}$  dada por (2.4.6) y  $L$  dada por (2.2.6). Entonces*

$$\max_{i=0,\dots,n} \text{cond}(x_i, d) = \text{Cond}(L^{-1}) = \|\hat{L}\|_\infty.$$

*Demostración.* Teniendo en cuenta que la base de Newton es la base dual de la base de diferencias divididas, deducimos de (2.1.4) y (2.4.5)

$$\max_{i=0,\dots,n} \text{cond}(x_i, d) = \text{Cond}(L^{-1}) = \|\hat{L}\|_\infty.$$

□

La matriz  $\tilde{L}$

$$\tilde{L} := M \begin{pmatrix} \tilde{\omega}_0, \dots, \tilde{\omega}_n \\ x_0, \dots, x_n \end{pmatrix} = L \text{diag}(1/\omega_0(x_0), \dots, 1/\omega_n(x_n)), \quad (2.4.7)$$

correspondiente a la representación de Newton con diferencias finitas tiene elementos diagonales iguales a 1 y

$$\tilde{l}_{ik} = \prod_{j=0}^{k-1} \frac{x_i - x_j}{x_k - x_j}, \quad i = 0, \dots, n, \quad k = i, \dots, n. \quad (2.4.8)$$

Por (2.4.7), tenemos que  $\tilde{L}^{-1} = \text{diag}(\omega_0(x_0), \dots, \omega_n(x_n))L^{-1}$ . Sean  $\tilde{l}_{ik}^{(-1)}$  los elementos de  $\tilde{L}^{-1}$ . Usando (2.2.8), se tiene

$$\tilde{l}_{ik}^{(-1)} = \frac{\omega_i(x_i)}{\omega'_{i+1}(x_k)}. \quad (2.4.9)$$

Por la optimalidad de  $\hat{L}$  descrita en la Proposición 2.4.2, tenemos

$$\kappa_\infty(\hat{L}) \leq \kappa_\infty(\tilde{L}).$$

## 2.5 Matrices totalmente positivas

Las matrices totalmente positivas son un caso particular de matrices interesantes tanto por su importancia teórica como por sus amplias aplicaciones en estadística, economía, biología, combinatoria, sistemas dinámicos, teoría de la aproximación y diseño geométrico asistido por ordenador. I.J. Schoenberg, M. G. Krein, F.R. Gantmacher y S. Karlin fueron pioneros en algunas de las contribuciones más importantes en la historia de la total positividad (véase [50]).

Veamos primero la notación usada para denotar submatrices. Denotamos por  $Q_{k,n}$  al conjunto de sucesiones estrictamente crecientes de  $k$  enteros elegidos en  $\{1, 2, \dots, n\}$

$$Q_{k,n} := \{\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k \mid 1 \leq \alpha_1 < \alpha_2 < \dots < \alpha_k \leq n\}.$$

Sean  $A$  una matriz  $n \times m$ ,  $\alpha \in Q_{k,n}$  y  $\beta \in Q_{l,m}$ . Se denota por  $A[\alpha|\beta]$  a la submatriz  $k \times l$  de  $A$  que se obtiene usando las filas que ocupan los lugares  $\alpha$  y columnas que ocupan los lugares  $\beta$ . Cuando  $\alpha = \beta$ , la submatriz principal  $A[\alpha|\beta]$  se denota simplemente por  $A[\alpha]$ .

**Definición 2.5.1.** *Sea  $A$  una matriz real  $n \times m$ .  $A$  se dice totalmente positiva (TP) si todos sus menores son no negativos, es decir, si se verifica*

$$\det A[\alpha|\beta] \geq 0, \tag{2.5.1}$$

para todo  $\alpha \in Q_{k,n}$ ,  $\beta \in Q_{k,m}$ ,  $k = 1, 2, \dots, \min(n, m)$ . *A se dice estrictamente totalmente positiva si todos sus menores son positivos, es decir, si en (2.5.1)  $\geq$  se reemplaza por  $>$ .*

El siguiente resultado corresponde al Corolario 3.8 de [3] de T. Ando sobre matrices TP no singulares.

**Proposición 2.5.2.** *Si una matriz  $A n \times n$  TP es no singular entonces todos sus menores principales son estrictamente positivos, es decir,*

$$\det A[\alpha] > 0,$$

para todo  $\alpha \in Q_{k,n}$ ,  $k = 1, \dots, n$ .

Como consecuencia de la Proposición 2.5.2, podemos deducir que la eliminación gaussiana se puede aplicar a una matriz TP no singular  $A$  hasta triangularizarla sin necesidad de cambios de filas y, por tanto,  $A$  posee factorización triangular  $LU$ . Más adelante, en el Teorema 2.5.4 veremos que dicha factorización caracteriza estas matrices.

Los menores del producto de dos matrices  $A$  y  $B$  de tamaño  $n \times n$  verifican la fórmula de Cauchy-Binet

$$\det(AB)[\alpha|\beta] = \sum_{\omega \in Q_{k,n}} \det A[\alpha|\omega] \det B[\omega|\beta], \quad \alpha, \beta \in Q_{k,n}.$$

Como consecuencia directa de la identidad de Cauchy-Binet se deduce que la total positividad se hereda con el producto de matrices cuadradas. De hecho, esto sucede con dos matrices cualesquiera que se puedan multiplicar.

**Proposición 2.5.3.** *Si  $A$  y  $B$  son dos matrices  $n \times n$  TP entonces  $AB$  es TP.*

Cryer probó (véase [23]) una caracterización sobre matrices TP y una factorización triangular  $LU$ .

**Teorema 2.5.4.** *Una matriz cuadrada  $A$  es TP si y sólo si tiene una factorización  $LU$  tal que  $L$  y  $U$  son ambas TP, con  $L$  y  $U$  matrices triangulares inferior y superior, respectivamente.*

Las siguientes operaciones preservan la clase de matrices totalmente positivas (ver [50]):

1. Multiplicar una fila (columna) por un escalar positivo.
2. Sumar un múltiplo positivo de una fila (columna) a la fila (columna) anterior o siguiente.
3. Sumar un valor positivo al elemento  $(1, 1)$  de la matriz (y al elemento  $(n, m)$  de una matriz  $n \times m$ ).

Para describir ejemplos de matrices TP, usaremos matrices de colocación (véase la Definición 2.1.4)

$$M \begin{pmatrix} u_0, u_1, \dots, u_n \\ t_0, t_1, \dots, t_n \end{pmatrix} = (u_j(t_i))_{0 \leq i, j \leq n}.$$

Presentamos ahora unos ejemplos de matrices TP.

**Ejemplos** 1. Matriz de Vandermonde.

Sea  $(m_0, \dots, m_n)$ ,  $m_j(x) = x^j$ ,  $j = 0, \dots, n$ , la base de monomios. La matriz de colocación de la base de monomios en nodos  $x_0, \dots, x_n$

$$V := M \begin{pmatrix} m_0, m_1, \dots, m_n \\ x_0, x_1, \dots, x_n \end{pmatrix} = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}$$

es la matriz de Vandermonde. Si los nodos son distintos, positivos y están ordenados crecientemente, es decir,  $0 < x_0 < x_1 < \dots < x_n$ , entonces la matriz de Vandermonde es una matriz estrictamente totalmente positiva (véase [30] y la Sección 3 del Capítulo 2 de [33]). Si  $0 \leq x_0 < x_1 < \dots < x_n$  entonces la matriz de Vandermonde es TP.

2. Matriz de Bernstein-Vandermonde.

La base de Bernstein de los polinomios de grado menor o igual que  $n$  en el intervalo  $[0, 1]$  es  $\mathcal{B}_n = (b_{0,n}, \dots, b_{n,n})$ , donde

$$b_{i,n}(x) = \binom{n}{i} (1-x)^{n-i} x^i, \quad i = 0, \dots, n.$$

La siguiente matriz de colocación se conoce como matriz de Bernstein-Vandermonde

$$B := M \begin{pmatrix} b_{0,n}, b_{1,n}, \dots, b_{n,n} \\ x_0, x_1, \dots, x_n \end{pmatrix} = \begin{pmatrix} \binom{n}{0} (1-x_0)^n & \binom{n}{1} (1-x_0)^{n-1} x_0 & \dots & \binom{n}{n} x_0^n \\ \binom{n}{0} (1-x_1)^n & \binom{n}{1} (1-x_1)^{n-1} x_1 & \dots & \binom{n}{n} x_1^n \\ \vdots & & \ddots & \vdots \\ \binom{n}{0} (1-x_n)^n & \binom{n}{1} (1-x_n)^{n-1} x_n & \dots & \binom{n}{n} x_n^n \end{pmatrix},$$

donde  $x_0, \dots, x_n \in [0, 1]$  son nodos distintos. Las matrices de Bernstein-Vandermonde son una generalización de las matrices de Vandermonde cuando se considera la base de Bernstein en lugar de la base de monomios. Es bien conocido (véase [16]) que una matriz de Bernstein-Vandermonde es una matriz totalmente positiva cuando los nodos verifican  $0 \leq x_0 < x_1 < \dots < x_n \leq 1$  (si las desigualdades son estrictas la matriz es estrictamente totalmente positiva).

3. Matrices de Pascal.

Una matriz de Pascal,  $P = (p_{ij})_{0 \leq i, j \leq n}$ , es una matriz simétrica cuyo elemento  $(i, j)$  viene dado por  $\binom{i+j}{j}$ ,  $i, j = 0, \dots, n$ . La matriz de Pascal es TP (véase el Capítulo 0 de [30] o Sección 4.7 de [50]). La matriz de Pascal triangular inferior  $L_P$  cuyo elemento  $(i, j)$  viene dado por  $\binom{i}{j}$ , también es TP (véase la Sección 2.5 de [30]).

4. Matriz de Cauchy.

Una matriz de Cauchy  $C = (c_{i,j})$  viene dada por  $c_{i,j} = 1/(x_i + y_j)$  y estrictamente totalmente positiva siempre que  $0 < x_0 < \dots < x_n$  y  $0 < y_0 < \dots < y_n$  (véase [30], la Sección 3 del Capítulo 2 de [33] o la Sección 4.3 de [50]). En el caso particular de que  $x_i = y_i = i + 1/2$ , tenemos la matriz de Hilbert.

La matriz de Hilbert es un típico ejemplo de matriz muy mal condicionada. En general, las matrices TP pueden estar muy mal condicionadas. Sin embargo, para matrices TP

no singulares adecuadamente parametrizadas, se pueden obtener (véase [39]) métodos que permiten realizar cálculos algebraicos con alta precisión relativa (para este concepto véase la Sección 2.3).

## 2.6 Matrices de Vandermonde, factorizaciones triangulares y alta precisión relativa

En esta sección vamos a considerar otra representación del operador de interpolación de Lagrange dada por

$$L[f](x) = \sum_{i=0}^n c_i f m_i(x),$$

donde  $\mathbf{m} = (m_0, \dots, m_n)^T$ ,  $m_j(x) = x^j$ ,  $j = 0, \dots, n$ , es la base de monomios y  $c_i$  es el funcional que asocia a cada  $f$  su coeficiente en  $m_i$  al expresar  $L[f]$  como combinación lineal de la base de monomios. Así, el problema de interpolación de Lagrange se reduce a resolver el sistema lineal  $V\mathbf{c} = \mathbf{f}$ , donde

$$V = V(x_0, \dots, x_n) := M \begin{pmatrix} m_0, m_1, \dots, m_n \\ x_0, x_1, \dots, x_n \end{pmatrix} = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix} \quad (2.6.1)$$

es la *matriz de Vandermonde* en los nodos  $x_0, \dots, x_n$ . Notemos que  $V$  es una matriz no singular ya que los nodos  $x_0, \dots, x_n$  son todos distintos.

De la fórmula (1.1.3) se deduce que la matriz de cambio de base entre la base de Lagrange  $(l_0, \dots, l_n)^T$  y la base de monomios  $(m_0, \dots, m_n)^T$  es

$$(\lambda_i m_j)_{0 \leq i, j \leq n} = (m_j(x_i))_{0 \leq i, j \leq n} = (x_i^j)_{0 \leq i, j \leq n} = V, \quad (2.6.2)$$

es decir,

$$(l_0, \dots, l_n) = (m_0, \dots, m_n) V^{-1}.$$

Así, la matriz de cambio de base entre la base de Lagrange y la base de monomios es la inversa de la matriz de Vandermonde. Expresando los elementos de la base de Lagrange en términos de los monomios obtenemos el elemento  $(i, j)$  de  $V^{-1}$

$$v_{ij}^{(-1)} = \frac{(-1)^{n-i} \sum_{k_1 < \dots < k_{n-i} \in \{0, \dots, n\} \setminus \{j\}} x_{k_1} \cdots x_{k_{n-i}}}{\prod_{k \neq j} (x_j - x_k)}. \quad (2.6.3)$$

Si aplicamos la fórmula de Newton

$$L[f] = \sum_{i=0}^n d_i f \omega_i(x) = (\omega_0(x), \dots, \omega_n(x)) \begin{pmatrix} [x_0]f \\ [x_0, x_1]f \\ \vdots \\ [x_0, \dots, x_n]f \end{pmatrix},$$

a  $f(x) = 1, x, \dots, x^n$ , obtenemos

$$\mathbf{m}^T = \boldsymbol{\omega}^T M \begin{pmatrix} m_0, m_1, \dots, m_n \\ [x_0], [x_0, x_1], \dots, [x_0, \dots, x_n] \end{pmatrix}. \quad (2.6.4)$$

Teniendo en cuenta que  $[x_0, \dots, x_i]m_j = 0$  si  $i > j$ , deducimos que la matriz de cambio de base entre la base de Newton y la base de monomios es la matriz triangular superior

$$U := M \begin{pmatrix} m_0, m_1, \dots, m_n \\ [x_0], [x_0, x_1], \dots, [x_0, \dots, x_n] \end{pmatrix} = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 0 & 1 & [x_0, x_1]x^2 & \cdots & [x_0, x_1]x^n \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & [x_0, \dots, x_{n-1}]x^n \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \quad (2.6.5)$$

Aplicando la regla de Leibniz para diferencias divididas a  $m_j(x) = xm_{j-1}(x)$ , se tiene

$$[x_0, \dots, x_i]m_j = x_i[x_0, \dots, x_i]m_{j-1} + [x_0, \dots, x_{i-1}]m_{j-1},$$

es decir,  $u_{ij} = u_{i-1,j-1} + x_i u_{i,j-1}$ . Esta fórmula da lugar al siguiente algoritmo para calcular los elementos de  $U$

$$\begin{aligned} u_{00} &= 1, & u_{0j} &= x_0 u_{0,j-1}, & j &= 1, \dots, n, \\ u_{ii} &= 1, & u_{ij} &= u_{i-1,j-1} + x_i u_{i,j-1}, & j &= i+1, \dots, n, & i &= 1, \dots, n. \end{aligned} \quad (2.6.6)$$

De (2.6.6) se deduce la siguiente fórmula para la fila  $i$ -ésima en términos de la fila  $(i-1)$ -ésima

$$u_{ij} = u_{i-1,j-1} + x_i u_{i-1,j-2} + x_i^2 u_{i-1,j-3} + \cdots + x_i^{j-i} u_{i-1,i-1},$$

o equivalentemente

$$[x_0, \dots, x_i]m_j = \sum_{k=0}^{j-i} x_i^k [x_0, \dots, x_{i-1}]m_{j-1-k}, \quad j \geq i.$$

Por inducción se demuestra que

$$u_{ij} = [x_0, \dots, x_i]m_j = \sum_{\alpha_0 + \dots + \alpha_i = j-i} x_0^{\alpha_0} \dots x_i^{\alpha_i}. \quad (2.6.7)$$

Considerando matrices de colocación en la fórmula (2.6.4) en los nodos  $x_0, \dots, x_n$ , se deduce

$$M \begin{pmatrix} m_0, m_1, \dots, m_n \\ x_0, x_1, \dots, x_n \end{pmatrix} = M \begin{pmatrix} \omega_0, \omega_1, \dots, \omega_n \\ x_0, x_1, \dots, x_n \end{pmatrix} U,$$

lo que implica que

$$V = LU, \quad (2.6.8)$$

donde  $L$  es la matriz dada por (2.2.6). Por tanto, las matrices  $L$  y  $U$  forman la (única) factorización de Crout de la matriz de Vandermonde, donde  $L$  es una matriz triangular inferior cuya diagonal tiene los pivotes de la eliminación de Gauss y  $U$  es una matriz triangular superior con unos en la diagonal.

Para obtener la inversa de  $U$ , empecemos con la relación entre la base de Newton y la base de monomios

$$(\omega_0(x), \dots, \omega_n(x)) = (m_0(x), \dots, m_n(x))U^{-1},$$

con  $U^{-1} = (u_{ij}^{(-1)})_{0 \leq i, j \leq n}$ . Teniendo en cuenta que los elementos  $u_{ij}^{(-1)}, i = 0, \dots, j$ , son los coeficientes de  $\omega_j$  respecto a la base de monomios  $(m_0, \dots, m_j)$ , podemos comparar los coeficientes en la relación  $\omega_j(x) = (x - x_{j-1})\omega_{j-1}(x)$  y obtener el algoritmo

$$\begin{aligned} u_{00}^{(-1)} &= 1, & u_{0j}^{(-1)} &= -x_{j-1}u_{0,j-1}^{(-1)}, & j &= 1, \dots, n, \\ u_{ii}^{(-1)} &= 1, & u_{ij}^{(-1)} &= u_{i-1,j-1}^{(-1)} - x_{j-1}u_{i,j-1}^{(-1)}, & j &= i+1, \dots, n, & i &= 1, \dots, n. \end{aligned} \quad (2.6.9)$$

Es bien conocido que la inversa de la matriz de Vandermonde con nodos no negativos y en orden creciente puede calcularse con HRA ya que la matriz es totalmente positiva y su factorización bidiagonal puede obtenerse con HRA (ver [27] and [39]). El siguiente resultado demuestra que el cálculo de  $V^{-1}$  con HRA es posible cuando los nodos son todos distintos con el mismo signo no estricto y para cualquier orden.

**Teorema 2.6.1.** *Sea  $V$  la matriz de Vandermonde dada por (2.6.1) correspondiente a una sucesión de nodos distintos con el mismo signo no estricto. Entonces,  $V^{-1}$  puede calcularse con HRA.*

*Demostración.* La fórmula (2.6.3) expresa los elementos de  $V^{-1}$  como un cociente. El denominador es un producto de diferencias de nodos (datos iniciales). El numerador es una suma de un producto de nodos todos con el mismo signo.  $\square$

El siguiente resultado demuestra que  $U$  y  $U^{-1}$  pueden calcularse con HRA bajo las mismas condiciones sobre los nodos que en el teorema anterior.

**Teorema 2.6.2.** *Sea  $U$  la matriz triangular superior dada por (2.6.5) correspondiente a una sucesión de nodos distintos con el mismo signo no estricto. Entonces  $U$  y  $U^{-1}$  pueden calcularse con HRA, usando los algoritmos (2.6.6) y (2.6.9), respectivamente.*

*Demostración.* Supongamos primero que todos los nodos son no negativos, es decir,  $x_i \geq 0$  para  $i = 0, \dots, n$ .

Por el algoritmo (2.6.6), los elementos de la primera fila y la diagonal son no negativos y pueden calcularse usando sólo productos de los datos iniciales. Los elementos  $u_{ij}$ ,  $j = i + 1, \dots, n$ ,  $i = 1, \dots, n$ , se calculan sumando dos términos no negativos. Así, puede asegurarse que  $U$  puede calcularse con HRA.

Análogamente, por (2.6.9) el signo de los elementos  $(i, j)$  de la primera fila y de la diagonal de  $U^{-1}$  coincide con  $(-1)^{i+j}$  y dichos elementos pueden calcularse con HRA. Por tanto, vemos que los  $u_{ij}^{(-1)}$ ,  $j = i + 1, \dots, n$ ,  $i = 1, \dots, n$  son suma de dos términos que pueden calcularse con HRA, ambos con el mismo signo que  $(-1)^{i+j}$ .

Ahora consideremos el caso en que todos los nodos son no positivos. Entonces, se tiene

$$V(x_0, \dots, x_n) = V(-x_0, \dots, -x_n)S, \quad S := \begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & (-1)^n \end{pmatrix},$$

es decir, la matriz de Vandermonde coincide con una matriz de Vandermonde con nodos no negativos salvo un cambio de signo en las columnas. Sean  $LU$  y  $L_2U_2$  las factorizaciones de Crout de  $V(x_0, \dots, x_n)$  and  $V(-x_0, \dots, -x_n)$ , respectivamente. Teniendo en cuenta que  $LU = (L_2S)(SU_2S)$  y la unicidad de la factorización de Crout, se tiene

$$L = L_2S, \quad U = SU_2S. \tag{2.6.10}$$

Así,  $(-1)^{i+j}u_{ij}$  y  $u_{ij}^{(-1)}$  son no negativos. La demostración ahora sigue de manera análoga al caso de nodos no negativos teniendo en cuenta que los algoritmos (2.6.6) y (2.6.9) suman términos del mismo signo.  $\square$

En esta sección hemos estudiado la factorización de Crout de la matriz de Vandermonde. Sin embargo, la factorización triangular más común usada con la eliminación de

Gauss es  $V = \tilde{L}\tilde{U}$ , donde  $\tilde{L}$  es una matriz triangular inferior con unos en la diagonal y  $\tilde{U}$  una matriz triangular superior. Esta factorización se conoce también como factorización de Doolittle. Sea  $D = \text{diag}(\omega_0(x_0), \dots, \omega_n(x_n))$ . Se tiene

$$\tilde{L} = LD^{-1}, \quad \tilde{U} = DU. \quad (2.6.11)$$

Notemos que  $\tilde{L}$  es la matriz dada por (2.4.7), correspondiente a la representación de Newton con diferencias finitas. Por el Teorema 2.3.1,  $\tilde{L}$  y  $\tilde{L}^{-1}$  pueden calcularse con HRA. Por el Teorema 2.6.2,  $\tilde{U}$  y  $\tilde{U}^{-1}$  pueden calcularse con HRA cuando los nodos tienen el mismo signo no estricto.

# 3

## Estabilidad de la fórmula de Newton para diferentes elecciones de nodos con orden creciente

---

### 3.1 Estabilidad de la fórmula de Newton para nodos equidistantes ordenados

En esta sección vamos a tomar nodos equidistantes  $x_0, \dots, x_n$  en  $[a, b]$  ordenados de menor a mayor

$$x_i = a + \frac{i}{n}(b - a), \quad i = 0, \dots, n.$$

Un simple cambio de variable  $T(x) = a + b - x$  permite reducir el caso en que los nodos sean decrecientes al anterior y se tiene

$$c(a + b - x; T^*(d)) = \text{cond}(x; d).$$

La constante de Lebesgue en nodos equidistantes tiene un crecimiento exponencial al incrementarse el grado  $n$ . Schönhage demuestra en [58] que la constante de Lebesgue admite la siguiente expresión asintótica

$$\frac{2^{n+1}}{e n \log(n + \gamma)}, \quad (3.1.1)$$

donde  $\gamma \approx 0.5772156649$  es la constante de Euler–Mascheroni. Para nosotros es suficiente la siguiente cota, cuya demostración está basada en una comunicación privada de Jia Rong-Qing a Carl de Boor (véase el ejercicio 1.3.22 del libro de T. J. Rivlin [52]).

**Proposición 3.1.1.** *La función de Lebesgue para nodos equidistantes verifica*

$$\lambda_L(x) \leq 2^n.$$

*Demostración.* Si  $x \in [a, b]$ , entonces  $x = a + th$  para algún  $t \in [0, n]$ . Los polinomios de Lagrange verifican

$$|l_i(a + th)| = \frac{\prod_{j \in \{0, \dots, n\} \setminus \{i\}} |t - j|}{i!(n - i)!}.$$

Veamos por inducción sobre  $n$  que  $\prod_{j \in \{0, \dots, n\} \setminus \{i\}} |t - j| \leq n!$ . Si  $n = 1$ , tenemos que

$$\max_{t \in [0, 1]} |t| = \max_{t \in [0, 1]} |t - 1| = 1.$$

Supongamos que se cumple para  $n - 1$  y veamos el caso  $n$ . Tenemos la siguiente propiedad de simetría

$$l_{n-i}(a + th) = l_i(a + (n - t)h),$$

por lo que podemos suponer que el máximo se alcanza en  $[0, n - 1]$ . Supongamos que  $t \in [0, n - 1]$ . Si  $i = n$

$$\prod_{j=0}^{n-1} |t - j| = (n - 1 - t) \prod_{j=0}^{n-2} |t - j| \leq (n - 1)!.$$

Si  $i \leq n - 1$  tenemos

$$\prod_{j \in \{0, \dots, n\} \setminus \{i\}} |t - j| = (n - t) \prod_{j \in \{0, \dots, n-1\} \setminus \{i\}} |t - j| \leq (n - t)(n - 1)! \leq n!.$$

Por lo tanto, tenemos

$$|l_i(x)| \leq \frac{n!}{i!(n - i)!} = \binom{n}{i}, \quad x \in [a, b],$$

y así,

$$\lambda_L(x) = \sum_{i=0}^n |l_i(x)| \leq \sum_{i=0}^n \binom{n}{i} = 2^n.$$

□

Pretendemos acotar  $\text{cond}(x; d) = \sum_{k=0}^n \|d_k\|_\infty |\omega_k(x)|$ , donde  $d_k$  es el funcional diferencia dividida en los nodos  $x_0, \dots, x_k$  y  $\omega_k(x)$  el polinomio mónico de Newton de grado  $k$ . Para ello veamos el siguiente lema.

**Lema 3.1.2.** Sean  $x_i = a + \frac{i}{n}(b - a)$ ,  $i = 0, \dots, n$ , y sean  $\omega_k(x)$ ,  $k = 0, \dots, n + 1$ , definidos en (2.2.2), y  $d_k$ ,  $k = 0, \dots, n$ , definidos en (2.2.3). Entonces

$$(i) \quad \omega_k(x_i) = \left(\frac{b-a}{n}\right)^k \frac{i!}{(i-k)!}, \quad 0 \leq k \leq i \leq n,$$

$$(ii) \quad \omega'_{k+1}(x_i) = (-1)^{k-i} \left(\frac{b-a}{n}\right)^k i!(k - i)!, \quad 0 \leq i \leq k \leq n,$$

$$(iii) \|d_k\|_\infty = \frac{1}{k!} \left(\frac{b-a}{n}\right)^{-k} 2^k, \quad 0 \leq k \leq n.$$

*Demostración.* Veamos primero (i). Como  $x_i - x_j = \frac{b-a}{n}(i-j)$ , tenemos

$$\omega_k(x_i) = \prod_{j=0}^{k-1} (x_i - x_j) = \prod_{j=0}^{k-1} \frac{b-a}{n}(i-j) = \left(\frac{b-a}{n}\right)^k i \cdots (i+1-k) = \left(\frac{b-a}{n}\right)^k \frac{i!}{(i-k)!}.$$

Ahora probemos (ii)

$$\omega'_{k+1}(x_i) = \prod_{j \neq i} (x_i - x_j) = \left(\frac{b-a}{n}\right)^k \prod_{j \neq i} (i-j) = (-1)^{k-i} \left(\frac{b-a}{n}\right)^k i!(k-i)!.$$

Usando la Proposición 2.2.2 y (ii) se tiene (iii). En efecto

$$\begin{aligned} \|d_k\|_\infty &= \sum_{j=0}^k \frac{1}{|\omega'_{k+1}(x_j)|} = \sum_{j=0}^k \left(\frac{b-a}{n}\right)^{-k} \frac{1}{j!(k-j)!} \\ &= \frac{1}{k!} \left(\frac{b-a}{n}\right)^{-k} \sum_{j=0}^k \binom{k}{j} = \frac{1}{k!} \left(\frac{b-a}{n}\right)^{-k} 2^k. \end{aligned}$$

□

La norma de los funcionales diferencia dividida depende de la longitud del intervalo. Sin embargo, la norma de los funcionales diferencia finita, definidos en (2.4.2), no depende del intervalo y puede calcularse utilizando el lema anterior.

**Proposición 3.1.3.** Sean  $x_i = a + \frac{i}{n}(b-a)$ ,  $i = 0, \dots, n$ , y sea  $\tilde{d}_k = \Delta(x_0, \dots, x_k)$  el funcional diferencia finita,  $k = 0, \dots, n$ . Entonces

$$\|\Delta(x_0, \dots, x_k)\|_\infty = 2^k.$$

*Demostración.* Por el Lema 3.1.2 (i) y (iii) se tiene

$$\omega_k(x_k) = \left(\frac{b-a}{n}\right)^k k!, \quad \|d_k\|_\infty = \frac{1}{k!} \left(\frac{b-a}{n}\right)^{-k} 2^k.$$

Por la definición (2.4.2)

$$\|\Delta(x_0, \dots, x_k)\|_\infty = |\omega_k(x_k)| \|d_k\|_\infty = 2^k.$$

□

El siguiente resultado nos da el valor de  $\text{Cond}(L^{-1}) = \max_{i=0, \dots, n} \text{cond}(x_i; d)$  en el caso de nodos equidistantes, donde  $L$  viene dada por (2.2.6).

**Teorema 3.1.4.** Si  $x_i = a + \frac{i}{n}(b - a)$ ,  $i = 0, \dots, n$ , entonces

$$\text{cond}(x_i; d) = 3^i$$

y

$$\max_{i=0, \dots, n} \text{cond}(x_i; d) = \text{cond}(x_n; d) = 3^n.$$

*Demostración.* Como  $\omega_k(x_i) = 0$  si  $k > i$ , se tiene

$$\text{cond}(x_i; d) = \sum_{k=0}^i \|d_k\|_{\infty} |\omega_k(x_i)|.$$

Usando el Lema 3.1.2,

$$\text{cond}(x_i; d) = \sum_{k=0}^i \frac{i!}{(i-k)! k!} 2^k = \sum_{k=0}^i \binom{i}{k} 2^k = 3^i$$

y se sigue el resultado. □

Los resultados anteriores nos muestran que la evaluación en puntos equidistantes mediante la fórmula de Newton puede presentar mayores problemas que la evaluación directa de la fórmula de Lagrange. Mientras que la función de Lebesgue toma un valor próximo a 1 en un entorno pequeño de los nodos, la función  $\text{cond}(x; d)$  alcanza valores  $3^i$  en el nodo  $x_i$ . Esto implica una significativa pérdida de estabilidad en los nodos. El Teorema 2.2.4 nos indica que el valor máximo de  $\frac{\text{cond}(x; d)}{\lambda_L(x)}$  se alcanza en los nodos. Por tanto, en el resto de los puntos tenemos

$$\text{cond}(x; d) \leq 3^n \lambda_L(x),$$

y de la Proposición 3.1.1 deducimos la cota uniforme

$$\text{cond}(x; d) \leq 6^n.$$

Esta cota es demasiado pesimista y, como veremos a continuación, no se supera el valor  $3^n$ . Para demostrarlo necesitamos los siguientes lemas. Para ello, transformamos los nodos equidistantes en  $[a, b]$

$$x_i = a + ih, \quad i = 0, \dots, n, \quad \text{con } h = \frac{b-a}{n},$$

en los primeros enteros no negativos  $0, 1, \dots, n$  haciendo el cambio de variable  $x = a + th$ . Observemos que

$$\omega_k(a + th) = h^k t(t-1) \cdots (t-k+1). \quad (3.1.2)$$

Veamos el primer lema necesario para mejorar la cota de  $\text{cond}(x; d)$ .

**Lema 3.1.5.** *Sea  $i$  un entero tal que  $1 \leq i \leq n$  y  $\omega_k$  dado por (3.1.2). Entonces, para  $t \in [i-1, i]$ , se tiene*

$$|\omega_k(a+th)| \leq h^k \frac{i!}{(i-k)!}, \quad \text{si } 0 \leq k \leq i,$$

$$|\omega_k(a+th)| \leq \frac{h^k}{4} i!(k-i)!, \quad \text{si } i < k \leq n.$$

*Demostración.* Como  $\omega_k(a+th)$  es creciente en  $[k-1, \infty]$ , entonces  $\omega_k(a+th) \leq \omega_k(a+ih)$ , para  $k \leq i$ . Luego

$$|\omega_k(a+th)| \leq |\omega_k(a+ih)| = h^k i(i-1) \cdots (i-k+1) = h^k \frac{i!}{(i-k)!}.$$

Si  $k \geq i+1$ , tenemos que  $[i-1, i] \subseteq [0, k-1]$  y

$$|\omega_k(a+th)| = h^k |t| \cdots |t-i+2| |t-i+1| |t-i| |t-i-1| \cdots |t-k+1|.$$

Usamos las siguientes cotas para deducir el resultado

$$|t-i+1| |t-i| \leq \frac{1}{4},$$

$$|t| |t-1| \cdots |t-i+2| \leq i(i-1) \cdots 2 = i!,$$

$$|t-i-1| \cdots |t-k+1| \leq 2 \cdots (k-i) = (k-i)!.$$

□

**Lema 3.1.6.** *Si  $0 < i \leq \frac{k+1}{2}$ , entonces*

$$\binom{k}{i} \geq \binom{k}{i-1}$$

y si  $\frac{k+1}{2} \leq i \leq k$ ,

$$\binom{k}{i-1} \geq \binom{k}{i}.$$

*Demostración.* Teniendo en cuenta que

$$\frac{\binom{k}{i}}{\binom{k}{i-1}} = \frac{k-i+1}{i}$$

deducimos que  $\binom{k}{i} \geq \binom{k}{i-1}$  si y sólo si  $k-i+1 \geq i$ , lo que equivale a afirmar que  $i \leq \frac{k+1}{2}$ . □

**Lema 3.1.7.** Si  $0 < i < k$ , entonces

$$\binom{k}{i} \geq k.$$

*Demostración.* Si  $1 \leq i \leq \frac{k+1}{2}$ , por el Lema anterior, la sucesión

$$\binom{k}{1} \leq \binom{k}{2} \leq \dots \leq \binom{k}{i}$$

es creciente. Luego

$$\binom{k}{i} \geq \binom{k}{1} = k.$$

Si  $\frac{k+1}{2} \leq i \leq k-1$ , entonces

$$\binom{k}{i} \geq \binom{k}{i+1} \geq \dots \geq \binom{k}{k-1}$$

es una sucesión decreciente y

$$\binom{k}{i} \geq \binom{k}{k-1} = k.$$

□

**Lema 3.1.8.** Sea la sucesión  $f_i = (1+a)^i + \frac{1}{4} \sum_{k=i+1}^n \frac{a^k}{k}$ ,  $i = 1, \dots, n$ , con  $a \geq 0$ . Se tiene

$$f_1 \leq f_2 \leq \dots \leq f_n = (1+a)^n,$$

es decir,  $\{f_i\}_{i=1}^n$  es creciente.

*Demostración.* Veamos que  $\Delta f_i \geq 0$  para todo  $i$

$$\Delta f_i = f_{i+1} - f_i = (1+a)^{i+1} - (1+a)^i - \frac{a^{i+1}}{4(i+1)} = a(1+a)^i - \frac{a^{i+1}}{4(i+1)}$$

Como  $a \geq 0$  se tiene

$$\Delta f_i \geq a^{i+1} - \frac{a^{i+1}}{4(i+1)} \geq 0$$

□

**Lema 3.1.9.** Sea  $1 \leq i \leq n-1$  y  $a > 3/2$ . Entonces

$$\frac{1}{a-1} \left( \frac{a^{n+1}}{n} - \frac{a^{i+1}}{i} \right) \leq \sum_{k=i+1}^n \frac{a^k}{k} \leq \frac{1}{2a-3} \left( \frac{2a^{n+1}}{n} - \frac{3a^{i+1}}{i+1} \right).$$

*Demostración.* Sea  $c_k := a^k/k$ ,  $k \geq 1$ . Entonces

$$\Delta c_k = c_{k+1} - c_k = \frac{a^{k+1}}{k+1} - \frac{a^k}{k} = \frac{(a-1)k-1}{k(k+1)} a^k$$

y así

$$\Delta c_k = \frac{(a-1)k-1}{ak} c_{k+1}, \quad k \geq 1.$$

Por tanto, se tiene que

$$\Delta c_k \leq \frac{a-1}{a} c_{k+1}, \quad k \geq 1.$$

Sumando y aplicando la desigualdad anterior, obtenemos

$$\sum_{k=i+1}^n \frac{a^k}{k} = \sum_{k=i}^{n-1} c_{k+1} \geq \frac{a}{a-1} \sum_{k=i}^{n-1} \Delta c_k = \frac{a}{a-1} (c_n - c_i) = \frac{1}{a-1} \left( \frac{a^{n+1}}{n} - \frac{a^{i+1}}{i} \right).$$

Para la otra desigualdad

$$\begin{aligned} (2a-3) \sum_{k=i+1}^n \frac{a^k}{k} &= \frac{2a^{n+1}}{n} - \frac{3a^{i+1}}{i+1} + \sum_{k=i+1}^{n-1} \left( \frac{2}{k} - \frac{3}{k+1} \right) a^{k+1} \\ &= \frac{2a^{n+1}}{n} - \frac{3a^{i+1}}{i+1} - \sum_{k=i+1}^{n-1} \frac{k-2}{k(k+1)} a^{k+1}. \end{aligned}$$

Como  $i \geq 1$ , todos los términos  $\frac{k-2}{k(k+1)} a^{k+1} \geq 0$ . Luego

$$(2a-3) \sum_{k=i+1}^n \frac{a^k}{k} \leq \frac{2a^{n+1}}{n} - \frac{3a^{i+1}}{i+1}$$

y, como  $2a-3 > 0$ , se tiene el resultado.  $\square$

Veamos ahora el siguiente teorema que nos da una cota para  $\text{cond}(x; d)$  dependiendo del subintervalo  $[x_{i-1}, x_i]$  en el que se encuentra  $x$  y demostraremos que el máximo condicionamiento de la representación de Newton es  $3^n$  y se alcanza en  $x_n$ .

**Teorema 3.1.10.** *Sean nodos equidistantes en  $[a, b]$ ,  $x_i = a + \frac{i}{n}(b-a)$ ,  $i = 0, \dots, n$ . Si  $x \in [x_{n-1}, x_n]$ , se tiene*

$$\text{cond}(x; d) \leq 3^n$$

y si  $x \in [x_{i-1}, x_i]$

$$\text{cond}(x; d) \leq 3^i + \frac{2^n}{n} - 3 \frac{2^{i-1}}{i+1}, \quad i = 1, \dots, n-1.$$

Además, se tiene la siguiente cota global

$$\max_{x \in [a, b]} \text{cond}(x; d) \leq 3^n = \text{cond}(x_n; d).$$

*Demostración.* Por el Lema 3.1.5 tenemos

$$|\omega_k(x)| \leq h^k \frac{i!}{(i-k)!}, \quad \text{si } k \leq i,$$

$$|\omega_k(x)| \leq \frac{h^k}{4} i!(k-i)!, \quad \text{si } k \geq i+1.$$

Por el Teorema 2.2.4

$$\text{cond}(x; d) = \sum_{k=0}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} = \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} + \sum_{k=i+1}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|}.$$

Aplicando las desigualdades anteriores y el Lema 3.1.2 (ii) tenemos

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} + \sum_{k=i+1}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} \\ &\leq \sum_{k=0}^i \sum_{j=0}^k \frac{h^k i! / (i-k)!}{h^k j! (k-j)!} + \frac{1}{4} \sum_{k=i+1}^n \sum_{j=0}^k \frac{h^k i! (k-i)!}{h^k j! (k-j)!} \\ &= \sum_{k=0}^i \sum_{j=0}^k \binom{i}{k} \binom{k}{j} + \frac{1}{4} \sum_{k=i+1}^n \sum_{j=0}^k \frac{\binom{k}{j}}{\binom{k}{i}} = \sum_{k=0}^i \binom{i}{k} 2^k + \frac{1}{4} \sum_{k=i+1}^n \frac{2^k}{\binom{k}{i}} \\ &= 3^i + \frac{1}{4} \sum_{k=i+1}^n \frac{2^k}{\binom{k}{i}}. \end{aligned}$$

Por el Lema 3.1.7

$$\text{cond}(x; d) \leq 3^i + \frac{1}{4} \sum_{k=i+1}^n \frac{2^k}{k}. \quad (3.1.3)$$

Si  $i = n$ ,  $x$  se encuentra en el intervalo  $[x_{n-1}, x_n]$  y tenemos

$$\text{cond}(x; d) \leq 3^n.$$

Si  $i < n$ , podemos aplicar el Lema 3.1.9 y se tiene

$$\text{cond}(x; d) \leq 3^i + \frac{2^n}{n} - 3 \frac{2^{i-1}}{i+1}.$$

Por último, aplicando el Lema 3.1.8 a la cota (3.1.3) con  $a = 2$ , y el Teorema 3.1.4, obtenemos

$$\text{cond}(x; d) \leq 3^i + \frac{1}{4} \sum_{k=i+1}^n \frac{2^k}{k} \leq 3^n = \text{cond}(x_n; d),$$

para todo  $x \in [a, b]$ . □

En la Figura 3.1.1 podemos ver las gráficas de la función de Lebesgue y la función  $\text{cond}(x; d)$  en puntos equidistantes para  $n = 10$  en el intervalo  $[0, 1]$ . El eje de ordenadas está en escala logarítmica en base 3 para poder visualizar más fácilmente el resultado del Teorema 3.1.10. El crecimiento de  $\text{cond}(x; d)$  aparece como lineal en la escala logarítmica y corresponde a un crecimiento exponencial en consonancia con el Teorema 3.1.10.

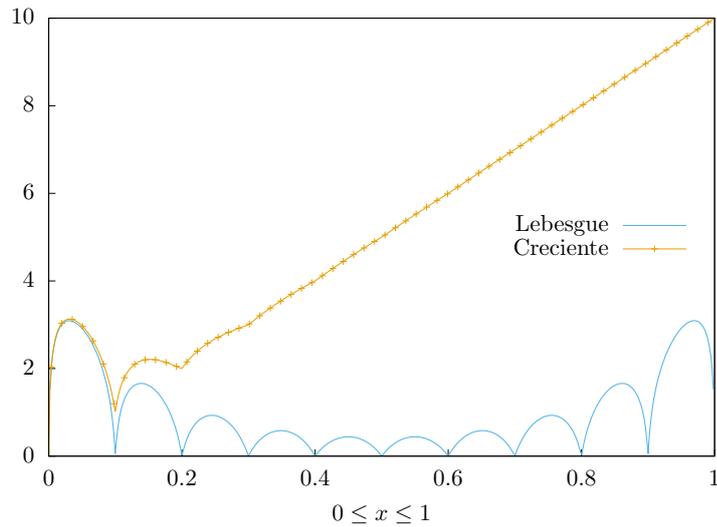


Figura 3.1.1: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x; d))$  para puntos en orden creciente con  $n = 10$ .

Vamos a interpretar los resultados obtenidos usando la fórmula (1.2.1). Debido a los errores de redondeo, podemos suponer que los datos provienen de  $\hat{f}$ , una perturbación de  $f$ , tal que

$$\|f - \hat{f}\|_\infty \leq u\|f\|_\infty,$$

donde  $u$  es la unidad de redondeo. Así, el error en el cálculo del interpolante puede acotarse superiormente por

$$|L[f](x) - L[\hat{f}](x)| \leq u \text{cond}(x; d)\|f\|_\infty \leq 3^n u\|f\|_\infty,$$

para la representación de Newton con nodos equidistantes. Si  $3^n \leq u^{-1}$ , entonces el error de redondeo puede ser menor que  $\|f\|_\infty$ . Por lo tanto, la fórmula de Newton puede tener errores de redondeo mayores que  $\|f\|_\infty$  solamente si  $n \geq 24 \log(2)/\log(3) \approx 15.14$ , para simple precisión aritmética ( $u^{-1} = 2^{24}$ ), o  $n \geq 53 \log(2)/\log(3) \approx 33.44$ , para doble precisión ( $u^{-1} = 2^{53}$ ). La constante de Lebesgue toma valores mayores que el inverso de la unidad de redondeo si  $n \geq 32$ , para simple precisión, o  $n \geq 62$ , para doble precisión.

Así, para doble precisión aritmética y valores de  $n$  entre 33 y 62, la fórmula de Lagrange da resultados correctos mientras que la fórmula de Newton puede fallar (especialmente cuando el punto  $x$  está cerca del último nodo  $x_n$ ).

Para ilustrar los resultados de esta sección con experimentos numéricos, hemos considerado el problema de interpolación

$$\begin{aligned} x_i &= -1 + \frac{2i}{n}, \quad i = 0, \dots, n, \\ p(x_0) &= 2, \quad p(x_i) = 1, \quad i = 1, \dots, n, \end{aligned}$$

cuya solución es

$$p(x) = 1 + \frac{(x_1 - x) \cdots (x_n - x)}{(1 + x_1) \cdots (1 + x_n)}.$$

Hemos evaluado el polinomio con la fórmula de Newton para diferentes grados y hemos encontrado que los errores son cercanos a 1 para  $n = 54$ . Este grado es mucho mayor que 34, el grado que correspondería a la cota  $3^n$ , ya que los errores de redondeo pueden tener valores menores que la cota y signos distintos que pueden compensarse. Esto explica que la inestabilidad sólo pueda detectarse en algunos casos para grados más altos que los predichos por la cota. La Figura 3.1.2 muestra que el logaritmo decimal del valor absoluto del error de redondeo en la fórmula de Newton crece linealmente. Esto concuerda con el crecimiento exponencial del condicionamiento previsto en el Teorema 3.1.10. El valor máximo del error se alcanza cerca del último nodo,  $x_n = 1$ , de acuerdo con el Teorema 3.1.10.

## 3.2 Comparación de las normalizaciones de la fórmula de Newton en nodos equidistantes ordenados

En esta sección comparamos el condicionamiento  $\kappa_\infty$  de las diferentes normalizaciones  $L$ ,  $\tilde{L}$  y  $\hat{L}$  presentadas en la Sección 2.4 cuando los nodos son equidistantes

$$x_i = a + \frac{i}{n}(b - a), \quad i = 0, \dots, n,$$

en el intervalo  $[a, b]$ . Recordemos las relaciones entre estas matrices. La matriz  $L$  es la matriz de colocación  $M \begin{pmatrix} x_0, \dots, x_n \\ \omega_0, \dots, \omega_n \end{pmatrix}$ . En la fórmula (2.4.7) se vio la relación entre  $L$  y  $\tilde{L}$

$$\tilde{L} = L \operatorname{diag}(1/\omega_0(x_0), \dots, 1/\omega_n(x_n)). \quad (3.2.1)$$

Las matrices  $L$  y  $\hat{L}$  también se relacionan a través de una matriz diagonal. Recordemos (2.4.5)

$$\hat{L} = L \operatorname{diag}(\|d_0\|_\infty, \dots, \|d_n\|_\infty) = LD_d.$$

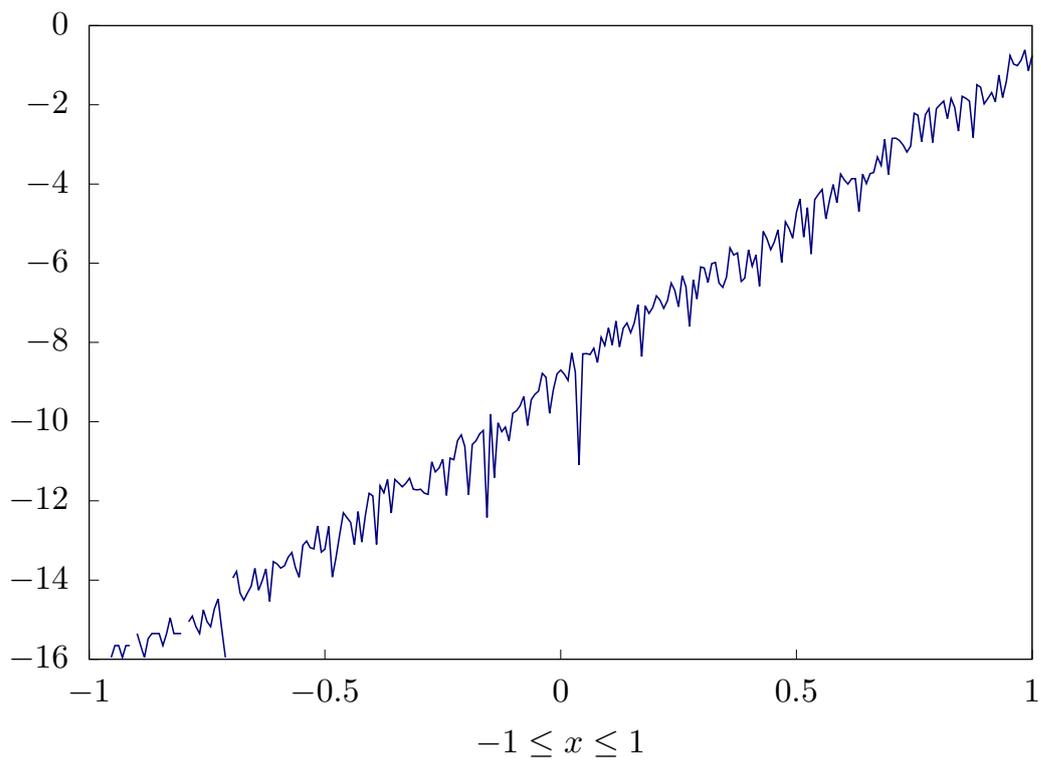


Figura 3.1.2: Logaritmo decimal del valor absoluto del error de redondeo de la fórmula de Newton para puntos equidistantes en  $[-1, 1]$  con  $n = 54$ .

El condicionamiento más bajo se presenta para la matriz de cambio de base de la representación de Newton normalizada  $\hat{L}$ . Por la Proposición 2.4.3 y el Teorema 3.1.4 tenemos

$$\kappa_\infty(\hat{L}) = \text{Cond}(L^{-1}) = 3^n. \quad (3.2.2)$$

El siguiente resultado nos da el valor del condicionamiento de Skeel de las inversas de estas matrices, que coinciden por la relación que hay entre ellas.

**Proposición 3.2.1.** Sean  $x_i = a + \frac{i}{n}(b - a)$ ,  $i = 0, \dots, n$ , nodos equidistantes en el intervalo  $[a, b]$ . Sean las matrices  $L$ ,  $\tilde{L}$  y  $\hat{L}$  dadas por (2.2.6), (3.2.1) y (2.4.5), respectivamente. Entonces

$$\text{Cond}(L^{-1}) = \text{Cond}(\hat{L}^{-1}) = \text{Cond}(\tilde{L}^{-1}) = 3^n$$

*Demostración.* Basta tener en cuenta las relaciones (3.2.1), (2.4.5), la Nota 1.3.7 y la fórmula (3.2.2).  $\square$

Analicemos ahora el comportamiento de la matriz  $\tilde{L}$  asociada a la representación de Newton con diferencias finitas.

**Proposición 3.2.2.** Sean  $x_i = a + \frac{i}{n}(b - a)$ ,  $i = 0, \dots, n$ , nodos equidistantes en el intervalo  $[a, b]$ . Entonces la matriz  $\tilde{L}$  dada por (3.2.1) es la matriz de Pascal triangular inferior  $L_P$  cuyo elemento  $(i, j)$  viene dado por  $\binom{i}{j}$ . Además

$$\|L_P\|_\infty = \|L_P^{-1}\|_\infty = 2^n, \quad \kappa_\infty(L_P) = 4^n, \quad \text{Cond}(L_P) = \text{Cond}(L_P^{-1}) = 3^n.$$

*Demostración.* Denotemos por  $\tilde{l}_{ij}$  y  $\tilde{l}_{ij}^{(-1)}$  los elementos  $(i, j)$  de  $\tilde{L}$  y  $\tilde{L}^{-1}$ , respectivamente.

Usando el Lema 3.1.2

$$\tilde{l}_{ij} = \frac{\omega_j(x_i)}{\omega_j(x_j)} = \frac{i!}{(i-j)!j!} = \binom{i}{j} \quad j = 0, \dots, i, \quad i = 0, \dots, n.$$

Por tanto,  $\tilde{L} = L_P$ .

Ahora vamos a calcular las normas de  $L_P$  y  $L_P^{-1}$ . Para  $L_P$ , tenemos

$$\|L_P\|_\infty = \max_{i=0, \dots, n} \sum_{j=0}^i \binom{i}{j} = \max_{i=0, \dots, n} 2^i = 2^n.$$

Volviendo a usar el Lema 3.1.2 y que  $\tilde{l}_{ij}^{(-1)} = \frac{1}{\omega_i'(x_j)}$  se deduce que el elemento  $(i, j)$  de  $L_P^{-1}$  es

$$\frac{\omega_i'(x_i)}{\omega_i'(x_j)} = (-1)^{i+j} \binom{i}{j}. \quad (3.2.3)$$

En consecuencia,

$$\|L_P^{-1}\|_\infty = \max_{i=0,\dots,n} \sum_{j=0}^i \binom{i}{j} = \max_{i=0,\dots,n} 2^i = 2^n.$$

Así

$$\kappa_\infty(L_P) = \|L_P\|_\infty \|L_P^{-1}\|_\infty = 4^n.$$

Teniendo en cuenta que  $|L_P| = |L_P^{-1}|$  se tiene

$$\text{Cond}(L_P) = \| |L_P^{-1}| |L_P| \|_\infty = \| |L_P| |L_P^{-1}| \|_\infty = \text{Cond}(L_P^{-1}),$$

y, por la Proposición 3.2.1,

$$\text{Cond}(L_P) = \text{Cond}(L_P^{-1}) = 3^n.$$

□

Vamos a analizar el comportamiento de la matriz  $L$ . Notemos que la matriz  $L$  no es invariante por transformaciones afines de los nodos. Esto crea una dificultad adicional, dando lugar a condicionamientos muy altos si la longitud del intervalo es demasiado grande o demasiado pequeña. Para analizar el comportamiento de  $\|L\|_\infty$ , introducimos la función gamma incompleta

$$\Gamma(a, x) := \int_x^{+\infty} t^{a-1} e^{-t} dt.$$

De la definición se desprende que

$$\Gamma(a, x) < \Gamma(a, 0) = \Gamma(a), \quad \forall x > 0.$$

Como es bien conocido, la función gamma incompleta permite obtener una representación integral del polinomio de Taylor de la función exponencial.

**Lema 3.2.3.** *El polinomio de Taylor de grado  $n$  de la función exponencial admite la representación siguiente*

$$\sum_{k=0}^n \frac{x^k}{k!} = e^x \frac{\Gamma(n+1, x)}{\Gamma(n+1)} = \frac{1}{n!} \int_0^{+\infty} e^{x-t} t^n dt.$$

*Demostración.* La función gamma incompleta verifica la siguiente ecuación funcional

$$\Gamma(n+1, x) = e^{-x} x^n + n\Gamma(n, x),$$

que puede demostrarse integrando por partes.

Sumando relaciones del tipo

$$\frac{\Gamma(n+1, x)}{\Gamma(n+1)} - \frac{\Gamma(n, x)}{\Gamma(n)} = \frac{e^{-x}x^n}{n!},$$

obtenemos

$$\frac{\Gamma(n+1, x)}{\Gamma(n+1)} - \frac{\Gamma(1, x)}{\Gamma(1)} = e^{-x} \sum_{k=1}^n \frac{x^k}{k!}.$$

El resultado se deduce teniendo en cuenta que

$$\frac{\Gamma(1, x)}{\Gamma(1)} = e^{-x}.$$

□

Denotamos por  $\ell := b - a$  a la longitud del intervalo. El siguiente resultado nos da una desigualdad para la función gamma incompleta.

**Lema 3.2.4.** *Si  $\ell > 1$ , entonces*

$$\Gamma(n+1, \frac{n}{\ell}) \geq \Gamma(n+1) - e^{-n/\ell} \left(\frac{n}{\ell}\right)^{n+1}$$

y

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+1, \frac{n}{\ell})}{\Gamma(n+1)} = 1.$$

*Demostración.* Como  $\ell > 1$  se tiene

$$0 < \frac{\Gamma(n+1) - \Gamma(n+1, \frac{n}{\ell})}{\Gamma(n+1)} = \int_0^{n/\ell} \frac{t^n}{n!} e^{-t} dt.$$

La función  $f(t) = t^n e^{-t}$  es creciente si  $0 \leq t \leq n$ . Por lo tanto

$$\frac{\Gamma(n+1) - \Gamma(n+1, \frac{n}{\ell})}{\Gamma(n+1)} \leq \int_0^{n/\ell} \left(\frac{n}{\ell}\right)^n \frac{1}{n!} e^{-n/\ell} dt = \left(\frac{n}{\ell}\right)^{n+1} \frac{1}{n!} e^{-n/\ell},$$

de donde se deduce la desigualdad del enunciado

$$\Gamma(n+1, \frac{n}{\ell}) \geq \Gamma(n+1) - e^{-n/\ell} \left(\frac{n}{\ell}\right)^{n+1}.$$

Veamos que si  $n \rightarrow \infty$ , entonces  $\left(\frac{n}{\ell}\right)^{n+1} \frac{1}{n!} e^{-n/\ell} \rightarrow 0$ . Denotemos por

$$c_n := \left(\frac{n}{\ell}\right)^{n+1} \frac{1}{n!} e^{-n/\ell}.$$

Se tiene

$$\begin{aligned} \frac{c_{n+1}}{c_n} &= \left(\frac{n+1}{\ell}\right)^{n+2} \frac{1}{(n+1)!} e^{-(n+1)/\ell} \left(\frac{n}{\ell}\right)^{-(n+1)} n! e^{n/\ell} \\ &= \left(\frac{n+1}{n}\right)^{n+1} \frac{n+1}{\ell} \frac{1}{n+1} e^{-1/\ell} \rightarrow \frac{1}{\ell} e^{1-1/\ell}. \end{aligned}$$

Veamos que  $\frac{1}{\ell} e^{1-1/\ell} < 1$ . Sea  $g(x) = x e^{1-x}$ . Derivando

$$g'(x) = e^{1-x}(1-x) > 0, \quad \text{si } x < 1.$$

Por tanto  $g(x)$  es estrictamente creciente si  $x < 1$ , luego

$$\frac{1}{\ell} e^{1-1/\ell} = g\left(\frac{1}{\ell}\right) < g(1) = 1.$$

Por tanto,  $\lim_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} < 1$  y  $\lim_{n \rightarrow \infty} c_n = 0$ .

□

El siguiente resultado nos proporciona una fórmula compacta para  $\|L\|_\infty$ .

**Proposición 3.2.5.** *Sea  $L$  dada por (2.2.6) en nodos  $x_i = a + \frac{i}{n}\ell$ ,  $i = 0, \dots, n$ , equidistantes en  $[a, b]$  donde  $\ell = b - a$ . Entonces*

$$\|L\|_\infty = n! \left(\frac{n}{\ell}\right)^{-n} \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k = \left(\frac{n}{\ell}\right)^{-n} e^{n/\ell} \Gamma\left(n+1, \frac{n}{\ell}\right).$$

*Demostración.* Por (2.2.7), la norma de  $L$  es

$$\|L\|_\infty = \max_{i=0, \dots, n} \sum_{k=0}^i |\omega_k(x_i)|,$$

pero como

$$|\omega_k(x_i)| \leq |\omega_k(x_n)|, \quad k = 0, \dots, i, \quad i = 0, \dots, n,$$

se tiene que el máximo se alcanza en  $n$ . Así, por el Lema 3.1.2 (ii)

$$\|L\|_\infty = \sum_{k=0}^n |\omega_k(x_n)| = n! \sum_{k=0}^n \frac{1}{(n-k)!} \left(\frac{\ell}{n}\right)^k.$$

Por tanto,

$$\begin{aligned} \|L\|_\infty &= \sum_{k=0}^n \frac{n!}{(n-k)!} \left(\frac{\ell}{n}\right)^k = n! \sum_{k=0}^n \frac{1}{(n-k)!} \left(\frac{\ell}{n}\right)^k = n! \sum_{k=0}^n \frac{1}{k!} \left(\frac{\ell}{n}\right)^{n-k} \\ &= n! \left(\frac{n}{\ell}\right)^{-n} \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k. \end{aligned}$$

Usando el Lema 3.2.3

$$\|L\|_\infty = n! \left(\frac{n}{\ell}\right)^{-n} e^{n/\ell} \frac{\Gamma(n+1, \frac{n}{\ell})}{n!} = \left(\frac{n}{\ell}\right)^{-n} e^{n/\ell} \Gamma\left(n+1, \frac{n}{\ell}\right).$$

□

De este resultado se deduce que la norma infinito de  $L$  para un valor de  $n$  prefijado es una función creciente de la longitud del intervalo.

El cálculo de la norma de  $L^{-1}$  presenta diferentes casos dependiendo de la longitud del intervalo. Al contrario que en el caso de  $\hat{L}$ , la fila en la que la suma de valores absolutos es mayor no es necesariamente la última, a no ser que la longitud del intervalo sea menor o igual que 2. También vemos que si el intervalo es de gran de longitud (mayor o igual que  $2n$ ), la norma de  $\|L^{-1}\|$  es 1, lo que dará lugar a un mal condicionamiento de  $L$ .

**Proposición 3.2.6.** *Sea  $L$  dada por (2.2.6) en nodos equidistantes  $x_i = a + \frac{i}{n}\ell$ ,  $i = 0, \dots, n$ , con  $\ell = b - a$ . Entonces*

$$\|L^{-1}\|_\infty = \begin{cases} \frac{1}{n!} \left(\frac{2n}{\ell}\right)^n, & \text{si } \ell \leq 2, \\ \frac{1}{i_n!} \left(\frac{2n}{\ell}\right)^{i_n}, & \text{si } 2 \leq \ell \leq 2n, \\ 1, & \text{si } \ell \geq 2n. \end{cases} \quad i_n = \lfloor \frac{2n}{\ell} \rfloor,$$

*Demostración.* Por la fórmula (2.2.8) y la Proposición 2.2.2 tenemos que

$$\|L^{-1}\|_\infty = \max_{i=0, \dots, n} \|d_i\|_\infty \quad (3.2.4)$$

y, por el Lema 3.1.2 (iii), se tiene

$$\|d_i\|_\infty = \frac{1}{i!} \left(\frac{2n}{\ell}\right)^i, \quad i = 0, \dots, n.$$

Sea

$$c_i := \frac{1}{i!} \left(\frac{2n}{\ell}\right)^i, \quad i = 0, \dots, n.$$

Veamos dónde alcanza el máximo la sucesión  $\{c_i\}_{i=0, \dots, n}$ . Para ello, consideramos el cociente entre dos elementos consecutivos

$$\frac{c_{i+1}}{c_i} = \frac{2n}{(i+1)\ell}.$$

Si  $\frac{2}{\ell} \geq 1$ , la sucesión es creciente y por tanto

$$\|L^{-1}\|_\infty = \frac{1}{n!} \left(\frac{2n}{\ell}\right)^n.$$

Si  $\frac{2n}{\ell} \leq 1$ , entonces la sucesión es decreciente y

$$\|L^{-1}\|_{\infty} = 1.$$

Por último, si  $1 \leq 2n/\ell \leq n$ , el máximo se alcanza en  $i_n := \lfloor \frac{2n}{\ell} \rfloor$ , es decir,

$$\|L^{-1}\|_{\infty} = \frac{1}{i_n!} \left(\frac{2n}{\ell}\right)^{i_n}, \quad i_n = \lfloor \frac{2n}{\ell} \rfloor.$$

□

Como consecuencia de las proposiciones anteriores obtenemos el siguiente resultado sobre  $\kappa_{\infty}(L)$ .

**Proposición 3.2.7.** *Sea  $L$  dada por (2.2.6) en nodos equidistantes  $x_i = a + \frac{i}{n}\ell$ ,  $i = 0, \dots, n$ , con  $\ell = b - a$ . Entonces*

$$\kappa_{\infty}(L) = \begin{cases} 2^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k, & \text{si } \ell \leq 2, \\ \frac{n!}{i_n!} 2^{i_n} \left(\frac{\ell}{n}\right)^{n-i_n} \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k, & \text{si } 2 \leq \ell \leq 2n, \\ n! \left(\frac{\ell}{n}\right)^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k, & \text{si } \ell \geq 2n. \end{cases} \quad i_n = \lfloor \frac{2n}{\ell} \rfloor,$$

*Demostración.* Basta usar las Proposiciones 3.2.5 y 3.2.6. □

En el caso  $\ell \leq 2$  tenemos que  $\kappa_{\infty}(L)$  es una función decreciente de la longitud del intervalo. Por lo tanto, el mejor condicionamiento posible se obtendrá cuando  $\ell = 2$ . En este caso,

$$\kappa_{\infty}(L) = 2^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{2}\right)^k \leq (2\sqrt{e})^n, \quad b - a = 2.$$

Si  $\ell \geq 2n$ ,

$$\kappa_{\infty}(L) = n! \sum_{k=0}^n \frac{1}{k!} \left(\frac{\ell}{n}\right)^{n-k}$$

es una función creciente de  $\ell$  y el menor condicionamiento lo obtendremos cuando  $\ell = 2n$

$$\kappa_{\infty}(L) = n! 2^n \sum_{k=0}^n \frac{1}{k!} 2^{-k} \geq n! 2^n, \quad \ell = 2n.$$

Como  $n! \geq e^{n/2}$  si  $n \geq 3$ , tenemos  $\kappa_{\infty}(L)$  toma valores mayores si  $b - a \geq 2n$  que en el caso  $\ell = 2$ . Por tanto, podemos afirmar que, para un  $n$  fijo, la longitud que da lugar a un menor condicionamiento de  $L$  verifica  $2 \leq \ell \leq 2n$ .

Vamos a analizar si existe una longitud  $b - a$  tal que el crecimiento de  $\kappa_{\infty}(L)$  sea el menor posible. Demostraremos que en todos los casos  $\lim_{n \rightarrow \infty} \kappa_{\infty}(L)^{1/n}$  es una constante mayor que 1 revelando un crecimiento de carácter exponencial. También demostraremos que la longitud que corresponde a una tasa asintótica de crecimiento mínima es  $b - a = 3$ .

**Teorema 3.2.8.** Sea  $L$  dada por (2.2.6) en nodos equidistantes  $x_i = a + \frac{i}{n}\ell$ ,  $i = 0, \dots, n$ , con  $\ell = b - a$ . Entonces

$$\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} = \begin{cases} \ell e^{3/\ell-1}, & \ell \geq 2, \\ 2e^{1/\ell}, & 1 < \ell \leq 2, \\ \frac{2e}{\ell}, & \ell \leq 1. \end{cases}$$

El menor valor del límite anterior es 3 para  $\ell = 3$  y se tiene

$$\lim_{n \rightarrow \infty} \frac{\kappa_\infty(L)}{3^n} = \sqrt{\frac{3}{2}}, \quad \ell = 3.$$

*Demostración.* Partimos de la expresión de  $\kappa_\infty(L)$  deducida en la Proposición 3.2.7. Si  $\ell \geq 2$  existirá un  $n$  suficientemente grande tal que  $2 \leq \ell \leq 2n$ . Por el Lema 3.2.3 tenemos

$$\kappa_\infty(L) = \ell^n \frac{n! n^{i_n}}{i_n! n^n} \left(\frac{2}{\ell}\right)^{i_n} e^{n/\ell} \frac{\Gamma(n+1, \frac{n}{\ell})}{n!} = \left(\ell e^{1/\ell}\right)^n \left(\frac{2}{\ell}\right)^{i_n} \frac{n! n^{i_n}}{i_n! n^n} \frac{\Gamma(n+1, \frac{n}{\ell})}{n!}.$$

Además,

$$\frac{i_n}{n} = \frac{\lfloor \frac{2n}{\ell} \rfloor}{n} \rightarrow \frac{2}{\ell}, \quad n \rightarrow \infty, \quad (3.2.5)$$

y por tanto,

$$\lim_{n \rightarrow \infty} \left(\frac{n^{i_n}}{i_n!}\right)^{1/n} = \lim_{n \rightarrow \infty} \left(\frac{n^{i_n}}{e^{-i_n} \sqrt{2\pi i_n} i_n^{i_n}}\right)^{1/n} = \lim_{n \rightarrow \infty} \left(\frac{e}{i_n/n}\right)^{i_n/n} = \left(\frac{\ell e}{2}\right)^{2/\ell}.$$

También tenemos

$$\lim_{n \rightarrow \infty} \left(\frac{n!}{n^n}\right)^{1/n} = e^{-1}.$$

Aplicando el Lema 3.2.4 se deduce

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+1, \frac{n}{\ell})}{n!} = 1.$$

Por tanto

$$\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} = \ell e^{1/\ell} \left(\frac{2}{\ell}\right)^{2/\ell} e^{-1} \left(\frac{\ell e}{2}\right)^{2/\ell} = \ell e^{3/\ell-1}.$$

Si  $1 < \ell \leq 2$ , tenemos por el Lema 3.2.3

$$\kappa_\infty(L) = 2^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k = \left(2e^{1/\ell}\right)^n \frac{\Gamma(n+1, \frac{n}{\ell})}{n!}.$$

Por el Lema 3.2.4 se tiene

$$\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} = 2e^{1/\ell}, \quad n \rightarrow \infty.$$

En el caso  $\ell < 1$ , tenemos

$$\begin{aligned}\kappa_\infty(L) &= 2^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k = 2^n \frac{n^n}{n! \ell^n} \left(1 + \frac{n\ell}{n} + \frac{n(n-1)\ell^2}{n^2} + \cdots + \frac{n! \ell^n}{n^n}\right) \\ &\leq 2^n \frac{n^n}{n! \ell^n} (1 + \ell + \cdots + \ell^n) = 2^n \frac{n^n}{n! \ell^n} \frac{1 - \ell^{n+1}}{1 - \ell} \leq 2^n \frac{n^n}{n! \ell^n} \frac{1}{1 - \ell}.\end{aligned}$$

Tomamos ahora el límite cuando  $n \rightarrow \infty$  de la raíz  $n$ -ésima

$$\lim_{n \rightarrow \infty} \sup \kappa_\infty(L)^{1/n} \leq \lim_{n \rightarrow \infty} \left(2^n \frac{n^n}{n! \ell^n} \frac{1}{1 - \ell}\right)^{1/n} = \frac{2}{\ell} \lim_{n \rightarrow \infty} \left(\frac{n^n}{n!}\right)^{1/n} = \frac{2e}{\ell}.$$

Si  $\ell = 1$ ,

$$\kappa_\infty(L) = 2^n \sum_{k=0}^n \frac{n^k}{k!} = 2^n \frac{n^n}{n!} \left(1 + \frac{n}{n} + \frac{n(n-1)}{n^2} + \cdots + \frac{n!}{n^n}\right)$$

y, como cada término de la suma entre paréntesis es menor que 1, tenemos

$$\lim_{n \rightarrow \infty} \sup \kappa_\infty(L)^{1/n} \leq \lim_{n \rightarrow \infty} \left(2^n \frac{n^n(n+1)}{n!}\right)^{1/n} = 2e.$$

Por otro lado, teniendo en cuenta que todos los sumandos son no negativos y tomando el sumando correspondiente a  $k = n$ , obtenemos

$$\kappa_\infty(L) = 2^n \sum_{k=0}^n \frac{1}{k!} \left(\frac{n}{\ell}\right)^k \geq \left(\frac{2n}{\ell}\right)^n \frac{1}{n!}, \quad \ell \leq 1.$$

Así, si  $\ell \leq 1$

$$\lim_{n \rightarrow \infty} \inf \kappa_\infty(L)^{1/n} \geq \lim_{n \rightarrow \infty} \left[\left(\frac{2n}{\ell}\right)^n \frac{1}{n!}\right]^{1/n} = \frac{2}{\ell} \lim_{n \rightarrow \infty} \left(\frac{n^n}{n!}\right)^{1/n} = \frac{2e}{\ell}.$$

Teniendo en cuenta los límites obtenidos anteriormente, tenemos

$$\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} = \frac{2e}{\ell}, \quad \ell \leq 1.$$

Observemos que, si  $\ell \leq 2$ , el  $\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n}$  es una función decreciente de  $\ell$ . Como la función  $\ell e^{3/\ell-1}$  alcanza su mínimo para  $\ell = 3$ , tenemos que el menor crecimiento exponencial se produce para  $\ell = 3$ . En cuyo caso, usando el Lema 3.2.4 se tiene

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{\kappa_\infty(L)}{3^n} &= \lim_{n \rightarrow \infty} \frac{1}{3^n} \frac{3^n n! n^{i_n}}{i_n! n^n} \left(\frac{2}{3}\right)^{i_n} \frac{\Gamma(n+1, \frac{n}{3})}{n!} e^{n/3} \\ &= \lim_{n \rightarrow \infty} \frac{n^n e^{-n} \sqrt{2\pi n n^{i_n}}}{i_n^{i_n} e^{-i_n} \sqrt{2\pi i_n n^n}} \left(\frac{2}{3}\right)^{i_n} \frac{\Gamma(n+1, \frac{n}{3})}{n!} e^{n/3} = \lim_{n \rightarrow \infty} \left(\frac{2n}{3i_n}\right)^{i_n} e^{i_n - 2n/3} \sqrt{\frac{n}{i_n}}.\end{aligned}$$

Por la fórmula (3.2.5) tenemos que  $\frac{i_n}{n} \rightarrow \frac{2}{3}$ , cuando  $n \rightarrow \infty$ . Denotemos por

$$c_n := e^{i_n - 2n/3} \left( \frac{2n}{3i_n} \right)^{i_n}$$

y veamos que  $\lim_{n \rightarrow \infty} c_n = 1$  o, lo que es equivalente,  $\lim_{n \rightarrow \infty} \log c_n = 0$ . Para ello, vamos a utilizar la siguiente propiedad

$$\lim_{x \rightarrow 0} \frac{\log(1-x) + x}{x} = 0.$$

Entonces

$$f(x) := \begin{cases} \frac{\log(1-x) + x}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

es continua en  $x = 0$ . Como  $i_n - 2n/3$  está acotado, se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \log c_n &= \lim_{n \rightarrow \infty} \left( i_n - \frac{2n}{3} \right) + i_n \log \left( \frac{2n}{3i_n} \right) \\ &= \lim_{n \rightarrow \infty} f \left( 1 - \frac{2n}{3i_n} \right) \left( i_n - \frac{2n}{3} \right) = 0. \end{aligned}$$

En consecuencia

$$\lim_{n \rightarrow \infty} \frac{\kappa_\infty(L)}{3^n} = \sqrt{\frac{3}{2}}.$$

□

Analicemos algunos casos importantes del resultado anterior. En el caso  $\ell \leq 1$ , el menor valor del límite  $\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n}$  se alcanza para  $\ell = 1$  y este valor es  $2e \approx 5.4366$ . Si estamos en el caso  $1 \leq \ell \leq 2$ , el Teorema 3.2.8 nos dice que el menor valor se alcanza para  $\ell = 2$  que es  $2\sqrt{e} \approx 3.2974$ . Por último, si  $\ell \geq 2$  se tiene que el mínimo se alcanza cuando  $\ell = 3$  obteniéndose el valor 3.

Vamos a comparar  $\kappa_\infty(L)$  con  $\kappa_\infty(L_P)$  para diferentes longitudes del intervalo. En el caso particular en que  $1/\log 2 \leq \ell \leq 2$ , tenemos

$$\kappa_\infty(L) = \left( 2e^{1/\ell} \right)^n \frac{\Gamma(n+1, \frac{n}{\ell})}{n!} \leq (2e^{1/\ell})^n \leq 4^n = \kappa_\infty(L_P).$$

En el caso  $\ell \geq 2$  tenemos que  $\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} \leq \lim_{n \rightarrow \infty} \kappa_\infty(L_P)^{1/n}$  si y sólo si se verifica

$$\ell e^{3/\ell-1} \leq 4.$$

Las longitudes enteras para las que se verifica esta desigualdad son las comprendidas entre 2 y 7 (ver Figura 3.2.1). Observamos que  $\kappa_\infty(L_P)^{1/n} = 4$  y por tanto, tenemos

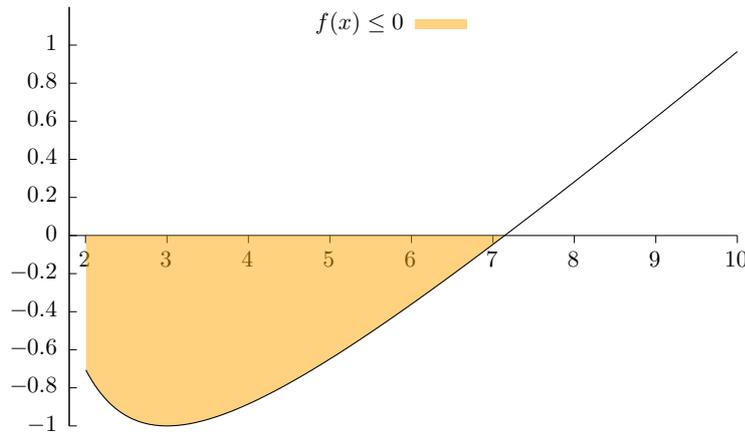


Figura 3.2.1: La zona sombreada corresponde a los valores de  $x \geq 2$  en los que  $f(x) = xe^{3/x-1} - 4 \leq 0$ .

mejores comportamientos asintóticos con  $L$  que con  $L_P$  en intervalos de longitudes 2 y 3. En el caso  $\ell = 1$ , se tiene peor comportamiento asintótico que con  $L_P$  ya que  $\lim_{n \rightarrow \infty} \kappa_\infty(L)^{1/n} = 2e \geq 4 = \kappa_\infty(L_P)^{1/n}$ .

La Tabla 3.2.1 muestra los condicionamientos de las matrices  $L$ ,  $\tilde{L} = L_P$ , y  $\hat{L}$  en nodos equidistantes para distintos valores de  $n$ . Las matrices  $L_P$ , y  $\hat{L}$  no dependen de la longitud del intervalo en el que estén los nodos. Sin embargo,  $L$  sí depende. Hemos analizado  $\kappa_\infty(L)$  en los intervalos  $[0, 1]$ ,  $[-1, 1]$  y  $[0, 3]$ , de longitudes respectivas 1, 2 y 3.

n	$\kappa_\infty(L)$ en $[0, 1]$	$\kappa_\infty(L)$ en $[-1, 1]$	$\kappa_\infty(L)$ en $[0, 3]$	$\kappa_\infty(L_P)$	$\kappa_\infty(\hat{L})$
3	104	33.5	32	64	27
4	549.3333	112	101.2222	256	81
5	$2.9253 \times 10^3$	373.4583	302.7358	$1.0240 \times 10^3$	243
9	$2.4370 \times 10^6$	$4.5301 \times 10^4$	$2.3969 \times 10^4$	$2.6214 \times 10^5$	$1.9683 \times 10^4$
14	$1.124 \times 10^{10}$	$1.7865 \times 10^7$	$5.9094 \times 10^6$	$2.6844 \times 10^8$	$4.7830 \times 10^6$
19	$5.2459 \times 10^{13}$	$6.9906 \times 10^9$	$1.4329 \times 10^9$	$2.7488 \times 10^{11}$	$1.1623 \times 10^9$

Cuadro 3.2.1: Condicionamientos de las matrices  $L$ ,  $L_P$  y  $\hat{L}$  en nodos equidistantes.

En esta tabla podemos ver reflejados los resultados que hemos obtenido en esta sección sobre los condicionamientos de  $L$ ,  $L_P$  y  $\hat{L}$ . Vemos que  $\kappa_\infty(\hat{L}) = 3^n$  y  $\kappa_\infty(L_P) = 4^n$ . También podemos ver que en el intervalo de longitud 3 el condicionamiento de  $L$  es menor que en los otros dos casos estudiados y éste no dista mucho de  $\kappa_\infty(\hat{L}) = \text{Cond}(L^{-1})$ . Otra de las propiedades que hemos analizado también la vemos reflejada en la tabla: el intervalo de longitud 2 da mejores resultados que el intervalo estándar  $[0, 1]$ , de longitud 1.

Hemos mencionado anteriormente que los funcionales  $d_0, \dots, d_n$  son invariantes por traslaciones pero no por transformaciones afines. Para conseguir el mismo efecto de crecimiento asintótico que en los intervalos de longitud 3 introducimos las normalizaciones

$$d_i^{(3)} := \left(\frac{b-a}{3}\right)^i d_i, \quad \omega_i^{(3)}(x) := \left(\frac{3}{b-a}\right)^i \omega_i(x).$$

Entonces la matriz

$$L^{(3)} := L \text{diag} \left(1, \frac{b-a}{3}, \dots, \left(\frac{b-a}{3}\right)^n\right)$$

tiene un condicionamiento próximo al condicionamiento de Skeel, en el sentido de que el cociente  $\kappa_\infty(L^{(3)}) / \text{Cond}(L^{-1})$  está acotado. Por el Teorema 3.2.8, el condicionamiento de  $L^{(3)}$  presenta el siguiente comportamiento asintótico

$$\kappa_\infty(L^{(3)}) \sim \sqrt{\frac{3}{2}} 3^n.$$

### 3.3 Mallas generales con nodos ordenados

En esta sección vamos a realizar un análisis similar al de la Sección 3.1 con nodos  $x_0 < \dots < x_n$  ordenados en un intervalo  $[a, b]$ . Definimos la distancia mínima y máxima entre nodos consecutivos

$$h := \min_{i=0, \dots, n-1} |x_{i+1} - x_i|, \quad r := \max_{i=0, \dots, n-1} |x_{i+1} - x_i|. \quad (3.3.1)$$

Los nodos son casi equidistantes si  $r$  y  $h$  están próximos, situación frecuente en la práctica. Comencemos dando una cota de  $\max_{i=0, \dots, n} \text{cond}(x_i; d) = \text{Cond}(L^{-1})$ , con  $L$  dada por (2.2.6).

**Teorema 3.3.1.** Sean  $x_0 < \dots < x_n$  nodos ordenados en  $[a, b]$  y sean  $h$  y  $r$  dados por (3.3.1). Entonces

$$\text{cond}(x_i; d) \leq \left(1 + \frac{2r}{h}\right)^i.$$

Además,

$$\max_{i=0, \dots, n} \text{cond}(x_i; d) \leq \left(1 + \frac{2r}{h}\right)^n.$$

*Demostración.* Se tiene

$$|\omega_k(x_i)| = \prod_{j=0}^{k-1} |x_i - x_j| \leq r^k \prod_{j=0}^{k-1} (i - j) = r^k \frac{i!}{(i - k)!}$$

y

$$|\omega'_{k+1}(x_j)| = \prod_{\substack{i=0 \\ i \neq j}}^k |x_j - x_i| \geq h^k \prod_{\substack{i=0 \\ i \neq j}}^k |j - i| = h^k j!(k - j)!. \quad (3.3.2)$$

Por tanto

$$\begin{aligned} \text{cond}(x_i; d) &= \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x_i)|}{|\omega'_{k+1}(x_j)|} \leq \sum_{k=0}^i \sum_{j=0}^k \left(\frac{r}{h}\right)^k \frac{i!}{k!(i - k)!} \frac{k!}{j!(k - j)!} \\ &= \sum_{k=0}^i \binom{i}{k} \left(\frac{r}{h}\right)^k \sum_{j=0}^k \binom{k}{j} = \sum_{k=0}^i \binom{i}{k} \left(\frac{2r}{h}\right)^k = \left(1 + \frac{2r}{h}\right)^i. \end{aligned}$$

□

Para acotar  $\text{cond}(x; d)$  para cualquier  $x$  vamos a generalizar el Teorema 3.1.10.

**Teorema 3.3.2.** Sean  $x_0 < \dots < x_n$  en  $[a, b]$  y sean  $h$  y  $r$  dados por (3.3.1). Si  $x \in [x_{n-1}, x_n]$ , se tiene

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^n$$

y si  $x \in [x_{i-1}, x_i]$ ,  $i = 1, \dots, n - 1$ , se tiene

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^i + \frac{1}{4(4r/h - 3)} \left( \frac{2(2r/h)^{n+1}}{n} - \frac{3(2r/h)^{i+1}}{i + 1} \right).$$

En general,

$$\max_{x \in [x_0, x_n]} \text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^n,$$

*Demostración.* Sea  $x \in [x_{i-1}, x_i]$ , con  $i \in \{1, \dots, n\}$ . Se tiene

$$|\omega_k(x)| = \prod_{j=0}^{k-1} |x - x_j| \leq r^k \prod_{j=0}^{k-1} |i - j| = r^k \frac{i!}{(i - k)!}, \quad \text{si } k \leq i,$$

y

$$|\omega_k(x)| = \prod_{j=0}^{k-1} |x - x_j| = |x - x_{i-1}| |x_i - x| \prod_{j=0}^{i-2} |x - x_j| \prod_{j=i+1}^{k-1} |x_j - x| \leq r^k \frac{1}{4} i!(k - i)!, \quad \text{si } k \geq i + 1.$$

Aplicando las desigualdades anteriores al Teorema 2.2.4 y (3.3.2) tenemos

$$\begin{aligned}
\text{cond}(x; d) &= \sum_{k=0}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} = \sum_{k=0}^i \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} + \sum_{k=i+1}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} \\
&\leq \sum_{k=0}^i \sum_{j=0}^k \left(\frac{r}{h}\right)^k \frac{i!}{(i-k)! j!(k-j)!} + \frac{1}{4} \sum_{k=i+1}^n \sum_{j=0}^k \left(\frac{r}{h}\right)^k \frac{i!(k-i)!}{j!(k-j)!} \\
&= \sum_{k=0}^i \sum_{j=0}^k \binom{i}{k} \binom{k}{j} \left(\frac{r}{h}\right)^k + \frac{1}{4} \sum_{k=i+1}^n \sum_{j=0}^k \frac{\binom{k}{j}}{\binom{k}{i}} \left(\frac{r}{h}\right)^k \\
&= \left(1 + \frac{2r}{h}\right)^i + \frac{1}{4} \sum_{k=i+1}^n \frac{1}{\binom{k}{i}} \left(\frac{2r}{h}\right)^k.
\end{aligned}$$

Usando el Lema 3.1.7

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^i + \frac{1}{4} \sum_{k=i+1}^n \frac{1}{k} \left(\frac{2r}{h}\right)^k.$$

Si  $i = n$ , es decir,  $x \in [x_{n-1}, x_n]$ , se tiene

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^n.$$

Si  $x \in [x_{i-1}, x_i]$ ,  $i = 1, \dots, n-1$ , aplicando el Lema 3.1.9 se tiene

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^i + \frac{1}{4(4r/h-3)} \left( \frac{2(2r/h)^{n+1}}{n} - \frac{3(2r/h)^{i+1}}{i+1} \right).$$

Finalmente, aplicando el Lema 3.1.8 deducimos la cota para  $x \in [x_0, x_n]$

$$\text{cond}(x; d) \leq \left(1 + \frac{2r}{h}\right)^n.$$

□

Hasta ahora hemos analizado el caso en el que  $x \in [x_0, x_n]$ . A continuación, vamos a considerar que  $a < x_0 < \dots < x_n < b$  y vamos a estudiar el caso en el que  $x \in [a, x_0]$  ó  $x \in [x_n, b]$ . Para ello veamos el siguiente lema.

**Lema 3.3.3.** Sean  $n, k \in \mathbb{N}$  con  $0 \leq k \leq n$  y  $t \in (0, +\infty)$ . Entonces

$$\frac{\binom{t+n}{k}}{\binom{n}{k}} \leq (t+1)^k$$

y

$$\frac{\binom{t+k-1}{k}}{\binom{n}{k}} \leq (t+1)^k$$

donde  $\binom{t}{k} := \frac{t(t-1)\dots(t-k+1)}{k!}$ .

*Demostración.* Veamos primero

$$\frac{\binom{t+n}{k}}{\binom{n}{k}} = \frac{(t+n)(t+n-1)\cdots(t+n-k+1)}{n(n-1)\cdots(n-k+1)} \leq (t+1)^k.$$

Consideramos cada uno de los factores

$$\frac{t+n-j}{n-j}, \quad j = 0, \dots, k-1 \leq n-1.$$

Como cada factor es menor que  $t+1$

$$\frac{t+n-j}{n-j} = 1 + \frac{t}{n-j} \leq 1+t,$$

tenemos probada la primera desigualdad. Para la segunda desigualdad

$$\frac{\binom{t+k-1}{k}}{\binom{n}{k}} = \frac{t(t+1)\cdots(t+k-1)}{n(n-1)\cdots(n-k+1)} \leq (t+1)^k$$

procedemos análogamente. Se tiene

$$\frac{t+k-1-j}{n-j} \leq \frac{t+n-j}{n-j} = 1 + \frac{t}{n-j} \leq 1+t, \quad j = 0, \dots, k-1 \leq n-1,$$

y se deduce el resultado.  $\square$

**Teorema 3.3.4.** Sean  $a < x_0 < \cdots < x_n < b$  nodos ordenados en  $[a, b]$  y sean  $h$  y  $r$  dados por (3.3.1). Si  $x \in (x_n, b]$  entonces

$$\text{cond}(x; d) \leq \left(1 + \frac{2}{h}(r + x - x_n)\right)^n,$$

y si  $x \in [a, x_0)$  entonces

$$\text{cond}(x; d) \leq \left(1 + \frac{2}{h}(r + x_0 - x)\right)^n.$$

*Demostración.* Sea  $x \in (x_n, b]$ . Entonces podemos escribir  $x = x_n + tr$  con  $t > 0$  y se tiene

$$\begin{aligned} |\omega_k(x)| &= \prod_{j=0}^{k-1} (x - x_j) \leq r^k \prod_{j=0}^{k-1} (t+n-j) \\ &= r^k (t+n)(t+n-1)\cdots(t+n-k+1) = r^k k! \binom{t+n}{k}. \end{aligned}$$

Aplicando las desigualdades anteriores y (3.3.2) al Teorema 2.2.4, se tiene

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} \leq \sum_{k=0}^n \sum_{j=0}^k \left(\frac{r}{h}\right)^k \binom{t+n}{k} \frac{k!}{j!(k-j)!} \\ &= \sum_{k=0}^n \left(\frac{2r}{h}\right)^k \binom{t+n}{k} = \sum_{k=0}^n \left(\frac{2r}{h}\right)^k \binom{n}{k} \frac{\binom{t+n}{k}}{\binom{n}{k}}. \end{aligned}$$

Usando la primera desigualdad del Lema 3.3.3

$$\text{cond}(x; d) \leq \sum_{k=0}^n \left(\frac{2r}{h}(1+t)\right)^k \binom{n}{k} = \left(1 + \frac{2r}{h}(1+t)\right)^n = \left(1 + \frac{2}{h}(r+x-x_n)\right)^n.$$

Si estamos en el caso en el que  $x \in [a, x_0)$ , entonces  $x = x_0 - tr$  con  $t > 0$  y se tiene

$$|\omega_k(x)| = \prod_{j=0}^{k-1} (x_j - x) \leq r^k \prod_{j=0}^{k-1} (j+t) = r^k k! \binom{t+k-1}{k}.$$

Por el Teorema 2.2.4

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \sum_{j=0}^k \frac{|\omega_k(x)|}{|\omega'_{k+1}(x_j)|} \leq \sum_{k=0}^n \sum_{j=0}^k \left(\frac{r}{h}\right)^k \binom{t+k-1}{k} \frac{k!}{j!(k-j)!} \\ &= \sum_{k=0}^n \left(\frac{2r}{h}\right)^k \binom{t+k-1}{k} = \sum_{k=0}^n \left(\frac{2r}{h}\right)^k \binom{n}{k} \frac{\binom{t+k-1}{k}}{\binom{n}{k}}. \end{aligned}$$

Usando ahora la segunda desigualdad del Lema 3.3.3 obtenemos

$$\text{cond}(x; d) \leq \sum_{k=0}^n \left(\frac{2r}{h}(1+t)\right)^k \binom{n}{k} = \left(1 + \frac{2r}{h}(1+t)\right)^n = \left(1 + \frac{2}{h}(r+x_0-x)\right)^n.$$

□

**Corolario 3.3.5.** Sean  $x_0 < \dots < x_n$  en  $[a, b]$  y sean  $h$  y  $r$  dados por (3.3.1) y  $x \in [a, b]$ .

Entonces

$$\text{cond}(x; d) \leq \left(1 + \frac{2R}{h}\right)^n,$$

donde

$$R := \max\{r, r+x_0-a, r+b-x_n\}.$$

### 3.4 Nodos en progresión geométrica

Dado  $q > 0$  y  $k \in \mathbb{N}$  se define el  $q$ -entero  $[k]$  como

$$[k] := \sum_{i=0}^{k-1} q^i = \begin{cases} (1-q^k)/(1-q), & \text{si } q \neq 1, \\ k, & \text{si } q = 1. \end{cases}$$

Por tanto, la sucesión de  $q$ -enteros es

$$0, 1, 1 + q, 1 + q + q^2, 1 + q + q^2 + q^3, \dots$$

Podemos definir también el  $q$ -factorial

$$[k]! := \begin{cases} [k][k-1] \cdots [1], & \text{si } k \geq 1, \\ 1, & \text{si } k = 0. \end{cases}$$

El  $q$ -coeficiente binomial se define como

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{[n][n-1] \cdots [n-k+1]}{[k]!} = \frac{[n]!}{[k]![n-k]!},$$

con  $k, n$  enteros tales que  $n \geq k$ . Considerados como polinomios en  $q$  también son llamados *polinomios Gaussianos* (véase [49]).

La definición de  $q$ -entero puede extenderse a cualquier número real. Dado  $t \in \mathbb{R}$  definimos

$$[t] := \begin{cases} (1 - q^t)/(1 - q), & \text{si } q \neq 1, \\ t, & \text{si } q = 1. \end{cases}$$

De la definición se deduce que  $[t] > 0$  si  $t > 0$  y  $[t] < 0$  si  $t < 0$ . Además, la función  $t \mapsto [t]$  es creciente.

También podemos extender el  $q$ -coeficiente binomial al caso real

$$\begin{bmatrix} t \\ k \end{bmatrix} := \frac{[t][t-1] \cdots [t-k+1]}{[k]!}.$$

Observemos que  $\begin{bmatrix} t \\ k \end{bmatrix} > 0$  si  $t > k - 1$ .

En el siguiente resultado presentamos las relaciones de Pascal (véase (8.7) y (8.8) de [49]) de los  $q$ -coeficientes binomiales.

**Proposición 3.4.1.** *Para cualesquiera  $n, k \in \mathbb{N}$ , con  $n > k$ , y  $q > 0$  se cumple*

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + q^k \begin{bmatrix} n-1 \\ k \end{bmatrix} \tag{3.4.1}$$

y

$$\begin{bmatrix} n \\ k \end{bmatrix} = q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + \begin{bmatrix} n-1 \\ k \end{bmatrix}. \tag{3.4.2}$$

*Demostración.* Veamos la primera fórmula. Se tiene

$$\begin{aligned} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + q^k \begin{bmatrix} n-1 \\ k \end{bmatrix} &= \frac{[n-1]!}{[k-1]![n-k]!} + q^k \frac{[n-1]!}{[k]![n-1-k]!} \\ &= \frac{[n-1]!}{[k]![n-k]!} \left( [k] + q^k [n-k] \right). \end{aligned}$$

Además,

$$[k] + q^k [n-k] = \frac{1-q^k}{1-q} + \frac{q^k(1-q^{n-k})}{1-q} = \frac{1-q^n}{1-q} = [n].$$

Por tanto

$$\begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + q^k \begin{bmatrix} n-1 \\ k \end{bmatrix} = \frac{[n-1]!}{[k]![n-k]!} \left( [k] + q^k [n-k] \right) = \frac{[n-1]!}{[k]![n-k]!} [n] = \begin{bmatrix} n \\ k \end{bmatrix}.$$

Para la segunda fórmula procedemos de manera análoga

$$\begin{aligned} q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + \begin{bmatrix} n-1 \\ k \end{bmatrix} &= q^{n-k} \frac{[n-1]!}{[k-1]![n-k]!} + \frac{[n-1]!}{[k]![n-1-k]!} \\ &= \frac{[n-1]!}{[k]![n-k]!} \left( q^{n-k} [k] + [n-k] \right) \end{aligned}$$

y

$$q^{n-k} [k] + [n-k] = \frac{q^{n-k}(1-q^k)}{1-q} + \frac{1-q^{n-k}}{1-q} = \frac{1-q^n}{1-q} = [n].$$

En consecuencia

$$q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + \begin{bmatrix} n-1 \\ k \end{bmatrix} = \frac{[n-1]!}{[k]![n-k]!} \left( q^{n-k} [k] + [n-k] \right) = \frac{[n-1]!}{[k]![n-k]!} [n] = \begin{bmatrix} n \\ k \end{bmatrix}.$$

□

Necesitamos deducir desigualdades similares a los lemas 3.1.6 y 3.1.7.

**Lema 3.4.2.** Si  $0 < i \leq \frac{n+1}{2}$ , entonces

$$\begin{bmatrix} n \\ i \end{bmatrix} \geq \begin{bmatrix} n \\ i-1 \end{bmatrix},$$

y si  $\frac{n+1}{2} \leq i \leq n$

$$\begin{bmatrix} n \\ i-1 \end{bmatrix} \geq \begin{bmatrix} n \\ i \end{bmatrix}.$$

*Demostración.* Se tiene

$$\frac{\begin{bmatrix} n \\ i \end{bmatrix}}{\begin{bmatrix} n \\ i-1 \end{bmatrix}} = \frac{\begin{bmatrix} n-i+1 \\ i \end{bmatrix}}{\begin{bmatrix} i \\ i \end{bmatrix}} = \frac{1 - q^{n-i+1}}{1 - q^i}.$$

Si  $q = 1$  ya está probado en el Lema 3.1.6. Si  $q > 1$ , entonces  $\begin{bmatrix} n \\ i \end{bmatrix} \geq \begin{bmatrix} n \\ i-1 \end{bmatrix}$  si y sólo si  $q^{n-i+1} \geq q^i$ , lo que equivale a  $\frac{n+1}{2} \geq i$ .

Si  $q < 1$ , entonces  $\begin{bmatrix} n \\ i \end{bmatrix} \geq \begin{bmatrix} n \\ i-1 \end{bmatrix}$  si y sólo si  $q^{n-i+1} \leq q^i$ , es decir,  $\frac{n+1}{2} \geq i$  ya que  $q < 1$ .  $\square$

**Lema 3.4.3.** *Si  $0 < i < n$ , entonces*

$$\begin{bmatrix} n \\ i \end{bmatrix} \geq [n].$$

*Demostración.* Si  $1 \leq i \leq \frac{n+1}{2}$ , por el Lema 3.4.2, se tiene

$$[n] = \begin{bmatrix} n \\ 1 \end{bmatrix} \leq \begin{bmatrix} n \\ 2 \end{bmatrix} \leq \dots \leq \begin{bmatrix} n \\ i \end{bmatrix}.$$

Si  $\frac{n+1}{2} \leq i \leq n-1$ , aplicando el Lema 3.4.2 de nuevo, se deduce

$$\begin{bmatrix} n \\ i \end{bmatrix} \geq \begin{bmatrix} n \\ i+1 \end{bmatrix} \geq \dots \geq \begin{bmatrix} n \\ n-1 \end{bmatrix} = [n]$$

y se tiene el resultado.  $\square$

**Definición 3.4.4.** *Se dice que una sucesión de nodos  $x_0, \dots, x_n$  está en progresión geométrica si es de la forma*

$$x_j = x_0 + [j]h, \quad j = 0, \dots, n,$$

donde  $h \neq 0$ . Si  $x_0 = 1$ ,  $q \neq 1$  y  $h = q - 1$  tenemos el importante caso  $x_i = q^i$ ,  $i = 0, \dots, n$ . El caso  $q = 1$  corresponde a nodos equidistantes ordenados creciente o decrecientemente.

En esta sección vamos a considerar nodos en progresión geométrica. En la Proposición 2.2.5 vimos que el condicionamiento de la fórmula de Newton  $\text{cond}(x; d)$  es invariante por transformaciones afines. Por tanto, el estudio de  $\text{cond}(x; d)$  con nodos en progresión geométrica equivale al caso en que los nodos coinciden con los  $q$ -enteros.

El siguiente lema nos va a permitir calcular la diferencia entre  $q$ -enteros.

**Lema 3.4.5.** *Sean  $t, s \in \mathbb{R}$ . Entonces*

$$[t] - [s] = q^s [t - s].$$

*Demostración.* Se tiene

$$[t] - [s] = \frac{1 - q^t}{1 - q} - \frac{1 - q^s}{1 - q} = \frac{q^s - q^t}{1 - q} = q^s \frac{1 - q^{t-s}}{1 - q} = q^s [t - s].$$

□

Observemos que tomando  $t = 0$  se deduce que

$$[-s] = -q^{-s}[s], \quad s \in \mathbb{R}.$$

**Lema 3.4.6.** Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . Entonces

$$\omega'_{k+1}(x_j) = \prod_{i \in \{0, \dots, k\} \setminus \{j\}} (x_j - x_i) = (-1)^{k-j} h^k q^{j(2k-j-1)/2} [j]! [k-j]!, \quad k = 0, \dots, n.$$

*Demostración.* Aplicando el Lema 3.4.5, la diferencia entre dos nodos es

$$x_j - x_i = ([j] - [i])h = q^i [j - i]h.$$

Entonces tenemos

$$\begin{aligned} \omega'_{k+1}(x_j) &= \prod_{i \in \{0, \dots, k\} \setminus \{j\}} (x_j - x_i) = (-1)^{k-j} \prod_{i=0}^{j-1} (x_j - x_i) \prod_{i=j+1}^k (x_i - x_j) \\ &= (-1)^{k-j} h^k \prod_{i=0}^{j-1} q^i [j - i] \prod_{i=j+1}^k q^j [i - j] \\ &= (-1)^{k-j} h^k q^{j(j-1)/2} q^{j(k-j)} \prod_{i=0}^{j-1} [j - i] \prod_{i=j+1}^k [i - j] \\ &= (-1)^{k-j} h^k q^{j(2k-j-1)/2} [j]! [k-j]!. \end{aligned}$$

□

En el Capítulo 8 de [49] se demuestra una generalización de la fórmula del binomio de Newton con  $q$ -coeficientes binomiales

$$\prod_{j=1}^k (1 + q^{j-1}x) = \sum_{j=0}^k q^{j(j-1)/2} \begin{bmatrix} k \\ j \end{bmatrix} x^j. \quad (3.4.3)$$

Vamos a usar esta fórmula para dar una cota de la función de Lebesgue en el caso de que los nodos sean  $q$ -enteros.

**Teorema 3.4.7.** Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$  y  $x \in [x_i, x_{i+1}]$  para algún  $i \in \{0, \dots, n-1\}$ . Entonces, si  $q \leq 1$ ,

$$\lambda_L(x) \leq q^{-\frac{(n-i)(n-i+1)}{2}} \frac{[n+1]}{[i+1]} \prod_{j=1}^n (1+q^j),$$

y si  $q \geq 1$ , se tiene

$$\lambda_L(x) \leq q^{-\frac{(n-i)(n-i-1)}{2}} \frac{[n+1]}{[i+1]} \prod_{j=1}^n (1+q^j).$$

*Demostración.* Por el Lema 3.4.6, los polinomios de Lagrange verifican

$$|l_j(x_0 + [t]h)| = \prod_{k \in \{0, \dots, n\} \setminus \{j\}} \frac{|[t] - [k]|}{|[j] - [k]|} = \frac{1}{q^{j(2n-j-1)/2} [j]! [n-j]!} \prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t] - [k]|.$$

Usando el Lema 3.4.5, se tiene

$$\begin{aligned} |l_j(x_0 + [t]h)| &= \frac{1}{q^{j(2n-j-1)/2} [j]! [n-j]!} \prod_{k \in \{0, \dots, n\} \setminus \{j\}} q^k |[t-k]| \\ &= \frac{q^{n(n+1)/2-j}}{q^{j(2n-j-1)/2} [j]! [n-j]!} \prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t-k]| \\ &= \frac{q^{(n-j)(n-j+1)/2}}{[j]! [n-j]!} \prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t-k]|. \end{aligned}$$

Si  $x = x_0 + [t]h \in [x_i, x_{i+1}]$ , entonces  $t \in [i, i+1]$ . Vamos a acotar  $\prod_{k \neq j} |[t-k]|$  distinguiendo dos casos. Si  $j \leq i$ , se tiene

$$\prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t-k]| = \prod_{\substack{k=0 \\ k \neq j}}^i [t-k] \prod_{k=i+1}^n |[t-k]| \leq \prod_{\substack{k=0 \\ k \neq j}}^i [i+1-k] \prod_{k=i+1}^n |[i-k]|.$$

Usando el Lema 3.4.5, obtenemos

$$\prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t-k]| \leq \frac{[i+1]!}{[j]} \prod_{k=i+1}^n q^{-(k-i)} [k-i] \leq [i+1]! q^{-(n-i)(n-i+1)/2} [n-i]!,$$

y, por tanto,

$$|l_j(x_0 + [t]h)| \leq q^{\frac{(n-j)(n-j+1) - (n-i)(n-i+1)}{2}} \frac{[i+1]! [n-i]!}{[j]! [n-j]!}, \quad j \leq i. \quad (3.4.4)$$

Para el caso  $j \geq i+1$  procedemos de manera análoga

$$\prod_{k \in \{0, \dots, n\} \setminus \{j\}} |[t-k]| = \prod_{k=0}^i [t-k] \prod_{\substack{k=i+1 \\ k \neq j}}^n |[t-k]| \leq \prod_{k=0}^i [i+1-k] \prod_{\substack{k=i+1 \\ k \neq j}}^n |[i-k]|.$$

Aplicando de nuevo el Lema 3.4.5, se tiene

$$\prod_{k \in \{0, \dots, n\} \setminus \{j\}} | [t - k] | \leq [i + 1]! \prod_{\substack{k=i+1 \\ k \neq j}}^n q^{-(k-i)} [k - i] \leq [i + 1]! q^{-\frac{(n-i)(n-i+1)}{2} + (j-i)} [n - i]!.$$

Así,

$$|l_j(x_0 + [t]h)| \leq q^{\frac{(n-j)(n-j+1) - (n-i)(n-i+1)}{2} + (j-i)} \frac{[i + 1]! [n - i]!}{[j]! [n - j]!}, \quad j \geq i + 1. \quad (3.4.5)$$

Para  $q \leq 1$  tenemos la siguiente cota para todos los polinomios fundamentales de Lagrange

$$|l_j(x_0 + [t]h)| \leq q^{\frac{(n-j)(n-j+1) - (n-i)(n-i+1)}{2}} \frac{[i + 1]! [n - i]!}{[j]! [n - j]!}, \quad j = 0, \dots, n,$$

y, por tanto,

$$\begin{aligned} \lambda_L(x) &= \sum_{j=0}^n |l_j(x_0 + [t]h)| \leq q^{-\frac{(n-i)(n-i+1)}{2}} \frac{[n + 1]}{[i + 1]} \sum_{j=0}^n q^{\frac{(n-j)(n-j+1)}{2}} \begin{bmatrix} n \\ j \end{bmatrix} \\ &= q^{-\frac{(n-i)(n-i+1)}{2}} \frac{[n + 1]}{[i + 1]} \sum_{j=0}^n q^{\frac{j(j-1)}{2}} \begin{bmatrix} n \\ j \end{bmatrix} q^j. \end{aligned}$$

Usando (3.4.3), obtenemos la cota

$$\lambda_L(x) \leq q^{-\frac{(n-i)(n-i+1)}{2}} \frac{[n + 1]}{[i + 1]} \prod_{j=1}^n (1 + q^j).$$

Si  $q \geq 1$ , entonces

$$|l_j(x_0 + [t]h)| \leq q^{\frac{(n-j)(n-j+1) - (n-i)(n-i+1)}{2} + (n-i)} \frac{[i + 1]! [n - i]!}{[j]! [n - j]!}, \quad j = 0, \dots, n.$$

Así,

$$\begin{aligned} \lambda_L(x) &= \sum_{j=0}^n |l_j(x_0 + [t]h)| \leq q^{-\frac{(n-i)(n-i-1)}{2}} \frac{[n + 1]}{[i + 1]} \sum_{j=0}^n q^{\frac{(n-j)(n-j+1)}{2}} \begin{bmatrix} n \\ j \end{bmatrix} \\ &= q^{-\frac{(n-i)(n-i-1)}{2}} \frac{[n + 1]}{[i + 1]} \sum_{j=0}^n q^{\frac{j(j-1)}{2}} \begin{bmatrix} n \\ j \end{bmatrix} q^j. \end{aligned}$$

Aplicando la fórmula (3.4.3), se tiene

$$\lambda_L(x) \leq q^{-\frac{(n-i)(n-i-1)}{2}} \frac{[n + 1]}{[i + 1]} \prod_{j=1}^n (1 + q^j).$$

□

Si  $q = 1$ , el teorema anterior recupera la cota obtenida en la Proposición 3.1.1 teniendo en cuenta que  $n + 1 \leq \binom{n+1}{i+1}$ . El siguiente corolario proporciona una cota uniforme de la constante de Lebesgue para  $q \neq 1$ .

**Corolario 3.4.8.** *Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$  y  $x \in [x_0, x_n]$ . Si  $q \leq 1$*

$$\lambda_L(x) \leq \prod_{j=1}^n (1 + q^{-j}).$$

y si  $q \geq 1$

$$\lambda_L(x) \leq \prod_{j=1}^n (1 + q^j).$$

*Demostración.* Sea  $i \in \{0, \dots, n-1\}$ . Por el Lema 3.4.3, tenemos que  $\frac{\binom{n+1}{i+1}}{\binom{n+1}{i+1}} \leq 1$ . Si  $q \leq 1$  tenemos

$$q^{-\frac{(n-i)(n-i+1)}{2}} \leq q^{-\frac{n(n+1)}{2}},$$

y, por tanto, por el Teorema 3.4.7,

$$\lambda_L(x) \leq q^{-\frac{n(n+1)}{2}} \prod_{j=1}^n (1 + q^j) = \prod_{j=1}^n q^{-j} (1 + q^j) = \prod_{j=1}^n (1 + q^{-j}).$$

Si  $q \geq 1$

$$q^{-\frac{(n-i)(n-i-1)}{2}} \leq 1$$

y, usando el Teorema 3.4.7, tenemos el resultado,

$$\lambda_L(x) \leq \prod_{j=1}^n (1 + q^j).$$

□

A continuación, vamos a dar una fórmula específica para la representación de Newton con nodos en progresión geométrica. Para ello, definimos la *q-diferencia progresiva*

$$\Delta_q f(x_j) := f(x_{j+1}) - f(x_j),$$

que coincide con la diferencia progresiva  $\Delta f(x_j)$ . Por tanto, la diferencia dividida en los nodos  $x_j, x_{j+1}$  es

$$[x_j, x_{j+1}]f = \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j} = \frac{\Delta_q f(x_j)}{q^j h}.$$

La  $q$ -diferencia de orden  $k + 1$  se define como

$$\Delta_q^{k+1} f(x_j) := \Delta_q^k f(x_{j+1}) - q^k \Delta_q^k f(x_j), \quad k \geq 0,$$

partiendo de  $\Delta_q^0 f(x_j) = f(x_j)$ .

En el siguiente resultado se expresa la diferencia dividida en función de las  $q$ -diferencias y los  $q$ -enteros.

**Teorema 3.4.9.** *Dados los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ , se tiene*

$$[x_j, \dots, x_{j+k}]f = \frac{\Delta_q^k f(x_j)}{h^k q^{k(2j+k-1)/2} [k]}, \quad k = 0, \dots, n - j.$$

*Demostración.* Veámoslo por inducción sobre  $k$ . El caso  $k = 0$  y  $0 \leq j \leq n$  ya se ha visto. Supongamos cierto el resultado hasta algún  $k$  y cualquier  $0 \leq j \leq n - k$ . Entonces

$$\begin{aligned} [x_j, \dots, x_{j+k+1}]f &= \frac{[x_{j+1}, \dots, x_{j+k+1}]f - [x_j, \dots, x_{j+k}]f}{x_{j+k+1} - x_j} \\ &= \frac{1}{hq^j [k+1]} \left( \frac{\Delta_q^k f(x_{j+1})}{h^k q^{k(2(j+1)+k-1)/2} [k]} - \frac{\Delta_q^k f(x_j)}{h^k q^{k(2j+k-1)/2} [k]} \right) \\ &= \frac{\Delta_q^k f(x_{j+1}) - q^k \Delta_q^k f(x_j)}{h^{k+1} q^{(k+1)(2j+k)/2} [k+1]} = \frac{\Delta_q^{k+1} f(x_j)}{h^{k+1} q^{(k+1)(2j+k)/2} [k+1]}, \end{aligned}$$

y se cumple el resultado.  $\square$

Del resultado anterior se deduce el siguiente corolario que relaciona las  $q$ -diferencias con las diferencias finitas.

**Corolario 3.4.10.** *Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . El funcional que asigna a cada  $f$  el valor  $\Delta_q^k f(x_j)$  coincide con el funcional diferencia finita definido en (2.4.2) dado por*

$$\begin{aligned} \Delta(x_j, \dots, x_{j+k})f &= (x_{j+k} - x_j) \cdots (x_{j+k} - x_{j+k-1}) [x_j, \dots, x_{j+k}]f \\ &= h^k q^{k(2j+k-1)/2} [k]! [x_j, \dots, x_{j+k}]f. \end{aligned}$$

*Demostración.* Ambos funcionales son múltiplos del funcional  $[x_j, \dots, x_{j+k}]f$ , normalizados para que, al expresarlo en términos de  $f(x_j), \dots, f(x_{j+k})$ , el coeficiente en  $f(x_{j+k})$  sea 1.  $\square$

**Proposición 3.4.11.** *Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . Sea  $\omega_k(x) = (x - x_0) \cdots (x - x_{k-1})$ ,  $k = 1, \dots, n + 1$ . Entonces*

$$\omega_k(x_0 + [t]h) = h^k q^{k(k-1)/2} [k]! \begin{bmatrix} t \\ k \end{bmatrix}.$$

En particular,

$$\omega_k(x_j) = h^k q^{k(k-1)/2} [k]! \begin{bmatrix} j \\ k \end{bmatrix}. \quad (3.4.6)$$

*Demostración.* Usando el Lema 3.4.5, se tiene

$$\begin{aligned} \omega_k(x_0 + [t]h) &= h^k \prod_{j=0}^{k-1} ([t] - [j]) = h^k \prod_{j=0}^{k-1} q^j [t - j] = h^k q^{k(k-1)/2} [t][t-1] \cdots [t-k+1] \\ &= h^k q^{k(k-1)/2} [k]! \begin{bmatrix} t \\ k \end{bmatrix}. \end{aligned}$$

La fórmula (3.4.6), se obtiene tomando  $t = j$ . □

Gracias al Teorema 3.4.9 y la Proposición 3.4.11, tenemos que la representación de Newton en el caso de nodos en progresión geométrica puede expresarse como

$$\begin{aligned} L[f](x_0 + [t]h) &= \sum_{k=0}^n [x_0, \dots, x_k] f \omega_k(x_0 + [t]h) = \sum_{k=0}^n \frac{h^k \Delta_q^k f(x_0)}{h^k q^{k(k-1)/2} [k]!} q^{k(k-1)/2} [k]! \begin{bmatrix} t \\ k \end{bmatrix} \\ &= \sum_{k=0}^n \begin{bmatrix} t \\ k \end{bmatrix} \Delta_q^k f(x_0). \end{aligned} \quad (3.4.7)$$

De la representación de Newton (3.4.7), se deduce que, si  $p$  es el polinomio de grado  $n$  que interpola una sucesión cualquiera  $f_0, \dots, f_n$ , entonces

$$f_n = p(x_0 + [n]h) = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \Delta_q^k f_0. \quad (3.4.8)$$

Vamos a calcular la norma del funcional diferencia dividida  $d_k = [x_0, \dots, x_k]$  en el caso de tener nodos en progresión geométrica.

**Teorema 3.4.12.** *Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . La norma del funcional diferencia dividida es*

$$\|d_k\|_\infty = \frac{1}{h^k [k]!} \prod_{j=0}^{k-1} (1 + q^{-j}), \quad k = 0, \dots, n.$$

*Demostración.* Por la Proposición 2.2.2 y el Lema 3.4.6 se tiene

$$\|d_k\|_\infty = \sum_{j=0}^k \frac{1}{|\omega'_{k+1}(x_j)|} = \sum_{j=0}^k \frac{1}{h^k q^{j(2k-j-1)/2} [j]! [k-j]!} = \frac{1}{h^k [k]!} \sum_{j=0}^k \begin{bmatrix} k \\ j \end{bmatrix} q^{j(j+1-2k)/2}.$$

Usando la fórmula (3.4.3) para  $x = q^{1-k}$ , obtenemos

$$\begin{aligned} \|d_k\|_\infty &= \frac{1}{h^k [k]!} \sum_{j=0}^k \begin{bmatrix} k \\ j \end{bmatrix} q^{j(j-1)/2} q^{j(1-k)} = \frac{1}{h^k [k]!} \prod_{j=1}^k (1 + q^{j-1} q^{1-k}) \\ &= \frac{1}{h^k [k]!} \prod_{j=1}^k (1 + q^{j-k}). \end{aligned}$$

El resultado se deduce teniendo en cuenta que los factores  $1 + q^{j-k}$ ,  $j = 1, \dots, k$ , son los mismos que  $1 + q^{-j}$ ,  $j = 0, \dots, k-1$ , cambiados de orden.  $\square$

Utilizando el Corolario 3.4.10 y el Teorema 3.4.12 vemos que

$$\|\Delta(x_0, \dots, x_k)\|_\infty = q^{k(k-1)/2} \prod_{j=0}^{k-1} (1 + q^{-j}) = \prod_{j=0}^{k-1} (1 + q^j).$$

Podemos dar una expresión de  $\text{cond}(x; d)$  usando el teorema anterior, la Proposición 3.4.11 y la fórmula anterior

$$\begin{aligned} \text{cond}(x_0 + [t]h; d) &= \sum_{k=0}^n \|d_k\|_\infty |\omega_k(x_0 + [t]h)| = \sum_{k=0}^n q^{k(k-1)/2} [k]! \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \frac{1}{[k]!} \prod_{j=1}^k (1 + q^{j-k}) \\ &= \sum_{k=0}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1 + q^j), \end{aligned} \quad (3.4.9)$$

con  $t \in \mathbb{R}$ . Por tanto,

$$\text{cond}(x_i; d) = \sum_{k=0}^i \begin{bmatrix} i \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1 + q^j). \quad (3.4.10)$$

Vamos a considerar la sucesión

$$f_n := \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1 + q^j), \quad n = 0, 1, \dots$$

Comparando con la fórmula (3.4.8), vemos que  $\Delta_q^k f_0$  debe coincidir con  $\prod_{j=0}^{k-1} (1 + q^j)$ . Veamos el siguiente lema para demostrarlo.

**Lema 3.4.13.** *Sea*

$$f_n := \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1 + q^j), \quad n = 0, 1, \dots$$

*Entonces*

$$\Delta_q^l f_n = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} q^{l(n-k)} \prod_{j=0}^{k-1+l} (1 + q^j).$$

*Demostración.* Vamos a demostrarlo por inducción sobre  $l$ . Para  $l = 1$ , por la definición de  $q$ -diferencia, se tiene

$$\begin{aligned} \Delta_q f_n &= f_{n+1} - f_n = \sum_{k=0}^{n+1} \begin{bmatrix} n+1 \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1+q^j) - \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1+q^j) \\ &= \prod_{j=0}^n (1+q^j) + \sum_{k=1}^n \left( \begin{bmatrix} n+1 \\ k \end{bmatrix} - \begin{bmatrix} n \\ k \end{bmatrix} \right) \prod_{j=0}^{k-1} (1+q^j). \end{aligned}$$

Por la fórmula (3.4.2)

$$\Delta_q f_n = \prod_{j=0}^n (1+q^j) + \sum_{k=1}^n \begin{bmatrix} n \\ k-1 \end{bmatrix} q^{n+1-k} \prod_{j=0}^{k-1} (1+q^j) = \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} q^{n-k} \prod_{j=0}^k (1+q^j).$$

Supongamos que se cumple para  $l$  y veámoslo para  $l+1$ .

$$\begin{aligned} \Delta_q^{l+1} f_n &= \Delta_q^l f_{n+1} - q^l \Delta_q^l f_n \\ &= \sum_{k=0}^{n+1} \begin{bmatrix} n+1 \\ k \end{bmatrix} q^{l(n+1-k)} \prod_{j=0}^{k-1+l} (1+q^j) - q^l \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} q^{l(n-k)} \prod_{j=0}^{k-1+l} (1+q^j) \\ &= \prod_{j=0}^{n+l} (1+q^j) + \sum_{k=1}^n \left( \begin{bmatrix} n+1 \\ k \end{bmatrix} - \begin{bmatrix} n \\ k \end{bmatrix} \right) q^{l(n+1-k)} \prod_{j=0}^{k-1+l} (1+q^j). \end{aligned}$$

Aplicando (3.4.2), obtenemos

$$\begin{aligned} \Delta_q^{l+1} f_n &= \prod_{j=0}^{n+l} (1+q^j) + \sum_{k=1}^n \begin{bmatrix} n \\ k-1 \end{bmatrix} q^{(l+1)(n+1-k)} \prod_{j=0}^{k-1+l} (1+q^j) \\ &= \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} q^{(l+1)(n-k)} \prod_{j=0}^{k+l} (1+q^j). \end{aligned}$$

□

En el siguiente teorema vamos a acotar las  $q$ -diferencias progresivas para luego dar una cota superior de  $\max_{i=0, \dots, n} \text{cond}(x_i; d)$ .

**Teorema 3.4.14.** Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, 1, \dots$ . Definimos

$$f_n := \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1+q^j).$$

Definimos también  $f_{n,k} := \Delta_q^k f_n$ . Si  $q > 1$  se tiene

$$f_{j,k-j} \leq f_{0,k-j} (1 + 2q^{k-j}) (1 + 2q^{k-j+1}) \cdots (1 + 2q^{k-1}),$$

y si  $q < 1$  entonces

$$f_{j,k-j} \leq f_{0,k-j} \left(1 + 2q^{k-j}\right) \left(2 + q^{k-j+1}\right)^{j-1}.$$

En el caso particular  $j = k$ , se tiene que, si  $q > 1$ ,

$$f_k \leq \left(1 + 2q^0\right) \left(1 + 2q^1\right) \cdots \left(1 + 2q^{k-1}\right)$$

y si  $q < 1$

$$f_k \leq 3(2 + q)^{k-1}.$$

*Demostración.* Por el Lema 3.4.13, se tiene que  $f_{0,k} = \prod_{j=0}^{k-1} (1 + q^j)$ . Por la definición de  $q$ -diferencia se tiene

$$f_{1,k-1} = f_{0,k} + q^{k-1} f_{0,k-1} = f_{0,k-1} \left(1 + q^{k-1} + qk - 1\right) = f_{0,k-1} \left(1 + 2q^{k-1}\right).$$

Empecemos por el caso  $q > 1$  por inducción sobre  $j$ . Si  $j = 1$  tenemos que se cumple. Supongamos que se cumple para  $j$  y veamos que se verifica para  $j + 1$ :

$$\begin{aligned} f_{j+1,k-j-1} &= f_{j,k-j} + q^{k-j-1} f_{j,k-j-1} \\ &\leq f_{0,k-j-1} \left(1 + q^{k-j-1}\right) \left(1 + 2q^{k-j}\right) \left(1 + 2q^{k-j+1}\right) \cdots \left(1 + 2q^{k-1}\right) + \\ &\quad + q^{k-j-1} f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \cdots \left(1 + 2q^{k-2}\right) \end{aligned}$$

Como  $q > 1$  se tiene

$$\begin{aligned} f_{j+1,k-j-1} &\leq f_{0,k-j-1} \left(1 + q^{k-j-1}\right) \left(1 + 2q^{k-j}\right) \left(1 + 2q^{k-j+1}\right) \cdots \left(1 + 2q^{k-1}\right) + \\ &\quad + q^{k-j-1} f_{0,k-j-1} \left(1 + 2q^{k-j}\right) \cdots \left(1 + 2q^{k-1}\right) \\ &= f_{0,k-j-1} \left(1 + 2q^{k-j}\right) \cdots \left(1 + 2q^{k-1}\right) \left(1 + q^{k-j-1} + q^{k-j-1}\right) \\ &= f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(1 + 2q^{k-j}\right) \cdots \left(1 + 2q^{k-1}\right). \end{aligned}$$

Para el caso  $q < 1$  procedemos del mismo modo. Para  $j = 1$  se verifica la desigualdad. Supongamos cierto hasta  $j$  y veamos que se cumple para  $j + 1$ :

$$\begin{aligned} f_{j+1,k-j-1} &= f_{j,k-j} + q^{k-j-1} f_{j,k-j-1} \\ &\leq f_{0,k-j-1} \left(1 + q^{k-j-1}\right) \left(1 + 2q^{k-j}\right) \left(2 + q^{k-j+1}\right)^{j-1} + \\ &\quad + q^{k-j-1} f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(2 + q^{k-j}\right)^{j-1}. \end{aligned}$$

Teniendo en cuenta que

$$\left(1 + q^{k-j-1}\right) \left(1 + 2q^{k-j}\right) \leq \left(1 + 2q^{k-j-1}\right) \left(1 + q^{k-j}\right),$$

se tiene

$$\begin{aligned}
 f_{j+1,k-j-1} &\leq f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(1 + q^{k-j}\right) \left(2 + q^{k-j+1}\right)^{j-1} + \\
 &\quad + q^{k-j-1} f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(2 + q^{k-j}\right)^{j-1} \\
 &\leq f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(2 + q^{k-j}\right)^{j-1} \left(1 + q^{k-j} + q^{k-j-1}\right) \\
 &\leq f_{0,k-j-1} \left(1 + 2q^{k-j-1}\right) \left(2 + q^{k-j}\right)^j.
 \end{aligned}$$

□

Gracias al siguiente teorema podemos acotar  $\max_{i=0,\dots,n} \text{cond}(x_i; d)$ .

**Teorema 3.4.15.** *Sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . Entonces, si  $q > 1$ , se tiene*

$$\text{cond}(x_i; d) \leq \left(1 + 2q^0\right) \left(1 + 2q^1\right) \cdots \left(1 + 2q^{i-1}\right),$$

y si  $q < 1$

$$\text{cond}(x_i; d) \leq 3(2 + q)^{i-1}.$$

Además

$$\max_{i=0,\dots,n} \text{cond}(x_i; d) = \text{cond}(x_n; d).$$

*Demostración.* Por (3.4.10),  $\text{cond}(x_i; d) = f_i$ . Usando el Teorema 3.4.14, se obtienen las cotas para  $\text{cond}(x_i; d)$ . El máximo valor de  $\text{cond}(x_i; d)$  se alcanza en  $i = n$  ya que la expresión (3.4.10) contiene el mayor número de sumandos y, por la Proposición 3.4.1,  $[i] \leq [n]$ ,  $i = 0, \dots, n$ . □

Vemos que, si  $q = 1$ , el resultado anterior recupera la cota obtenida en el Teorema 3.1.4.

Recordemos que podemos dar una cota uniforme para  $\text{cond}(x; d)$  usando la desigualdad (2.1.6)

$$\text{cond}(x; d) \leq \lambda_L(x) \max_{i=0,\dots,n} \text{cond}(x_i; d).$$

Por tanto, por el Corolario 3.4.8 y el Teorema 3.4.15, en el resto de los puntos tenemos,

$$\text{cond}(x; d) \leq 3(2 + q)^{n-1} \prod_{j=1}^n (1 + q^{-j}), \quad \text{si } q < 1,$$

y

$$\text{cond}(x; d) \leq \prod_{j=1}^n (1 + 2q^{j-1}) \prod_{j=1}^n (1 + q^j), \quad \text{si } q > 1.$$

En el resultado siguiente mejoramos las cotas anteriores.

**Teorema 3.4.16.** Sea  $t \in (i, i + 1)$ ,  $i = 0, \dots, n - 1$  y sean los nodos  $x_j = x_0 + [j]h$ ,  $j = 0, \dots, n$ . Entonces, si  $q > 1$ , se tiene

$$\text{cond}(x_0 + [t]h; d) \leq \text{cond}(x_{i+1}; d) + (n - i - 1) \prod_{j=0}^{n-1} (1 + q^j), \quad (3.4.11)$$

y si  $q < 1$ ,

$$\text{cond}(x_0 + [t]h; d) \leq \text{cond}(x_{i+1}; d) + q^{-(n-(i+1))(n-i)/2} (n - i - 1) \prod_{j=0}^{n-1} (1 + q^j). \quad (3.4.12)$$

*Demostración.* Como  $t \in (i, i + 1)$ , por (3.4.9), se tiene

$$\begin{aligned} \text{cond}(x_0 + [t]h; d) &= \sum_{k=0}^{i+1} \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1 + q^j) + \sum_{k=i+2}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1 + q^j) \\ &\leq \sum_{k=0}^{i+1} \begin{bmatrix} i+1 \\ k \end{bmatrix} \prod_{j=0}^{k-1} (1 + q^j) + \sum_{k=i+2}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1 + q^j) \\ &= \text{cond}(x_{i+1}; d) + \sum_{k=i+2}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1 + q^j). \end{aligned}$$

Tenemos  $t < i + 1 \leq k - 1$ . Así,

$$\begin{aligned} \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| &= \frac{|[t][t-1] \cdots [t-(k-1)]|}{[k]!} = \frac{[t][t-1] \cdots [t-i][t-(i+1)] \cdots [t-(k-1)]}{[k]!} \\ &= q^{-(k-i-1)(k+i-2t)/2} \frac{[t][t-1] \cdots [t-i][i+1-t] \cdots [k-1-t]}{[k]!} \\ &\leq q^{-(k-i-1)(k+i-2t)/2} \frac{[i+1]![k-1-i]!}{[k]!} = \frac{q^{-(k-(i+1))\left(\frac{k-(i+2)}{2}+(i+1-t)\right)}}{[k]_{i+1}}. \end{aligned}$$

Usando el Lema 3.4.3, tenemos

$$\left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \leq \frac{q^{-(k-(i+1))\left(\frac{k-(i+2)}{2}+(i+1-t)\right)}}{[k]}.$$

Observemos que el exponente anterior es siempre negativo. Por tanto, si  $q > 1$ , se tiene

$$q^{-(k-(i+1))\left(\frac{k-(i+2)}{2}+(i+1-t)\right)} < 1$$

y si  $q < 1$ ,

$$q^{-(k-(i+1))\left(\frac{k-(i+2)}{2}+(i+1-t)\right)} < q^{-(n-(i+1))\left(\frac{n-(i+2)}{2}+(i+1-t)\right)} \leq q^{-(n-(i+1))(n-i)/2}.$$

Así, usando las cotas anteriores, para  $q > 1$ , tenemos

$$\begin{aligned} \sum_{k=i+2}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1+q^j) &\leq \sum_{k=i+2}^n \frac{\prod_{j=0}^{k-1} (1+q^j)}{[k]} = \left( \prod_{j=0}^{n-1} (1+q^j) \right) \sum_{k=i+2}^n \frac{1}{[k] \prod_{j=k}^n (1+q^j)} \\ &\leq (n-i-1) \prod_{j=0}^{n-1} (1+q^j). \end{aligned}$$

Por tanto, tenemos (3.4.11).

Para el caso  $q < 1$ , procedemos de manera análoga. Se tiene

$$\begin{aligned} \sum_{k=i+2}^n \left| \begin{bmatrix} t \\ k \end{bmatrix} \right| \prod_{j=0}^{k-1} (1+q^j) &\leq q^{-(n-(i+1))(n-i)/2} \sum_{k=i+2}^n \frac{\prod_{j=0}^{k-1} (1+q^j)}{[k]} \\ &\leq q^{-(n-(i+1))(n-i)/2} (n-i-1) \prod_{j=0}^{n-1} (1+q^j). \end{aligned}$$

Finalmente, obtenemos (3.4.12). □

Recordemos que la matriz  $L$  del cambio de base entre la base de Lagrange y la base de Newton, dada por (2.2.6), coincide con el factor triangular inferior de la factorización  $V = LU$  de la matriz de Vandermonde  $V$ . El factor triangular superior  $U$  corresponde a la matriz dada por (2.6.5). Podemos dar una fórmula explícita para los elementos de estas matrices triangulares cuando los nodos son de la forma  $x_i = [i]$ ,  $i = 0, \dots, n$ . Usando (2.2.7) y (3.4.6) obtenemos que los elementos de la matriz triangular inferior  $L$  son

$$l_{ij} = \omega_j(x_i) = q^{j(j-1)/2} [j]! \begin{bmatrix} i \\ j \end{bmatrix}, \quad j \leq i,$$

y para la matriz triangular superior  $U$ , por (2.6.7), se tiene

$$u_{ij} = \sum_{\alpha_0 + \dots + \alpha_i = j-i} [0]^{\alpha_0} \dots [i]^{\alpha_i}, \quad i \leq j.$$



# Estabilidad de la fórmula de Newton para otras ordenaciones de nodos

## 4

---

### 4.1 Orden de Leja

En [41] Leja definió una sucesión de puntos en un conjunto compacto del plano complejo con notables propiedades para la aproximación de la función de Green con un polo en el infinito. Reichel en [51] muestra que un condicionamiento asociado a la fórmula de Newton del operador de interpolación en los puntos de Leja tiene un crecimiento subexponencial. También sugiere que el orden de Leja, que corresponde a extraer la sucesión de Leja de un conjunto finito, tiene buen comportamiento respecto a la propagación de errores de redondeo (ver Sección 5.3 de [37]). Taylor y Totik demuestran en [61] el crecimiento subexponencial de la constante de Lebesgue en los nodos de Leja.

**Definición 4.1.1.** *Se dice que una sucesión finita  $x_0, \dots, x_n$  está ordenada según el orden de Leja si verifica*

$$|\omega_k(x_k)| \geq |\omega_k(x_i)|, \quad i \geq k, \quad (4.1.1)$$

donde  $\omega_k(x) = \prod_{j=0}^{k-1} (x - x_j)$ .

El orden de Leja se consigue a partir de la siguiente estrategia, como señala Reichel en [51].

**Orden de Leja.** Sea  $X$  un conjunto finito de nodos. Extraemos de  $X$  una sucesión ordenada  $x_0, \dots, x_n$ , de elementos mediante el siguiente criterio

- (i) Inicialmente, se elige  $x_0$  un nodo cualquiera. No obstante, para maximizar  $x_1 - x_0$  en el segundo paso conviene elegir un punto extremo, el mínimo o el máximo de  $X$ .

(ii) En el segundo paso se elige  $x_1$  tal que

$$|x_1 - x_0| = \max_{i=1, \dots, n} |x_i - x_0|.$$

El segundo punto  $x_1$  es el otro extremo (mínimo o máximo) de  $X$ .

(iii) El nodo  $x_j$ ,  $j \geq 2$ , se selecciona entre los nodos de  $X$  no seleccionados previamente tal que

$$\prod_{k=0}^{j-1} |x_j - x_k| = \max_{i=j, \dots, n} \prod_{k=0}^{j-1} |x_i - x_k|.$$

El orden de Leja no es único, pero podemos establecer un criterio para que lo sea (por ejemplo, si dos puntos maximizan ese producto, coger el máximo de los dos).

De (4.1.1) se deduce

$$\text{cond}(x_i; d) = \sum_{k=0}^i \|d_k\|_\infty |\omega_k(x_i)| \leq \sum_{k=0}^i \|d_k\|_\infty |\omega_k(x_k)|.$$

Luego, por el Teorema 2.2.4

$$\text{Cond}(L^{-1}) = \max_{i=0, \dots, n} \text{cond}(x_i; d) \leq \sum_{k=0}^n \|d_k\|_\infty |\omega_k(x_k)|.$$

Recordemos que, por (2.4.2),

$$\Delta(x_0, \dots, x_k) = \omega_k(x_k) d_k$$

y la desigualdad anterior puede expresarse en la forma

$$\text{Cond}(L^{-1}) = \max_{i=0, \dots, n} \text{cond}(x_i; d) \leq \sum_{k=0}^n \|\Delta(x_0, \dots, x_k)\|_\infty.$$

El siguiente teorema nos da una cota para  $\text{Cond}(L^{-1})$  para una distribución arbitraria de nodos.

**Teorema 4.1.2.** *Sean  $x_0, \dots, x_n$  nodos siguiendo el orden de Leja. Sea  $L$  dada por (2.2.6) y  $\tilde{L}$  la matriz asociada a la representación de Newton con diferencias finitas. Entonces, se tienen las siguientes desigualdades*

$$\|\Delta(x_0, \dots, x_k)\|_\infty \leq 2^k,$$

$$\max_{i=0, \dots, n} \text{cond}(x_i; d) = \text{Cond}(L^{-1}) = \text{Cond}(\tilde{L}^{-1}) \leq 2^{n+1} - 1,$$

$$\|\tilde{L}\|_\infty \leq n, \quad \|\tilde{L}^{-1}\|_\infty \leq 2^n$$

y

$$\kappa_\infty(\tilde{L}) \leq n2^n.$$

*Demostración.* Por la Nota 1.3.7 y la fórmula (2.4.7),  $\text{Cond}(\tilde{L}^{-1}) = \text{Cond}(L^{-1})$ . Para cada vector  $s = (s_0, \dots, s_n)^T$  cuyas componentes verifican  $s_i \in \{-1, 0, 1\}$ ,  $i = 0, \dots, n$ , sea  $c$  la solución del sistema  $\tilde{L}c = s$ .

Veamos por inducción que  $|c_k| \leq 2^k$  para  $k = 0, \dots, n$ . Para  $k = 0$  es trivial ya que la diagonal de  $\tilde{L}^{-1}$  tiene unos y por tanto

$$c_0 = s_0 \implies |c_0| \leq 1.$$

Supongamos que es cierto hasta  $k - 1$  y veámoslo para  $k$ . Primero, recordemos que los elementos de  $\tilde{L}$  son (véase (2.4.8))

$$\tilde{l}_{ij} = \frac{\omega_j(x_i)}{\omega_j(x_j)},$$

y, como los nodos siguen el orden de Leja, se tiene

$$|\tilde{l}_{ij}| \leq 1, \quad j \leq i.$$

Como

$$c_k = s_k + \sum_{j=0}^{k-1} \tilde{l}_{kj} c_j,$$

tenemos

$$|c_k| \leq |s_k| + \sum_{j=0}^{k-1} |\tilde{l}_{kj}| |c_j| \leq 1 + (1 + 2 + \dots + 2^{k-1}) = 2^k.$$

Sea  $s = (\text{sign}(\tilde{l}_{kj}^{(-1)}))_{j=0, \dots, n}$ . Entonces la componente  $k$ -ésima del vector  $c := \tilde{L}^{-1}s$  es

$$\left( |\tilde{L}|^{-1} e \right)_k = \sum_{j=0}^k \tilde{l}_{kj}^{(-1)} \text{sign}(\tilde{l}_{kj}^{(-1)}) \leq 2^k. \quad (4.1.2)$$

Teniendo en cuenta (2.4.9) y (2.4.3), se tiene que  $\|\Delta(x_0, \dots, x_k)\|_\infty = |\tilde{L}|^{-1} e \leq 2^k$ . De (4.1.2), se deduce

$$|\tilde{L}| |\tilde{L}|^{-1} e \leq |\tilde{L}| \begin{pmatrix} 1 \\ 2 \\ \vdots \\ 2^n \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1+2 \\ \vdots \\ 1+2+\dots+2^n \end{pmatrix},$$

donde  $e = (1, \dots, 1)^T$ . Por tanto

$$\text{Cond}(L^{-1}) = \|\tilde{L}| |\tilde{L}|^{-1} e\|_\infty \leq 2^{n+1} - 1.$$

Para la segunda parte

$$\|\tilde{L}^{-1}\|_{\infty} = \|\tilde{L}^{-1}e\|_{\infty} \leq 2^n,$$

y

$$\|\tilde{L}\|_{\infty} \leq n$$

porque sus elementos son menores o iguales que 1. Por tanto,

$$\kappa_{\infty}(\tilde{L}) = \|\tilde{L}\|_{\infty}\|\tilde{L}^{-1}\|_{\infty} \leq n2^n$$

□

Comparando el resultado anterior y la Proposición 3.1.3, vemos que el orden de Leja para nodos equidistantes ordenados siempre da funcionales diferencia finita con norma menor o igual que si usáramos el orden creciente.

La siguiente proposición demuestra que los condicionamientos de Skeel y en norma  $\infty$  de  $\tilde{L}^{-1}$  en nodos equidistantes son menores con el orden de Leja que con el orden creciente. Por la Proposición 3.2.2, la matriz  $\tilde{L}$  en nodos equidistantes con orden creciente es la matriz de Pascal triangular inferior (véase el Ejemplo 3 de la Sección 2.5).

**Teorema 4.1.3.** *Para cada sucesión de nodos ordenados, consideramos la matriz  $\tilde{L}$  dada por (2.4.7), asociada a la representación de Newton con diferencias finitas. Se tiene:*

- (i) *El condicionamiento de Skeel de la matriz  $\tilde{L}^{-1}$  con nodos equidistantes con el orden de Leja es menor que con el orden creciente.*
- (ii) *El condicionamiento en norma  $\infty$  de  $\tilde{L}$  con nodos equidistantes siguiendo el orden de Leja es menor que con el orden creciente.*

*Demostración.* (i) Por la Proposición 3.2.1 y el Teorema 4.1.2, tenemos que probar que

$$2^{n+1} - 1 \leq 3^n.$$

Veámoslo por inducción sobre  $n$ . Para  $n = 1$  se cumple

$$2^2 - 1 \leq 3.$$

Supongámoslo cierto para  $n$  y veámoslo para  $n + 1$ . Así

$$2^{n+2} - 1 = 2^{n+1} + 2^{n+1} - 1 \leq 2^{n+1} + 3^n = 2 \cdot 2^n + 3^n \leq 2 \cdot 3^n + 3^n \leq 3^{n+1}.$$

n	Creciente		Leja	
	$\kappa_\infty(\tilde{L})$	$\text{Cond}(\tilde{L}^{-1})$	$\kappa_\infty(\tilde{L})$	$\text{Cond}(\tilde{L}^{-1})$
3	64	27	8.8889	6.3333
4	256	81	16	9.25
5	$1.0240 \times 10^3$	$2.43 \times 10^2$	14.88	10.3
9	$2.6214 \times 10^5$	$1.9683 \times 10^4$	46.1569	23.8833
19	$2.7488 \times 10^{11}$	$1.1622 \times 10^9$	91.9665	50.9247
29	$2.8823 \times 10^{17}$	$6.8630 \times 10^{13}$	$1.3374 \times 10^2$	75.9232
39	$3.0223 \times 10^{23}$	$4.0525 \times 10^{18}$	$3.0334 \times 10^2$	$1.1567 \times 10^2$
49	$3.1691 \times 10^{29}$	$2.393 \times 10^{23}$	$2.9423 \times 10^2$	$1.4789 \times 10^2$

Cuadro 4.1.1:  $\kappa_\infty(\tilde{L})$  y  $\text{Cond}(\tilde{L}^{-1})$  para orden creciente y orden de Leja

(ii) Por la Proposición 3.2.2 y el Teorema 4.1.2 es suficiente probar

$$n2^n \leq 4^n, \quad n = 1, 2, 3, \dots$$

Esto equivale a demostrar

$$n \leq 2^n,$$

y esta última desigualdad se verifica para todo  $n$ .

□

En la Tabla 4.1.1 podemos ver experimentos numéricos que confirman el resultado anterior. Por la Nota 2.4.1 y la fórmula (2.4.7),  $\kappa_\infty(\tilde{L})$  y  $\text{Cond}(\tilde{L}^{-1})$  no dependen del intervalo pero sí del orden de los nodos. También vemos que los condicionamientos de  $\tilde{L}$  para el orden de Leja son considerablemente menores que las cotas teóricas obtenidas:  $n2^n$ , para el condicionamiento en norma  $\infty$  y  $2^{n+1} - 1$ , para el condicionamiento de Skeel. Creemos que este comportamiento guarda relación con el crecimiento subexponencial de un número de condición asociado a la fórmula de Newton con nodos de Leja, ya observado por Reichel [51]. También hemos incluido la Figura 4.1.1 comparando  $\kappa_\infty(\tilde{L})$  con la funciones  $2n \log n$  y  $8n \log n/3$  y la Figura 4.1.2 comparando  $\text{Cond}(\tilde{L}^{-1})$  con la funciones  $n \log n$  y  $4n \log n/5$ .

Hemos experimentado también con la medida puntual del condicionamiento de la representación de Newton. En la Figura 4.1.3 comparamos el orden creciente y el orden

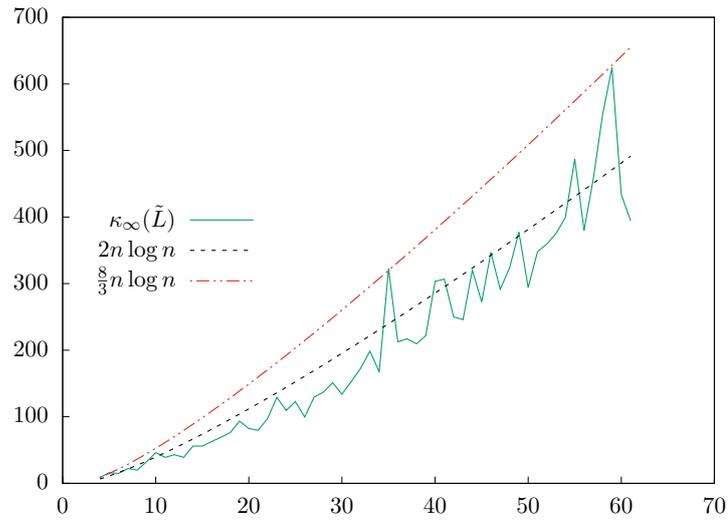


Figura 4.1.1:  $\kappa_\infty(\tilde{L})$  siguiendo el orden de Leja y gráficas de  $2n \log n$  y  $8n \log n/3$

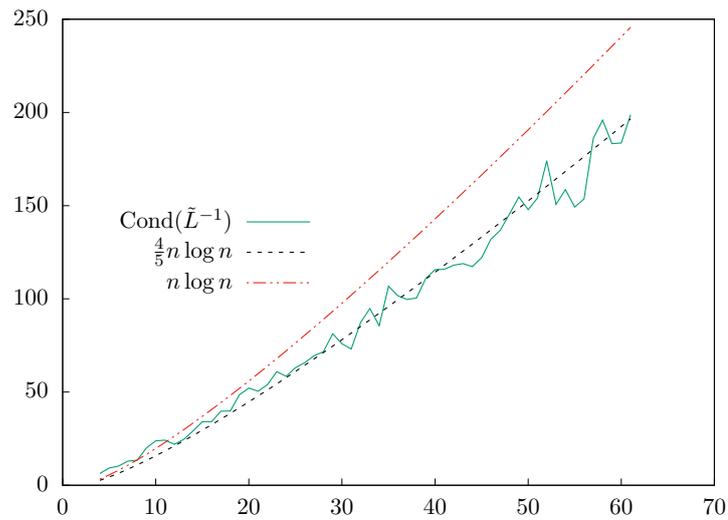


Figura 4.1.2:  $\text{Cond}(\tilde{L}^{-1})$  para nodos siguiendo el orden de Leja y gráficas de  $n \log n$  y  $4n \log n/5$

de Leja. El orden creciente presenta los valores más altos para  $\text{cond}(x; d)$  cerca del extremo derecho del intervalo. Para el orden de Leja, los valores más altos de  $\text{cond}(x; d)$  se obtienen en los subintervalos extremos, al igual que la función de Lebesgue. Además,

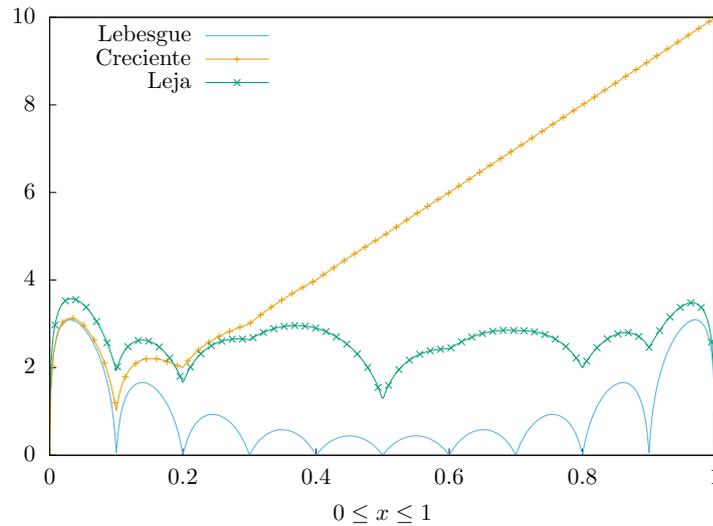


Figura 4.1.3: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; d))$  para orden creciente y para el orden de Leja con  $n = 10$ .

se puede observar que  $\text{cond}(x; d)$  con el orden de Leja recuerda a la forma de la gráfica de la función de Lebesgue.

## 4.2 Orden central

Consideramos otra forma de ordenar los nodos que llamaremos orden central. Con este orden, los nodos están ordenados de menor a mayor según su distancia a un centro  $c$ .

**Orden central.** Dada una sucesión de puntos  $x_0, \dots, x_n$  está ordenada siguiendo un orden central de centro  $c$  si verifica

$$|x_0 - c| \leq |x_1 - c| \leq \dots \leq |x_n - c|.$$

*Nota 4.2.1.* Si ordenamos los nodos siguiendo un orden central, la cápsula convexa  $[x_0, \dots, x_k]$  de la sucesión de nodos  $x_0, \dots, x_k$  es un intervalo que no contiene ningún nodo  $x_j$  con  $j > k$ .

El orden central no es único. La no unicidad puede darse solamente en el caso en que haya nodos menores y mayores que  $c$  situados ambos a la misma distancia. En el paso  $i$ -ésimo, si hay dos puntos con la misma distancia buscamos la elección de  $x_i$  que proporcione menor condicionamiento.

Para una mayor simplicidad, consideraremos el caso especial en que  $c = 0$ . Al final de esta sección ilustraremos la influencia de la elección de diferentes centros.

En el siguiente resultado acotamos un número combinatorio generalizado.

**Proposición 4.2.2.** *Sea  $k$  un entero y  $\binom{x}{k} := x(x-1)\cdots(x-k+1)/k!$ . Entonces*

$$\left| \binom{x}{k} \right| \leq \binom{m}{k}, \quad k-1-m \leq x \leq m,$$

para cualquier  $m \geq k$ .

*Demostración.* Por inducción sobre  $k$ . Para  $k = 0, 1$  es trivial. Supongamos que la desigualdad se verifica para  $k-1$  y veámosla para  $k$ . Si  $x \in [0, k-1]$ ,

$$\left| \binom{x}{k} \right| = \left| \frac{x-k+1}{k} \right| \left| \binom{x}{k-1} \right| \leq \frac{k-1}{k} \binom{m-1}{k-1} \leq \binom{m}{k}.$$

Observemos que  $\binom{x}{k}$  es una función creciente en  $[k-1, +\infty)$ . Así,

$$0 \leq \binom{x}{k} \leq \binom{m}{k}, \quad x \in [k-1, m].$$

Si  $x \leq 0$ , entonces  $\left| \binom{x}{k} \right| = \binom{k-1-x}{k}$  es una función decreciente en  $x$ . Por lo tanto,

$$\left| \binom{x}{k} \right| \leq \binom{m}{k}, \quad x \in [k-1-m, 0].$$

□

A lo largo de esta sección consideraremos sucesiones de nodos tales que los subconjuntos de elementos consecutivos sean equidistantes. En el siguiente resultado obtenemos la norma del funcional diferencia dividida para cualquier ordenación de nodos equidistantes.

**Lema 4.2.3.** *Sean  $x_i = x_0 + ih$ ,  $i = 0, \dots, n$ , nodos equidistantes. Entonces*

$$\| [x_{\sigma(0)}, \dots, x_{\sigma(n)}] \|_{\infty} = \frac{1}{n!} \left( \frac{2}{h} \right)^n,$$

para cualquier permutación  $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$ .

*Demostración.* Por la fórmula (2.2.3), el funcional diferencia dividida  $[x_{\sigma(0)}, \dots, x_{\sigma(n)}]$  coincide con  $[x_0, \dots, x_n]$ , para cualquier permutación  $\sigma$ . Por lo tanto, basta calcular

$\| [x_0, \dots, x_n] \|_\infty$  para nodos equidistantes. Por la Proposición 2.2.2, se tiene

$$\begin{aligned} \| [x_0, \dots, x_n] \|_\infty &= \sum_{k=0}^n \frac{1}{\prod_{j \in \{0, \dots, n\} \setminus \{k\}} |x_k - x_j|} \\ &= \frac{1}{h^n} \sum_{k=0}^n \frac{1}{\prod_{j \in \{0, \dots, n\} \setminus \{k\}} |k - j|} = \frac{1}{n! h^n} \sum_{k=0}^n \binom{n}{k} = \frac{1}{n!} \left( \frac{2}{h} \right)^n. \end{aligned}$$

□

Ahora presentamos cotas superiores para  $\text{cond}(x; d)$  para el orden central.

**Teorema 4.2.4.** *Sea  $x_0, \dots, x_n \in \{-a + 2ai/n; i = 0, \dots, n\}$  una sucesión de nodos que siguen un orden central respecto a  $c = 0$ . Entonces se tiene*

$$\text{cond}(x; d) \leq \frac{(1 + \sqrt{2})^{n+1} + (1 - \sqrt{2})^{n+1}}{2}.$$

*Demostración.* Sea  $h := 2a/n$ . Por el Lema 4.2.3

$$\|d_k\|_\infty = \left( \frac{h}{2} \right)^{-k} \frac{1}{k!}. \quad (4.2.1)$$

Recordemos que  $\omega_k(x) = \prod_{i=0}^{k-1} (x - x_i)$ ,  $k = 1, \dots, n+1$  and  $\omega_0(x) = 1$ . Sea  $n$  par. Si  $k$  es impar,  $k < n$ , los  $k$  primeros nodos son  $jh$ ,  $j = -(k-1)/2, \dots, (k-1)/2$ . Para acotar  $\omega_k$ , aplicamos la Proposición 4.2.2,

$$|\omega_k(x)| = \prod_{j=-(k-1)/2}^{(k-1)/2} |x - jh| = h^k k! \left| \binom{\frac{x}{h} + \frac{k-1}{2}}{k} \right| \leq h^k k! \binom{\frac{n+k-1}{2}}{k},$$

para  $x \in [-a, a]$ . Si  $k$  par, aplicando la cota anterior, obtenemos

$$|\omega_k(x)| = |\omega_{k-1}(x)| \left| x \pm \frac{kh}{2} \right| \leq h^k k! \binom{\frac{n+k}{2}}{k},$$

para  $x \in [-a, a]$ .

Por el Teorema 2.2.4, la fórmula (4.2.1) y lo anterior se tiene

$$\text{cond}(x; d) = \sum_{k=0}^n \|d_k\|_\infty |\omega_k(x)| \leq \sum_{k=0}^n \binom{n/2 + \lfloor \frac{k}{2} \rfloor}{k} 2^k = s_n + 2t_n,$$

donde

$$s_n := \sum_{\substack{k=0 \\ k \text{ par}}}^n \binom{n/2 + \frac{k}{2}}{k} 2^k = \sum_{k=0}^{n/2} \binom{n/2 + k}{2k} 2^{2k}$$

y

$$t_n := \frac{1}{2} \sum_{\substack{k=1 \\ k \text{ impar}}}^{n-1} \binom{n/2 + \frac{k-1}{2}}{k} 2^k = \sum_{k=0}^{n/2-1} \binom{n/2 + k}{2k+1} 2^{2k}.$$

Entonces, se tiene

$$\begin{aligned} s_n + t_n &= \sum_{k=0}^{n/2} \binom{n/2 + k}{2k} 2^{2k} + \sum_{k=0}^{n/2-1} \binom{n/2 + k}{2k+1} 2^{2k} \\ &= \sum_{k=0}^{n/2} \binom{n/2 + k + 1}{2k+1} 2^{2k} = t_{n+2}, \end{aligned}$$

y

$$\begin{aligned} s_n + 4t_{n+2} &= \sum_{k=0}^{n/2} \binom{n/2 + k}{2k} 2^{2k} + 4 \sum_{k=0}^{n/2} \binom{n/2 + k + 1}{2k+1} 2^{2k} = \\ &= \sum_{k=0}^{n/2} \binom{n/2 + k}{2k} 2^{2k} + \sum_{k=1}^{n/2+1} \binom{n/2 + k}{2k-1} 2^{2k} = \sum_{k=0}^{n/2+1} \binom{n/2 + k + 1}{2k} 2^{2k} = s_{n+2}. \end{aligned}$$

Por lo tanto,  $s_{n+2} = 5s_n + 4t_n$ . Así, tenemos que resolver la siguiente ecuación en diferencias

$$\begin{pmatrix} s_{n+2} \\ t_{n+2} \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s_n \\ t_n \end{pmatrix}, \quad \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Los valores propios son  $(1 + \sqrt{2})^2$  y  $(1 - \sqrt{2})^2$  y los vectores propios asociados son respectivamente  $(2(1 + \sqrt{2}), 1)$  y  $(2(1 - \sqrt{2}), 1)$ . Así

$$\begin{pmatrix} 5 & 4 \\ 1 & 1 \end{pmatrix}^{n/2} \begin{pmatrix} 2(1 \pm \sqrt{2}) \\ 1 \end{pmatrix} = (1 \pm \sqrt{2})^n \begin{pmatrix} 2(1 \pm \sqrt{2}) \\ 1 \end{pmatrix}.$$

Además,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{4\sqrt{2}} \left[ \begin{pmatrix} 2(1 + \sqrt{2}) \\ 1 \end{pmatrix} - \begin{pmatrix} 2(1 - \sqrt{2}) \\ 1 \end{pmatrix} \right].$$

Así,

$$\begin{pmatrix} s_n \\ t_n \end{pmatrix} = \frac{1}{4\sqrt{2}} \left[ (1 + \sqrt{2})^n \begin{pmatrix} 2(1 + \sqrt{2}) \\ 1 \end{pmatrix} - (1 - \sqrt{2})^n \begin{pmatrix} 2(1 - \sqrt{2}) \\ 1 \end{pmatrix} \right]. \quad (4.2.2)$$

Por lo tanto,  $\text{cond}(x; d)$  para  $n$  par puede acotarse por el siguiente valor

$$\begin{aligned} s_n + 2t_n &= \frac{(1 + \sqrt{2})^{n+1}}{2\sqrt{2}} - \frac{(1 - \sqrt{2})^{n+1}}{2\sqrt{2}} + \frac{(1 + \sqrt{2})^n}{2\sqrt{2}} - \frac{(1 - \sqrt{2})^n}{2\sqrt{2}} \\ &= \frac{1}{2} (1 + \sqrt{2})^{n+1} + \frac{1}{2} (1 - \sqrt{2})^{n+1}. \end{aligned}$$

Para el caso  $n$  impar, procedemos del mismo modo. Si  $k$  es par,  $k < n$ , los  $k$  primeros nodos son  $(j - 1/2)h$ ,  $j = -k/2 + 1, \dots, k/2$ . Usando la Proposición 4.2.2, se tiene

$$|\omega_k(x)| = \prod_{j=-k/2+1}^{k/2} |x - (j - \frac{1}{2})h| = h^k k! \left| \binom{x/h + \frac{k-1}{2}}{k} \right| \leq h^k k! \binom{\frac{n+k-1}{2}}{k},$$

para todo  $x \in [-a, a]$ . Para  $k$  impar,

$$|\omega_k(x)| = |\omega_{k-1}(x)| |x \pm \frac{kh}{2}| \leq h^k k! \binom{\frac{n+k}{2}}{k}$$

para todo  $x \in [-a, a]$ .

Por el Teorema 2.2.4, la fórmula (4.2.1) y lo anterior, se tiene

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \|d_k\|_{\infty} |\omega_k(x)| \leq \sum_{k=0}^n \left( \frac{n-1}{2} + \left\lfloor \frac{k+1}{2} \right\rfloor \right) 2^k \\ &= \sum_{k=0}^{(n-1)/2} \binom{(n-1)/2 + k}{2k} 2^{2k} + \sum_{k=0}^{(n-1)/2} \binom{(n-1)/2}{2k+1} 2^{2k+1}. \end{aligned}$$

De este modo, el condicionamiento para  $n$  impar puede acotarse por

$$\text{cond}(x; d) \leq s_{n-1} + 2t_{n+1}.$$

Usando (4.2.2), se tiene

$$\begin{aligned} s_{n-1} + 2t_{n+1} &= \frac{(1 + \sqrt{2})^n}{2\sqrt{2}} - \frac{(1 - \sqrt{2})^n}{2\sqrt{2}} + \frac{(1 + \sqrt{2})^{n+1}}{2\sqrt{2}} - \frac{(1 - \sqrt{2})^{n+1}}{2\sqrt{2}} \\ &= \frac{1}{2} (1 + \sqrt{2})^{n+1} + \frac{1}{2} (1 - \sqrt{2})^{n+1}. \end{aligned}$$

□

Esta cota muestra que el crecimiento exponencial de  $\text{cond}(x; d)$  con orden central es menor que la cota  $3^n$  para nodos equidistantes ordenados obtenida en el Teorema 3.1.10.

*Nota 4.2.5.* Recordemos que en el orden central puede ocurrir que dos nodos estén a la misma distancia del centro. En este caso, si elegimos primero el nodo menor, entonces la cota del Teorema 4.2.4 se alcanza en  $x = a$ . Análogamente, la cota se alcanza en  $x = -a$  si elegimos primero el nodo mayor.

Por el Teorema 3.1.10, el teorema anterior muestra que  $C_n := \max_{[-a,a]} \text{cond}(x; d)$  es más pequeño para nodos equidistantes siguiendo cualquier orden central que con el orden creciente. En ambos casos,  $C_n$  tiene un crecimiento exponencial que puede medirse con  $r := \limsup_{n \rightarrow \infty} C_n^{1/n}$ . Para el orden creciente,  $r = 3$ , y para un orden central  $r \leq 1 + \sqrt{2}$ . Por la Proposición 2.1.7, la representación de Lagrange tiene un condicionamiento óptimo y, por la fórmula (3.1.1),  $r = \lim_{n \rightarrow \infty} \Lambda^{1/n} = 2$ .

Propongamos un criterio de elección en el caso de nodos equidistantes. La simetría del problema nos permite ordenar indistintamente el primer par de puntos con la misma distancia al centro. Si tenemos dos puntos con la misma distancia al centro, una elección que parece conveniente consiste en escoger  $x_i$ ,  $i \geq 2$ , tal que  $(x_i - c)(x_{i-1} - c) > 0$ . Tenemos que  $x_{i+1} = 2c - x_i$ . Veamos qué consecuencias tiene esta elección en el cálculo de  $\text{cond}(x_{i+1}; d)$ . En este caso,  $x_{i-1}$ ,  $x_{i-2}$  son un par de nodos simétricos respecto a  $c$  y a menor distancia que  $x_i$ ,  $x_{i+1}$ . Teniendo en cuenta que el polinomio  $\omega_i(x)$  toma en módulo los mismos valores para  $x_i$  y  $x_{i+1}$ , parece pertinente comparar  $\omega_{i-1}(x_i)$  y  $\omega_{i-1}(x_{i+1})$ . Notemos que el polinomio  $\omega_{i-2}(x)$  también verifica  $|\omega_{i-2}(x_i)| = |\omega_{i-2}(x_{i+1})|$ . Se tiene,

$$|\omega_{i-1}(x_j)| = |\omega_{i-2}(x_j)| |x_j - x_{i-2}|, \quad j = i, i+1.$$

Como  $x_i$  y  $x_{i-2}$  verifican  $(x_i - c)(x_{i-2} - c) < 0$ , se deduce que

$$|x_i - x_{i-2}| > |x_{i+1} - x_{i-2}|$$

y

$$|\omega_{i-1}(x_i)| > |\omega_{i-1}(x_{i+1})|.$$

Esta desigualdad implica que  $\|d_{i-1}\|_\infty |\omega_{i-1}(x_{i+1})|$  (un sumando de  $\text{cond}(x_{i+1}; d)$ ) es menor que si hubiéramos intercambiado los papeles de  $x_i$  y  $x_{i+1}$ .

Hemos realizado experimentos numéricos de la cota obtenida en el Teorema 4.2.4 para nodos equidistantes siguiendo un orden central respecto al origen en un intervalo  $[-a, a]$ . En el caso de tener dos nodos con la misma distancia al origen, seguimos el criterio mencionado en el párrafo anterior y elegimos como nodo siguiente el que se encuentra al mismo lado del origen que el nodo anterior. Se tiene además,  $h = 2a/n$ . Distinguiamos dos casos. Si  $n$  es par,  $i = 0, \dots, n$ , se tiene

$$x_i = \begin{cases} -(-1)^{i/2} \frac{ih}{2}, & \text{si } i \text{ es par,} \\ -(-1)^{(i-1)/2} \frac{(i+1)h}{2}, & \text{si } i \text{ es impar} \end{cases} \quad (4.2.3)$$

y si  $n$  es impar,  $i = 0, \dots, n$ ,

$$x_i = \begin{cases} -(-1)^{i/2} \frac{(i+1)h}{2}, & \text{si } i \text{ es par,} \\ (-1)^{(i-1)/2} \frac{ih}{2}, & \text{si } i \text{ es impar} \end{cases} \quad (4.2.4)$$

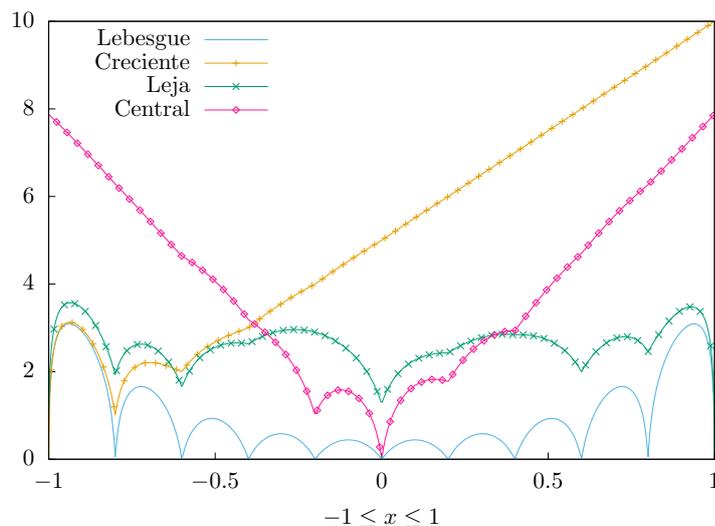


Figura 4.2.1: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; d))$  para puntos en orden creciente, orden de Leja y orden central con  $n = 10$ .

La Figura 4.2.1 compara  $\log_3(\text{cond}(x; d))$  para nodos equidistantes con el orden creciente, el de Leja y el central en  $[-1, 1]$  y  $n = 10$ . El orden creciente presenta los valores más altos para  $\text{cond}(x; d)$  cerca del extremo derecho del intervalo, aunque se comporta mucho mejor en un entorno del extremo izquierdo. El orden central tiene una cota global inferior al orden creciente, como hemos visto en el Teorema 4.2.4, y presenta un comportamiento mucho mejor que los otros dos en un entorno del centro. Por último, el orden de Leja es el que tiene un comportamiento global con cotas uniformes más bajas.

Los experimentos numéricos muestran que, si cambiamos la posición del centro, el orden central es muy próximo a la constante de Lebesgue en un entorno de dicho centro. En la gráficas 4.2.2 y 4.2.3 se comparan  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x; d))$  para puntos equidistantes con  $n = 19$  siguiendo el orden central con centros  $c = 2/3$  y  $c = -2/3$ , respectivamente. Observamos en ambos casos que  $\text{cond}(x; d)$  tiene un comportamiento similar al de la constante de Lebesgue en un entorno del centro y que crece rápidamente cuando nos alejamos de  $c$ .

### 4.3 Orden central respecto al punto de evaluación

En la sección anterior, el centro respecto al que se ordenaba era fijo, sin importar el punto en el que luego se evalúa  $\text{cond}(x; d)$ . Ahora vamos a ordenar los nodos según la

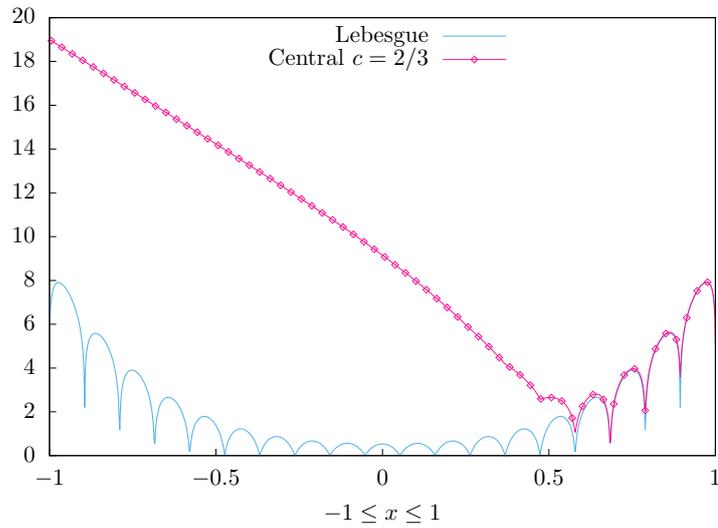


Figura 4.2.2: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; d))$  para puntos equidistantes con orden central para  $c = 2/3$  con  $n = 19$ .

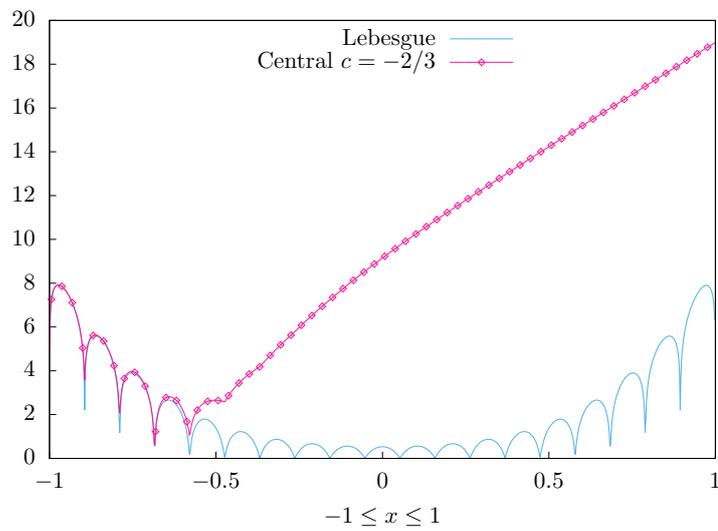


Figura 4.2.3: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; d))$  para puntos equidistantes con orden central para  $c = -2/3$  con  $n = 19$ .

distancia al punto  $x$  en el que se quiere evaluar  $\text{cond}(x; d)$ , es decir,

$$|x - x_0| \leq |x - x_1| \leq \dots \leq |x - x_n|. \quad (4.3.1)$$

El siguiente resultado muestra una ventaja del orden central respecto al punto de evaluación, que tiene repercusión en el condicionamiento.

**Proposición 4.3.1.** *Supongamos que hemos ordenado los nodos  $x_0, \dots, x_n$  con el orden central respecto al punto de evaluación  $x$ , como en (4.3.1). Entonces*

$$|\omega_k(x)| = |(x - x_0) \cdots (x - x_{k-1})| \leq |(x - x_{\sigma(0)}) \cdots (x - x_{\sigma(k-1)})|, \quad k = 0, \dots, n,$$

para cualquier permutación  $\sigma \in \Sigma_{\{0, \dots, n\}}$ .

*Demostración.* Ordenamos cada uno de los factores  $|x - x_{\sigma(i)}|$ ,  $i = 0, \dots, k - 1$ , de menor a mayor y comprobamos que cada uno de esos factores es mayor o igual que el correspondiente  $|x - x_i|$ ,  $i = 0, \dots, k - 1$ .  $\square$

La Proposición 4.3.1 sugiere que este orden da lugar a un buen condicionamiento. De hecho, el condicionamiento de la representación de Newton con nodos siguiendo el orden central respecto el punto de evaluación está más cerca de la función de Lebesgue que con otros órdenes, como el orden de Leja. La Figura 4.3.1 muestra que el orden de Leja está peor condicionado que el orden central respecto al punto de evaluación para nodos de Chebyshev de orden  $n = 10$ .

Una desventaja aparente de esta ordenación es el mayor coste computacional que supone reordenar los nodos al cambiar los puntos de evaluación. Sin embargo, para conjuntos de nodos concretos (por ejemplo, equidistantes) pueden diseñarse estrategias que disminuyan este coste.

*Nota 4.3.2.* El número de órdenes centrales respecto al punto de evaluación está acotado superiormente por  $n(n + 1)/2 + 1$ . Supongamos que los nodos satisfacen  $x_0 \leq x_1 \leq \dots \leq x_n$ , es decir, están ordenados crecientemente. Si  $x < (x_0 + x_1)/2$  entonces el orden central respecto al punto de evaluación coincide con el orden creciente. Sea  $x_{ij} = (x_i + x_j)/2$  el punto de bisección de los nodos  $x_i$  y  $x_j$ . Si el punto de evaluación se mueve de izquierda a derecha, es conveniente ordenar todos los puntos de bisección. Cada vez que  $x$  sobrepase un punto de bisección  $x_{ij}$  es necesario actualizar el orden intercambiando  $x_i$  y  $x_j$ . Si varios puntos de bisección coinciden, entonces debemos intercambiar los correspondientes pares de nodos. Al final, si  $x$  es mayor que el último punto de bisección, entonces el orden central coincide con el orden decreciente. En el caso particular de que los nodos sean equidistantes, sólo hay  $2n$  ordenaciones y  $2n - 1$  puntos de bisección: los puntos medios de cada par de nodos consecutivos y  $x_1, \dots, x_{n-1}$ .

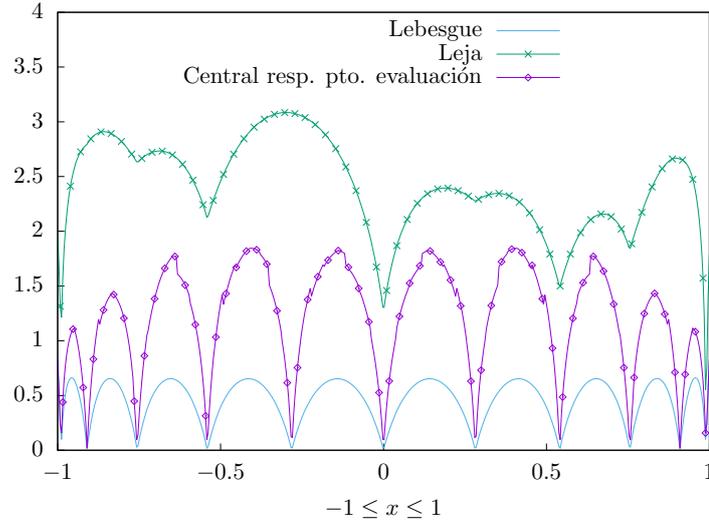


Figura 4.3.1: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x;d))$  para nodos de Chebyshev con orden central respecto al punto de evaluación y orden de Leja para  $n = 10$ .

*Nota 4.3.3.* Veamos la siguiente propiedad de optimalidad del orden central:  $\text{cond}(x;d)$  con nodos  $x_0, \dots, x_n$  siguiendo el orden central respecto al punto de evaluación (4.3.1) es más bajo que con cualquier otra ordenación  $\sigma : \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  que dé lugar a subconjuntos de nodos equidistantes  $\{x_{\sigma(0)}, \dots, x_{\sigma(k)}\}$ , para cualquier  $k \leq n$ . En efecto, aplicando el Lema 4.2.3 obtenemos el resultado

$$\| [x_0, \dots, x_k] \|_\infty = \| [x_{\sigma(0)}, \dots, x_{\sigma(k)}] \|_\infty = \frac{1}{k!} \left( \frac{2}{h} \right)^k.$$

El siguiente resultado nos da una cota superior de  $\text{cond}(x;d)$  cuando los nodos siguen el orden central respecto al punto de evaluación.

**Teorema 4.3.4.** Sean  $x_0, x_1, \dots, x_n \in \{a + (b-a)i/n; i = 0, \dots, n\}$  nodos equidistantes siguiendo el orden central respecto a  $x$  en  $[a, b]$ . Entonces

$$\text{cond}(x;d) \leq \frac{2^n}{n} + \frac{1}{2}.$$

*Demostración.* Sea  $h := (b-a)/n$  la distancia entre los nodos. Los nodos inducen una partición del intervalo  $[a, b]$  y  $x$  se encuentra en uno de esos subintervalos de longitud  $h$ . De la fórmula (4.3.1), se deduce que  $x$  se encuentra entre  $x_0$  y  $x_1$  con  $|x_1 - x_0| = h$ . Se tiene

$$|x - x_0| \leq \frac{h}{2}, \quad |x - x_1| \leq h$$

y

$$|\omega_2(x)| = |x - x_0| |x - x_1| \leq \frac{h^2}{4}.$$

Además,  $x_{i+1}$  dista  $h$  de la cápsula convexa de los nodos  $x_0, \dots, x_i$ . Sea  $j \leq i$  tal que  $x_j$  es un punto extremo de dicha cápsula convexa con  $|x_{i+1} - x_j| = h$ . Por lo tanto,

$$|x - x_{i+1}| \leq |x - x_j| + h \leq |x - x_i| + h.$$

Se deduce que

$$|x - x_i| \leq ih, \quad i \geq 1.$$

Por tanto,

$$|\omega_k(x)| = |(x - x_0) \cdots (x - x_{k-1})| \leq \frac{h^2}{4} \cdot 2h \cdots (k-1)h = \frac{h^k}{4} (k-1)!, \quad k = 2, \dots, n.$$

Por la Nota 4.2.1, la ordenación (4.3.1) siempre da lugar a nodos  $x_0, \dots, x_k$  equidistantes. Por los lemas 4.2.3 y 3.1.9,

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \|d_k\|_{\infty} |\omega_k(x)| = 2 + \sum_{k=2}^n \|d_k\|_{\infty} |\omega_k(x)| \leq 2 + \sum_{k=2}^n \left(\frac{h}{2}\right)^{-k} \frac{1}{k!} \frac{h^k}{4} (k-1)! \\ &= 2 + \frac{1}{4} \sum_{k=2}^n \frac{2^k}{k} \leq 2 + \frac{1}{4} \left( \frac{2^{n+2}}{n} - 3 \frac{2^2}{2} \right) = \frac{2^n}{n} + \frac{1}{2}. \end{aligned}$$

□

*Nota 4.3.5.* Notemos que la demostración del Teorema 4.3.4 permite demostrar la cota  $\text{cond}(x; d) \leq \frac{2^n}{n} + \frac{1}{2}$  para cualquier ordenación de nodos equidistantes que verifique las hipótesis más débiles  $|x - x_0| \leq h/2$  y  $|x - x_i| \leq ih$  para  $i \geq 1$ .

En el Teorema 2 de [62], Trefethen y Weideman demostraron la siguiente cota superior de la constante de Lebesgue

$$\Lambda_L \leq \frac{2^{n+3}}{n}.$$

El Teorema 4.3.4 nos permite mejorar tanto la cota anterior como la Proposición 3.1.1, ofreciendo además de manera simple una cota de la constante de Lebesgue más próxima a (3.1.1), la estimación asintótica obtenida por Schönhage en [58].

**Corolario 4.3.6.** *La función de Lebesgue para nodos equidistantes en  $[a, b]$  verifica*

$$\lambda_L(x) \leq \frac{2^n}{n} + \frac{1}{2}, \quad x \in [a, b].$$

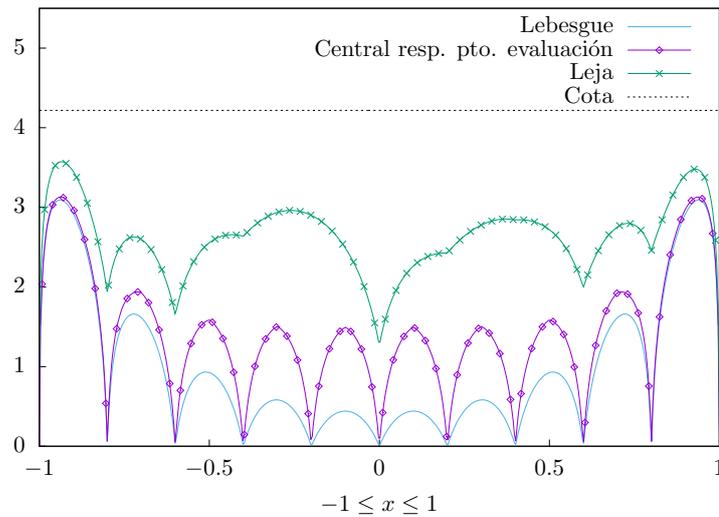


Figura 4.3.2: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; d))$  para puntos equidistantes con orden central respecto al punto de evaluación, para puntos equidistantes con orden de Leja y  $\log_3(2^n/n + 1/2)$  con  $n = 10$ .

*Demostración.* Para cada  $x \in [a, b]$  ordenamos los nodos según (4.3.1). Aplicando la Proposición 2.1.7 y el Teorema 4.3.4, se tiene

$$\lambda_L(x) \leq \text{cond}(x; d) \leq \frac{2^n}{n} + \frac{1}{2}.$$

□

Notemos que siempre se tiene

$$\frac{2^n}{n} + \frac{1}{2} \leq \frac{2^{n+3}}{n}.$$

En la Figura 4.3.2 podemos ver que esta forma de ordenar los nodos da lugar a que  $\text{cond}(x; d)$  se parezca a la función de Lebesgue en los extremos del intervalo mientras que en el centro del intervalo la diferencia es mayor.

## 4.4 Orden central inverso

Consideramos otra forma de ordenar los nodos que hemos llamado orden central inverso. Con este orden los nodos están ordenados de mayor a menor según su distancia a un centro  $c$ .

**Orden central inverso.** Decimos que  $x_0, \dots, x_n$  siguen un orden central inverso respecto de un centro  $c$ , si verifican

$$|x_0 - c| \geq |x_1 - c| \geq \dots \geq |x_n - c|.$$

Notemos que si  $x_0, \dots, x_n$  siguen un orden central inverso, entonces  $x_n, \dots, x_0$  siguen un orden central respecto al mismo centro.

Vamos a estudiar nodos equidistantes con mínima distancia  $h$  entre dos nodos en un intervalo simétrico con centro  $c$ . Sin pérdida de generalidad, podemos considerar  $c = 0$ . Por tanto, el intervalo será de la forma  $[-nh/2, nh/2]$ , donde  $n$  es el grado. En el caso de que tengamos dos nodos con la misma distancia al centro, elegimos como siguiente nodo el menor de los dos. Entonces, podemos escribir

$$x_{i,n} = \begin{cases} \frac{i-n}{2}h, & \text{si } i \equiv 0 \pmod{2}, \\ \frac{n-i+1}{2}h, & \text{si } i \equiv 1 \pmod{2}, \end{cases} \quad i = 0, \dots, n. \quad (4.4.1)$$

Usando (4.4.1) calculamos las diferencias entre dos nodos

$$\begin{aligned} x_{2j,n} - x_{2i,n} &= \left(j - \frac{n}{2}\right)h - \left(i - \frac{n}{2}\right)h = (j - i)h, & i, j = 0, \dots, \lfloor \frac{n}{2} \rfloor, \\ x_{2j+1,n} - x_{2i+1,n} &= \left(\frac{n}{2} - j\right)h - \left(\frac{n}{2} - i\right)h = (i - j)h, & i, j = 0, \dots, \lfloor \frac{n-1}{2} \rfloor, \\ x_{2j+1,n} - x_{2i,n} &= \left(\frac{n}{2} - j\right)h - \left(i - \frac{n}{2}\right)h = (n - j - i)h, & i = 0, \dots, \lfloor \frac{n}{2} \rfloor, \\ & & j = 0, \dots, \lfloor \frac{n-1}{2} \rfloor, \end{aligned} \quad (4.4.2)$$

Notemos que los primeros nodos cambian al variar el grado. Vamos a comparar funcionales diferencia dividida para diferentes grados  $n$ . Denotamos por

$$d_i^n := [x_{0,n}, \dots, x_{i,n}]f, \quad i = 0, \dots, n, \quad (4.4.3)$$

donde  $x_{0,n}, \dots, x_{i,n}$  denotan a los nodos (4.4.1). En los casos en que  $n$  sea fijo utilizaremos la notación simplificada  $x_i := x_{i,n}$ ,  $i = 0, \dots, n$ .

En el siguiente resultado calculamos la norma de los funcionales diferencia dividida cuando los nodos siguen el orden central inverso.

**Proposición 4.4.1.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1) y  $d_i^n$ ,  $i = 0, \dots, n$ , los funcionales diferencia dividida dados por (4.4.3). Se tiene

$$\|d_{2k}^n\|_\infty = \frac{1}{h^{2k}} \left( \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i)\cdots(n-i-k+1)} + \sum_{i=0}^{k-1} \frac{1}{i!(k-i-1)!(n-i)\cdots(n-i-k)} \right),$$

para  $k = 0, \dots, \lfloor \frac{n}{2} \rfloor$  y

$$\|d_{2k+1}^n\|_\infty = \frac{2}{h^{2k+1}} \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i)\cdots(n-i-k)},$$

para  $k = 0, \dots, \lfloor \frac{n-1}{2} \rfloor$ .

*Demostración.* Denotemos por  $x_i := x_{i,n}$ ,  $i = 0, \dots, n$ . Por la Proposición 2.2.2, tenemos

$$\|d_k^n\|_\infty = \sum_{i=0}^k \frac{1}{|\omega'_{k+1}(x_i)|}.$$

Calculamos  $|\omega'_{k+1}(x_j)|$  usando (4.4.2)

$$\begin{aligned} |\omega'_{2k}(x_{2i})| &= \prod_{\substack{j=0 \\ j \neq 2i}}^{2k-1} |x_{2i} - x_j| = h^{2k-1} \prod_{\substack{j=0 \\ j \neq i}}^{k-1} |i - j| \prod_{\substack{j=0 \\ j \neq i}}^{k-1} (n - j - i) \\ &= h^{2k-1} i!(k-1-i)!(n-i)\cdots(n-k-i+1), \end{aligned}$$

$$\begin{aligned} |\omega'_{2k}(x_{2i+1})| &= \prod_{\substack{j=0 \\ j \neq 2i}}^{2k-1} |x_{2i+1} - x_j| = h^{2k-1} \prod_{\substack{j=0 \\ j \neq i}}^{k-1} (n - j - i) \prod_{\substack{j=0 \\ j \neq i}}^{k-1} |i - j| \\ &= h^{2k-1} i!(k-1-i)!(n-i)\cdots(n-k-i+1), \end{aligned}$$

$$\begin{aligned} |\omega'_{2k+1}(x_{2i})| &= \prod_{\substack{j=0 \\ j \neq 2i}}^{2k} |x_{2i} - x_j| = h^{2k} \prod_{\substack{j=0 \\ j \neq i}}^k |i - j| \prod_{\substack{j=0 \\ j \neq i}}^{k-1} (n - j - i) \\ &= h^{2k} i!(k-i)!(n-i)\cdots(n-k-i+1), \end{aligned}$$

$$\begin{aligned} |\omega'_{2k+1}(x_{2i+1})| &= \prod_{\substack{j=0 \\ j \neq 2i}}^{2k} |x_{2i+1} - x_j| = h^{2k} \prod_{\substack{j=0 \\ j \neq i}}^k (n - j - i) \prod_{\substack{j=0 \\ j \neq i}}^{k-1} |i - j| \\ &= h^{2k} i!(k-1-i)!(n-i)\cdots(n-k-i). \end{aligned}$$

Por tanto

$$\begin{aligned} \|d_{2k}^n\|_\infty &= \sum_{i=0}^{2k} \frac{1}{|\omega'_{2k+1}(x_i)|} = \sum_{i=0}^k \frac{1}{|\omega'_{2k+1}(x_{2i})|} + \sum_{i=0}^{k-1} \frac{1}{|\omega'_{2k+1}(x_{2i+1})|} \\ &= \frac{1}{h^{2k}} \left( \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i)\cdots(n-i-k+1)} \right. \\ &\quad \left. + \sum_{i=0}^{k-1} \frac{1}{i!(k-i-1)!(n-i)\cdots(n-i-k)} \right). \end{aligned}$$

Análogamente

$$\begin{aligned} \|d_{2k+1}^n\|_\infty &= \sum_{i=0}^{2k+1} \frac{1}{|\omega'_{2k+2}(x_i)|} = \sum_{i=0}^k \frac{1}{|\omega'_{2k+2}(x_{2i})|} + \sum_{i=0}^k \frac{1}{|\omega'_{2k+2}(x_{2i+1})|} \\ &= \frac{2}{h^{2k+1}} \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i)\cdots(n-i-k)}. \end{aligned}$$

□

Para poder obtener una expresión explícita de la norma de los funcionales diferencia dividida, obtenemos una fórmula de recurrencia para estas normas.

**Proposición 4.4.2.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1) y  $d_i^m$ ,  $i = 0, \dots, n$ , los funcionales diferencia dividida dados por (4.4.3). Se tiene

$$\|d_{2k}^n\|_\infty = \frac{1}{kh} \|d_{2k-1}^{m-1}\|_\infty, \quad k = 1, \dots, \lfloor \frac{n}{2} \rfloor,$$

y

$$\|d_{2k+1}^n\|_\infty = \frac{2}{nh} \|d_{2k}^{m-1}\|_\infty, \quad k = 0, \dots, \lfloor \frac{n-1}{2} \rfloor.$$

*Demostración.* Por la Proposición 4.4.1, se tiene

$$k!h^{2k} \|d_{2k}^n\|_\infty = \sum_{i=0}^k \binom{k}{i} \frac{1}{(n-i)\cdots(n-i-k+1)} + \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{k}{(n-i)\cdots(n-i-k)}.$$

Usando la identidad de Pascal,  $\binom{k}{i} = \binom{k-1}{i} + \binom{k-1}{i-1}$ , obtenemos

$$\begin{aligned} k!h^{2k} \|d_{2k}^n\|_\infty &= \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{1}{(n-i)\cdots(n-i-k+1)} \\ &\quad + \sum_{i=1}^k \binom{k-1}{i-1} \frac{1}{(n-i)\cdots(n-i-k+1)} \\ &\quad + \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{k}{(n-i)\cdots(n-i-k)}. \end{aligned}$$

Haciendo un cambio de índice en la segunda suma, tenemos la siguiente expresión

$$\begin{aligned} k!h^{2k}\|d_{2k}^m\|_\infty &= \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{1}{(n-i)\cdots(n-i-k+1)} \\ &+ \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{1}{(n-i-1)\cdots(n-i-k)} \\ &+ \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{k}{(n-i)\cdots(n-i-k)}. \end{aligned}$$

Por tanto,

$$\begin{aligned} k!h^{2k}\|d_{2k}^m\|_\infty &= \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{n-i-k+n-i+k}{(n-i)\cdots(n-i-k)} \\ &= 2 \sum_{i=0}^{k-1} \frac{(k-1)!}{i!(k-1-i)!(n-1-i)\cdots(n-i-k)} = (k-1)!h^{2k-1}\|d_{2k-1}^{m-1}\|_\infty. \end{aligned}$$

Usamos la Proposición 4.4.1 para la norma de los funcionales diferencia dividida de orden impar

$$\begin{aligned} n \frac{h^{2k+1}}{2} \|d_{2k+1}^m\|_\infty &= \sum_{i=0}^k \frac{n}{i!(k-i)!(n-i)\cdots(n-i-k)} \\ &= \sum_{i=0}^k \frac{n-i+i}{i!(k-i)!(n-i)\cdots(n-i-k)} \\ &= \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i-1)\cdots(n-i-k)} \\ &+ \sum_{i=1}^k \frac{1}{(i-1)!(k-i)!(n-i)\cdots(n-i-k)}. \end{aligned}$$

Haciendo un cambio de índices en la segunda suma, se tiene

$$\begin{aligned} n \frac{h^{2k+1}}{2} \|d_{2k+1}^m\|_\infty &= \sum_{i=0}^k \frac{1}{i!(k-i)!(n-i-1)\cdots(n-i-k)} \\ &+ \sum_{i=0}^{k-1} \frac{1}{i!(k-i-1)!(n-1-i)\cdots(n-i-k-1)} \\ &= h^{2k}\|d_{2k}^{m-1}\|_\infty \end{aligned}$$

y se sigue el resultado. □

Ahora ya podemos dar una expresión explícita de las normas de los funcionales diferencia dividida.

**Teorema 4.4.3.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1) y  $d_i^n$ ,  $i = 0, \dots, n$ , los funcionales diferencia dividida dados por (4.4.3). Se tiene

$$\|d_{2k}^n\|_\infty = \frac{2^k}{h^{2k} k! (n-1)(n-3) \cdots (n-2k+1)}, \quad k = 1, \dots, \lfloor \frac{n}{2} \rfloor, \quad (4.4.4)$$

y

$$\|d_{2k+1}^n\|_\infty = \frac{2^{k+1}}{h^{2k+1} k! n(n-2) \cdots (n-2k)}, \quad k = 0, \dots, \lfloor \frac{n-1}{2} \rfloor. \quad (4.4.5)$$

*Demostración.* Veamos que se cumple por inducción sobre  $k$ . Para  $d_1^n$  se tiene

$$\|d_1^n\|_\infty = \frac{1}{|\omega_2'(x_{0,n})|} + \frac{1}{|\omega_2'(x_{1,n})|} = \frac{2}{|x_{0,n} - x_{1,n}|} = \frac{2}{nh},$$

y por tanto, se verifica (4.4.5) para  $k = 0$ . Para  $d_2^n$ , tenemos

$$\begin{aligned} \|d_2^n\|_\infty &= \frac{1}{|\omega_3'(x_{0,n})|} + \frac{1}{|\omega_3'(x_{1,n})|} + \frac{1}{|\omega_3'(x_{2,n})|} \\ &= \frac{1}{|x_{0,n} - x_{1,n}| |x_{0,n} - x_{2,n}|} + \frac{1}{|x_{1,n} - x_{0,n}| |x_{1,n} - x_{2,n}|} \\ &\quad + \frac{1}{|x_{2,n} - x_{0,n}| |x_{2,n} - x_{1,n}|} = \frac{1}{nh^2} + \frac{1}{n(n-1)h^2} + \frac{1}{(n-1)h^2} = \frac{2}{(n-1)h^2} \end{aligned}$$

y se verifica (4.4.4) para  $k = 1$ . Supongamos que se verifican (4.4.4) y (4.4.5) hasta  $k$  y veamos que se cumplen para  $k + 1$ . Por la Proposición 4.4.2 y la hipótesis de inducción, se tiene

$$\begin{aligned} \|d_{2(k+1)}^n\|_\infty &= \frac{1}{(k+1)h} \|d_{2k+1}^{n-1}\|_\infty = \frac{1}{(k+1)h} \frac{2^{k+1}}{h^{2k+1} k! (n-1)(n-3) \cdots (n-2k-1)} \\ &= \frac{2^{k+1}}{h^{2(k+1)} (k+1)! (n-1)(n-3) \cdots (n-2k-1)} \end{aligned}$$

y se cumple (4.4.4). Veamos que se verifica también (4.4.5) para  $k + 1$ . Por la hipótesis de inducción y la Proposición 4.4.2

$$\begin{aligned} \|d_{2(k+1)+1}^n\|_\infty &= \frac{2}{nh} \|d_{2(k+1)}^{n-1}\|_\infty = \frac{2}{nh} \frac{2^{k+1}}{h^{2(k+1)} (k+1)! (n-2)(n-4) \cdots (n-2(k+1))} \\ &= \frac{2^{(k+1)+1}}{h^{2(k+1)+1} (k+1)! n(n-2) \cdots (n-2(k+1))} \end{aligned}$$

y se sigue el resultado.  $\square$

En el siguiente resultado obtenemos cotas uniformes de la base de Newton correspondiente a nodos en orden central inverso.

**Proposición 4.4.4.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1) y  $x \in [-nh/2, nh/2]$ . Entonces

$$|\omega_0(x)| = 1, \quad |\omega_1(x)| \leq nh,$$

$$|\omega_{2k}(x)| \leq \frac{h^{2k}}{4} n^2 (n-1)^2 \cdots (n-k+1)^2, \quad k = 1, \dots, \lfloor \frac{n}{2} \rfloor, \quad (4.4.6)$$

y

$$|\omega_{2k+1}(x)| \leq \frac{h^{2k+1}}{4} n^2 (n-1)^2 \cdots (n-k+1)^2 (n-k), \quad k = 1, \dots, \lfloor \frac{n-1}{2} \rfloor, \quad (4.4.7)$$

*Demostración.* Denotemos por  $x_i := x_{i,n}$ ,  $i = 0, \dots, n$ . Se tiene  $|\omega_0(x)| = 1$  y

$$|\omega_1(x)| = |x - x_0| = |x + \frac{nh}{2}| \leq nh.$$

Probemos las otras dos desigualdades por inducción sobre  $k$ . Para  $k = 1$ , tenemos

$$|\omega_2(x)| = |x - x_0| |x - x_1| = |x + \frac{nh}{2}| |x - \frac{nh}{2}| \leq \frac{h^2}{4} n^2$$

y se cumple (4.4.6). Veamos (4.4.7) para  $k = 1$ . Usando la cota anterior se tiene

$$|\omega_3(x)| = |x - x_0| |x - x_1| |x - x_2| \leq \frac{h^2}{4} n^2 |x + \frac{h(n-2)}{2}| \leq \frac{h^3}{4} n^2 (n-1).$$

Supongamos que se verifican (4.4.6) y (4.4.7) para  $k$  y veamos que se cumplen para  $k+1$ . Usando la hipótesis de inducción se tiene

$$\begin{aligned} |\omega_{2(k+1)}(x)| &= |\omega_{2k+1}(x)| |x - x_{2k+1}| \\ &\leq \frac{h^{2k+1}}{4} \left( n^2 (n-1)^2 \cdots (n-k+1)^2 (n-k) \right) |x - \left( \frac{n}{2} - k \right) h| \\ &\leq \frac{h^{2k+1}}{4} n^2 (n-1)^2 \cdots (n-k+1)^2 (n-k) h (n-k) \end{aligned}$$

y, por tanto, se verifica (4.4.6). Para finalizar veamos que se cumple (4.4.7). Por la hipótesis de inducción

$$\begin{aligned} |\omega_{2(k+1)+1}(x)| &= |\omega_{2(k+1)}(x)| |x - x_{2k+2}| \\ &\leq \frac{h^{2(k+1)}}{4} \left( n^2 (n-1)^2 \cdots (n-k)^2 \right) |x + \left( \frac{n}{2} - (k+1) \right) h| \\ &\leq \frac{h^{2(k+1)}}{4} \left( n^2 (n-1)^2 \cdots (n-k)^2 \right) h (n-k-1) \end{aligned}$$

y se cumple el resultado. □

Antes de calcular una cota para  $\text{cond}(x; d)$  veamos el siguiente resultado auxiliar conocido como identidad de Vandermonde.

**Lema 4.4.5.** Sean  $r, q, i \in \mathbb{N} \cup \{0\}$ . Entonces

$$\binom{r+q}{k} = \sum_{i=0}^k \binom{r}{i} \binom{q}{k-i}.$$

*Demostración.* Se tiene

$$(1+x)^{r+q} = \sum_{k=0}^{r+q} \binom{r+q}{k} x^k.$$

Por otro lado

$$(1+x)^{r+q} = (1+x)^r (1+x)^q = \sum_{i=0}^r \binom{r}{i} x^i \sum_{j=0}^q \binom{q}{j} x^j = \sum_{k=0}^{r+q} \left( \sum_{i=0}^k \binom{r}{i} \binom{q}{k-i} \right) x^k.$$

Comparando los coeficientes de las potencias se tiene el resultado

$$\binom{r+q}{k} = \sum_{i=0}^k \binom{r}{i} \binom{q}{k-i}.$$

□

Veamos ahora una cota para  $\text{cond}(x; d)$  evaluado en los nodos.

**Teorema 4.4.6.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1). Entonces

$$\text{cond}(x_{2i,n}; d) \leq \binom{n+1}{i} + n, \quad i = 0, \dots, \lfloor \frac{n}{2} \rfloor, \quad (4.4.8)$$

y

$$\text{cond}(x_{2i+1,n}; d) \leq \binom{n+2}{i+1} - 1, \quad i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor. \quad (4.4.9)$$

*Demostración.* Empecemos acotando  $\text{cond}(x_{2i}; d)$ ,  $i = 0, \dots, \lfloor \frac{n}{2} \rfloor$ . Teniendo en cuenta que  $\omega_k(x_i) = 0$  si  $k > i$ , solamente necesitamos evaluar la base de Newton en los nodos con índices posteriores. Primero, calculemos  $\omega_{2k}$ ,  $k = 0, \dots, \lfloor \frac{n}{2} \rfloor$ . Por (4.4.2)

$$\begin{aligned} |\omega_{2k}(x_{2i})| &= \prod_{j=0}^{2k-1} |x_{2i} - x_j| = \prod_{j=0}^{k-1} |x_{2i} - x_{2j}| \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\ &= h^{2k} \prod_{j=0}^{k-1} |i-j| \prod_{j=0}^{k-1} (n-i-j) = h^{2k} \frac{i!(n-i)!}{(i-k)!(n-i-k)!}, \quad k \leq i, \end{aligned} \quad (4.4.10)$$

$$\begin{aligned}
|\omega_{2k+1}(x_{2i})| &= \prod_{j=0}^{2k} |x_{2i} - x_j| = \prod_{j=0}^k |x_{2i} - x_{2j}| \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\
&= h^{2k+1} \prod_{j=0}^k |i - j| \prod_{j=0}^{k-1} (n - i - j) \\
&= h^{2k+1} \frac{i!(n-i)!}{(i-k-1)!(n-i-k)!}, \quad k \leq i-1. \tag{4.4.11}
\end{aligned}$$

Comencemos calculando  $\text{cond}(x_n; d)$  con  $n$  par. Acotemos primero  $\text{cond}(x_{2i}; d)$ . Usando las fórmulas (4.4.4), (4.4.5), (4.4.10) y (4.4.11), se tiene

$$\begin{aligned}
\text{cond}(x_{2i}; d) &= \sum_{k=0}^{2i} \|d_k^n\|_\infty |\omega_k(x_{2i})| = \sum_{k=0}^i \|d_{2k}^m\|_\infty |\omega_{2k}(x_{2i})| + \sum_{k=0}^{i-1} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x_{2i})| \\
&= 1 + \sum_{k=1}^i \frac{2^k i!(n-i)!}{k!(n-1)(n-3)\cdots(n-2k+1)(i-k)!(n-i-k)!} \\
&\quad + \sum_{k=0}^{i-1} \frac{2^{k+1} i!(n-i)!}{k!n(n-2)\cdots(n-2k)(i-k-1)!(n-i-k)!} \\
&= 1 + \sum_{k=1}^i \binom{n-i}{k} \binom{i}{k} \frac{k!}{(n-1)/2((n-1)/2-1)\cdots((n-1)/2-k+1)} \\
&\quad + \sum_{k=0}^{i-1} \binom{n-i}{k} \binom{i}{k+1} \frac{(k+1)!}{n/2(n/2-1)\cdots(n/2-k)}.
\end{aligned}$$

Supongamos que  $n$  es par e  $i = n/2$ . Usando el Lema 4.4.5

$$\begin{aligned}
\text{cond}(x_n; d) &\leq 1 + \sum_{k=1}^{n/2-1} \frac{\binom{n/2}{k}^2}{\binom{(n-1)/2}{k}} + \sum_{k=0}^{n/2-1} \frac{\binom{n/2}{k} \binom{n/2}{k+1}}{\binom{n/2}{k}} + n \\
&\leq \sum_{k=0}^{n/2-1} \binom{n/2}{k} \binom{n/2+1}{k+1} + n \leq \binom{n+1}{n/2} + n.
\end{aligned}$$

Para el resto de casos, aplicando el Lema 4.4.5, se tiene

$$\begin{aligned} \text{cond}(x_{2i}; d) &= 1 + \sum_{k=1}^i \frac{\binom{n-i}{k} \binom{i}{k}}{\binom{(n-1)/2}{k}} + \sum_{k=0}^{i-1} \frac{\binom{n-i}{k} \binom{i}{k+1}}{\binom{n/2}{k}} \\ &\leq \sum_{k=0}^i \binom{n-i}{k} \binom{i}{k} + \sum_{k=0}^i \binom{n-i}{k} \binom{i}{k+1} = \sum_{k=0}^i \binom{n-i}{k} \binom{i+1}{k+1} \\ &= \sum_{k=0}^i \binom{n-i}{k} \binom{i+1}{i-k} = \binom{n+1}{i}. \end{aligned}$$

En el caso de nodos impares, por (4.4.2), se cumple

$$\begin{aligned} |\omega_{2k}(x_{2i+1})| &= \prod_{j=0}^{2k-1} |x_{2i+1} - x_j| = \prod_{j=0}^{k-1} (x_{2i+1} - x_{2j}) \prod_{j=0}^{k-1} |x_{2i+1} - x_{2j+1}| \\ &= h^{2k} \prod_{j=0}^{k-1} (n-i-j) \prod_{j=0}^{k-1} |i-j| = h^{2k} \frac{i!(n-i)!}{(i-k)!(n-i-k)!}, \quad k \leq i, \quad (4.4.12) \end{aligned}$$

$$\begin{aligned} |\omega_{2k+1}(x_{2i+1})| &= \prod_{j=0}^{2k} |x_{2i+1} - x_j| = \prod_{j=0}^k (x_{2i+1} - x_{2j}) \prod_{j=0}^{k-1} |x_{2i+1} - x_{2j+1}| \\ &= h^{2k+1} \prod_{j=0}^k (n-i-j) \prod_{j=0}^{k-1} |i-j| = h^{2k+1} \frac{i!(n-i)!}{(i-k-1)!(n-i-k-1)!}, \end{aligned} \quad (4.4.13)$$

para  $k \leq i$ . Para finalizar, usando las fórmulas (4.4.4), (4.4.5), (4.4.12), (4.4.13), se tiene

$$\begin{aligned} \text{cond}(x_{2i+1}; d) &= \sum_{k=0}^{2i+1} \|d_k^n\|_{\infty} |\omega_k(x_{2i+1})| \\ &= \sum_{k=0}^i \|d_{2k}^n\|_{\infty} |\omega_{2k}(x_{2i+1})| + \sum_{k=0}^i \|d_{2k+1}^n\|_{\infty} |\omega_{2k+1}(x_{2i+1})| \\ &= 1 + \sum_{k=1}^i \frac{2^k i!(n-i)!}{k!(n-1)(n-3) \cdots (n-2k+1)(i-k)!(n-i-k)!} \\ &\quad + \sum_{k=0}^i \frac{2^{k+1} i!(n-i)!}{k!n(n-2) \cdots (n-2k)(i-k-1)!(n-i-k-1)!} \\ &= 1 + \sum_{k=1}^i \binom{n-i}{k} \binom{i}{k} \frac{k!}{(n-1)/2((n-1)/2-1) \cdots ((n-1)/2-k+1)} \\ &\quad + \sum_{k=0}^{i-1} \binom{n-i}{k+1} \binom{i}{k+1} \frac{(k+1)!}{n/2(n/2-1) \cdots (n/2-k)}. \end{aligned}$$

Como  $i \leq (n-1)/2$ , tenemos

$$\begin{aligned} \text{cond}(x_{2i+1}; d) &\leq \sum_{k=0}^i \binom{n-i}{k} \binom{i}{k} + \sum_{k=0}^i \binom{n-i}{k+1} \binom{i}{k} = \sum_{k=0}^i \binom{n-i+1}{k+1} \binom{i}{k} \\ &= \sum_{k=1}^{i+1} \binom{n-i+1}{k} \binom{i}{k-1} = \sum_{k=1}^{i+1} \binom{n-i+1}{k} \frac{k}{i+1} \binom{i+1}{k} \\ &\leq \sum_{k=1}^{i+1} \binom{n-i+1}{k} \binom{i+1}{k}, \end{aligned}$$

y, aplicando el Lema 4.4.5, obtenemos el resultado

$$\text{cond}(x_{2i+1}; d) \leq \binom{n+2}{i+1} - 1.$$

□

Tenemos la siguiente cota uniforme para el condicionamiento en los nodos siguiendo el orden central inverso.

**Corolario 4.4.7.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1). Entonces

$$\max_{i=0, \dots, n} \text{cond}(x_{i,n}; d) \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1.$$

*Demostración.* Por el Teorema 4.4.6, se tiene

$$\text{cond}(x_{2i,n}; d) \leq \binom{n+1}{i} + n, \quad i = 0, \dots, \lfloor \frac{n}{2} \rfloor,$$

y

$$\text{cond}(x_{2i+1,n}; d) \leq \binom{n+2}{i+1} - 1, \quad i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor.$$

Por el Lema 3.1.6, la sucesión  $\binom{n+1}{i}$  es no decreciente,  $0 \leq i \leq \lfloor n/2 \rfloor$ . Así,

$$\text{cond}(x_{2i,n}; d) \leq \binom{n+1}{\lfloor \frac{n}{2} \rfloor} + n, \quad i = 0, \dots, \lfloor \frac{n}{2} \rfloor,$$

y

$$\text{cond}(x_{2i+1,n}; d) \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1, \quad i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor.$$

Veamos que  $\binom{n+1}{\lfloor \frac{n}{2} \rfloor} + n \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1$ . En efecto, por el Lema 3.1.6 y la identidad de Pascal

$$\binom{n+1}{\lfloor \frac{n}{2} \rfloor} + n + 1 = \binom{n+1}{\lfloor \frac{n}{2} \rfloor} + \binom{n+1}{1} \leq \binom{n+1}{\lfloor \frac{n}{2} \rfloor} + \binom{n+1}{\lfloor \frac{n}{2} \rfloor - 1} = \binom{n+2}{\lfloor \frac{n}{2} \rfloor}.$$

Teniendo en cuenta que

$$\lfloor \frac{n}{2} \rfloor \leq \frac{n}{2} = \frac{n-2}{2} + 1 \leq \lfloor \frac{n-1}{2} \rfloor + 1 \leq \frac{n+3}{2}$$

y el Lema 3.1.6, se obtiene el resultado

$$\binom{n+1}{\lfloor \frac{n}{2} \rfloor} + n + 1 \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1}.$$

□

En el siguiente teorema vamos a acotar  $\text{cond}(x; d)$  dependiendo de si el punto en el que se evalúa se encuentra en la mitad derecha o izquierda del intervalo.

**Teorema 4.4.8.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1). Si  $x \in (x_{2i,n}, x_{2(i+1),n})$ , se tiene

$$\text{cond}(x; d) \leq \binom{n+2}{i+1} + \frac{2^{n+1-i}}{i+2}, \quad i = 0, \dots, \lfloor \frac{n}{2} \rfloor - 1,$$

y, si  $x \in (x_{2(i+1)+1,n}, x_{2i+1,n})$ , entonces

$$\text{cond}(x; d) \leq \binom{n+3}{i+2} + \frac{2^{n+1-i}}{i+2}, \quad i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor - 1.$$

*Demostración.* Denotemos por  $x_i := x_{i,n}$ ,  $i = 0, \dots, n$ . Sea  $x \in (x_{2i}, x_{2(i+1)})$ ,  $i = 0, \dots, \lfloor \frac{n}{2} \rfloor - 1$ . Recordemos que

$$|\omega_{2k}(x)| = \prod_{j=0}^{2k-1} |x - x_j| = \prod_{j=0}^{k-1} |x - x_{2j}| \prod_{j=0}^{k-1} |x - x_{2j+1}|.$$

Así, si  $k \leq i+1$ , se tiene por (4.4.2)

$$\begin{aligned} |\omega_{2k}(x)| &= \prod_{j=0}^{k-1} (x - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x) \leq \prod_{j=0}^{k-1} (x_{2(i+1)} - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\ &= h^{2k} \prod_{j=0}^{k-1} (i+1-j) \prod_{j=0}^{k-1} (n-j-i) = h^{2k} \frac{(i+1)!(n-i)!}{(i+1-k)!(n-i-k)!}. \end{aligned} \quad (4.4.14)$$

Si  $k > i + 1$ ,

$$\begin{aligned}
|\omega_{2k}(x)| &= \prod_{j=0}^i (x - x_{2j}) \prod_{j=i+1}^{k-1} (x_{2j} - x) \prod_{j=0}^{k-1} (x_{2j+1} - x) \\
&\leq \prod_{j=0}^i (x_{2(i+1)} - x_{2j}) \prod_{j=i+1}^{k-1} (x_{2j} - x_{2i}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\
&= h^{2k} \prod_{j=0}^i (i+1-j) \prod_{j=i+1}^{k-1} (j-i) \prod_{j=0}^{k-1} (n-j-i) \\
&= h^{2k} (i+1)!(k-i-1)! \frac{(n-i)!}{(n-i-k)!}. \tag{4.4.15}
\end{aligned}$$

Ahora pasamos a acotar los polinomios  $\omega_{2k+1}$  correspondientes a grado impar. Si  $k \leq i$ ,

$$\begin{aligned}
|\omega_{2k+1}(x)| &= \prod_{j=0}^k (x - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x) \leq \prod_{j=0}^k (x_{2(i+1)} - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\
&= h^{2k+1} \prod_{j=0}^k (i+1-j) \prod_{j=0}^{k-1} (n-j-i) = h^{2k+1} \frac{(i+1)!(n-i)!}{(i-k)!(n-i-k)!}, \tag{4.4.16}
\end{aligned}$$

y si  $k > i$ , entonces

$$\begin{aligned}
|\omega_{2k+1}(x)| &= \prod_{j=0}^i (x - x_{2j}) \prod_{j=i+1}^k (x_{2j} - x) \prod_{j=0}^{k-1} (x_{2j+1} - x) \\
&\leq \prod_{j=0}^i (x_{2(i+1)} - x_{2j}) \prod_{j=i+1}^k (x_{2j} - x_{2i}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i}) \\
&= h^{2k+1} \prod_{j=0}^i (i+1-j) \prod_{j=i+1}^k (j-i) \prod_{j=0}^{k-1} (n-j-i) \\
&= h^{2k+1} (i+1)!(k-i)! \frac{(n-i)!}{(n-i-k)!}. \tag{4.4.17}
\end{aligned}$$

Descomponiendo  $\text{cond}(x; d)$  para  $x \in (x_{2i}, x_{2(i+1)})$

$$\begin{aligned}
\text{cond}(x; d) &= \sum_{k=0}^n \|d_k^m\|_\infty |\omega_k(x)| = \sum_{k=0}^{\lfloor n/2 \rfloor} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| + \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)| \\
&= \sum_{k=0}^{i+1} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| + \sum_{k=i+2}^{\lfloor n/2 \rfloor} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| \\
&\quad + \sum_{k=0}^i \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)| + \sum_{k=i+1}^{\lfloor (n-1)/2 \rfloor} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)|.
\end{aligned}$$

Usando las fórmulas (4.4.4), (4.4.5) del Teorema 4.4.3 y las cotas (4.4.14), (4.4.15), (4.4.16) y (4.4.17), se tiene

$$\begin{aligned} \text{cond}(x; d) &\leq \sum_{k=0}^{i+1} \frac{2^k(i+1)!(n-i)!}{k!(i+1-k)!(n-i-k)!(n-1)(n-3)\cdots(n-2k-1)} \\ &+ \sum_{k=i+2}^{\lfloor n/2 \rfloor} \frac{2^k(i+1)!(k-i-1)!(n-i)!}{k!(n-i-k)!(n-1)(n-3)\cdots(n-2k-1)} \\ &+ \sum_{k=0}^i \frac{2^{k+1}(i+1)!(n-i)!}{k!(i-k)!(n-i-k)!n(n-2)\cdots(n-2k)} \\ &+ \sum_{k=i+1}^{\lfloor (n-1)/2 \rfloor} \frac{2^{k+1}(i+1)!(k-i)!(n-i)!}{k!(n-i-k)!n(n-2)\cdots(n-2k)}. \end{aligned}$$

Expresándolo con números combinatorios, obtenemos

$$\begin{aligned} \text{cond}(x; d) &\leq \sum_{k=0}^{i+1} \binom{n-i}{k} \binom{i+1}{k} \frac{2^k k!}{(n-1)(n-3)\cdots(n-2k-1)} \\ &+ \sum_{k=i+2}^{\lfloor n/2 \rfloor} \frac{\binom{n-i}{k}}{\binom{k}{i+1}} \frac{2^k k!}{(n-1)(n-3)\cdots(n-2k-1)} \\ &+ \sum_{k=0}^i \binom{n-i}{k} \binom{i+1}{k} \frac{2^{k+1}(k+1)!}{n(n-2)\cdots(n-2k)} \\ &+ \sum_{k=i+1}^{\lfloor (n-1)/2 \rfloor} \frac{\binom{n-i}{k}}{\binom{k+1}{i+1}} \frac{2^{k+1}(k+1)!}{n(n-2)\cdots(n-2k)}. \end{aligned}$$

Teniendo en cuenta que  $k \leq \lfloor \frac{n}{2} \rfloor$  en las dos primeras sumas y  $k \leq \lfloor \frac{n-1}{2} \rfloor$  en las dos últimas, podemos acotar el condicionamiento de la siguiente manera

$$\begin{aligned} \text{cond}(x; d) &\leq \sum_{k=0}^{i+1} \binom{n-i}{k} \binom{i+1}{k} + \sum_{k=i+2}^{\lfloor n/2 \rfloor} \frac{\binom{n-i}{k}}{\binom{k}{i+1}} + \sum_{k=0}^i \binom{n-i}{k} \binom{i+1}{k+1} + \sum_{k=i+1}^{\lfloor (n-1)/2 \rfloor} \frac{\binom{n-i}{k}}{\binom{k+1}{i+1}} \\ &\leq \sum_{k=0}^{i+1} \binom{n-i}{k} \binom{i+1}{k} + \frac{1}{i+2} \sum_{k=i+2}^{\lfloor n/2 \rfloor} \binom{n-i}{k} \\ &+ \sum_{k=0}^i \binom{n-i}{k} \binom{i+1}{k+1} + \frac{1}{i+2} \sum_{k=i+1}^{\lfloor (n-1)/2 \rfloor} \binom{n-i}{k}. \end{aligned}$$

Como  $i < \lfloor \frac{n}{2} \rfloor$ , entonces  $\lfloor \frac{n-1}{2} \rfloor \leq \lfloor \frac{n}{2} \rfloor \leq \frac{n}{2} = n - \frac{n}{2} \leq n - \lfloor \frac{n}{2} \rfloor \leq n - i$ . Por tanto, extendiendo la segunda y cuarta suma desde  $k = 0$  hasta  $n - i$ , y usando el Lema 4.4.5, obtenemos

$$\text{cond}(x; d) \leq \binom{n+1}{i+1} + \binom{n+1}{i} + \frac{2^{n+1-i}}{i+2} = \binom{n+2}{i+1} + \frac{2^{n+1-i}}{i+2}.$$

Sea ahora  $x \in (x_{2(i+1)+1}, x_{2i+1})$ ,  $i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor - 1$ . Si  $k \leq i + 1$  se tiene, por (4.4.2),

$$\begin{aligned} |\omega_{2k}(x)| &= \prod_{j=0}^{k-1} (x - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x) \leq \prod_{j=0}^{k-1} (x_{2i+1} - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2i+1+1}) \\ &= h^{2k} \prod_{j=0}^{k-1} (n - j - i) \prod_{j=0}^{k-1} (i + 1 - j) = h^{2k} \frac{(i+1)!(n-i)!}{(i+1-k)!(n-i-k)!}. \end{aligned} \quad (4.4.18)$$

Si  $k > i + 1$

$$\begin{aligned} |\omega_{2k}(x)| &= \prod_{j=0}^{k-1} (x - x_{2j}) \prod_{j=0}^i (x_{2j+1} - x) \prod_{j=i+1}^{k-1} (x - x_{2j+1}) \\ &\leq \prod_{j=0}^{k-1} (x_{2i+1} - x_{2j}) \prod_{j=0}^i (x_{2j+1} - x_{2(i+1)+1}) \prod_{j=i+1}^{k-1} (x_{2i+1} - x_{2j+1}) \\ &= h^{2k} \prod_{j=0}^{k-1} (n - i - j) \prod_{j=0}^i (i + 1 - j) \prod_{j=i+1}^{k-1} (j - i) \\ &= h^{2k} (i+1)!(k-i-1)! \frac{(n-i)!}{(n-i-k)!}. \end{aligned} \quad (4.4.19)$$

Si  $k \leq i + 1$

$$\begin{aligned} |\omega_{2k+1}(x)| &= \prod_{j=0}^k (x - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x) \leq \prod_{j=0}^k (x_{2i+1} - x_{2j}) \prod_{j=0}^{k-1} (x_{2j+1} - x_{2(i+1)+1}) \\ &= h^{2k+1} \prod_{j=0}^k (n - i - j) \prod_{j=0}^{k-1} (i + 1 - j) = h^{2k+1} \frac{(i+1)!(n-i)!}{(i+1-k)!(n-i-k-1)!}, \end{aligned} \quad (4.4.20)$$

y si  $k > i + 1$ , entonces

$$\begin{aligned} |\omega_{2k+1}(x)| &= \prod_{j=0}^k (x - x_{2j}) \prod_{j=0}^i (x_{2j+1} - x) \prod_{j=i+1}^k (x - x_{2j+1}) \\ &\leq \prod_{j=0}^k (x_{2i+1} - x_{2j}) \prod_{j=0}^i (x_{2j+1} - x_{2(i+1)+1}) \prod_{j=i+1}^{k-1} (x_{2i+1} - x_{2j+1}) \\ &= h^{2k+1} \prod_{j=0}^i (n - i - j) \prod_{j=i+1}^k (i + 1 - j) \prod_{j=i+1}^{k-1} (j - i) \\ &= h^{2k+1} (i+1)!(k-i-1)! \frac{(n-i)!}{(n-i-k-1)!}. \end{aligned} \quad (4.4.21)$$

Descomponiendo  $\text{cond}(x; d)$  para  $x \in (x_{2(i+1)+1}, x_{2i+1})$

$$\begin{aligned} \text{cond}(x; d) &= \sum_{k=0}^n \|d_k^m\|_\infty |\omega_k(x)| = \sum_{k=0}^{\lfloor n/2 \rfloor} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| + \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)| \\ &= \sum_{k=0}^{i+1} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| + \sum_{k=i+2}^{\lfloor n/2 \rfloor} \|d_{2k}^m\|_\infty |\omega_{2k}(x)| \\ &\quad + \sum_{k=0}^{i+1} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)| + \sum_{k=i+2}^{\lfloor (n-1)/2 \rfloor} \|d_{2k+1}^m\|_\infty |\omega_{2k+1}(x)|. \end{aligned}$$

Usando las fórmulas (4.4.4), (4.4.5) del Teorema 4.4.3 y las cotas (4.4.18), (4.4.19), (4.4.20) y (4.4.21), se tiene

$$\begin{aligned} \text{cond}(x; d) &\leq \sum_{k=0}^{i+1} \frac{2^k (i+1)! (n-i)!}{k! (i+1-k)! (n-i-k-1)! (n-1)(n-3) \cdots (n-2k-1)} \\ &\quad + \sum_{k=i+2}^{\lfloor n/2 \rfloor} \frac{2^k (i+1)! (k-i-1)! (n-i)!}{k! (n-i-k)! (n-1)(n-3) \cdots (n-2k-1)} \\ &\quad + \sum_{k=0}^{i+1} \frac{2^{k+1} (i+1)! (n-i)!}{k! (i+1-k)! (n-i-k-1)! n(n-2) \cdots (n-2k)} \\ &\quad + \sum_{k=i+2}^{\lfloor (n-1)/2 \rfloor} \frac{2^{k+1} (i+1)! (k-i-1)! (n-i)!}{k! (n-i-k-1)! n(n-2) \cdots (n-2k)}. \end{aligned}$$

Procedemos de manera análoga al caso anterior: expresamos los sumandos con números combinatorios. Como  $k \leq \lfloor \frac{n}{2} \rfloor$  en las dos primeras sumas y  $k \leq \lfloor \frac{n-1}{2} \rfloor$  en las dos últimas, se tiene

$$\text{cond}(x; d) \leq \sum_{k=0}^{i+1} \binom{n-i}{k} \binom{i+1}{k} + \sum_{k=i+2}^{\lfloor n/2 \rfloor} \frac{\binom{n-i}{k}}{\binom{k}{i+1}} + \sum_{k=0}^{i+1} \binom{n-i}{k+1} \binom{i+1}{k} + \sum_{k=i+2}^{\lfloor (n-1)/2 \rfloor} \frac{\binom{n-i}{k+1}}{\binom{k}{i+1}}.$$

Recordemos que  $\lfloor \frac{n-1}{2} \rfloor \leq \lfloor \frac{n}{2} \rfloor \leq n-i$ . Por tanto, extendiendo la segunda y cuarta suma desde 0 hasta  $n-i$ , obtenemos la cota

$$\begin{aligned} \text{cond}(x; d) &\leq \sum_{k=0}^{i+1} \binom{n-i+1}{k+1} \binom{i+1}{k} + \frac{2^{n+1-i}}{i+2} = \sum_{k=1}^{i+2} \binom{n-i+1}{k} \binom{i+1}{k-1} + \frac{2^{n+1-i}}{i+2} \\ &= \sum_{k=1}^{i+2} \binom{n-i+1}{k+1} \binom{i+2}{k} \frac{k}{i+2} + \frac{2^{n+1-i}}{i+2} \\ &\leq \sum_{k=1}^{i+2} \binom{n-i+1}{k+1} \binom{i+2}{k} + \frac{2^{n+1-i}}{i+2}, \end{aligned}$$

y, finalmente, usando el Lema 4.4.5, se tiene el resultado

$$\text{cond}(x; d) \leq \binom{n+3}{i+2} + \frac{2^{n+1-i}}{i+2}.$$

□

Vamos a aplicar los resultados anteriores para dar una cota global de  $\text{cond}(x; d)$ . Por el Corolario 4.4.7, tenemos la siguiente cota para el condicionamiento en los nodos

$$\text{cond}(x_i; d) \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1, \quad i = 0, \dots, n.$$

Tenemos las siguientes desigualdades para  $n!$  (véase [53])

$$\sqrt{2\pi n} n^n e^{-n} < n! < \sqrt{2\pi n} n^n e^{-n + \frac{1}{12n}}. \quad (4.4.22)$$

**Corolario 4.4.9.** Sean  $x_{0,n}, \dots, x_{n,n}$  nodos dados por (4.4.1). Si  $x \in [-nh/2, nh/2]$ , se tiene

$$\text{cond}(x; d) \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n.$$

*Demostración.* Veamos que la cota se verifica para los nodos. Para ello, veamos que la cota obtenida en el Corolario 4.4.7 verifica

$$\text{cond}(x_{i,n}; d) \leq \binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1 \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n.$$

Si  $n$  es par, se tiene

$$\binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1 = \binom{n+2}{\frac{n}{2}} - 1 = \frac{(n+2)(n+1)}{(n/2+2)(n/2+1)} \frac{n!}{(n/2)!^2} - 1.$$

Usando las desigualdades (4.4.22),

$$\begin{aligned} \binom{n+2}{\frac{n}{2}} - 1 &\leq \frac{4(n+2)(n+1)}{(n+4)(n+3)} \frac{n^n e^{-n} \sqrt{2\pi n} e^{1/12n}}{((n/2)^{n/2} e^{-n/2} \sqrt{\pi n})^2} - 1 \\ &\leq 2^{n+5/2} \frac{e^{1/12n}}{\sqrt{\pi n}} - 1 \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n. \end{aligned} \quad (4.4.23)$$

Para el caso  $n$  impar, procedemos de una manera análoga

$$\binom{n+2}{\lfloor \frac{n-1}{2} \rfloor + 1} - 1 = \binom{n+2}{\frac{n+1}{2}} - 1 = \frac{n+2}{(n+1)/2+1} \frac{(n+1)!}{(n+1)/2!^2} - 1,$$

y aplicando (4.4.22), obtenemos

$$\binom{n+2}{\frac{n+1}{2}} \leq \frac{(n+1)^{n+1} e^{-(n+1)} \sqrt{2\pi(n+1)} e^{1/12(n+1)}}{\left(\left((n+1)/2\right)^{(n+1)/2} e^{-(n+1)/2} \sqrt{\pi(n+1)}\right)^2} - 1 \leq 2^{n+5/2} \frac{e^{1/12(n+1)}}{\sqrt{\pi(n+1)}} - 1,$$

y, claramente, se cumple la cota.

Obtengamos una cota uniforme para  $\text{cond}(x; d)$ ,  $x \in (x_{2i}, x_{2(i+1)})$ ,  $i = 0, \dots, \lfloor \frac{n}{2} \rfloor - 1$ .

Si  $n$  es par, por el Teorema 4.4.8 y las desigualdades de la fórmula (4.4.22), se tiene

$$\begin{aligned} \text{cond}(x; d) &\leq \binom{n+2}{i+1} + \frac{2^{n+1-i}}{i+2} \leq \binom{n+2}{\lfloor \frac{n}{2} \rfloor} + 2^n \leq \binom{n+2}{\frac{n}{2}} + 2^n \\ &= \frac{(n+2)(n+1)}{(n/2+2)(n/2+1)} \frac{n!}{(n/2)!^2} + 2^n \\ &\leq \frac{4(n+2)(n+1)}{(n+4)(n+3)} \frac{n^n e^{-n} \sqrt{2\pi n} e^{1/12n}}{\left((n/2)^{n/2} e^{-n/2} \sqrt{\pi n}\right)^2} + 2^n \\ &\leq 2^{n+5/2} \frac{e^{1/12n}}{\sqrt{\pi n}} + 2^n \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n. \end{aligned}$$

Para el caso  $n$  impar, usando el Teorema 4.4.8 y (4.4.22),

$$\begin{aligned} \text{cond}(x; d) &\leq \binom{n+2}{i+1} + \frac{2^{n+1-i}}{i+2} \leq \binom{n+2}{\lfloor \frac{n}{2} \rfloor} + 2^n \leq \binom{n+2}{\frac{n-1}{2}} + 2^n \\ &= \frac{(n+2)(n+1)n}{\left((n-1)/2+3\right)\left((n-1)/2+2\right)\left((n-1)/2+1\right)} \frac{(n-1)!}{\left((n-1)/2\right)!^2} + 2^n \\ &\leq 2^3 \frac{(n-1)^{n-1} e^{-(n-1)} \sqrt{2\pi(n-1)} e^{1/12(n-1)}}{\left(\left((n-1)/2\right)^{(n-1)/2} e^{-(n-1)/2} \sqrt{\pi(n-1)}\right)^2} + 2^n \\ &\leq 2^{n+5/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n. \end{aligned}$$

Finalmente, sea  $x \in (x_{2(i+1)+1}, x_{2i+1}, n)$ ,  $i = 0, \dots, \lfloor \frac{n-1}{2} \rfloor - 1$ . Si  $n$  es impar, por el Teorema 4.4.8 y las desigualdades de la fórmula (4.4.22), se tiene

$$\begin{aligned} \text{cond}(x; d) &\leq \binom{n+3}{i+2} + \frac{2^{n+1-i}}{i+2} \leq \binom{n+3}{\lfloor \frac{n-1}{2} \rfloor + 1} + 2^n = \binom{n+3}{\frac{n+1}{2}} + 2^n \\ &= \frac{(n+3)(n+2)}{\left((n+1)/2+2\right)\left((n+1)/2+1\right)} \frac{(n+1)!}{\left((n+1)/2\right)!^2} + 2^n \\ &\leq 2^2 \frac{(n+1)^{n+1} e^{-(n+1)} \sqrt{2\pi(n+1)} e^{1/12(n+1)}}{\left(\left((n+1)/2\right)^{(n+1)/2} e^{-(n+1)/2} \sqrt{\pi(n+1)}\right)^2} + 2^n \\ &\leq 2^{n+7/2} \frac{e^{1/12(n+1)}}{\sqrt{\pi(n+1)}} + 2^n \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n. \end{aligned}$$

Por último, si  $n$  es par, de manera análoga obtenemos

$$\begin{aligned} \text{cond}(x; d) &\leq \binom{n+3}{i+2} + \frac{2^{n+1-i}}{i+2} \leq \binom{n+3}{\lfloor \frac{n-1}{2} \rfloor + 1} + 2^n = \binom{n+3}{\frac{n}{2}} + 2^n \\ &= \frac{(n+3)(n+2)(n+1)}{(n/2+3)(n/2+2)(n/2+1)} \frac{n!}{(n/2)!^2} + 2^n \leq 2^3 \frac{n^n e^{-n} \sqrt{2\pi n} e^{1/12n}}{((n/2)^{n/2} e^{-n/2} \sqrt{\pi n})^2} + 2^n \\ &\leq 2^{n+7/2} \frac{e^{1/12n}}{\sqrt{\pi n}} + 2^n \leq 2^{n+7/2} \frac{e^{1/12(n-1)}}{\sqrt{\pi(n-1)}} + 2^n. \end{aligned}$$

□

Observemos que la cota del Corolario 4.4.9 es próxima a la estimación de la constante de Lebesgue obtenida por Schönhage en [58]

$$\frac{2^{n+1}}{e n \log(n + \gamma)}.$$

Por tanto, el orden central inverso es una estrategia de ordenación muy conveniente para la fórmula de Newton, además de ser de fácil implementación.

Las figuras 4.4.1, 4.4.2, 4.4.3 y 4.4.4 comparan la función de Lebesgue y  $\text{cond}(x; d)$  con nodos equidistantes siguiendo el orden central inverso. Podemos ver que si los nodos siguen este orden,  $\text{cond}(x; d)$  se parece a la constante de Lebesgue en los extremos del intervalo y, a medida que nos acercamos al origen, la diferencia entre ambas funciones aumenta. También vemos en estas gráficas que el máximo de  $\text{cond}(x; d)$  se alcanza en un entorno del centro para  $n \leq 13$  y para  $n \geq 14$  el valor en los extremos es mayor que en un entorno del origen. En la figura 4.4.5 representamos la función de Lebesgue y el condicionamiento con nodos equidistantes siguiendo el orden central inverso y el orden central respecto al punto de evaluación dado por (4.3.1) para  $n = 10$  en el intervalo  $[-1, 1]$ . Comprobamos que globalmente el orden central respecto al punto de evaluación tiene un mejor comportamiento que el orden central inverso.

En la sección siguiente veremos una interpretación del orden central inverso para nodos equidistantes en términos de una eliminación matricial llamada eliminación de Neville. Esta interpretación muestra ciertas analogías con las propiedades del orden de Leja respecto a la eliminación Gaussiana.

## 4.5 Matriz de Vandermonde, eliminación y pivotaje

El método de *eliminación de Gauss* es un algoritmo que transforma un sistema lineal de ecuaciones

$$Ax = b, \tag{4.5.1}$$

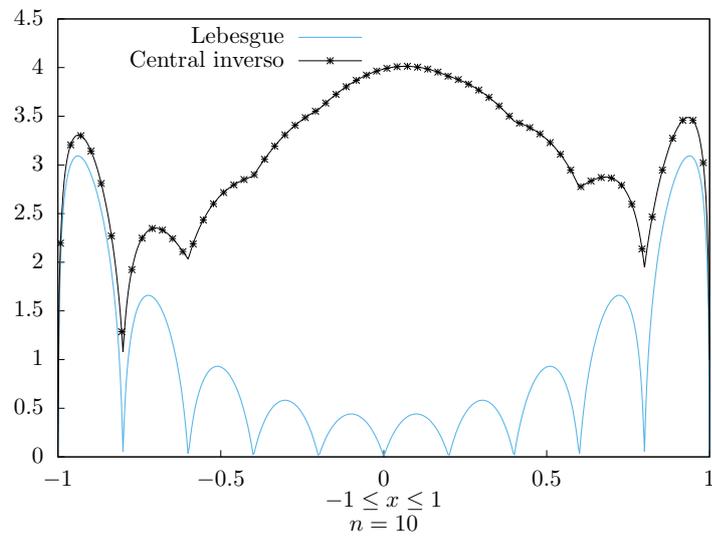


Figura 4.4.1: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x;d))$  para puntos equidistantes con orden central inverso con  $n = 10$ .

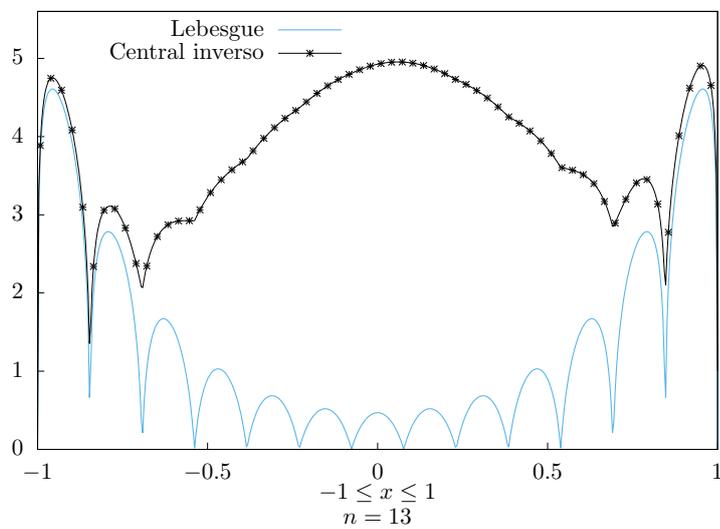


Figura 4.4.2: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x;d))$  para puntos equidistantes con orden central inverso con  $n = 13$ .

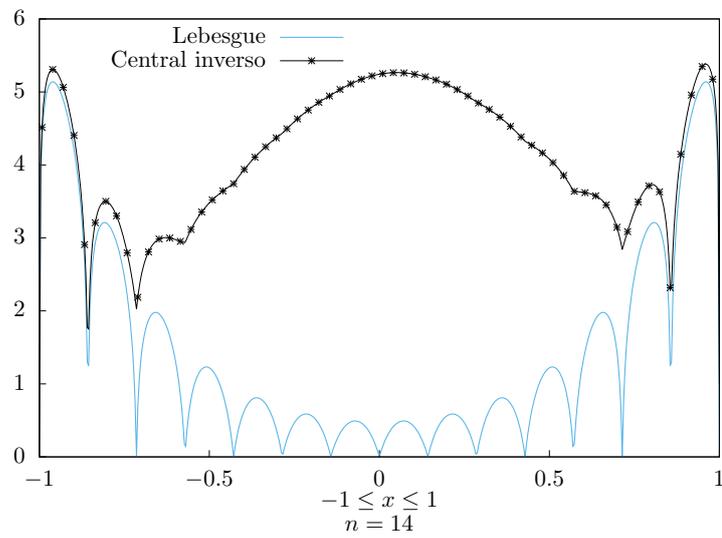


Figura 4.4.3: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x;d))$  para puntos equidistantes con orden central inverso con  $n = 14$ .

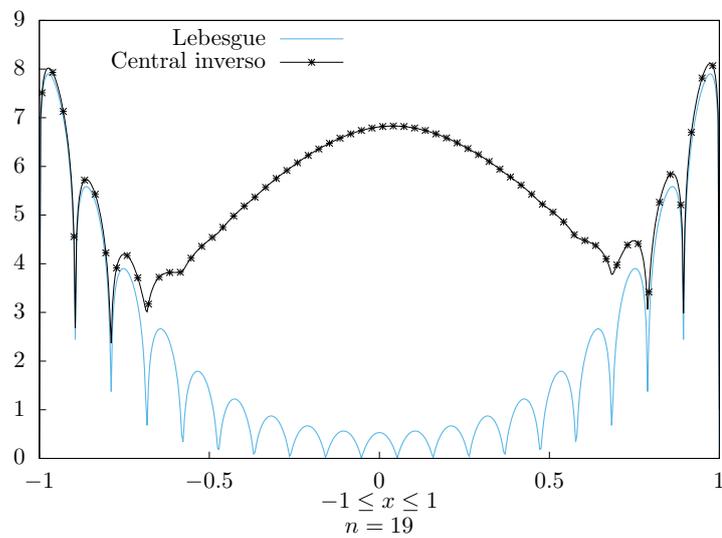


Figura 4.4.4: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x;d))$  para puntos equidistantes con orden central inverso con  $n = 19$ .

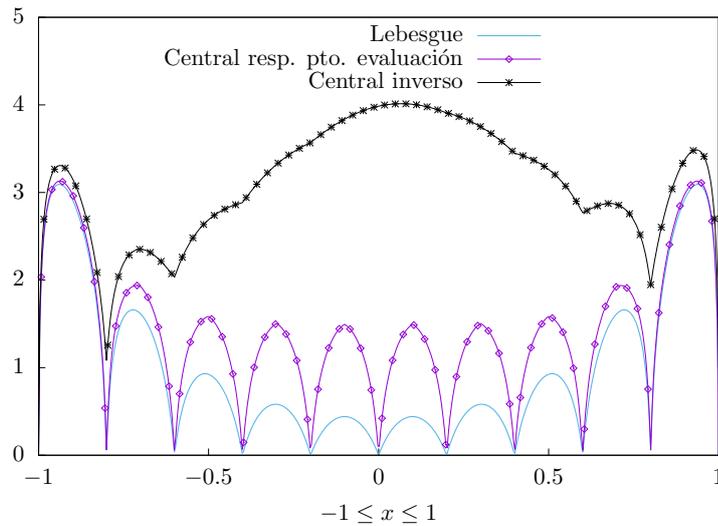


Figura 4.4.5: Gráfica de  $\log_3(\lambda_L(x))$  y  $\log_3(\text{cond}(x; d))$  para puntos equidistantes con orden central inverso y con el orden central respecto al punto de evaluación para  $n = 10$ .

con  $A = (a_{ij})_{0 \leq i, j \leq n}$  una matriz  $(n + 1) \times (n + 1)$  no singular y  $b \in \mathbb{R}^{n+1}$ , en otro sistema

$$Ux = c$$

que tiene la misma solución que (4.5.1), donde  $U$  es una matriz triangular superior. El sistema  $Ux = c$  se puede resolver fácilmente por sustitución regresiva. La eliminación de Gauss consiste en  $n$  pasos que dan lugar a una sucesión de matrices

$$A = \tilde{A}^{(0)} \rightarrow A^{(0)} \rightarrow \tilde{A}^{(1)} \rightarrow A^{(1)} \rightarrow \dots \rightarrow \tilde{A}^{(n)} = A^{(n)} = U$$

Si  $A = (a_{ij})_{0 \leq i, j \leq n}$ , para cada  $k$ ,  $1 \leq k \leq n$ , la matriz  $A^{(k)} = (a_{ij}^{(k)})_{0 \leq i, j \leq n}$  tiene ceros debajo de la diagonal principal en las primeras  $k - 1$  columnas. Para  $k = 0, \dots, n - 1$ , la matriz  $A^{(k)}$  se obtiene de  $\tilde{A}^{(k)}$  permutando las filas de  $\tilde{A}^{(k)}$  desde la diagonal de manera que  $a_{kk}^{(k)}$  sea distinto de cero. En este paso se pueden aplicar estrategias de pivotaje. Para calcular  $\tilde{A}^{(k+1)}$  a partir de  $A^{(k)}$  realizamos ceros en la columna  $k$  debajo de la diagonal principal, restando un múltiplo de la fila  $k$ -ésima a las filas posteriores

$$\tilde{a}_{ij}^{(k+1)} := \begin{cases} a_{ij}^{(k)}, & \text{si } i \leq k, \\ a_{ij}^{(k)} - (a_{ik}^{(k)} / a_{kk}^{(k)}) a_{kj}^{(k)}, & \text{si } i \geq k + 1 \text{ y } a_{kk}^{(k)} \neq 0. \end{cases}$$

El elemento  $p_j := a_{jj}^{(j)}$  se denomina pivote  $j$ -ésimo de la eliminación de Gauss y

$$m_{ij} := \frac{a_{ij}^{(j)}}{a_{jj}^{(j)}} = \frac{a_{ij}^{(j)}}{p_j}, \quad (4.5.2)$$

se llama multiplicador de la eliminación de Gauss de  $A$ , para  $0 \leq j < i \leq n$ .

Gracias a la eliminación de Gauss de una matriz  $A$  podemos calcular una factorización  $LU$  con  $U$  la matriz final que se obtiene en el proceso descrito anteriormente y  $L$  una matriz triangular inferior de la siguiente forma

$$L = \begin{pmatrix} 1 & & & & \\ m_{10} & 1 & & & \\ m_{20} & m_{21} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ m_{n0} & m_{n1} & \cdots & m_{n,n-1} & 1 \end{pmatrix}.$$

Es bien conocido que la eliminación de Gauss con una estrategia de pivotaje por filas calcula una factorización  $PA = LU$ , con  $P$  una matriz de permutación,  $L$  una matriz triangular inferior con unos en la diagonal principal y  $U$  triangular superior. Por tanto, la eliminación de Gauss con pivotaje parcial aplicada a una matriz  $A$  es equivalente a realizar primero una permutación de filas en  $A$  y realizar posteriormente eliminación de Gauss sin pivotaje.

La *eliminación de Gauss con pivotaje parcial* es una estrategia de pivotaje por filas que consiste en reordenar las filas de  $\tilde{A}^{(k)}$ , desde la diagonal en adelante, de forma que los elementos de la columna  $k$ -ésima de  $A^{(k)}$  verifiquen

$$|p_k| = |a_{kk}^{(k)}| \geq |a_{ik}^{(k)}|, \quad i = k + 1, \dots, n,$$

para  $k = 0, \dots, n - 1$ . Esta propiedad es equivalente a que todos los  $|m_{ij}|$  sean menores o iguales que 1, para todo  $j = 0, \dots, n - 1$ ,  $i = j + 1, \dots, n$ . Así, los elementos de la matriz triangular inferior  $L$  tienen valor absoluto menor o igual que 1.

Recordemos (véase la Sección 2.6) que la matriz de Vandermonde es la matriz de colocación de la base de monomios en los nodos  $x_0, \dots, x_n$ ,

$$V(x_0, \dots, x_n) := \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}. \quad (4.5.3)$$

En la Sección 2.6, se obtuvo la descomposición de la matriz de Vandermonde con nodos distintos como producto de dos factores triangulares, una matriz triangular inferior con unos en la diagonal,  $\tilde{L}$ , y una matriz triangular superior,  $\tilde{U}$ . Recordemos que los elementos de la matriz triangular inferior de  $\tilde{L}$  son los multiplicadores de la eliminación de Gauss y, por (2.4.8), se tiene

$$m_{ij} = \prod_{k=0}^{j-1} \frac{x_i - x_k}{x_j - x_k} = \frac{\omega_j(x_i)}{\omega_j(x_j)}, \quad i = 0, \dots, n, \quad j = i, \dots, n. \quad (4.5.4)$$

Higham ya observó que la eliminación de Gauss con pivotaje parcial conduce a un orden de filas correspondiente esencialmente al orden de Leja (véase Sección 22.3.3 de [37]).

**Proposición 4.5.1.** Sean  $x_0, \dots, x_n$  una sucesión de nodos que sigue el orden de Leja. Entonces la eliminación de Gauss con pivotaje parcial aplicado a la matriz de Vandermonde  $V(x_0, \dots, x_n)$  dada por (4.5.3) puede realizarse sin cambios de filas.

*Demostración.* Por la fórmula (4.1.1), los multiplicadores de la eliminación de Gauss, dados por (4.5.4), verifican  $|m_{ij}| \leq 1$ . Esta condición es equivalente al criterio de la estrategia de pivotaje parcial.  $\square$

La *eliminación de Neville* es un proceso de eliminación alternativo al de Gauss en el que para producir ceros en una columna de la matriz restamos a cada fila un múltiplo adecuado de la anterior (véase [34]). Esta eliminación ha sido muy útil al tratar ciertas clases de matrices estructuradas como las totalmente positivas (TP). Sea  $A = (a_{ij})_{0 \leq i, j \leq n}$  una matriz no singular. La eliminación de Neville de  $A$  consiste en  $n$  pasos sucesivos, dando lugar a una sucesión de matrices

$$A = \tilde{A}^{(0)} \rightarrow A^{(0)} \rightarrow \tilde{A}^{(1)} \rightarrow A^{(1)} \rightarrow \dots \rightarrow \tilde{A}^{(n)} = A^{(n)} = U,$$

donde  $U$  es una matriz triangular superior. Para cada  $k$ ,  $1 \leq k \leq n$ , la matriz  $A^{(k)} = (a_{ij}^{(k)})_{0 \leq i, j \leq n}$  tiene ceros debajo de la diagonal principal en las primeras  $k - 1$  columnas. Para  $k = 0, \dots, n - 1$ , la matriz  $A^{(k)}$  se obtiene de  $\tilde{A}^{(k)}$  permutando las filas de  $\tilde{A}^{(k)}$  desde la diagonal de manera que los elementos nulos de la columna  $k$  queden al final. En este paso se pueden aplicar estrategias de pivotaje. Para calcular  $\tilde{A}^{(k+1)}$  a partir de  $A^{(k)}$  realizamos ceros en la columna  $k$  debajo de la diagonal principal restando un múltiplo de la fila  $i$ -ésima a la fila  $(i + 1)$ -ésima

$$\tilde{a}_{ij}^{(k+1)} := \begin{cases} a_{ij}^{(k)}, & \text{si } i \leq k, \\ a_{ij}^{(k)} - \left( a_{ik}^{(k)} / a_{i-1,k}^{(k)} \right) a_{i-1,j}^{(k)}, & \text{si } i \geq k + 1 \text{ y } a_{i-1,k}^{(k)} \neq 0, \\ a_{ij}^{(k)}, & \text{si } i \geq k + 1 \text{ y } a_{i-1,k}^{(k)} = 0, \end{cases}$$

para  $i = k + 1, \dots, n$  y  $k = 0, 1, \dots, n - 1$ . El elemento

$$p_{ij} := a_{ij}^{(j)}, \quad 0 \leq j < i \leq n,$$

es el pivote  $(i, j)$  de la eliminación de Neville de  $A$ . Si todos los pivotes son no nulos, entonces la eliminación de Neville puede realizarse sin cambios de filas y por el Lema 2.6 de [34]

$$\begin{aligned} p_{i0} &= a_{i0}, & i &= 0, \dots, n \\ p_{ij} &= \frac{\det A[i - j + 1, \dots, i + 1 | 1, \dots, j + 1]}{\det A[i - j + 1, \dots, i | 1, \dots, j]}, & 0 < j \leq i \leq n. \end{aligned} \quad (4.5.5)$$

El número

$$m_{ij} := \begin{cases} \frac{a_{ij}^{(j)}}{a_{i-1,j}^{(j)}} = \frac{p_{ij}}{p_{i-1,j}}, & \text{si } a_{i-1,j}^{(j)} \neq 0 \\ 0, & \text{si } a_{i-1,j}^{(j)} = 0, \end{cases} \quad (4.5.6)$$

se llama multiplicador de la eliminación de Neville de  $A$ , para  $0 \leq j < i \leq n$ .

La *eliminación de Neville con pivotaje parcial* es una estrategia que consiste en reordenar las filas de  $\tilde{A}^{(k)}$  desde la  $k$ -ésima en adelante, de modo que los pivotes que aparecen en la columna  $k$  de  $A^{(k)}$  formen una sucesión decreciente:

$$|p_{kk}| \geq |p_{k+1,k}| \geq \dots \geq |p_{nk}|, \quad k = 0, \dots, n - 1.$$

Esto es equivalente a que todos los  $|m_{ij}|$  sean menores o iguales que 1, para todo  $j = 0, \dots, n - 1$ ,  $i = j + 1, \dots, n$ . Esta condición es análoga a la que cumplen los multiplicadores de la eliminación de Gauss con pivotaje parcial. Por el Teorema 1 de [2] se sabe que el factor de crecimiento de Wilkinson de una matriz  $n \times n$  a la que se aplica Neville con pivotaje parcial está acotado superiormente por  $2^{n-1}$ , la misma cota del factor de crecimiento que se cumple para la eliminación Gaussiana con pivotaje parcial. Recordemos que el factor de Wilkinson es una medida de la estabilidad regresiva o “backward” en la resolución del sistema de ecuaciones lineales.

Veamos cómo son los pivotes y los multiplicadores de la eliminación de Neville para la matriz de Vandermonde  $V(x_0, \dots, x_n)$  con nodos distintos. Por (2.6.2), se tiene  $V = (x_j^i)_{0 \leq i, j \leq n}$ . Por la fórmula (4.5.5),  $p_{i0} = v_{i0} = 1$  para  $i = 0, \dots, n$  y

$$p_{ij} = \frac{\det V(x_{i-j}, \dots, x_i)}{\det V(x_{i-j}, \dots, x_{i-1})} = \frac{\prod_{i-j \leq k < l \leq i} (x_l - x_k)}{\prod_{i-j \leq k < l \leq i-1} (x_l - x_k)}, \quad (4.5.7)$$

donde hemos usado la conocida fórmula del determinante para una matriz de Vandermonde en la última igualdad. Como los nodos son distintos, se tiene que la eliminación

de Neville de  $V$  puede hacerse sin cambios de filas. Por tanto, por la fórmula (4.5.6), se tiene  $m_{i0} = 1$ ,  $i = 1, \dots, n$ , y

$$\begin{aligned} m_{ij} &= \frac{p_{ij}}{p_{i-1,j}} = \frac{\prod_{i-j \leq k < l \leq i} (x_l - x_k)}{\prod_{i-j \leq k < l \leq i-1} (x_l - x_k)} \frac{\prod_{i-j-1 \leq k < l \leq i-2} (x_l - x_k)}{\prod_{i-j-1 \leq k < l \leq i-1} (x_l - x_k)} \\ &= \frac{\prod_{k=i-j}^{i-1} (x_i - x_k)}{\prod_{k=i-j-1}^{i-2} (x_{i-1} - x_k)} = \prod_{k=1}^j \frac{x_i - x_{i-k}}{x_{i-1} - x_{i-1-k}}, \end{aligned} \quad (4.5.8)$$

para  $0 \leq j < i \leq n$ .

*Nota 4.5.2.* Observemos que estos multiplicadores son invariantes por transformaciones afines de los nodos.

En los siguientes resultados probaremos que, para una matriz de Vandermonde con ciertas ordenaciones de los nodos, la eliminación de Neville con pivotaje parcial no requiere reordenación de filas. Esta propiedad es similar a la que cumple el orden de Leja para la eliminación de Gauss con pivotaje parcial.

**Teorema 4.5.3.** *Sea  $V(x_0, \dots, x_n)$  la matriz de Vandermonde dada por (4.5.3) con  $x_0, \dots, x_n$  nodos distintos ordenados crecientemente. Entonces la eliminación de Neville con pivotaje parcial de  $V(x_0, \dots, x_n)$  puede llevarse a cabo sin cambios de filas si y sólo si la sucesión de distancias  $d_i := x_{i+1} - x_i$ ,  $i = 0, \dots, n-1$ , es decreciente.*

*Demostración.* Supongamos que la eliminación de Neville con pivotaje parcial puede realizarse sin cambios de filas. Entonces, por (4.5.8),

$$|m_{i1}| = \frac{d_{i-1}}{d_{i-2}} \leq 1, \quad i = 2, \dots, n,$$

y se tiene una implicación. Para el recíproco, empecemos calculando la distancia entre dos nodos. Se tiene

$$x_i - x_{i-k} = \sum_{l=1}^k x_{i+1-l} - x_{i-l} = \sum_{l=1}^k d_{i-l}. \quad (4.5.9)$$

Por tanto, usando que los nodos están ordenados crecientemente y las fórmulas (4.5.8), (4.5.9), se tiene

$$|m_{ij}| = \prod_{k=1}^j \frac{x_i - x_{i-k}}{x_{i-1} - x_{i-1-k}} = \prod_{k=1}^j \frac{\sum_{l=1}^k d_{i-l}}{\sum_{l=1}^k d_{i-1-l}},$$

y, como la sucesión  $\{d_i\}_{i=0, \dots, n-1}$  es decreciente, se tiene  $\sum_{l=1}^k d_{i-l} \leq \sum_{l=1}^k d_{i-1-l}$ , de donde se deduce que  $|m_{ij}| \leq 1$ .  $\square$

Como consecuencia del Teorema 4.5.3 y la Nota 4.5.2 se pueden deducir las siguientes situaciones particulares en las que la eliminación de Neville con pivotaje parcial de  $V(x_0, \dots, x_n)$  no requiere reordenación de filas:

- $x_0, \dots, x_n$  equidistantes y en orden creciente o decreciente.
- $x_0, \dots, x_n$  en progresión geométrica con razón  $q < 1$  (véase la Sección 3.4).

El siguiente resultado ya no exige que los nodos estén ordenados crecientemente.

**Teorema 4.5.4.** *Sea  $V(x_0, \dots, x_n)$  la matriz de Vandermonde dada por (4.5.3) con  $x_0, \dots, x_n$  nodos tales que existe  $s \in \{-1, 1\}$  tal que*

$$d_i := |x_{i+1} - x_i| = s(-1)^i(x_{i+1} - x_i) > 0, \quad i = 0, \dots, n-1.$$

*Si  $\{d_i\}_{i=0, \dots, n-1}$  es una sucesión decreciente y  $\{d_i - d_{i+1}\}_{i=0, \dots, n-1}$  es decreciente entonces la eliminación de Neville con pivotaje parcial de  $V$  puede llevarse a cabo sin cambios de filas.*

*Demostración.* Observemos que

$$x_i - x_{i-k} = \sum_{l=0}^{k-1} (x_{i-l} - x_{i-l-1}) = s \sum_{l=0}^{k-1} (-1)^{i-l-1} d_{i-l-1} = s(-1)^{i-1} \sum_{l=0}^{k-1} (-1)^l d_{i-l-1}.$$

Teniendo en cuenta que las distancias  $d_i$  forman una sucesión decreciente, se tiene que

$$|x_i - x_{i-k}| = \sum_{l=0}^{k-1} (-1)^l d_{i-l-1} > 0. \quad (4.5.10)$$

Sea  $\Delta d_i := d_{i+1} - d_i$ ,  $i = 0, \dots, n-1$ . Como  $|\Delta d_i| = d_i - d_{i+1}$ , se tiene que  $\{|\Delta d_i|\}_{i=0, \dots, n-1}$  es una sucesión decreciente. Usando (4.5.10), se tiene

$$|x_i - x_{i-k}| = \begin{cases} d_{i-1} + \sum_{l=1}^{\lfloor k/2 \rfloor} |\Delta d_{i-1-2l}|, & k \text{ impar,} \\ \sum_{l=1}^{\lfloor k/2 \rfloor} |\Delta d_{i-2l}|, & k \text{ par.} \end{cases} \quad (4.5.11)$$

Ahora, por (4.5.11), se tiene que la sucesión  $\{|x_i - x_{i-k}|\}_{i=k, \dots, n}$  es decreciente en  $i$  para cada  $k$  fijo. Por tanto,

$$\prod_{k=1}^j |x_i - x_{i-k}| \leq \prod_{k=1}^j |x_{i-1} - x_{i-1-k}|.$$

Finalmente, usando (4.5.8), se deduce que  $|m_{ij}| \leq 1$ . □

En el siguiente resultado vamos a considerar nodos siguiendo el orden central inverso dado por (4.4.1). Veremos que la eliminación de Neville con pivotaje parcial se puede

hacer sin cambiar las filas y además, vamos a probar que la sucesión doble  $(|m_{ij}|)_{0 \leq j < i \leq n}$  es bimonótona decreciente o, equivalentemente,

$$|m_{j+1,j}| \geq |m_{j+2,j}| \geq \cdots \geq |m_{n,j}|, \quad j = 0, \dots, n-1, \quad (4.5.12)$$

y

$$|m_{i,0}| \geq |m_{i,1}| \geq \cdots \geq |m_{i,i-1}|, \quad i = 1, \dots, n. \quad (4.5.13)$$

Recordemos que una sucesión doble  $(a_{ij})_{i,j \in I}$ , con  $I \subset \mathbb{Z}^2$ , es *bimonótona decreciente* si  $a_{ij} \leq a_{i',j'}$  para todo par de índices  $(i,j), (i',j') \in I$  tales que  $i' \leq i, j' \leq j$ .

**Teorema 4.5.5.** *Sea  $V(x_0, \dots, x_n)$  la matriz de Vandermonde dada por (4.5.3) con  $x_0, \dots, x_n$  nodos equidistantes siguiendo el orden central inverso dado por (4.4.1). Entonces la eliminación de Neville con pivotaje parcial de  $V(x_0, \dots, x_n)$  puede llevarse a cabo sin cambios de filas y, además, la sucesión doble  $(|m_{ij}|)_{0 \leq j < i \leq n}$  es bimonótona decreciente.*

*Demostración.* En este caso, usando (4.4.1), tenemos que los  $d_i = (-1)^i(x_{i+1} - x_i) = (n-i)h$ ,  $i = 0, \dots, n-1$ , forman una sucesión decreciente. Además,  $|\Delta d_i| = h$ ,  $i = 0, \dots, n-1$ . Por tanto, por el Teorema 4.5.4, se tiene que la eliminación de Neville con pivotaje parcial de  $V(x_0, \dots, x_n)$  puede llevarse a cabo sin cambios de filas.

Veamos cuánto vale la distancia entre dos nodos cualesquiera. Si los dos nodos tienen la misma paridad se tiene, usando (4.4.1),

$$|x_i - x_k| = \frac{|i-k|}{2}h, \quad i \equiv k \pmod{2}, \quad (4.5.14)$$

y si tienen signo opuesto

$$|x_i - x_k| = \left(n - \frac{k+i-1}{2}\right)h, \quad i \not\equiv k \pmod{2}. \quad (4.5.15)$$

Vamos a calcular explícitamente el valor de los multiplicadores en este caso. Para ello, usando (4.5.14) y (4.5.15),

$$\begin{aligned} \prod_{k=i-j}^{i-1} |x_i - x_k| &= \prod_{k=1}^{\lfloor j/2 \rfloor} |x_i - x_{i-2k}| \prod_{k=1}^{\lfloor (j+1)/2 \rfloor} |x_i - x_{i+1-2k}| = h^j \prod_{k=1}^{\lfloor j/2 \rfloor} k \prod_{k=1}^{\lfloor (j+1)/2 \rfloor} (n-i+k) \\ &= h^j \frac{j!}{2^{\lfloor j/2 \rfloor}} \frac{(n-i + \lfloor (j+1)/2 \rfloor)!}{(n-i)!}. \end{aligned} \quad (4.5.16)$$

Por lo tanto, usando (4.5.16), podemos calcular explícitamente el valor absoluto de los multiplicadores

$$\begin{aligned} |m_{ij}| &= \frac{\prod_{k=i-j}^{i-1} |x_i - x_k|}{\prod_{k=i-1-j}^{i-2} |x_{i-1} - x_k|} = \frac{(n-i + \lfloor (j+1)/2 \rfloor)!}{(n-i)!} \frac{(n-i+1)!}{(n-i+1 + \lfloor (j+1)/2 \rfloor)!} \\ &= \frac{n-i+1}{n-i+1 + \lfloor (j+1)/2 \rfloor} \leq 1, \end{aligned} \quad (4.5.17)$$









## 4.6 Experimentos numéricos con el condicionamiento de las factorizaciones triangulares de una matriz de Vandermonde

Recordemos que el condicionamiento en norma  $\infty$  de una matriz no singular se define como

$$\kappa_{\infty}(A) := \|A\|_{\infty} \|A^{-1}\|_{\infty}.$$

Sea  $V$  la matriz de Vandermonde en los nodos  $x_0, \dots, x_n$  dada por (2.6.1) y  $V = LU$ , donde  $L$  es una matriz triangular inferior y  $U$  es una matriz triangular superior. El condicionamiento alto de la matriz de Vandermonde explica la dificultad de calcular los coeficientes del interpolante respecto a la base de monomios  $\mathbf{m} = (m_0, \dots, m_n)^T$ , donde  $m_i(x) = x^i, i = 0, \dots, n$ . En la Sección 22.1 de [37] encontramos una tabla con cotas y estimaciones de  $\kappa_{\infty}(V)$  para distintos nodos como nodos de Chebyshev, nodos equidistantes en distintos intervalos o nodos formando una sucesión armónica, es decir,  $x_i = 1/(i+1)$ . Si tenemos una factorización triangular de la matriz de Vandermonde, entonces podemos encontrar la solución resolviendo estos dos sistemas triangulares. Por lo tanto, el condicionamiento de cada una de estas matrices influyen en este método de resolución. Así, el producto de ambos condicionamientos proporciona una cota superior de la sensibilidad de la solución del sistema lineal respecto a las perturbaciones en los datos iniciales si se usa un algoritmo de factorización triangular.

Consideramos dos factorizaciones triangulares particulares de la matriz de Vandermonde  $V$ . La primera es la factorización de Crout  $V = LU$ , donde  $L$  y  $U$  están dadas por (2.2.6) y (2.6.5), respectivamente. La otra factorización que consideramos es la factorización de Doolittle,  $V = \tilde{L}\tilde{U}$ , con  $\tilde{L}, \tilde{U}$  dadas por (2.6.11). Aunque el operador de interpolación es el mismo, los coeficientes correspondientes a cada factorización están escalados de forma distinta y eso afecta al producto de condicionamiento de los factores triangulares.

Para nuestros experimentos hemos considerado dos tipos de intervalos para los nodos: centrados en el origen y empezando en cero. Hemos tomado los intervalos  $[-1, 1]$  y  $[-1.5, 1.5]$  como ejemplos del primer tipo y  $[0, 1]$  y  $[0, 3]$  como ejemplos del segundo tipo.

Vamos a tomar nodos equidistantes en  $[a, b]$

$$x_i := a + \sigma(i) \frac{b-a}{n}, \quad i = 0, \dots, n,$$

donde  $\sigma$  es una permutación del conjunto  $\{0, \dots, n\}$  correspondiente al orden asociado.

Consideramos cuatro órdenes para estos nodos:

- Creciente: los nodos ordenados crecientemente.
- Leja: los nodos ordenados de acuerdo a (i), (ii) y (iii) de la estrategia descrita en la Sección 4.1.
- Central: los nodos ordenados crecientemente según su distancia al centro del intervalo (el origen en el caso de intervalos simétricos).
- Central inverso: los nodos ordenados decrecientemente según su distancia al centro del intervalo (el origen en el caso de intervalos simétricos).

Veamos, antes de seguir con los experimentos, un ejemplo de nodos equidistantes con los distintos órdenes.

**Ejemplo 4.6.1.** *Seleccionemos  $[-2, 2]$ ,  $n = 6$  y  $h = 2/3$ . En ese caso, los diferentes órdenes dan lugar a las siguientes sucesiones de nodos.*

- *Creciente:*

$$x_0 = -2, \quad x_1 = -\frac{4}{3}, \quad x_2 = -\frac{2}{3}, \quad x_3 = 0, \quad x_4 = \frac{2}{3}, \quad x_5 = \frac{4}{3}, \quad x_6 = 2.$$

- *Leja:*

$$x_0 = 2, \quad x_1 = -2, \quad x_2 = 0, \quad x_3 = \frac{4}{3}, \quad x_4 = -\frac{4}{3}, \quad x_5 = \frac{2}{3}, \quad x_6 = -\frac{2}{3}.$$

- *Central:*

$$x_0 = 0, \quad x_1 = -\frac{2}{3}, \quad x_2 = \frac{2}{3}, \quad x_3 = -\frac{4}{3}, \quad x_4 = \frac{4}{3}, \quad x_5 = -2, \quad x_6 = 2.$$

- *Central inverso:*

$$x_0 = -2, \quad x_1 = 2, \quad x_2 = -\frac{4}{3}, \quad x_3 = \frac{4}{3}, \quad x_4 = -\frac{2}{3}, \quad x_5 = \frac{2}{3}, \quad x_6 = 0.$$

Ya vimos en la Sección 2.5 que la matriz de Vandermonde es una matriz totalmente positiva (TP) cuando  $0 \leq x_0 < \dots < x_n$ . Algunas propiedades de las matrices TP indican que la eliminación gaussiana sin reordenación de filas conduce a buenos resultados de estabilidad (ver [9] y [35]), lo que proporciona argumentos para trabajar con los nodos ordenados de menor a mayor.

Sin embargo, otras ordenaciones de los nodos también pueden dar lugar a estabilidad en los cálculos. Observamos que en la diagonal de la matriz  $L$  tenemos los pivotes de la eliminación gaussiana. Mediante una estrategia de pivotaje parcial, se intenta maximizar

en cada paso los pivotes. Notemos que el pivotaje parcial equivale a reordenar los nodos en el conjunto  $X$ . En [11],[36] y la Sección 4.5 se muestra que esta forma de ordenar los nodos es esencialmente el orden de Leja (la diferencia es que en el pivotaje parcial no se cambia el primer pivote). Es bien conocido que el pivotaje parcial controla el tamaño de los elementos de la matriz triangular inferior y, por tanto,  $\kappa_\infty(L)$  y  $\kappa_\infty(\tilde{L})$  serán menores con el orden de Leja.

En [46] se justifica el buen condicionamiento de  $\tilde{U}$ , y por tanto, por la Proposición 1.3.4, el buen condicionamiento de Skeel de  $U$ , para nodos positivos con el orden creciente. En dicho artículo se prueba que si existe una estrategia de pivotaje óptima para reducir el condicionamiento de Skeel de la matriz triangular superior  $U$ , entonces esta estrategia coincide con el pivotaje parcial escalado para una norma  $\|\cdot\|$  estrictamente monótona y en [35] se prueba que dicha estrategia aplicada a la eliminación gaussiana de una matriz TP no da lugar a cambios de filas. Una norma  $\|\cdot\|$  se dice estrictamente monótona si, para cualesquiera vectores  $\mathbf{u} = (u_1, \dots, u_n)$ ,  $\mathbf{v} = (v_1, \dots, v_n)$  con  $|u_j| \geq |v_j|$ ,  $\forall j = 1, \dots, n$  entonces  $\|\mathbf{u}\| \geq \|\mathbf{v}\|$  y, si para algún,  $j$ ,  $|u_j| > |v_j|$  entonces  $\|\mathbf{u}\| > \|\mathbf{v}\|$ . Los resultados mencionados proporcionan una justificación teórica del buen condicionamiento de  $U$  y  $\tilde{U}$  para el orden creciente.

En general, debemos tener en cuenta que la resolución del sistema con la matriz de Vandermonde pretende hallar los coeficientes respecto a la base de potencias centradas en el origen. Considerar otros centros distintos del origen puede mejorar el condicionamiento de la matriz  $L$ , pero no tiene por qué mejorar el de  $V$  o el producto de los condicionamientos de los factores triangulares. Por ello, los órdenes basados en el centro del intervalo pueden no ser apropiados para la matriz de Vandermonde cuando el intervalo no está centrado en el origen.

En la Tabla 4.6.1 podemos ver que el mejor orden de nodos para la factorización de Crout  $LU$  es el orden creciente. Por otro lado, la Tabla 4.6.2 muestra que el mejor condicionamiento para la factorización de Doolittle corresponde al orden de Leja. No se observan diferencias significativas entre la factorización  $\tilde{L}\tilde{U}$  con el orden de Leja y la factorización  $LU$  con el orden creciente.

Observamos en las tablas 4.6.3 y 4.6.4 que las diferencias entre el condicionamiento de la matriz de Vandermonde  $V$  y el producto de los condicionamientos asociados a las factorizaciones, con las distintas ordenaciones, no son tan pronunciadas como las correspondientes al intervalo  $[0, 1]$ . Además, las diferencias entre las dos factorizaciones y las distintas ordenaciones son menos significativas para el intervalo  $[0, 3]$ . El Teorema 3.2.8 tiene relación con este fenómeno.

Las Tablas 4.6.5 y 4.6.6 corresponden al intervalo  $[-1, 1]$  centrado en el origen. Vemos

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(L)\kappa_\infty(U)$			
3	216	416	672	657.3333	576
4	$1.7067 \times 10^3$	$3.4333 \times 10^3$	$5.7600 \times 10^3$	$6.3984 \times 10^3$	$6.804 \times 10^3$
5	$1.2500 \times 10^4$	$3.3700 \times 10^4$	$5.6386 \times 10^4$	$5.4831 \times 10^4$	$6.772 \times 10^4$
9	$4.8184 \times 10^7$	$3.3789 \times 10^8$	$7.6455 \times 10^8$	$6.3028 \times 10^8$	$6.6891 \times 10^8$
19	$5.0877 \times 10^{16}$	$5.3085 \times 10^{18}$	$2.7587 \times 10^{19}$	$2.2446 \times 10^{19}$	$3.3275 \times 10^{19}$

Cuadro 4.6.1:  $\kappa_\infty(L)\kappa_\infty(U)$  in  $[0, 1]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$			
3	216	576	800	640	720
4	$1.7067 \times 10^3$	$6.8267 \times 10^3$	$7.6800 \times 10^3$	$1.0271 \times 10^4$	$7.488 \times 10^3$
5	$1.2500 \times 10^4$	$9.0667 \times 10^4$	$6.0450 \times 10^4$	$6.3904 \times 10^4$	$6.0667 \times 10^4$
9	$4.8184 \times 10^7$	$3.6280 \times 10^9$	$4.1542 \times 10^8$	$1.2597 \times 10^9$	$9.7049 \times 10^8$
19	$5.0877 \times 10^{16}$	$1.9664 \times 10^{21}$	$6.3140 \times 10^{17}$	$1.2392 \times 10^{20}$	$3.0774 \times 10^{19}$

Cuadro 4.6.2:  $\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$  in  $[0, 1]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$			
3	266.6667	512	2240	3300	1040
4	$2.8681 \times 10^3$	$6.8661 \times 10^3$	$2.2511 \times 10^4$	$2.9200 \times 10^4$	$1.100 \times 10^4$
5	$2.6963 \times 10^4$	$7.2298 \times 10^4$	$3.2855 \times 10^5$	$3.0184 \times 10^5$	$1.2224 \times 10^5$
9	$4.0049 \times 10^8$	$1.4074 \times 10^9$	$5.5271 \times 10^9$	$7.0915 \times 10^9$	$6.0662 \times 10^9$
19	$2.0837 \times 10^{19}$	$1.6470 \times 10^{20}$	$3.2911 \times 10^{20}$	$3.2084 \times 10^{21}$	$2.2280 \times 10^{21}$

Cuadro 4.6.3:  $\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$  in  $[0, 3]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(\tilde{\tilde{L}})\kappa_\infty(\tilde{\tilde{U}})$			
3	266.6667	938.6667	$1.1852 \times 10^3$	$1.9413 \times 10^3$	$1.6756 \times 10^3$
4	$2.8681 \times 10^3$	$9.9000 \times 10^3$	$1.5775 \times 10^4$	$2.1088 \times 10^4$	$1.8000 \times 10^4$
5	$2.6963 \times 10^4$	$1.1612 \times 10^5$	$1.6111 \times 10^5$	$3.7502 \times 10^5$	$1.7919 \times 10^5$
9	$4.0049 \times 10^8$	$2.2396 \times 10^9$	$3.9113 \times 10^9$	$6.3144 \times 10^9$	$9.2406 \times 10^9$
19	$2.0837 \times 10^{19}$	$2.3145 \times 10^{20}$	$2.6917 \times 10^{20}$	$1.0480 \times 10^{21}$	$1.3634 \times 10^{22}$

Cuadro 4.6.4:  $\kappa_\infty(\tilde{\tilde{L}})\kappa_\infty(\tilde{\tilde{U}})$  in  $[0, 3]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(L)\kappa_\infty(U)$			
3	18	327.5556	193.3333	78	193.3333
4	53.3333	$2.2295 \times 10^3$	580	202.6667	750
5	187.5000	$1.2696 \times 10^4$	$2.0157 \times 10^3$	754.2080	$2.5037 \times 10^3$
9	$2.0562 \times 10^4$	$1.5169 \times 10^7$	$2.1156 \times 10^5$	$1.1113 \times 10^5$	$3.8812 \times 10^5$
19	$1.7511 \times 10^9$	$8.6376 \times 10^{14}$	$5.8935 \times 10^{10}$	$1.9575 \times 10^{10}$	$2.0915 \times 10^{11}$

Cuadro 4.6.5:  $\kappa_\infty(L)\kappa_\infty(U)$  in  $[-1, 1]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$			
3	18	784	113.3333	96	113.3333
4	53.3333	$7.4667 \times 10^3$	360	588	472.5
5	187.5000	$6.1512 \times 10^4$	$1.2788 \times 10^3$	$2.4920 \times 10^3$	$1.3300 \times 10^3$
9	$2.0562 \times 10^4$	$3.0868 \times 10^8$	$1.6768 \times 10^5$	$1.4509 \times 10^6$	$4.8092 \times 10^5$
19	$1.7511 \times 10^9$	$4.5246 \times 10^{17}$	$2.2233 \times 10^{10}$	$2.2128 \times 10^{13}$	$1.1110 \times 10^{12}$

Cuadro 4.6.6:  $\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$  in  $[-1, 1]$ .

que  $\kappa_\infty(L)\kappa_\infty(U)$  es menor con el orden central, seguido por el de Leja. Con el orden central inverso, se obtienen resultados similares aunque algo peores. La Tabla 4.6.6 muestra que la factorización de Doolittle tiene un mejor comportamiento con el orden de Leja.

La Tabla 4.6.7 muestra  $\kappa_\infty(L)\kappa_\infty(U)$  para la factorización  $LU$  con las cuatro ordenaciones en el intervalo  $[-1.5, 1.5]$ . Vemos que el orden con el cual se obtienen los valores más bajos es con el orden central seguido de cerca por el orden de Leja. En la Tabla 4.6.8 consideramos la factorización  $\tilde{L}\tilde{U}$  en el mismo intervalo. En este caso, el orden de Leja da mejores resultados para  $\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$ .

Las conclusiones de los experimentos podemos resumirlos como sigue. Para la factorización de Crout las mejores elecciones son el orden central, para intervalos centrados en el origen, y el orden creciente, para los intervalos que empiezan en cero. Para la fac-

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(L)\kappa_\infty(U)$			
3	18.9583	942.5000	381.8750	125	381.8750
4	62.5185	$7.1397 \times 10^3$	905.6909	4.2193	$1.6487 \times 10^3$
5	192.4190	$5.1245 \times 10^4$	$3.4541 \times 10^3$	$1.7291 \times 10^3$	$5.6985 \times 10^3$
9	$1.8257 \times 10^4$	$1.3326 \times 10^8$	$4.4569 \times 10^5$	$2.0864 \times 10^5$	$1.0871 \times 10^6$
19	$2.3696 \times 10^9$	$7.0152 \times 10^{16}$	$8.4964 \times 10^{10}$	$5.9198 \times 10^{10}$	$9.4327 \times 10^{11}$

Cuadro 4.6.7:  $\kappa_\infty(L)\kappa_\infty(U)$  in  $[-1.5, 1.5]$ .

		Creciente	Leja	Central	Central inverso
n	$\kappa_\infty(V)$	$\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$			
3	18.9583	$1.5275 \times 10^3$	276.2500	324	276.25
4	62.5185	$1.3504 \times 10^4$	633	$1.9845 \times 10^3$	801.1406
5	192.4190	$1.4203 \times 10^5$	$1.6537 \times 10^3$	$8.2685 \times 10^3$	$2.5032 \times 10^3$
9	$1.8257 \times 10^4$	$1.4643 \times 10^9$	$2.7830 \times 10^5$	$5.7955 \times 10^6$	$5.9971 \times 10^5$
19	$2.3696 \times 10^9$	$2.0495 \times 10^{19}$	$4.1203 \times 10^{10}$	$6.7926 \times 10^{13}$	$1.9753 \times 10^{12}$

Cuadro 4.6.8:  $\kappa_\infty(\tilde{L})\kappa_\infty(\tilde{U})$  in  $[-1.5, 1.5]$ .

torización de Doolittle la mejor elección corresponde al orden de Leja. Finalmente, para los intervalos centrados en el origen el orden creciente da los peores resultados.



# 5

## Representaciones con bases ortogonales

---

### 5.1 Aproximación por mínimos cuadrados

Presentemos a continuación el problema de la mejor aproximación por mínimos cuadrados en un conjunto discreto.

**Problema de la mejor aproximación por mínimos cuadrados.** Sea  $F$  un espacio de funciones,  $U$  un subespacio de dimensión  $n+1$  y  $f \in F$ . Dados nodos  $x_0, \dots, x_N$ ,  $n \leq N$ , y pesos  $w_0, \dots, w_N$ ,  $w_i \geq 0$ ,  $i = 0, \dots, N$ , se busca una función en  $U$  que minimice

$$\sum_{i=0}^N w_i (f(x_i) - u(x_i))^2, \quad u \in U.$$

Para resolverlo introducimos la forma bilineal simétrica y semidefinida positiva en  $F$

$$\langle u, v \rangle = \sum_{i=0}^N w_i u(x_i) v(x_i).$$

Tomando una base  $u_0, \dots, u_n$  de  $U$  podemos expresar cualquier función de  $U$  de la forma  $\sum_{i=0}^n c_i u_i$ . El problema de aproximación por mínimos cuadrados se transforma en la minimización de la función cuadrática en  $\mathbb{R}^n$

$$V(c) = a_0 - 2a^T c + c^T G c, \quad c \in \mathbb{R}^n,$$

donde  $a_0 = \langle f, f \rangle$ ,  $a = (\langle u_0, f \rangle, \dots, \langle u_n, f \rangle)^T$  y  $G$  es la matriz de Gram de la base  $u_0, \dots, u_n$  cuyo elemento  $(i, j)$  es  $\langle u_i, u_j \rangle$ . Es bien conocido que la función cuadrática admite un único mínimo si y sólo si  $G$  es una matriz definida positiva, lo que equivale a afirmar que la forma bilineal simétrica es definida positiva en  $U$ . La única solución  $u^*$  del problema puede obtenerse resolviendo las ecuaciones normales

$$\sum_{j=0}^n \langle u_i, u_j \rangle c_j = \langle u_i, f \rangle, \quad i = 0, \dots, n,$$

que equivalen a imponer la condición

$$\langle f - u^*, u \rangle = 0, \quad \forall u \in U.$$

Decimos que dos funciones  $f, g$  son ortogonales si  $\langle f, g \rangle = 0$ . Observemos que la condición anterior equivale a afirmar que el error es ortogonal a las funciones del espacio  $U$ .

Como la forma es definida positiva en el espacio  $U$ , existirá una base ortonormal  $q_0, \dots, q_n$ , es decir,

$$\langle q_i, q_j \rangle = \delta_{ij}$$

y la única solución del problema puede expresarse en la forma

$$u^* = \sum_{i=0}^n \langle q_i, f \rangle q_i.$$

Sea  $Q$  la aplicación que a cada función le asocia su mejor aproximación en  $U$ . Entonces

$$Q[f] = \sum_{i=0}^n \langle q_i, f \rangle q_i,$$

lo que implica que  $Q$  es un operador lineal. La suma del miembro derecho de la fórmula anterior recibe el nombre de suma de Fourier de  $f$  respecto a la base ortonormal  $(q_0, \dots, q_n)$ . Como  $Q[u] = u$ , para todo  $u \in U$ , se tiene que  $Q$  es una proyección y, por tanto, un operador de cuasi-interpolación en el sentido de la Sección 1.1.

En el caso de que  $n = N$  y exista siempre solución única para el problema de interpolación de Lagrange en  $U$  en los nodos  $x_0, \dots, x_n$ , la solución del problema de mínimos cuadrados coincide con el operador de interpolación del problema de Lagrange. En el caso de que  $N > n$ , la solución del problema no interpola propiamente en un sentido estricto y se dice que la solución es un hiperinterpolante, en el sentido de que se ajusta lo máximo posible en los nodos a la función.

La discusión anterior es válida en el caso continuo reemplazando la forma bilineal simétrica por

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x) dx,$$

donde  $w \geq 0$  es una función peso tal que  $|wfg|$  son integrables para todas las funciones  $f, g \in F$ .

Los ejemplos anteriores se generalizan considerando productos escalares de la forma

$$\langle f, g \rangle = \int_a^b f(x) g(x) d\mu(x),$$

donde  $\mu$  es una medida no negativa y finita, de modo que las funciones consideradas son integrables.

## 5.2 Representaciones con bases ortogonales y normas

Vamos a analizar los operadores de cuasi-interpolación asociados a problemas de mínimos cuadrados. El espacio  $U \subset F$  en el que buscamos la solución del problema de mínimos cuadrados debe tener una base ortonormal, ya que la forma bilineal  $\langle \cdot, \cdot \rangle$  es definida positiva en  $U$ . La representación de  $Q$  respecto a la base ortonormal  $(q_0, \dots, q_n)$  es

$$Q[f] = \sum_{i=0}^n \langle q_i, f \rangle q_i.$$

Para una función  $f \in F$  definimos

$$\|f\|_2 := \langle f, f \rangle^{1/2}.$$

Tenemos que el funcional  $\|\cdot\|_2$  es una seminorma en  $F$  y una norma en el subespacio  $U$  en el sentido de que si  $u \in U$ ,  $u \neq 0$ , entonces  $\|u\|_2 > 0$ .

En el caso en el que la base  $(p_0, \dots, p_n)$  de  $U$  es ortogonal, es decir,  $\langle p_i, p_j \rangle = 0$  cuando  $i \neq j$ , la representación del operador de cuasi-interpolación es de la forma

$$Q[f] := \sum_{i=0}^n \frac{\langle f, p_i \rangle}{\langle p_i, p_i \rangle} p_i.$$

Definamos los funcionales

$$\alpha_i f := \frac{\langle f, p_i \rangle}{\langle p_i, p_i \rangle}, \quad i = 0, \dots, n.$$

Podemos definir la norma del funcional en el sentido de mínimos cuadrados mediante

$$\|\alpha_i\|_2 := \sup_{\|f\|_2=1} |\alpha_i f| = \sup_{\|f\|_2=1} \left| \frac{\langle f, p_i \rangle}{\langle p_i, p_i \rangle} \right|.$$

De la desigualdad de Cauchy-Schwarz, se deduce que

$$\langle f, p_i \rangle^2 \leq \langle f, f \rangle \langle p_i, p_i \rangle$$

y por tanto,

$$\|\alpha_i\|_2 = \langle p_i, p_i \rangle^{-1/2} = \frac{1}{\|p_i\|_2}.$$

es decir, la norma del funcional coincide con el inverso de la norma del elemento de la base correspondiente.

En este contexto tiene sentido definir la norma del cuasi-interpolante en el sentido de mínimos cuadrados

$$\|Q\|_2 := \sup_{\|f\|_2=1} \|Q[f]\|_2.$$

Como  $Q[f]$  y  $f - Q[f]$  son ortogonales, tenemos

$$\|f\|_2^2 = \|Q[f]\|_2^2 + \|f - Q[f]\|_2^2,$$

de donde se deduce que

$$\|Q[f]\|_2 \leq \|f\|_2$$

y, como  $Q$  fija las funciones del espacio  $U$ , tenemos

$$\|Q\|_2 = 1.$$

El hecho de que la norma cuadrática sea la más baja posible es un indicio de que la norma  $\infty$  de  $Q$  puede tomar valores bajos.

### 5.3 Mínimos cuadrados en el espacio de funciones continuas

Un caso particular importante es el caso en el que  $F = C[a, b]$  dotado con una forma bilineal simétrica de la forma

$$\langle f, g \rangle = \int_a^b f(x)g(x) d\mu(x), \quad (5.3.1)$$

donde  $\mu$  es una medida regular de Borel no negativa tal que  $0 < \mu[a, b] < \infty$ , y  $U = P_n$ , el espacio de polinomios de grado menor o igual que  $n$ . Este producto escalar da lugar a la norma

$$\|f\|_2 = \left( \int_a^b f(x)^2 d\mu(x) \right)^{1/2}.$$

En la sección anterior hemos discutido las normas de las funciones de  $U$ , los funcionales y el operador de cuasi-interpolación de mínimos cuadrados

$$Q[f] = \sum_{i=0}^n \alpha_i f p_i,$$

donde

$$\alpha_i f = \frac{\langle p_i, f \rangle}{\langle p_i, p_i \rangle}, \quad i = 0, \dots, n,$$

y  $p_0, \dots, p_n$  es una base ortogonal de  $U$ . En el análisis numérico de la estabilidad de las aproximaciones conviene utilizar el error máximo en lugar del error cuadrático medio.

Vamos a proporcionar una fórmula de Lebesgue para este tipo de problemas.

**Teorema 5.3.1.** Sea  $q_0, \dots, q_n$  una base ortonormal de  $U$  y

$$Q[f] = \int_a^b \sum_{i=0}^n q_i(y)q_i(x)f(y) d\mu(y),$$

el operador de cuasi-interpolación y sea

$$\lambda(x) := \int_a^b \left| \sum_{i=0}^n q_i(y)q_i(x) \right| d\mu(y).$$

Entonces

$$\|Q\|_\infty = \Lambda := \max_{x \in [a,b]} \lambda(x).$$

*Demostración.* Tomando módulos, tenemos

$$|Q[f](x)| \leq \|f\|_\infty \int_a^b \left| \sum_{i=0}^n q_i(y)q_i(x) \right| d\mu(y) = \|f\|_\infty \lambda(x).$$

de donde se deduce que  $\|Q\|_\infty \leq \Lambda$ . Sea  $\xi$  el punto en el que  $\lambda(x)$  alcanza su valor máximo en  $[a, b]$ . Tomando una función  $f(y)$  de norma 1 que aproxime a la función  $\text{sign} \sum_{i=0}^n q_i(y)q_i(\xi)$ , obtenemos el resultado.  $\square$

Notemos que el Teorema 5.3.1 es una generalización al problema de mínimos cuadrados del Teorema 2.1.2.

Recordemos que  $\|\alpha_i\|_\infty = \sup_{\|f\|_\infty=1} |\alpha_i f|$ . Enunciamos a continuación la siguiente consecuencia del Teorema 1.2.2.

**Proposición 5.3.2.** Sea  $q \in C[a, b]$  y  $\alpha : C[a, b] \rightarrow \mathbb{R}$  el funcional  $\alpha f = \int_a^b f q d\mu$ , donde  $\mu$  es una medida de Borel regular no negativa. Entonces

$$\|\alpha\|_\infty = \|q\|_1 := \int_a^b |q| d\mu.$$

Aplicando el resultado anterior a

$$\alpha_i f = \frac{\int_a^b p_i f d\mu}{\int_a^b p_i^2 d\mu},$$

obtenemos

$$\|\alpha_i\|_\infty = \frac{\int_a^b |p_i| d\mu}{\int_a^b p_i^2 d\mu} = \frac{\|p_i\|_1}{\|p_i\|_2^2}, \quad (5.3.2)$$

y de la desigualdad de Cauchy-Schwarz

$$\|p_i\|_1 = \int_a^b |p_i| d\mu \leq \|p_i\|_2 \|1\|_2 = \|p_i\|_2 (\mu[a, b])^{1/2} \quad (5.3.3)$$

deducimos

$$\|\alpha_i\|_\infty = \frac{\|p_i\|_1}{\|p_i\|_2^2} \leq \frac{(\mu[a, b])^{1/2}}{\|p_i\|_2}. \quad (5.3.4)$$

Por tanto, podemos expresar el condicionamiento de la representación que usa los funcionales  $\alpha_0, \dots, \alpha_n$  del siguiente modo.

**Proposición 5.3.3.** *Sea  $(p_0, \dots, p_n)$  una base ortogonal y  $Q[f] = \sum_{i=0}^n \alpha_i f p_i$ , donde  $\alpha_i f = \langle p_i, f \rangle / \|p_i\|_2^2$ ,  $i = 0, \dots, n$ . Entonces*

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \frac{\|p_i\|_1 |p_i(x)|}{\|p_i\|_2^2} = \sum_{i=0}^n \|q_i\|_1 |q_i(x)|,$$

donde  $q_i := p_i / \|p_i\|_2$ ,  $i = 0, \dots, n$ , es la correspondiente base ortonormal. Además, se tiene la cota

$$\max_{x \in [a, b]} \text{cond}(x; \alpha) \leq (\mu[a, b])^{1/2} \sum_{i=0}^n \frac{\|p_i\|_\infty}{\|p_i\|_2} = (\mu[a, b])^{1/2} \sum_{i=0}^n \|q_i\|_\infty. \quad (5.3.5)$$

*Demostración.* Por la definición (1.2.2)

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \|\alpha_i\|_\infty |p_i(x)|,$$

y aplicando la fórmula (5.3.2), obtenemos el resultado.

Para acotar uniformemente  $\text{cond}(x; \alpha)$  usamos (5.3.4), obteniendo

$$\text{cond}(x; \alpha) \leq (\mu[a, b])^{1/2} \sum_{i=0}^n \frac{|p_i(x)|}{\|p_i\|_2} = (\mu[a, b])^{1/2} \sum_{i=0}^n |q_i(x)|,$$

de donde se deduce la desigualdad (5.3.5).  $\square$

En el siguiente resultado comparamos el condicionamiento de una representación  $Q[f] = \sum_{i=0}^n \beta_i f v_i$  con una representación respecto a una base ortogonal  $\sum_{i=0}^n \alpha_i f p_i$ .

**Teorema 5.3.4.** *Sea  $(p_0, \dots, p_n)$  una base ortogonal y  $Q[f] = \sum_{i=0}^n \alpha_i f p_i$ , donde  $\alpha_i f = \langle p_i, f \rangle / \|p_i\|_2^2$ ,  $i = 0, \dots, n$ . Sea  $\sum_{i=0}^n \beta_i f v_i$  otra representación del cuasi-interpolante  $Q$  y  $K = (\langle p_i, v_j \rangle / \|p_i\|_1)_{0 \leq i, j \leq n}$ . Entonces*

$$\text{cond}(x; \beta, v) \leq \text{Cond}(K^{-1}) \text{cond}(x; \alpha, p).$$

*Demostración.* Por el Corolario 1.3.8,  $\text{cond}(x; \beta, v) \leq \text{Cond}(M^{-1}D_\alpha) \text{cond}(x; \alpha, p)$ , donde  $M$  es la matriz de cambio de base cuyo elemento  $(i, j)$  es

$$m_{ij} = \alpha_i v_j = \frac{\langle p_i, v_j \rangle}{\|p_i\|_2^2}$$

y  $D_\alpha := \text{diag}(\|\alpha_0\|_\infty, \dots, \|\alpha_n\|_\infty)$ . Por la fórmula (5.3.2),  $\|\alpha_i\|_\infty = \|p_i\|_1 / \|p_i\|_2^2$ . Por tanto,  $K = D_\alpha^{-1}M$ .  $\square$

## 5.4 Condicionamiento con polinomios ortogonales clásicos

Los polinomios ortogonales clásicos se usan en problemas de mínimos cuadrados con productos escalares del tipo (5.3.1), donde  $d\mu(x) = w(x) dx$ . Sin pérdida de generalidad se considera el caso estándar de funciones continuas en el intervalo  $[-1, 1]$ . Se define el polinomio ortogonal de grado  $n$  como aquel polinomio de grado exactamente  $n$  que es ortogonal a todos los de grado inferior. Las bases de polinomios ortogonales son bases con excelentes propiedades de estabilidad y vamos a confirmar esta afirmación calculando cotas para los condicionamientos.

Vamos a aplicar la Proposición 5.3.3 y la cota (5.3.5) para algunos de los polinomios clásicos ortogonales.

### Polinomios de Legendre

Los polinomios de Legendre  $P_n(x)$  son los polinomios ortogonales en el intervalo  $(-1, 1)$  respecto a la función  $w(x) = 1$  que verifican la condición de normalización  $P_n(1) = 1$ .

Veamos el siguiente resultado que nos permitirá acotar el condicionamiento usando la representación con polinomios de Legendre.

**Lema 5.4.1.** *Sea  $n$  un entero no negativo. Se tienen las siguientes desigualdades*

$$\frac{2}{3}\sqrt{2(n+1)^3} \leq \sum_{k=0}^n \sqrt{2k+1} \leq \frac{2}{3}\sqrt{2n(n+1)(n+2)} + 1.$$

*Demostración.* Sea  $f \in C^2[a, b]$  y  $x_0, x_1 \in [a, b]$ . Si  $f'' \leq 0$  se tiene, al aplicar la fórmula del punto medio,

$$\int_{x_0}^{x_1} f(x) dx \leq (x_1 - x_0) f\left(\frac{x_1 + x_0}{2}\right).$$

Tomando  $f = \sqrt{x}$ ,  $x_0 = 2k$ ,  $x_1 = 2k + 2$ , se tiene

$$\int_{x_0}^{x_1} \sqrt{x} dx = \frac{(2k+2)^{3/2} - (2k)^{3/2}}{3/2} \leq 2\sqrt{2k+1}, \quad k = 0, 1, \dots,$$

es decir,

$$\frac{2\sqrt{2(k+1)^3} - 2\sqrt{2k^3}}{3} \leq \sqrt{2k+1}, \quad k = 0, 1, \dots$$

Definimos  $g_k := 2\sqrt{2k^3}/3$ ,  $k = 0, 1, \dots$ , y sumando obtenemos la cota inferior

$$\frac{2}{3}\sqrt{2(n+1)^3} = g_{n+1} - g_0 = \sum_{k=0}^n (g_{k+1} - g_k) \leq \sum_{k=0}^n \sqrt{2k+1}.$$

Para la cota superior, definimos  $f_k := \sqrt{k(k+1)(k+2)}$ ,  $k = 0, 1, \dots$ . Se tiene

$$\begin{aligned} f_{k+1} - f_k &= \sqrt{(k+1)(k+2)(k+3)} - \sqrt{k(k+1)(k+2)} \\ &= \sqrt{(k+1)(k+2)(\sqrt{k+3} - \sqrt{k})} = \sqrt{(k+1)(k+2)} \frac{3}{\sqrt{k+3} + \sqrt{k}}. \end{aligned}$$

Sea  $p := k + 3/2$ , entonces

$$f_{k+1} - f_k = \frac{3\sqrt{p^2 - 1/4}}{\sqrt{p - 3/2} + \sqrt{p + 3/2}}.$$

Veamos que  $f_{k+1} - f_k \geq 3\sqrt{p}/2$ ,  $k = 0, 1, \dots$ . En efecto, la desigualdad

$$\frac{3\sqrt{p^2 - 1/4}}{\sqrt{p - 3/2} + \sqrt{p + 3/2}} \geq 3\sqrt{p}/2$$

equivale a

$$2\sqrt{p^2 - 1/4} \geq \sqrt{p^2 - 3p/2} + \sqrt{p^2 + 3p/2}.$$

Elevando al cuadrado

$$4(p^2 - 1/4) \geq 2p^2 + 2\sqrt{p^4 - 9p^2/4},$$

aislando el radical y elevando al cuadrado de nuevo, obtenemos

$$4p^4 - 4p^2 + 1 \geq 4p^4 - 9p^2,$$

y claramente la diferencia entre los dos términos es positiva.

Aplicando lo anterior, obtenemos

$$\begin{aligned} \sqrt{n(n+1)(n+2)} &= f_n - f_0 = \sum_{k=0}^{n-1} (f_{k+1} - f_k) \geq \frac{3}{2} \sum_{k=0}^{n-1} \sqrt{p} = \frac{3}{2\sqrt{2}} \sum_{k=0}^{n-1} \sqrt{2k+3} \\ &= \frac{3}{2\sqrt{2}} \sum_{k=1}^n \sqrt{2k+1}, \end{aligned}$$

y por tanto, se tiene el resultado

$$\sum_{k=0}^n \sqrt{2k+1} \leq \frac{2}{3} \sqrt{2n(n+1)(n+2)} + 1.$$

□

Es bien conocido (cf. Teorema 5.6.1 [5]) que el máximo valor de los polinomios de Legendre en  $[-1, 1]$  se alcanza en  $x = 1$  y que

$$\|P_n\|_\infty = 1. \tag{5.4.1}$$

La norma 2 de los polinomios de Legendre es (cf. Sección 5.1 de [5])

$$\|P_n\|_2 = \sqrt{\int_{-1}^1 (P_n(x))^2 dx} = \sqrt{\frac{2}{2n+1}}. \quad (5.4.2)$$

Además,

$$\|1\|_2^2 = \mu(-1, 1) = \int_{-1}^1 dx = 2. \quad (5.4.3)$$

**Proposición 5.4.2.** *El condicionamiento de la representación  $\sum_{i=0}^n \alpha_i f P_i$ , donde  $P_i$  es el polinomio de Legendre de grado  $i$  y  $\alpha_i f = \int_{-1}^1 P_i(x) f(x) dx$ ,  $i = 0, \dots, n$ , verifica*

$$\text{cond}(x; \alpha) \leq \frac{2}{3} \sqrt{2n(n+1)(n+2)} + 1.$$

*Demostración.* Por las fórmulas (5.3.5), (5.4.1), (5.4.2), (5.4.3) y por el Lema 5.4.1

$$\begin{aligned} \text{cond}(x; \alpha) &\leq (\mu(-1, 1))^{1/2} \sum_{i=0}^n \frac{\|P_i\|_\infty}{\|P_i\|_2} = \sqrt{2} \sum_{i=0}^n \sqrt{\frac{2i+1}{2}} \\ &= \sum_{i=0}^n \sqrt{2i+1} \leq \frac{2}{3} \sqrt{2n(n+1)(n+2)} + 1. \end{aligned}$$

□

La cota anterior presupone que la base de polinomios de Legendre puede calcularse con un algoritmo fiable. Sin embargo, la relación de recurrencia a tres términos que se utiliza normalmente para su cálculo

$$\begin{cases} P_{n+1}(x) = \frac{2n-1}{n} P_n(x) - \frac{n-1}{n} P_{n-1}(x), & n \geq 1 \\ P_0(x) = 1, & P_1(x) = x, \end{cases}$$

es inestable. Vamos a calcular una cota de la propagación del error al aplicar la fórmula de recurrencia a tres términos. Sea  $\hat{P}_n$  una aproximación numérica del polinomio de Legendre  $P_n$  calculada con la fórmula de recurrencia. Denotamos por  $\varepsilon_{n+1}$  al error acumulado en el paso  $n$ , es decir,

$$\hat{P}_{n+1}(x) = \frac{2n-1}{n} \hat{P}_n(x) - \frac{n-1}{n} \hat{P}_{n-1}(x) + \varepsilon_{n+1}.$$

Definimos  $e_n(x) := \hat{P}_n(x) - P_n(x)$ . Entonces, se tiene

$$\begin{cases} e_{n+1}(x) = \frac{2n-1}{n} e_n(x) - \frac{n-1}{n} e_{n-1}(x) + \varepsilon_{n+1}, & n \geq 1 \\ e_0 = \varepsilon_0, & e_1 = \varepsilon_1. \end{cases} \quad (5.4.4)$$

Sea  $\varepsilon$  una cota para  $|\varepsilon_j|$ ,  $j \geq 0$ . Usando (5.4.4), se deduce

$$|e_{n+1}(x)| \leq \frac{2n-1}{n}|e_n(x)| + \frac{n-1}{n}|e_{n-1}(x)| + \varepsilon \leq 2|e_n(x)| + |e_{n-1}(x)| + \varepsilon. \quad (5.4.5)$$

Veamos por inducción que se tiene la siguiente cota

$$|e_n(x)| \leq \frac{3\varepsilon}{2} (1 + \sqrt{2})^n - \frac{\varepsilon}{2}.$$

Para  $n = 0, 1$  es fácil comprobar que se cumple la desigualdad. Supongamos que se verifica para  $n$  y demostrémoslo para  $n + 1$ . Usando (5.4.5) y la hipótesis de inducción, se tiene

$$\begin{aligned} |e_{n+1}(x)| &\leq 2|e_n(x)| + |e_{n-1}(x)| \leq 2\frac{3\varepsilon}{2} (1 + \sqrt{2})^n - \varepsilon + \frac{3\varepsilon}{2} (1 + \sqrt{2})^{n-1} - \frac{\varepsilon}{2} + \varepsilon \\ &= \frac{3\varepsilon}{2} (1 + \sqrt{2})^{n-1} (3 + 2\sqrt{2}) - \frac{\varepsilon}{2}, \end{aligned}$$

y, teniendo en cuenta que  $3 + 2\sqrt{2} = (1 + \sqrt{2})^2$ , se deduce la cota para cualquier  $n$ .

Una posible estrategia para evitar la inestabilidad del cálculo a través de la fórmula de recurrencia consiste en expresar los polinomios de Legendre en término de los polinomios de Chebyshev. Por tanto, parece conveniente estudiar directamente representaciones en términos de los polinomios de Chebyshev, que pueden obtenerse a través de la fórmula explícita  $T_n(x) = \cos(n \arccos x)$ , fórmula que ofrece cálculos estables.

### Polinomios de Chebyshev de primera especie

Los polinomios de Chebyshev de primera especie  $T_n(x)$  son los polinomios ortogonales en el intervalo  $(-1, 1)$  respecto a la función  $w(x) = 1/\sqrt{1-x^2}$  que verifican la condición de normalización  $T_n(1) = 1$ .

Es bien conocido que  $T_n(x) = \cos(n \arccos x)$ , de donde se deduce

$$\|T_n\|_{\infty} = 1.$$

La norma 1 viene dada por

$$\|T_0\|_1 = \pi,$$

y

$$\|T_n\|_1 = \int_0^{\pi} |\cos(n\theta)| d\theta = \sum_{i=0}^{n-1} \int_{(2i-1)\pi/2n}^{(2i+1)\pi/2n} (-1)^i \cos(n\theta) d\theta = \sum_{i=0}^{n-1} \frac{2}{n} = 2, \quad n \geq 1. \quad (5.4.6)$$

La norma 2 de estos polinomios viene dada por

$$\|T_n\|_2^2 = \int_{-1}^1 \frac{(T_n(x))^2}{\sqrt{1-x^2}} dx = \int_0^\pi \cos^2(n\theta) d\theta = \begin{cases} \pi, & \text{si } n = 0 \\ \pi/2, & \text{si } n \neq 0. \end{cases} \quad (5.4.7)$$

Utilizaremos la Proposición 5.3.3 para calcular el máximo del condicionamiento en el siguiente resultado.

**Proposición 5.4.3.** *El condicionamiento de la representación  $\sum_{i=0}^n \tau_i f T_i$ , donde  $\tau_i f = \int_{-1}^1 (T_i(x) f(x) / \sqrt{1-x^2}) dx$ ,  $i = 0, \dots, n$ , verifica*

$$\max_{x \in [-1,1]} \text{cond}(x; \tau) = 1 + \frac{4n}{\pi}$$

y este valor máximo se alcanza en los extremos del intervalo  $-1$  y  $1$ .

*Demostración.* Usando la Proposición 5.3.3 y las fórmulas (5.4.6) y (5.4.7), se tiene

$$\text{cond}(x; \tau) = \sum_{i=0}^n \frac{\|T_i\|_1 |T_i(x)|}{\|T_i\|_2^2} = 1 + \frac{4}{\pi} \sum_{i=1}^n |T_i(x)|.$$

En particular,  $\text{cond}(\pm 1, \tau) = 1 + 4n/\pi$ . De donde se deduce que

$$\max_{x \in [-1,1]} \text{cond}(x; \tau) = 1 + \frac{4n}{\pi}.$$

□

En la Figura 5.4.1 podemos ver el  $\text{cond}(x; \tau)$  y la cota obtenida en la Proposición (5.4.5) para  $n = 10$ . Vemos que el máximo de este condicionamiento se alcanza en los extremos del intervalo  $[-1, 1]$ , como demuestra el resultado anterior.

## Polinomios de Chebyshev de segunda especie

Los polinomios de Chebyshev de segunda especie  $U_n(x)$  son los polinomios ortogonales en el intervalo  $(-1, 1)$  respecto a la función peso  $w(x) = \sqrt{1-x^2}$  con la condición de normalización  $U_n(1) = n + 1$ .

Es bien conocido que

$$U_n(x) = \frac{\sin((n+1) \arccos x)}{\sqrt{1-x^2}}.$$

Evalutando en  $\cos \theta$ , se tiene

$$\begin{aligned} |U_n(\cos \theta)| &= \left| \frac{\sin((n+1)\theta)}{\sin \theta} \right| = \left| \frac{e^{i(n+1)\theta} - e^{-i(n+1)\theta}}{e^{i\theta} - e^{-i\theta}} \right| = \left| \sum_{k=0}^n e^{i(n-k)\theta} e^{-ik\theta} \right| \\ &= \left| \sum_{k=0}^n e^{i(n-2k)\theta} \right|. \end{aligned}$$

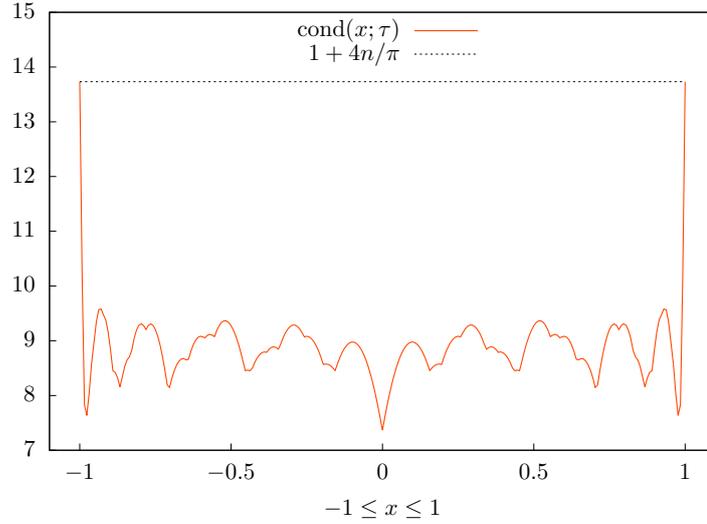


Figura 5.4.1:  $\text{cond}(x; \tau)$  del problema de mínimos cuadrados con polinomios ortogonales de Chebyshev de primera especie  $T_n$  en  $[-1, 1]$  y la cota  $1 + 4n/\pi$  para  $n = 10$ .

Así, obtenemos la siguiente cota uniforme

$$|U_n(x)| \leq \sum_{k=0}^n |e^{i(n-2k)\theta}| \leq n + 1, \quad \forall x \in [-1, 1], \quad (5.4.8)$$

y como  $U_n(1) = n + 1$ , se deduce que  $\|U_n\|_\infty = n + 1$ .

Veamos ahora la norma 2 de estos polinomios

$$\|U_n\|_2^2 = \int_{-1}^1 \sqrt{1-x^2} (U_n(x))^2 dx = \int_0^\pi \sin^2((n+1)\theta) d\theta = \frac{\pi}{2}. \quad (5.4.9)$$

El siguiente resultado muestra la norma uno de estos polinomios.

**Proposición 5.4.4.** *Sea  $U_n$ ,  $n \geq 0$ , el polinomio de Chebyshev de segunda especie de grado  $n$ . Entonces*

$$\|U_n\|_1 = \begin{cases} \frac{\pi}{2}, & \text{si } n = 0, \\ \frac{2(n+1)}{n(n+2)} \cotg \frac{\pi}{2(n+1)}, & \text{si } n \geq 1. \end{cases}$$

*Demostración.* Primero vamos a calcular la norma 1 de  $U_0(x) = 1$ .

$$\|U_0\|_1 = \int_{-1}^1 \sqrt{1-x^2} dx = \int_0^\pi \sin^2 \theta d\theta = \int_0^\pi \frac{1 - \cos 2\theta}{2} d\theta = \frac{\pi}{2}.$$

Para el resto de polinomios de Chebyshev  $U_n$ ,  $n \geq 1$ , tenemos

$$\begin{aligned} \|U_n\|_1 &= \int_{-1}^1 |U_n| \sqrt{1-x^2} dx = \int_0^\pi |\sin((n+1)\theta)| \sin \theta d\theta \\ &= \sum_{k=0}^n \int_{k\pi/(n+1)}^{(k+1)\pi/(n+1)} (-1)^k \sin((n+1)\theta) \sin \theta d\theta. \end{aligned} \quad (5.4.10)$$

Vamos a calcular el valor de estas integrales.

$$\begin{aligned} &\int_{k\pi/(n+1)}^{(k+1)\pi/(n+1)} (-1)^k \sin((n+1)\theta) \sin \theta d\theta \\ &= \int_{k\pi/(n+1)}^{(k+1)\pi/(n+1)} (-1)^k \frac{\cos n\theta - \cos((n+2)\theta)}{2} d\theta \\ &= \frac{(-1)^k}{2} \left( \frac{\sin \frac{n(k+1)\pi}{n+1} - \sin \frac{nk\pi}{n+1}}{n} - \frac{\sin \frac{(n+2)(k+1)\pi}{n+1} - \sin \frac{(n+2)k\pi}{n+1}}{n+2} \right). \end{aligned}$$

Usando las siguientes igualdades

$$\sin(k\pi - \theta) = (-1)^{k-1} \sin \theta, \quad \sin(k\pi + \theta) = (-1)^k \sin \theta,$$

y operando, obtenemos

$$\begin{aligned} &\int_{k\pi/(n+1)}^{(k+1)\pi/(n+1)} (-1)^k \sin((n+1)\theta) \sin \theta d\theta \\ &= \frac{1}{2} \left( \frac{\sin \frac{(k+1)\pi}{n+1} + \sin \frac{k\pi}{n+1}}{n} + \frac{\sin \frac{(k+1)\pi}{n+1} + \sin \frac{k\pi}{n+1}}{n+2} \right) \\ &= \frac{n+1}{n(n+2)} \left( \sin \frac{(k+1)\pi}{n+1} + \sin \frac{k\pi}{n+1} \right). \end{aligned}$$

Por tanto, sustituyendo en (5.4.10), se tiene

$$\|U_n\|_1 = \sum_{k=0}^n \frac{n+1}{n(n+2)} \left( \sin \frac{(k+1)\pi}{n+1} + \sin \frac{k\pi}{n+1} \right) = \frac{2(n+1)}{n(n+2)} \sum_{k=0}^n \sin \frac{k\pi}{n+1}.$$

Así,

$$\begin{aligned} \sum_{k=0}^n \sin \frac{k\pi}{n+1} &= \operatorname{Im} \sum_{k=0}^n e^{ik\pi/(n+1)} = \operatorname{Im} \left( \frac{e^{i(n+1)\pi/(n+1)} - 1}{e^{i\pi/(n+1)} - 1} \right) \\ &= \operatorname{Im} \left( \frac{-2}{\cos \frac{\pi}{n+1} - 1 + i \sin \frac{\pi}{n+1}} \right). \end{aligned}$$

Multiplicamos por el conjugado

$$\sum_{k=0}^n \sin \frac{k\pi}{n+1} = \operatorname{Im} \left( \frac{-2 \left( \cos \frac{\pi}{n+1} - 1 - i \sin \frac{\pi}{n+1} \right)}{\left( \cos \frac{\pi}{n+1} - 1 \right)^2 + \sin^2 \frac{\pi}{n+1}} \right) = \frac{2 \sin \frac{\pi}{n+1}}{\left( \cos \frac{\pi}{n+1} - 1 \right)^2 + \sin^2 \frac{\pi}{n+1}}.$$

Usando

$$\cos 2\theta = 1 - 2\sin^2 \theta, \quad \sin 2\theta = 2\cos \theta \sin \theta,$$

se deduce

$$\sum_{k=0}^n \sin \frac{k\pi}{n+1} = \frac{4 \sin \frac{\pi}{2(n+1)} \cos \frac{\pi}{2(n+1)}}{4 \sin^4 \frac{\pi}{2(n+1)} + 4 \sin^2 \frac{\pi}{2(n+1)} \cos^2 \frac{\pi}{2(n+1)}} = \frac{\cos \frac{\pi}{2(n+1)}}{\sin \frac{\pi}{2(n+1)}} = \cotg \frac{\pi}{2(n+1)}.$$

Así, se obtiene el resultado

$$\|U_n\|_1 = \frac{2(n+1)}{n(n+2)} \cotg \frac{\pi}{2(n+1)}, \quad n \geq 1.$$

□

**Proposición 5.4.5.** *El condicionamiento de la representación  $\sum_{i=0}^n \alpha_i f U_i$ , donde  $\alpha_i f = \int_{-1}^1 (\sqrt{1-x^2} U_i(x) f(x)) dx$ ,  $i = 0, \dots, n$ , viene dado por*

$$\text{cond}(x; \alpha) = 1 + \frac{4}{\pi} \sum_{i=1}^n \frac{i+1}{i(i+2)} \cotg \frac{\pi}{2(i+1)} |U_i(x)|,$$

y verifica

$$\text{cond}(x; \alpha) \leq 1 + \frac{4n(n+3)}{\pi^2} + \frac{24 - \pi^2}{3\pi^2} H_{n+1} - \frac{(12 - \pi^2)(3n^2 + 6n + 8)}{6\pi^2(n+1)(n+2)}, \quad (5.4.11)$$

donde  $H_{n+1} := \sum_{i=1}^{n+1} 1/i$ , la suma  $(n+1)$ -ésima de la serie armónica.

*Demostración.* Usando la fórmula (5.4.9) y la Proposición 5.3.3, se tiene

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \frac{\|U_i\|_1 |U_i(x)|}{\|U_i\|_2^2} = 1 + \frac{4}{\pi} \sum_{i=1}^n \frac{i+1}{i(i+2)} \cotg \frac{\pi}{2(i+1)} |U_i(x)|.$$

Vamos dar una cota uniforme para este condicionamiento. Teniendo en cuenta que  $\cotg x \leq 1/x - x/3$  y  $\|U_n\|_\infty = n+1$ , se tiene

$$\begin{aligned} \text{cond}(x; \alpha) &\leq 1 + \frac{4}{\pi} \sum_{i=1}^n \frac{(i+1)^2}{i(i+2)} \left( \frac{2(i+1)}{\pi} - \frac{\pi}{6(i+1)} \right) \\ &= 1 + \frac{4}{\pi} \sum_{i=1}^n \left( \frac{2(i+1)^3}{\pi i(i+2)} - \frac{\pi(i+1)}{6i(i+2)} \right) \\ &= 1 + \frac{4}{\pi} \sum_{i=0}^n \left( \frac{2}{\pi} - \frac{\pi}{6} \right) \frac{i+1}{i(i+2)} + \frac{4}{\pi} \sum_{i=0}^n \frac{2(i+1)}{\pi} \\ &= 1 + \frac{4n(n+3)}{\pi^2} + \frac{12 - \pi^2}{3\pi^2} \sum_{i=1}^n \left( \frac{1}{i} + \frac{1}{i+2} \right). \end{aligned}$$

Operando obtenemos la cota

$$\text{cond}(x; \alpha) \leq 1 + \frac{4n(n+3)}{\pi^2} + \frac{24 - \pi^2}{3\pi^2} H_{n+1} - \frac{(12 - \pi^2)(3n^2 + 6n + 8)}{6\pi^2(n+1)(n+2)}. \quad (5.4.12)$$

□

En la Figura (5.4.12) podemos ver el condicionamiento de la representación que usa los polinomios de Chebyshev de segunda especie para  $n = 10$  en el intervalo  $[-1, 1]$  y la cota obtenida en (5.4.12).

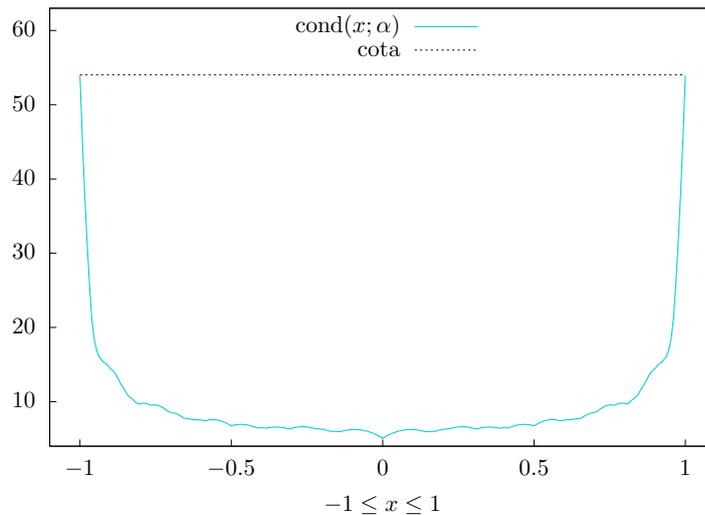


Figura 5.4.2:  $\text{cond}(x; \alpha)$  del problema de mínimos cuadrados con polinomios ortogonales de Chebyshev de segunda especie  $U_n$  en  $[-1, 1]$  y la cota obtenida en (5.4.12) para  $n = 10$ .

## 5.5 Condicionamiento de problemas de mínimos cuadrados discretos

El problema de mínimos cuadrados discretos corresponde al caso en el que la forma bilineal simétrica y semidefinida positiva es de la forma

$$\langle u, v \rangle_N = \sum_{i=0}^N w_i u(x_i) v(x_i),$$

donde  $x_0, \dots, x_n \in [a, b]$ . Esta forma bilineal es definida positiva en el espacio de polinomios de grado menor o igual que  $N$  dando lugar a una norma

$$\|p\|_{N,2} = \langle p, p \rangle_N^{1/2}, \quad p \in P_N.$$

Por tanto, tiene sentido definir polinomios ortogonales hasta grado  $N$  correspondientes a esta forma bilineal sobre  $P_N$ .

En este caso, la representación del operador respecto a una base ortogonal  $p_0, \dots, p_n$  es

$$Q[f](x) = \sum_{i=0}^n \frac{\sum_{j=0}^N w_j f(x_j) p_i(x_j)}{\sum_{j=0}^N w_j p_i(x_j)^2} p_i(x).$$

Por el Teorema 5.3.1

$$\|Q\|_\infty = \max_{x \in [a,b]} \sum_{j=0}^N \left| \sum_{i=0}^n \frac{w_j p_i(x_j) p_i(x)}{\sum_{k=0}^N w_k p_i(x_k)^2} \right|.$$

El siguiente resultado es una consecuencia directa de la Proposición 5.3.3.

**Proposición 5.5.1.** *Sea  $x_0, \dots, x_N$  una sucesión de nodos. Sea  $p_0, \dots, p_n$ ,  $n \leq N$ , una sucesión de polinomios ortogonales respecto al producto escalar discreto  $\langle u, v \rangle_N := \sum_{i=0}^N w_i u(x_i) v(x_i)$ . Sea  $Q[f](x) = \sum_{i=0}^n \alpha_i f p_i(x)$  el operador de cuasi-interpolación que asocia a cada función su mejor aproximación en  $P_n$ , donde*

$$\alpha_i f = \frac{\langle p_i, f \rangle_N}{\langle p_i, p_i \rangle_N} = \frac{\sum_{j=0}^N w_j f(x_j) p_i(x_j)}{\sum_{j=0}^N w_j p_i(x_j)^2}. \quad (5.5.1)$$

Entonces

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \frac{\sum_{j=0}^N w_j |p_i(x_j)|}{\sum_{j=0}^N w_j p_i(x_j)^2} |p_i(x)| = \sum_{i=0}^n \frac{\|p_i\|_{N,1}}{\langle p_i, p_i \rangle_N} |p_i(x)|, \quad (5.5.2)$$

donde  $\|p\|_{N,1} := \sum_{j=0}^N w_j |p(x_j)|$ .

*Demostración.* Por la Proposición 5.3.2, se tiene

$$\|\alpha_i\|_\infty = \frac{\sum_{j=0}^N w_j |p_i(x_j)|}{\sum_{j=0}^N w_j p_i(x_j)^2} = \frac{\|p_i\|_{N,1}}{\langle p_i, p_i \rangle_N}, \quad i = 0, \dots, n. \quad (5.5.3)$$

Usando la Proposición 5.3.3, obtenemos el resultado.  $\square$

Observemos que la relación (5.5.3) es un caso particular de la Proposición 2.1.3, donde los coeficientes son

$$c_j = \frac{w_j p_i(x_j)}{\sum_{j=0}^N w_j p_i(x_j)^2}.$$

En el caso de que  $n = N$  tenemos un número de nodos mínimo y el operador de cuasi-interpolación coincide con el operador de interpolación en los nodos  $x_0, \dots, x_N$ . Recordemos que la función de Lebesgue es el condicionamiento respecto a la base de

funcionales evaluación y corresponde a un condicionamiento mínimo (véase la Proposición 2.1.7). A lo largo de esta sección, veremos distintos ejemplos en los que se muestra que el uso de una representación respecto a una base ortogonal no difiere notablemente de la representación de Lagrange, que es la óptima.

**Ejemplo 5.5.2.** *Veamos la fórmula (5.5.2) en el caso particular de que los pesos sean todos iguales a 1,  $w_i = 1$ ,  $i = 0, \dots, N$ , y nodos equidistantes en el intervalo  $[-1, 1]$ , es decir,*

$$x_i = -1 + \frac{2i}{N}, \quad i = 0, \dots, N.$$

*Para calcular los polinomios ortogonales vamos a utilizar la bien conocida fórmula de recurrencia a tres términos para los polinomios ortogonales mónicos*

$$\begin{cases} p_{n+1}(x) = (x - a_n)p_n(x) - b_n p_{n-1}(x), & n = 1, 2, \dots, \\ p_0(x) = 1, & p_1(x) = x - a_0, \end{cases}$$

donde

$$a_n := \frac{\langle p_n, xp_n \rangle}{\langle p_n, p_n \rangle}, \quad b_n := \frac{\langle p_n, p_n \rangle}{\langle p_{n-1}, p_{n-1} \rangle}.$$

Notar que  $b_n > 0$ , para todo  $n$ . Recordemos

$$\|p\|_{N,1} = \sum_{j=0}^N w_j |p(x_j)| = \sum_{j=0}^N |p(x_j)|,$$

y

$$\langle p, p \rangle_N = \sum_{j=0}^N w_j p(x_j)^2 = \sum_{j=0}^N p(x_j)^2.$$

*En la Figura 5.5.1 podemos ver  $\text{cond}(x; \alpha)$  para  $n = 16$  y  $N = 16$ ,  $\text{cond}(x; \alpha)$  para  $n = 16$  y  $N = 32$  y la función de Lebesgue en puntos equidistantes para  $n = 16$  en el intervalo  $[-1, 1]$ . En este ejemplo vemos que, al considerar muestras de tamaño mayor que el grado, el condicionamiento disminuye notablemente con respecto al muestreo crítico, es decir, con el mínimo número de nodos  $n = N$ . En el caso  $n = N$  podemos comparar el condicionamiento de dos representaciones del mismo problema de interpolación: con respecto a la base ortogonal y con respecto a la base de Lagrange. Vemos que la función de Lebesgue es algo menor que el condicionamiento respecto a la base ortogonal.*

La construcción de la base de polinomios ortogonales a través de un proceso de ortogonalización o una relación de recurrencia a tres términos puede presentar fenómenos de inestabilidad. Por tanto, una dificultad que plantea el cálculo del condicionamiento

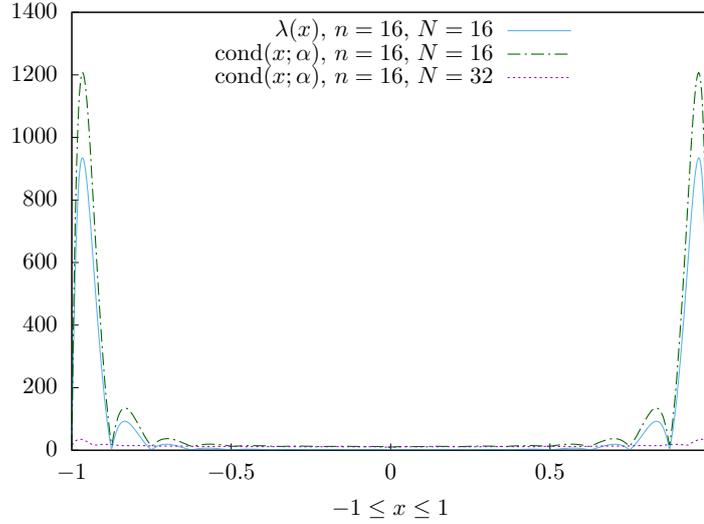


Figura 5.5.1:  $\text{cond}(x; \alpha)$  del problema de mínimos cuadrados discretos en nodos equidistantes con pesos  $w_i = 1$ ,  $i = 0, \dots, N$ , en  $[-1, 1]$  para  $n = 16$  y  $N = 16$ ,  $n = 16$  y  $N = 32$  y la función de Lebesgue para  $n = 16$ .

es que no siempre resulta fácil disponer de la base de polinomios ortogonales. A continuación, veremos que podemos utilizar polinomios ortogonales clásicos para resolver problemas discretos.

Consideremos el caso en el que  $x_0, \dots, x_N$  son los ceros de un polinomio ortogonal  $p_{N+1}$  respecto a un producto escalar continuo correspondiente a una función peso  $w(x)$ . Es bien conocido que los pesos de la regla de cuadratura Gaussiana  $w_0, \dots, w_N$  proporcionan una regla de cuadratura exacta para polinomios de grado menor o igual  $2N$ . Por tanto,

$$\langle p, q \rangle = \int_a^b p(x)q(x)w(x) dx = \sum_{i=0}^N w_i p(x_i)q(x_i) = \langle p, q \rangle_N, \quad p, q \in P_N,$$

es decir, coinciden los productos escalares continuo y discreto en el espacio de polinomios  $P_N$ . Por tanto, los polinomios ortogonales correspondientes a esos nodos y a esos pesos coinciden con los polinomios ortogonales continuos correspondientes. Así, aplicando la desigualdad de Cauchy-Schwarz a (5.5.3), se tiene una cota para la norma de los funcionales

$$\|\alpha_i\|_\infty = \frac{\|p_i\|_{N,1}}{\langle p_i, p_i \rangle_N} \leq \frac{\langle p_i, p_i \rangle_N^{1/2} \langle 1, 1 \rangle_N^{1/2}}{\langle p_i, p_i \rangle_N} = \frac{\|p_i\|_2 \|1\|_2}{\|p_i\|_2^2} = \frac{\|1\|_2}{\|p_i\|_2}, \quad i = 0, \dots, n. \quad (5.5.4)$$

### Mínimos cuadrados en nodos de Legendre

En la regla de cuadratura Gaussiana para los polinomios de Legendre,  $P_N(x)$ , el nodo  $x_j$  corresponde a la  $j$ -ésima raíz de  $P_N$ ,  $j = 0, \dots, N$ , y los pesos son (véase 25.4.29 de [1])

$$w_j = \frac{2}{(1-x_j^2)(P'_N(x_j))^2}, \quad j = 0, \dots, N.$$

Por (5.5.4), (5.4.2) y (5.4.3), podemos acotar la norma de los funcionales del siguiente modo

$$\|\alpha_i\|_\infty \leq \frac{\|1\|_2}{\|P_i\|_2} \leq \sqrt{2i+1}, \quad i = 0, \dots, n.$$

Teniendo en cuenta que  $\|P_i\|_\infty = 1$  y el Lema 5.4.1, llegamos a la misma cota para el condicionamiento del problema discreto que la obtenida en el caso continuo

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \|\alpha_i\|_\infty |P_i(x)| \leq \sum_{i=0}^n \sqrt{2i+1} \leq \frac{2}{3} \sqrt{2n(n+1)(n+2)} + 1.$$

### Mínimos cuadrados en nodos de Chebyshev de primera especie

Como coinciden el producto escalar discreto y continuo sobre los polinomios de grado menor o igual que  $n$ , tenemos, por la fórmula (5.4.7)

$$\langle T_i, T_i \rangle_N = \|T_i\|_2^2 = \begin{cases} \pi, & \text{si } i = 0 \\ \pi/2, & \text{si } i \neq 0. \end{cases}$$

La regla de cuadratura Gaussiana tiene pesos iguales en todos los nodos, es decir,  $w_j = \pi/(N+1)$ ,  $j = 0, \dots, N$ , y los nodos son de la forma  $x_j = \cos\left(\frac{2j+1}{2(N+1)}\pi\right)$ ,  $j = 0, \dots, N$  (véase 25.4.38 de [1]). Por tanto, por la fórmula (5.5.1), los funcionales son de la forma

$$\alpha_0 f = \frac{1}{N+1} \sum_{j=0}^N f(x_j),$$

$$\alpha_i f = \frac{2}{N+1} \sum_{j=0}^N T_i(x_j) f(x_j), \quad i = 1, \dots, N,$$

cuyas normas son

$$\|\alpha_0\|_\infty = 1, \quad \|\alpha_i\|_\infty = \frac{2}{N+1} \sum_{j=0}^N |T_i(x_j)|, \quad i = 1, \dots, N.$$

Utilizando las fórmulas (5.4.7) y (5.5.4), se tiene

$$\|\alpha_i\|_\infty \leq \frac{\|1\|_2}{\|T_i\|_2} = \frac{\sqrt{\pi}}{\sqrt{\pi/2}} = \sqrt{2}, \quad i = 1, \dots, N. \quad (5.5.5)$$

Por tanto, por la Proposición 5.3.3 y la fórmula (5.5.5), se tiene que el condicionamiento verifica la siguiente cota

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \|\alpha_i\|_\infty |T_i(x)| \leq 1 + \sqrt{2} \sum_{i=1}^n |T_i(x)| \leq 1 + \sqrt{2} n.$$

Como  $\sqrt{2} \geq 4/\pi$ , la cota obtenida es mayor que la proporcionada por la Proposición 5.4.5 para el caso continuo. Finalmente, observemos que cuando  $N \rightarrow \infty$ , la aproximación discreta converge a la continua. De hecho,

$$\begin{aligned} \lim_{N \rightarrow \infty} \|\alpha_i\|_\infty &= \lim_{N \rightarrow \infty} \frac{2}{N+1} \sum_{j=0}^N |T_i(x_j)| = \lim_{N \rightarrow \infty} \frac{2}{N+1} \sum_{j=0}^N |\cos(i\theta_j)| = \frac{2}{\pi} \int_0^\pi |\cos(i\theta)| d\theta \\ &= \frac{4}{\pi}, \quad i = 1, \dots, N. \end{aligned}$$

Así,

$$\lim_{N \rightarrow \infty} \text{cond}(x; \alpha) = 1 + \frac{4}{\pi} \sum_{i=1}^n |T_i(x)| = \text{cond}(x; \tau).$$

En la Figura 5.5.2 está representado el condicionamiento de los problemas continuo para  $n = 5$  y discreto para  $n = 5, N = 5$  respecto a los polinomios de Chebyshev de primera especie. Podemos ver que el condicionamiento del problema discreto es similar al condicionamiento del caso continuo.

### Mínimos cuadrados en nodos de Chebyshev de segunda especie

Los pesos de la regla de cuadratura Gaussiana para los polinomios de Chebyshev de segunda especie son

$$w_j = \frac{\pi}{N+2} \sin^2 \left( \frac{(j+1)\pi}{N+2} \right), \quad j = 0, \dots, N,$$

y los nodos son de la forma (véase 25.4.40 de [1])

$$x_j = \cos \left( \frac{j+1}{N+2} \pi \right), \quad j = 0, \dots, N.$$

Por tanto, usando las fórmulas (5.4.9) y (5.5.1), los funcionales son de la forma

$$\alpha_i f = \frac{\sum_{j=0}^N w_j U_i(x_j) f(x_j)}{\langle U_i, U_i \rangle_N} = \frac{2}{N+2} \sum_{j=0}^N \sin^2 \left( \frac{(j+1)\pi}{N+2} \right) U_i(x_j) f(x_j), \quad i = 0, \dots, n,$$

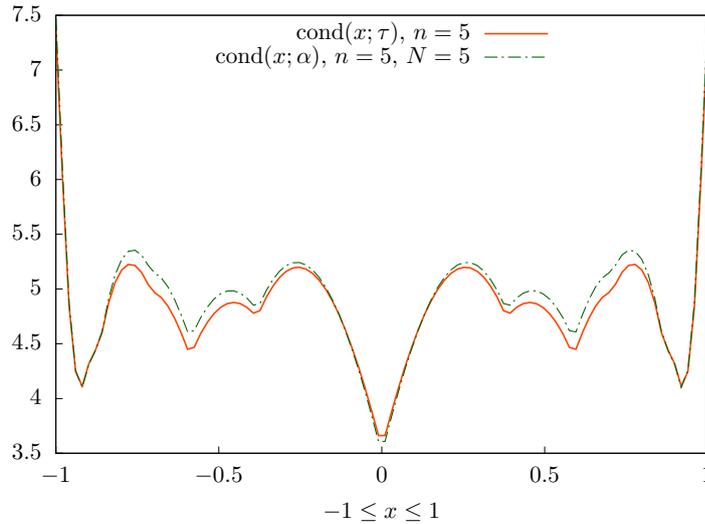


Figura 5.5.2:  $\text{cond}(x; \alpha)$  del problema de mínimos cuadrados discretos para los polinomios de Chebyshev de primera especie en  $[-1, 1]$  para  $n = 5$  y  $N = 5$  y  $\text{cond}(x; \tau)$  del problema de mínimos cuadrados continuo para los polinomios de Chebyshev de primera especie para  $n = 5$ .

y sus normas son

$$\|\alpha_i\|_\infty = \frac{2}{N+2} \sum_{j=0}^N \sin^2 \left( \frac{(j+1)\pi}{N+2} \right) |U_i(x_j)|, \quad i = 0, \dots, n.$$

Por (5.4.9) y (5.5.4), tenemos la siguiente cota para las normas de los funcionales

$$\|\alpha_i\|_\infty \leq \frac{\|1\|_2}{\|U_i\|_2} = \frac{\sqrt{\pi/2}}{\sqrt{\pi/2}} = 1, \quad i = 0, \dots, n.$$

Por tanto, usando (5.4.8), podemos acotar  $\text{cond}(x; \alpha)$  del siguiente modo

$$\text{cond}(x; \alpha) = \sum_{i=0}^n \|\alpha_i\|_\infty |U_i(x)| \leq \sum_{i=0}^n |U_i(x)| \leq \sum_{i=0}^n (i+1) = \frac{(n+1)(n+2)}{2}, \quad x \in [a, b].$$

Veamos que la cota en el caso continuo es menor que la cota del discreto. Para ello, veamos que la cota para la norma de los funcionales del caso continuo (ver Proposición 5.4.4) verifican

$$\|\alpha_i\|_\infty \leq \frac{4(i+1)}{\pi i(i+2)} \left( \frac{2(i+1)}{\pi} - \frac{\pi}{6(i+1)} \right) \leq 1, \quad i = 1, \dots, n.$$

En efecto,

$$\frac{2(i+1)}{\pi} - \frac{\pi}{6(i+1)} \leq \frac{\pi i(i+2)}{4(i+1)} = \frac{\pi}{4}(i+1) - \frac{\pi}{4(i+1)},$$

es equivalente a

$$(i+1)^2 \left( \frac{2}{\pi} - \frac{\pi}{4} \right) + \frac{\pi}{12} \leq 0,$$

y como esa sucesión es decreciente en  $i$ , basta ver que se cumple para  $i = 1$

$$4 \left( \frac{2}{\pi} - \frac{\pi}{4} \right) + \frac{\pi}{12} = \frac{8}{\pi} - \frac{11\pi}{12} < 0.$$

El Corolario 14.4.7 de [24] es un resultado de Stieltjes que muestra que las reglas de cuadraturas de Gauss-Jacobi discretas convergen a las integrales ponderadas correspondientes cuando  $N \rightarrow \infty$ . Por tanto, la norma de los funcionales del problema discreto converge a la norma del correspondiente funcional del caso continuo

$$\begin{aligned} \lim_{N \rightarrow \infty} \|\alpha_i\|_\infty &= \lim_{N \rightarrow \infty} \frac{2}{\pi} \sum_{j=0}^N w_j |U_i(x_j)| = \frac{2}{\pi} \int_{-1}^1 |U_i(x)| \sqrt{1-x^2} dx \\ &= \begin{cases} 1, & i = 0, \\ \frac{4(i+1)}{\pi i(i+2)} \cotg \frac{\pi}{2(i+1)}, & i \geq 1. \end{cases} \end{aligned}$$

Por tanto, el condicionamiento de la aproximación discreta converge a la continua

$$\lim_{N \rightarrow \infty} \text{cond}(x; \alpha) = 1 + \frac{4}{\pi} \sum_{i=0}^n \frac{i+1}{i(i+2)} \cotg \frac{\pi}{2(i+1)} |U_i(x)|.$$

## 5.6 Condicionamiento del problema de interpolación de Lagrange en los ceros de los polinomios ortogonales

Cuando resolvemos un problema de mínimos cuadrados con  $n = N$ , la solución es el interpolante de Lagrange en los nodos. En la sección anterior hemos elegido como nodos los ceros de los polinomios ortogonales de grado  $n+1$  y hemos representado el interpolante respecto a la misma base de polinomios ortogonales. En el caso de polinomios de Chebyshev de primera especie  $T_n$ , hemos obtenido

$$\text{cond}(x; \alpha) \leq 1 + \sqrt{2}n.$$

Para polinomios de Chebyshev de segunda especie  $U_n$ , tenemos

$$\text{cond}(x; \alpha) \leq \frac{(n+1)(n+2)}{2}.$$

Y, por último, en el caso de polinomios de Legendre  $P_n$

$$\text{cond}(x; \alpha) \leq \frac{2}{3} \sqrt{2n(n+1)(n+2)} + 1.$$

Notemos que el crecimiento del condicionamiento es subexponencial, lo que muestra buenas propiedades del problema de interpolación en ceros de polinomios ortogonales y el beneficio del uso de bases ortogonales apropiadas.

Como la función de Lebesgue corresponde a la base de funcionales evaluación  $\lambda_i f = f(x_i)$  y tiene condicionamiento óptimo (véase la Proposición 2.1.7), podemos deducir inmediatamente que la constante de Lebesgue satisface las cotas anteriores cuando los nodos son los ceros de los correspondientes polinomios ortogonales.

Cabe preguntarse cuáles son las distribuciones de nodos que proporcionan las constantes de Lebesgue más bajas. En el Capítulo 3 de [22] se presenta una discusión, que comentamos a continuación, de este problema clásico documentado con relevantes referencias. Se sabe que existe un conjunto de nodos óptimo que verifica

$$\Lambda_n > \frac{2}{\pi} \left( \log n + \log \frac{4}{\pi} + \gamma \right) \approx \frac{2}{\pi} \log n + 0.5212,$$

donde  $\gamma \approx 0.5772156649$  es la constante de Euler–Mascheroni. Es bien conocido que el operador de interpolación en los nodos de Chebyshev de primera especie verifica (véase también la fórmula (1.38) de la página 13 de [52])

$$\|L\|_{\infty} \leq \frac{2}{\pi} \log n + 1,$$

por lo que podemos considerar a los nodos de Chebyshev casi óptimos y útiles para la mayoría de propósitos básicos. Además, vemos que la representación respecto a la base de Lagrange tiene un crecimiento logarítmico. Por tanto, para valores muy grandes de  $n$ , es preferible trabajar con la base de Lagrange que con la base de polinomios ortogonales de Chebyshev, que tiene un crecimiento lineal. En la Figura 5.6.1, vemos una comparación de la función de Lebesgue en los nodos de Chebyshev para  $n = 6$  y el condicionamiento del problema de mínimos cuadrados discreto respecto a los polinomios de Chebyshev de primera especie para  $n = 6, N = 6$ .

La constante de Lebesgue en los ceros del polinomio de Chebyshev de segunda especie de grado  $n+1$  es  $n+1$  (véase [10]). En este caso, el crecimiento de la constante de Lebesgue es lineal y el condicionamiento respecto a la base de polinomios ortogonales presenta un crecimiento cuadrático.

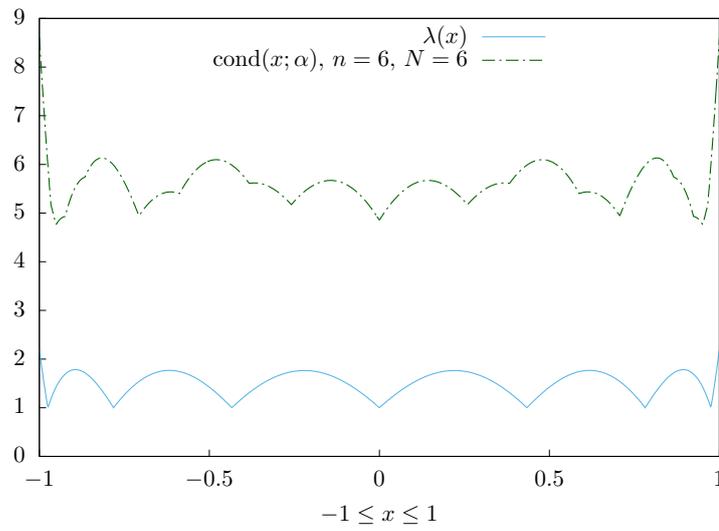


Figura 5.6.1:  $\lambda(x)$  en nodos de Chebyshev para  $n = 6$  y  $\text{cond}(x; \alpha)$  del problema de mínimos cuadrados discreto respecto a los polinomios  $T_n$ , para  $n = 6, N = 6$  en el intervalo  $[-1, 1]$ .

# 6

## Bases y funcionales positivos

---

### 6.1 Caracterización de funcionales no negativos

Un funcional  $\alpha : C[a, b] \rightarrow \mathbb{R}$  se dice no negativo si

$$\alpha f \geq 0, \quad \forall f \geq 0,$$

y se denota por  $\alpha \geq 0$ . Veamos un resultado que caracteriza los funcionales no negativos según su norma.

**Teorema 6.1.1.** *Sea  $\alpha : C[a, b] \rightarrow \mathbb{R}$  un funcional. Se tiene que  $\alpha \geq 0$  si y sólo si  $\alpha 1 = \|\alpha\|_\infty$ .*

*Demostración.* Supongamos primero que  $\alpha \geq 0$ . Entonces

$$\alpha 1 = |\alpha 1| \leq \|\alpha\|_\infty \|1\|_\infty = \|\alpha\|_\infty.$$

Para la otra desigualdad, sea  $f$  tal que  $\|f\|_\infty = 1$ . Se tiene que

$$\begin{aligned} 1 - f &\geq 0 \Rightarrow \alpha 1 \geq \alpha f, \\ 1 + f &\geq 0 \Rightarrow -\alpha 1 \leq \alpha f. \end{aligned}$$

Por tanto

$$|\alpha f| \leq \alpha 1,$$

para toda  $f$  tal que  $\|f\|_\infty = 1$ . Así,

$$\|\alpha\|_\infty \leq \alpha 1.$$

Supongamos para el recíproco que  $\alpha 1 = \|\alpha\|_\infty$  y sea  $f \geq 0$  cualquiera. Entonces

$$\alpha f \leq \|\alpha\|_\infty \|f\|_\infty = \alpha 1 \|f\|_\infty$$

y por tanto

$$\alpha(\|f\|_\infty - f) \geq 0. \quad (6.1.1)$$

Sea  $g \geq 0$  y  $m := \min_{x \in [a,b]} g(x)$ . Se tiene

$$\|g - m\|_\infty = \|g\|_\infty - m.$$

Definimos  $f := \|g\|_\infty - g = \|g - m\|_\infty - (g - m) \geq 0$  y calculamos su norma

$$\|f\|_\infty = \|g - m\|_\infty = \|g\|_\infty - m.$$

Así,

$$\alpha g = \alpha(\|g\|_\infty - f) = \alpha(\|f\|_\infty - f + m) = m\alpha 1 + \alpha(\|f\|_\infty - f) = m\|\alpha\|_\infty + \alpha(\|f\|_\infty - f),$$

y usando (6.1.1) se tiene

$$\alpha g = m\|\alpha\|_\infty + \alpha(\|f\|_\infty - f) \geq 0.$$

□

Podemos aplicar el Teorema 6.1.1 para caracterizar los funcionales positivos que son combinación lineal de funcionales evaluación.

**Proposición 6.1.2.** *Sea  $\alpha$  un funcional de la forma  $\alpha f = \sum_{i=0}^n c_i f(x_i)$ . Entonces  $\alpha \geq 0$  si y sólo si  $c_i \geq 0$  para todo  $i = 0, \dots, n$ .*

*Demostración.* Por el Teorema 6.1.1  $\alpha \geq 0$  si y sólo si  $\|\alpha\|_\infty = \alpha 1$ . Por la Proposición 2.1.3 se tiene

$$\sum_{i=0}^n |c_i| = \|\alpha\|_\infty.$$

Por tanto,

$$0 \leq \sum_{i=0}^n |c_i| = \|\alpha\|_\infty = \alpha 1 = \sum_{i=0}^n c_i \quad \text{si y sólo si} \quad c_i \geq 0 \quad \forall i = 0, \dots, n.$$

□

El resultado anterior puede considerarse como un caso particular del teorema de representación de Riesz en el espacio de las funciones continuas, que caracteriza a los funcionales no negativos. En el Teorema 2.14 del Capítulo 2 de [54] se considera el teorema en un contexto más general de espacios Hausdorff localmente compactos.

**Teorema 6.1.3** (Teorema de representación de Riesz para funcionales no negativos).  
Sea  $\alpha$  un funcional no negativo de  $C[a, b] \rightarrow \mathbb{R}$ . Entonces

$$\alpha f = \int_a^b f(x) d\mu(x),$$

donde  $\mu$  es una medida de Borel regular no negativa.

Una consecuencia de este resultado es el Teorema 1.2.2 sobre funcionales acotados, también llamado Teorema de representación de Riesz, enunciado en el Capítulo 1.

Como caso particular importante de funcionales definidos en  $C[a, b]$ , tenemos los funcionales de la forma

$$\alpha f = \int_a^b f(x)w(x)dx + \sum_{i=0}^n w_i f(x_i),$$

con  $x_0, \dots, x_n$  distintos y  $w$  integrable en  $[a, b]$ . Entonces  $\alpha$  es no negativo si y sólo si  $w \geq 0$  y  $w_i \geq 0$ ,  $i = 0, \dots, n$ . Por tanto, podemos ver la Proposición 6.1.2 como caso particular del Teorema de representación de Riesz.

## 6.2 Comparación entre sistemas de funcionales no negativos

Se dice que un sistema de funcionales  $\alpha = (\alpha_0, \dots, \alpha_n)$  definido en un espacio  $F$  dotado de la norma del supremo es no negativo si todos los funcionales que lo componen son no negativos y se denota por  $\alpha \geq 0$ . Recordemos que una matriz es no negativa,  $K \geq 0$ , si todos sus elementos son no negativos. La matriz  $D_\alpha = \text{diag}(\|\alpha_0\|_\infty, \dots, \|\alpha_n\|_\infty)$  nos permite expresar el vector de las normas de los funcionales  $\alpha = (\alpha_0, \dots, \alpha_n)$  como

$$\begin{pmatrix} \|\alpha_0\|_\infty \\ \vdots \\ \|\alpha_n\|_\infty \end{pmatrix} = D_\alpha e,$$

donde  $e = (1, \dots, 1)^T$ .

El condicionamiento de una representación  $\sum_{i=0}^n \alpha_i u_i$  es

$$\text{cond}(x; \alpha, u) = \sum_{i=0}^n \|\alpha_i\|_\infty |u_i(x)| = |u(x)|^T D_\alpha e,$$

donde  $|u(x)|$  denota el sistema de funciones

$$|u(x)| = (|u_0(x)|, \dots, |u_n(x)|)^T.$$

Esta anotación abreviada va a ser útil para poder realizar operaciones matriciales en los siguientes resultados.

**Teorema 6.2.1.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases de funcionales  $\alpha = (\alpha_0, \dots, \alpha_n)^T$  y  $\beta = (\beta_0, \dots, \beta_n)^T$  no negativas. Sea la matriz de cambio de base  $K = (\alpha_i b_j)_{0 \leq i, j \leq n}$  tal que  $\alpha = K\beta$ . Si  $K \geq 0$  entonces

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u).$$

*Demostración.* Como  $\alpha = K\beta$  tenemos que  $b^T = u^T K$ , luego

$$|b|^T \leq |u|^T K$$

y tenemos

$$\text{cond}(x; \beta, b) = \sum_{j=0}^n \|\beta_j\|_{\infty} |b_j(x)| = |b(x)|^T D_{\beta} e \leq |u(x)|^T K D_{\beta} e.$$

Aplicando el Teorema 6.1.1 se tiene

$$\begin{aligned} \text{cond}(x; \beta, b) &\leq |u(x)|^T K D_{\beta} e = |u(x)|^T K \beta(1) = |u(x)|^T \alpha(1) = |u(x)|^T D_{\alpha} e \\ &= \text{cond}(x; \alpha, u). \end{aligned}$$

□

Podemos dar una doble desigualdad para la comparación de los condicionamientos de dos representaciones de un operador de cuasi-interpolación.

**Teorema 6.2.2.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases de funcionales  $\alpha = (\alpha_0, \dots, \alpha_n)^T$  y  $\beta = (\beta_0, \dots, \beta_n)^T$  no negativas. Sea la matriz de cambio de base  $K = (\alpha_i b_j)_{0 \leq i, j \leq n}$  tal que  $\alpha = K\beta$ . Si  $K \geq 0$  entonces

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u) \leq \text{Cond}(\hat{K}^{-1}) \text{cond}(x; \beta, b),$$

donde  $\hat{K} = D_{\alpha}^{-1} K D_{\beta}$ .

*Demostración.* La primera desigualdad es el Teorema 6.2.1. Para la segunda desigualdad, basta aplicar el Teorema 1.2.11 y el Corolario 1.3.8. □

Definimos una relación de orden entre bases de funcionales no negativos.

**Definición 6.2.3.** Sean  $\alpha$  y  $\beta$  bases de funcionales no negativos. Decimos que  $\beta \preceq \alpha$  si existe una matriz  $K \geq 0$  tal que  $\alpha = K\beta$ .

Como el producto de matrices no negativas es no negativo, esto define un preorden.

**Lema 6.2.4.** Sean  $\alpha$  y  $\beta$  bases de funcionales no negativos. Si  $\alpha \preceq \beta$  y  $\beta \preceq \alpha$  entonces  $\alpha$  y  $\beta$  coinciden, salvo reordenación y escalado positivo.

*Demostración.* De la definición se deduce que  $\alpha = K\beta$ , con  $K^{-1} \geq 0$  y  $K \geq 0$ . Por el Lema 3.2 de [18],  $K$  es una matriz de permutación generalizada y, por tanto,  $K = DP$ , con  $D$  una matriz diagonal con elementos diagonales positivos y  $P$  una matriz de permutación. Por tanto, una permutación apropiada y un escalado positivo transforma una base en otra.  $\square$

El Teorema 6.2.1 nos permite afirmar que dos representaciones del mismo cuasi-interpolante con  $\alpha, \beta \geq 0$  tales que  $\beta \preceq \alpha$  verifican  $\text{cond}(x; \beta) \leq \text{cond}(x; \alpha)$ , para todo  $x$ . Pretendemos encontrar bases de funcionales no negativos tales que la representación asociada del cuasi-interpolante tenga el menor condicionamiento posible. Por ello es importante encontrar elementos minimales bajo la relación de orden anterior, ya que los condicionamientos correspondientes serán minimales en el sentido de que no podemos encontrar funcionales no negativos con menor condicionamiento en todos los puntos.

Para los funcionales de evaluación correspondientes al problema de interpolación de Lagrange se tiene el siguiente resultado.

**Teorema 6.2.5.** Sean  $\lambda_0, \dots, \lambda_n$  los funcionales de evaluación correspondientes al problema de interpolación de Lagrange en los nodos  $x_0, \dots, x_n$  en el espacio de funciones  $U$  de dimensión  $n+1$ . Sea  $\alpha$  una base de funcionales no negativos del espacio generado por  $\lambda_0, \dots, \lambda_n$ . Entonces  $\lambda \preceq \alpha$ . Además,  $\lambda$  es la única (salvo escalado positivo y permutación) base mínima entre las bases de funciones no negativos para la relación de orden  $\preceq$ .

*Demostración.* Sea  $K = (k_{ij})_{0 \leq i, j \leq n}$  la matriz tal que  $\alpha = K\lambda$ . Entonces, se tiene  $\alpha_i f = \sum_{j=0}^n k_{ij} f(x_j)$ . Como  $\alpha_i$  es no negativo, por la Proposición 6.1.2,  $k_{ij} \geq 0$ ,  $j = 0, \dots, n$ . Esto debe suceder para todo  $i = 0, \dots, n$ . Por tanto,  $K \geq 0$  y  $\lambda \preceq \alpha$ , lo que demuestra la minimalidad de  $\lambda$ . Si  $\alpha$  fuera otra base de funcionales no negativos mínima para la relación  $\preceq$ , tendríamos que  $\lambda \preceq \alpha$  y  $\alpha \preceq \lambda$ , luego por el Lema 6.2.4, se tiene el resultado.  $\square$

*Nota 6.2.6.* Por los teoremas 6.2.1 y 6.2.5, concluimos que la base  $\lambda_0, \dots, \lambda_n$  proporciona la representación más estable del operador de interpolación de Lagrange entre las bases

de funcionales no negativas en el sentido de que

$$\text{cond}(x; \lambda) \leq \text{cond}(x; \alpha),$$

para todo  $x$  y para toda base  $\alpha \geq 0$ . Esta propiedad también es consecuencia de la Proposición 2.1.7, donde se tiene un resultado análogo para cualquier base de funcionales  $\alpha$  sin la restricción de ser no negativas.

### 6.3 Comparación entre bases no negativas

Se dice que un sistema de funciones  $(u_0, \dots, u_n)$  definido en  $I \subseteq \mathbb{R}$  es no negativo si todas las funciones que la componen son no negativas, es decir,  $u_i(x) \geq 0$ , para todo  $x \in I, i = 0, \dots, n$ .

Sea  $F$  un espacio normado dotado de la norma del supremo y  $U \subset F$ . Sean  $u = (u_0, \dots, u_n)^T$  y  $b = (b_0, \dots, b_n)^T$  dos bases del espacio  $U$ . Sean los funcionales  $\alpha = (\alpha_0, \dots, \alpha_n)^T$  y  $\beta = (\beta_0, \dots, \beta_n)^T$  bases de  $F'$  de modo que  $u, \alpha$  y  $b, \beta$  forman dos pares de bases duales. Sea  $K$  la matriz de cambio de base entre  $u$  y  $b$ . Se tiene

$$(u_0, \dots, u_n) = (b_0, \dots, b_n)K,$$

donde  $K$  es una matriz no singular. Tenemos la siguiente relación entre las normas de los funcionales de las correspondientes representaciones.

**Lema 6.3.1.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases  $u = (u_0, \dots, u_n)^T$  y  $b = (b_0, \dots, b_n)^T$ . Sea la matriz de cambio de base  $K = (\beta_i u_j)_{0 \leq i, j \leq n}$  tal que  $u^T = b^T K$ . Si  $K \geq 0$  entonces

$$D_\beta e \leq K D_\alpha e.$$

*Demostración.* De (1.1.4) se tiene

$$\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = K \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix}. \quad (6.3.1)$$

Tomando normas en la fórmula (6.3.1) se deduce

$$D_\beta e \leq |K| D_\alpha e, \quad (6.3.2)$$

y, como  $K \geq 0$ , obtenemos  $D_\beta e \leq K D_\alpha e$ .  $\square$

Vamos a comparar los condicionamientos de dos representaciones con bases no negativas utilizando una condición sobre la matriz de cambio de base.

**Teorema 6.3.2.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases  $u = (u_0, \dots, u_n)^T$  y  $b = (b_0, \dots, b_n)^T$  no negativas. Sea la matriz de cambio de base  $K = (\beta_i u_j)_{0 \leq i, j \leq n}$  tal que  $u^T = b^T K$ . Si  $D_\beta e \leq K D_\alpha e$ , entonces

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u).$$

*Demostración.* Por la definición de condicionamiento

$$\text{cond}(x; \beta, b) = \sum_{j=0}^n \|\beta_j\|_\infty b_j(x) = b(x)^T D_\beta e \leq b(x)^T K D_\alpha e = u(x)^T D_\alpha e = \text{cond}(x; \alpha, u).$$

□

El resultado anterior nos permite comparar condicionamientos de dos representaciones con matriz de cambio de bases no negativa.

**Corolario 6.3.3.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases  $u = (u_0, \dots, u_n)^T$  y  $b = (b_0, \dots, b_n)^T$  no negativas. Sea la matriz de cambio de base  $K = (\beta_i u_j)_{0 \leq i, j \leq n}$  tal que  $u^T = b^T K$ . Si  $K \geq 0$ , entonces

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u).$$

*Demostración.* Basta aplicar el Lema 6.3.1 y el Teorema 6.3.2. □

Otro punto de vista alternativo del Corolario 6.3.3 proviene del uso del número de condición introducido por Farouki-Rajan (relacionado con (1.2.3)) y de una relación de orden entre bases no negativas que han sido desarrollados en [18], [31] y [42], entre otros artículos.

**Definición 6.3.4.** Sean  $b$  y  $u$  bases no negativas. Decimos que  $b \preceq u$  si existe una matriz  $K \geq 0$  tal que  $u^T = b^T K$ .

Con la misma demostración que la del Lema 6.2.4 se puede probar el siguiente resultado.

**Lema 6.3.5.** Sean  $b$  y  $u$  bases no negativas. Si  $u \preceq b$  y  $b \preceq u$  entonces  $b$  y  $u$  coinciden, salvo reordenación y escalado positivo.

Por tanto, el Corolario 6.3.3 afirma que si  $b$  y  $u$  son dos bases no negativas tales que  $b \preceq u$ , entonces  $\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u)$ . Si una base es minimal respecto a esta relación de orden  $\preceq$  entre las bases no negativas, entonces su condicionamiento también es minimal en el sentido de que ninguna otra base no negativa tiene menor condicionamiento en todos los puntos, salvo permutación y escalado.

La condición  $D_\beta e \leq K D_\alpha e$  equivale a  $\hat{K}e \geq e$ , donde  $\hat{K} = D_\beta^{-1} K D_\alpha$  es la matriz de cambio de base de las representaciones normalizadas. En el siguiente teorema obtenemos un resultado análogo a la Proposición 2.1.7.

**Teorema 6.3.6.** Sean  $\sum_{i=0}^n \alpha_i f u_i = \sum_{i=0}^n \beta_i f b_i$  dos representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a dos bases  $u = (u_0, \dots, u_n)^T$  y  $b = (b_0, \dots, b_n)^T$  no negativas. Sea la matriz de cambio de base  $K = (\beta_i u_j)_{0 \leq i, j \leq n}$  tal que  $u^T = b^T K$ . Si  $\hat{K}e \geq e$ , entonces

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u) \leq \text{Cond}(\hat{K}^{-1}) \text{cond}(x; \beta, b). \quad (6.3.3)$$

*Demostración.* Por el Teorema 6.3.2

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u),$$

y, aplicando el Teorema 1.2.11 y el Corolario 1.3.8, se tiene (6.3.3).  $\square$

## 6.4 Bases totalmente positivas

Recordemos que una matriz se dice totalmente positiva (TP) si todos sus menores son no negativos (véase la Sección 2.5). Por tanto, tomando los menores de orden 1, una matriz TP es no negativa. Dado un sistema de funciones  $(u_0, \dots, u_n)^T$  definido en un intervalo  $I \subseteq \mathbb{R}$ , la matriz de colocación de  $(u_0, \dots, u_n)^T$  en los puntos  $t_0 < \dots < t_m$  en  $I$  viene dada por

$$M \begin{pmatrix} u_0, u_1, \dots, u_n \\ t_0, t_1, \dots, t_m \end{pmatrix} := (u_j(t_i))_{i=0, \dots, m}^{j=0, \dots, n}.$$

Decimos que el sistema de funciones  $(u_0, \dots, u_n)^T$  es TP si todas sus matrices de colocación son TP. Es inmediato ver que un sistema de funciones TP es un sistema de funciones no negativo. Por tanto, vamos a poder aplicar algunos de los resultados de la sección anterior en este apartado.

Sean  $u = (u_0, \dots, u_n)^T$  un sistema de funciones linealmente independientes TP en  $I \subseteq \mathbb{R}$  y  $U$  el espacio generado por  $u$ . Como el producto de matrices TP es TP (véase la Proposición 2.5.3), se tiene que cualquier base  $v = (v_0, \dots, v_n)^T$  de la forma

$$v^T = u^T A, \quad A \text{ matriz no singular TP,}$$

es también TP.

Una base  $b = (b_0, \dots, b_n)^T$  TP de un espacio  $U$  de funciones definidas en  $I \subseteq \mathbb{R}$  se llama B-base, si todas las bases  $v$  TP de  $U$  verifican

$$v^T = b^T A, \quad A \text{ matriz no singular TP.}$$

Si existe una base TP de  $U$ , entonces el espacio  $U$  admite una B-base (véase el Teorema 3.6 de [17] o el Teorema 2.2 de [18]). Además, podemos establecer la siguiente relación entre cualquier base  $v$  TP con una B-base  $b$ ,

$$b \preceq v,$$

donde  $\preceq$  es el preorden dado por la Definición 6.3.4. Por tanto, una B-base es mínima para la relación  $\preceq$  entre todas las bases TP (véase también la Proposición 3.7 de [18]). En consecuencia, usando el Corolario 6.3.3, se tiene el siguiente resultado que muestra que una B-base tiene el condicionamiento mínimo respecto a todas las bases TP.

**Teorema 6.4.1.** *Sean  $\sum_{i=0}^n \beta_i f b_i$  la representaciones de un cuasi-interpolante en  $C[a, b]$  con respecto a la B-base  $b = (b_0, \dots, b_n)^T$ . Si  $\sum_{i=0}^n \alpha_i f u_i$  es otra representación del cuasi-interpolante respecto a una base TP  $u = (u_0, \dots, u_n)^T$ . Entonces*

$$\text{cond}(x; \beta, b) \leq \text{cond}(x; \alpha, u),$$

*es decir, la representación respecto a la B-base tiene el condicionamiento mínimo entre todas las representaciones con bases TP.*

Como ejemplos de B-bases podemos citar la base de Bernstein en el espacio  $P_n$  (véase la Sección 2.5), los polinomios de grado menor o igual que  $n$ , en un intervalo compacto (véase la Sección 4 de [16]), y la base de B-splines (véase [17]).

## 6.5 Base de Bernstein

En esta sección consideramos la siguiente representación del operador de interpolación de Lagrange

$$L[f](x) = \sum_{i=0}^n \mu_i f b_{i,n}(x), \quad x \in [0, 1],$$

donde  $\mathcal{B}_n = (b_{0,n}, \dots, b_{n,n})$ ,

$$b_{i,n}(x) = \binom{n}{i} (1-x)^{n-i} x^i, \quad i = 0, \dots, n,$$

es la base de Bernstein en el intervalo  $[0, 1]$ .

Como ya hemos dicho en la sección anterior, la base de Bernstein es una B-base del espacio de polinomios de grado menor o igual que  $n$ ,  $P_n$ , en el intervalo  $[0, 1]$ . Por el Teorema 6.4.1, el condicionamiento respecto a la base de Bernstein del interpolante es menor o igual que el condicionamiento respecto a cualquier otra base TP. La base de monomios  $(m_0, \dots, m_n)$ ,  $m_j(x) = x^j$ ,  $j = 0, \dots, n$  es TP si los nodos verifican  $0 \leq x_0 < x_1 < \dots < x_n$  (véase el Ejemplo 2 de la Sección 2.5). Por tanto, la base de monomios en  $[0, 1]$  da lugar a una representación con peor condicionamiento que la representación con respecto a la base de Bernstein. Por otro lado, recordemos que la base de Lagrange es la base que da lugar a la representación con mínimo condicionamiento (véase la Proposición 2.1.7). Sin embargo, no es TP ya que las funciones básicas cambian de signo en los nodos interiores y, por tanto, no podemos aplicar el Teorema 6.4.1.

Para ilustrar estos comportamientos consideremos el problema de interpolación con nodos equidistantes en el intervalo  $[0, 1]$  y grado  $n = 10$ . La Figura 6.5.1 muestra una comparación entre la representaciones del polinomio de interpolación de Lagrange respecto a la base de Bernstein, de monomios y de Lagrange. En la gráfica se confirma que la base de Lagrange es la que tiene menor condicionamiento, la de monomios la que tiene mayor condicionamiento, mientras que la de Bernstein tiene un comportamiento intermedio.

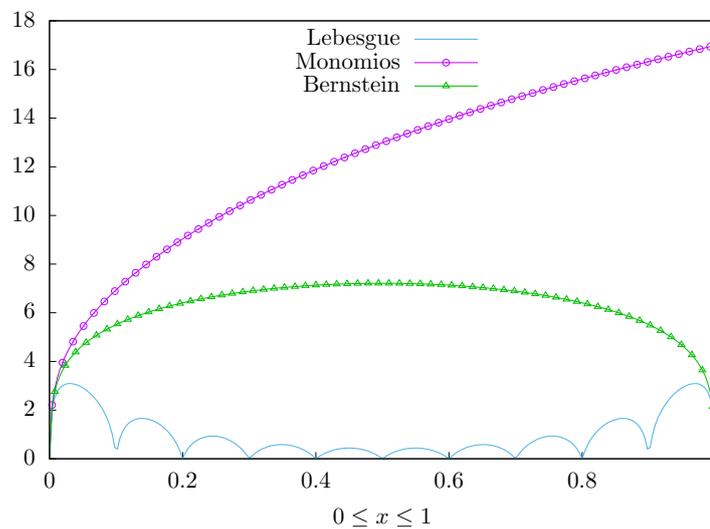


Figura 6.5.1: Gráfica de  $\log_3(\lambda_L(x))$ ,  $\log_3(\text{cond}(x; \alpha))$  respecto a la base de monomios y respecto a la base de Bernstein en nodos equidistantes en  $[0, 1]$  para  $n = 10$ .

Para calcular el condicionamiento hemos utilizado el Teorema 2.1.6 tomando como matriz  $M$  la matriz de colocación de las bases de Bernstein y de monomios, respectivamente. Veamos que podemos obtener los elementos de la inversa con alta precisión relativa, HRA (véase la Sección 2.6), lo que permite obtener, de manera fiable, el condicionamiento de estas representaciones. Por el Teorema 2.1.5, la matriz de cambio de base entre la base de Lagrange y la base  $\mathcal{B}_n$  es la matriz de colocación de la base de Bernstein en los nodos  $x_0, \dots, x_n$

$$B = (b_{j,n}(x_i))_{0 \leq i,j \leq n} = \begin{pmatrix} \binom{n}{0} (1-x_0)^n & \binom{n}{1} (1-x_0)^{n-1} x_0 & \cdots & \binom{n}{n} x_0^n \\ \binom{n}{0} (1-x_1)^n & \binom{n}{1} (1-x_1)^{n-1} x_1 & \cdots & \binom{n}{n} x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n}{0} (1-x_n)^n & \binom{n}{1} (1-x_n)^{n-1} x_n & \cdots & \binom{n}{n} x_n^n \end{pmatrix}, \quad (6.5.1)$$

donde  $x_0, \dots, x_n \in [0, 1]$  son nodos distintos. Esta matriz se conoce como matriz de Bernstein-Vandermonde (véase también la Sección 2.5). Así,

$$(b_{0,n}, \dots, b_{n,n}) = (l_0, \dots, l_n)B,$$

donde  $(l_0, \dots, l_n)$  es la base de Lagrange.

En [43] los autores obtuvieron una descomposición bidiagonal explícita de la matriz de Bernstein-Vandermonde que permite calcular su inversa con HRA. Más aún, los factores  $L$  y  $U$  de su descomposición triangular se pueden obtener a través de dichos factores bidiagonales, y por tanto, con HRA. A continuación, damos una construcción alternativa para obtener la inversa, similar a la que hemos utilizado en la Sección 2.6 para la matriz de Vandermonde.

**Teorema 6.5.1.** *Sea  $B$  la matriz de Bernstein-Vandermonde dada por (6.5.1) correspondiente a una sucesión de nodos distintos en  $[0, 1]$ . Entonces  $B^{-1}$  puede calcularse con HRA.*

*Demostración.* Haciendo el cambio de variable  $x = t/(t+1)$ , se tiene  $\tilde{l}_i(t) := l_i\left(\frac{t}{t+1}\right)$ ,  $\tilde{b}_{i,n}(t) := b_{i,n}\left(\frac{t}{t+1}\right)$ ,  $i = 0, \dots, n$ , de donde se obtiene

$$\tilde{b}_{i,n}(t) = \binom{n}{i} \frac{t^i}{(1+t)^n}, \quad i = 0, \dots, n.$$

Por tanto,  $(\tilde{l}_0, \dots, \tilde{l}_n) = (\tilde{b}_{0,n}, \dots, \tilde{b}_{n,n})B^{-1}$ . Expresando los elementos  $\tilde{l}_j$  respecto a  $\tilde{b}_{0,n}, \dots, \tilde{b}_{n,n}$ , obtenemos

$$\tilde{l}_j(t) = \prod_{k \in \{0, \dots, n\} \setminus \{j\}} \frac{t/(t+1) - x_k}{x_j - x_k} = \frac{1}{(t+1)^n} \prod_{k \in \{0, \dots, n\} \setminus \{j\}} \frac{(1-x_k)t - x_k}{(x_j - x_k)}.$$

Desarrollando

$$\prod_{k \in \{0, \dots, n\} \setminus \{j\}} (1-x_k)t - x_k = \sum_{i=0}^n c_{ij} t^i,$$

obtenemos

$$c_{ij} = (-1)^{n-i} \sum_{k_1 < \dots < k_{n-i} \in \{0, \dots, n\} \setminus \{j\}} x_{k_1} \cdots x_{k_{n-i}} \prod_{k \in \{0, \dots, n\} \setminus \{k_1, \dots, k_{n-i}, j\}} (1-x_k).$$

Tenemos

$$\tilde{l}_j(t) = \sum_{i=0}^n \frac{c_{ij}}{\binom{n}{i}} \tilde{b}_{j,n}(t), \quad j = 0, \dots, n.$$

Por tanto, si  $B^{-1} = (b_{ij}^{-1})_{0 \leq i, j \leq n}$ , se tiene que el elemento  $(i, j)$  de esta matriz es

$$b_{ij}^{-1} = \frac{c_{ij}}{\binom{n}{i}}$$

Como  $c_{ij}$  es suma de productos de expresiones que pueden calcularse con HRA, deducimos que  $c_{ij}$  y, por tanto, los elementos de la inversa de  $B$  pueden calcularse con HRA.  $\square$

Veamos que también podemos obtener una descomposición en factores triangulares de la matriz  $B$  con HRA. El elemento  $(i, j)$  de la matriz de Bernstein-Vandermonde en nodos  $0 \leq x_0 < \dots < x_n \leq 1$  es

$$b_{j,n}(x_i) = \binom{n}{j} (1-x_i)^{n-j} x_i^j = \binom{n}{j} (1-x_i)^n \left( \frac{x_i}{1-x_i} \right)^j.$$

Supongamos primero que  $x_n < 1$ . Sean  $t_i := x_i/(1-x_i)$ ,  $i = 0, \dots, n$ . Entonces

$$b_{j,n}(x_i) = \binom{n}{j} (1-x_i)^n t_i^j.$$

Por lo tanto, podemos expresar la matriz de Bernstein-Vandermonde como el producto de la matriz de Vandermonde en nodos  $t_0, \dots, t_n$  y dos matrices diagonales, es decir,

$$B = \text{diag}((1-x_0)^n, \dots, (1-x_n)^n) V(t_0, \dots, t_n) \text{diag} \left( \binom{n}{0}, \dots, \binom{n}{n} \right). \quad (6.5.2)$$

Sea  $V(t_0, \dots, t_n) = LU$ , con  $L$  la matriz triangular inferior dada por (2.2.6) y  $U$  la matriz triangular superior definida en (2.6.5). Definimos las matrices

$$K := \text{diag}((1 - x_0)^n, \dots, (1 - x_n)^n) L \quad (6.5.3)$$

y

$$W := U \text{diag} \left( \binom{n}{0}, \dots, \binom{n}{n} \right), \quad (6.5.4)$$

matrices triangulares inferior y superior, respectivamente. Así,  $B = KW$  es una factorización  $LU$  de la matriz de Bernstein-Vandermonde.

En el caso en el que  $x_n = 1$ , no tienen sentido, cuando  $t_n \rightarrow \infty$ , ni la matriz de Vandermonde ni su factor triangular  $L$ . Sin embargo,  $K$  y  $W$  están bien definidas. Comencemos por la matriz triangular inferior  $K = (k_{ij})_{0 \leq i, j \leq n}$ . Los elementos no nulos de esta matriz son

$$k_{ij} = (1 - x_i)^n l_{ij} = (1 - x_i)^n (t_i - t_0) \cdots (t_i - t_{j-1}) = \frac{(1 - x_i)^{n-j} (x_i - x_0) \cdots (x_i - x_{j-1})}{(1 - x_0) \cdots (1 - x_{j-1})},$$

para  $j \leq i$ . La simplificación del factor  $(1 - x_n)^j$  permite deducir que existe el límite cuando  $x_n \rightarrow 1$  de  $k_{nj}$ ,  $j = 0, \dots, n$ . La definición de  $U$  y  $W$  no ofrece ningún problema. Por tanto, tomando límites cuando  $x_n \rightarrow 1$ , podemos concluir que la factorización  $B = KW$  es válida incluso si  $x_n = 1$ .

Veamos que podemos calcular estas matrices triangulares y sus inversas con HRA. Para ello, vamos a utilizar los algoritmos (2.3.1), (2.3.2), (2.6.6) y (2.6.9) que nos permiten calcular  $L$ ,  $U$  y sus inversas con HRA.

Claramente,  $K$  puede calcularse con HRA porque

$$k_{ij} = \frac{(1 - x_i)^{n-j} (x_i - x_0) \cdots (x_i - x_{j-1})}{(1 - x_0) \cdots (1 - x_{j-1})}, \quad j \leq i,$$

son productos de diferencias de datos iniciales. Podemos expresar los elementos  $k_{ij}$  como

$$k_{ij} = (1 - x_i)^{n-j} r_{ij},$$

donde

$$r_{ij} = \frac{(x_i - x_0) \cdots (x_i - x_{j-1})}{(1 - x_0) \cdots (1 - x_{j-1})}, \quad j \leq i,$$

y los  $r_{ij}$  podemos calcularlos por recurrencia mediante el siguiente algoritmo

$$\begin{aligned} r_{i0} &= 1, \quad i = 0, \dots, n, \\ r_{ij} &= r_{i,j-1} \frac{x_i - x_{j-1}}{1 - x_{j-1}}, \quad i = j, \dots, n, \quad j = 1, \dots, n. \end{aligned} \quad (6.5.5)$$

Por la fórmula (6.5.3), se tiene que  $K^{-1} = L^{-1} \text{diag}(1/(1-x_0)^n, \dots, 1/(1-x_n)^n)$ . Aplicando la expresión explícita de los elementos de  $L^{-1}$  (2.2.8) con nodos  $t_0, \dots, t_n$ , obtenemos que el elemento  $(i, j)$  de  $K^{-1}$  es

$$k_{ij}^{(-1)} = \frac{l_{ij}^{(-1)}}{(1-x_j)^n} = \frac{1}{(1-x_j)^n \omega'_{i+1}(t_j)} = \frac{1}{(1-x_j)^{n-i}} \prod_{k \in \{0, \dots, i\} \setminus \{j\}} \frac{1-x_k}{x_j-x_k}, \quad j \leq i.$$

Podemos expresar los elementos de  $K^{-1}$  como

$$k_{ij}^{(-1)} = \frac{1}{(1-x_j)^{n-i}} r'_{ij},$$

donde

$$r'_{ij} = \prod_{k \in \{0, \dots, i\} \setminus \{j\}} \frac{1-x_k}{x_j-x_k}, \quad j \leq i.$$

Por tanto, podemos calcular los  $r'_{ij}$  con el siguiente algoritmo

$$r'_{ij} = -r'_{i-1,j} \frac{1-x_i}{x_i-x_j}, \quad j = 0, \dots, i-1, \quad r'_{ii} = \prod_{j=0}^{i-1} \frac{1-x_j}{x_i-x_j}, \quad i = 0, \dots, n. \quad (6.5.6)$$

Para la matriz  $W = (w_{ij})_{0 \leq i, j \leq n}$  se tiene

$$w_{ij} = \binom{n}{j} u_{ij},$$

donde  $U = (u_{ij})_{0 \leq i, j \leq n}$  es la matriz dada por (2.6.5) en los nodos  $\frac{x_0}{1-x_0}, \dots, \frac{x_n}{1-x_n}$ . Aplicando (2.6.6), obtenemos un algoritmo para calcular los elementos de  $U$

$$\begin{aligned} u_{00} &= 1, & u_{0j} &= \frac{x_0}{1-x_0} u_{0,j-1}, & j &= 1, \dots, n, \\ u_{ii} &= 1, & u_{ij} &= u_{i-1,j-1} + \frac{x_i}{1-x_i} u_{i,j-1}, & j &= i+1, \dots, n, & i &= 1, \dots, n-1, \\ u_{nn} &= 1. \end{aligned} \quad (6.5.7)$$

Sea  $W^{-1} = (w_{ij}^{(-1)})_{0 \leq i, j \leq n}$ . Por (6.5.4), los elementos de  $W^{-1}$  son de la forma

$$w_{ij}^{(-1)} = \frac{1}{\binom{n}{i}} u_{ij}^{(-1)},$$

y, usando el algoritmo (2.6.9), podemos calcular la matriz  $U^{-1}$  como sigue

$$\begin{aligned} u_{00}^{(-1)} &= 1, & u_{0j}^{(-1)} &= -\frac{x_{j-1}}{1-x_{j-1}} u_{0,j-1}^{(-1)}, & j &= 1, \dots, n, \\ u_{ii}^{(-1)} &= 1, & u_{ij}^{(-1)} &= u_{i-1,j-1}^{(-1)} - \frac{x_{j-1}}{1-x_{j-1}} u_{i,j-1}^{(-1)}, & j &= i+1, \dots, n, & i &= 1, \dots, n. \end{aligned} \quad (6.5.8)$$

# Conclusiones

---

Esta memoria trata del estudio de la estabilidad de diferentes aproximaciones polinómicas, con especial énfasis en interpolación. A pesar de que diversos autores han establecido comparaciones entre las fórmulas de Lagrange y de Newton y mencionado su diferente comportamiento en estabilidad, no hemos encontrado en la literatura un análisis cuantitativo que permita explicar este diferente comportamiento. Esto ha motivado nuestro estudio, que hemos extendido además a otras representaciones y a problemas de mínimos cuadrados polinómicos.

A continuación señalaremos las principales aportaciones de esta memoria.

- Hemos definido un condicionamiento asociado a cada representación de un operador de cuasi-interpolación. Hemos comparado los condicionamientos de diferentes representaciones del mismo operador a través de los cambios de base, mostrando el importante papel que juega en este estudio el condicionamiento de Skeel de una matriz.
- En el caso del problema de interpolación polinómica de Lagrange, hemos demostrado que la representación de Lagrange es óptima y, además, su condicionamiento coincide con la norma del operador.
- Al establecer comparaciones entre las fórmulas de Lagrange y de Newton aparece una matriz de cambio de base triangular inferior  $L$  que relaciona los valores en los nodos y las diferencias divididas. Esta matriz  $L$  es el factor triangular inferior de una descomposición triangular de la matriz de Vandermonde  $V = LU$ . Hemos probado que tanto  $V$ ,  $L$  y  $U$  como sus inversas pueden calcularse bajo ciertas condiciones con alta precisión relativa.
- Hemos descrito el condicionamiento de la fórmula de Newton, que depende del orden de los nodos.

- En el caso de la fórmula de Newton en el caso de nodos equidistantes ordenados de forma creciente hemos probado que el máximo condicionamiento es  $3^n$  y se alcanza en el último nodo. Este valor es mucho mayor que el condicionamiento de la representación de Lagrange, que tiene un crecimiento asintótico de  $\frac{2^{n+1}}{e n \log(n+\gamma)}$ . También hemos realizado este estudio para nodos casi equidistantes y nodos en progresión geométrica.
- Como consecuencia del análisis anterior, se obtiene el condicionamiento de Skeel de la matriz de Pascal triangular inferior cuyo valor es  $3^n$ .
- Hemos obtenido el condicionamiento en norma  $\infty$  de la matriz triangular  $L$  y demostrado su crecimiento exponencial con una razón  $r$  que depende de la longitud del intervalo. El menor crecimiento asintótico se alcanza para intervalos de longitud 3 y, en este caso, el condicionamiento admite la siguiente expresión asintótica  $\sqrt{\frac{3}{2}}3^n$ .
- Analizamos el orden de Leja para la fórmula de Newton. Demostramos que la fórmula de Newton con nodos siguiendo el orden de Leja tiene máximo condicionamiento en los nodos acotado por  $2^{n+1} - 1$ . A la vista de los experimentos numéricos, esta cota parece pesimista.
- Hemos introducido y analizado la fórmula de Newton con nodos siguiendo un orden central respecto a un centro. En el caso particular de nodos equidistantes, acotamos el condicionamiento por  $(1 + \sqrt{2})^{n+1}$ , una cota menor que el máximo condicionamiento  $3^n$ , correspondiente al orden creciente.
- Presentamos la fórmula de Newton con nodos siguiendo el orden central respecto al punto de evaluación que, a pesar de su mayor coste computacional, presenta ciertas propiedades de optimalidad. En el caso particular de nodos equidistantes, se obtienen cotas del condicionamiento muy próximas al condicionamiento de la fórmula de Lagrange.
- Presentamos y analizamos la fórmula de Newton con nodos siguiendo el orden central inverso. Para nodos equidistantes también se obtienen cotas próximas al condicionamiento de la fórmula de Lagrange.
- Interpretamos el orden de Leja y el orden central inverso en términos de estrategias de pivotaje parcial para las eliminaciones de Gauss y de Neville, respectivamente.
- A la vista de nuestro resultados teóricos y experimentaciones numéricas, podemos concluir que el orden de Leja da buenos resultados globales pero es difícil de

analizar teóricamente. En contraste, el orden central respecto al punto de evaluación y el orden central inverso permiten realizar un análisis teórico cuantitativo y proporcionan un condicionamiento máximo muy ajustado.

- Estudiamos representaciones con bases ortogonales, muy utilizadas en la práctica porque corresponden a sumas de Fourier.
- En el caso de polinomios ortogonales clásicos probamos que el condicionamiento del problema de mínimos cuadrados continuo crece de forma subexponencial. Para polinomios de Legendre es del orden  $n^{3/2}$ , para polinomios de Chebyshev de primera especie es un crecimiento lineal y para polinomios de Chebyshev de segunda especie es un crecimiento cuadrático. En el caso de mínimos cuadrados discretos se tienen las mismas propiedades.
- En el caso de que el número de nodos sea igual al  $n+1$ , se obtienen representaciones ortogonales del interpolante en nodos que son ceros de polinomios ortogonales clásicos.
- Estudiamos las consecuencias que tienen la positividad de bases y funcionales en la obtención de bases que tengan buen condicionamiento. Confirmamos la optimalidad de la representación de Lagrange con respecto a otras con funcionales positivos y la base de Bernstein con respecto a otras bases totalmente positivas.



# Conclusions

---

This work studies the stability of different polynomial approximations, with special emphasis on interpolation. Although several authors have performed comparisons between the Lagrange and the Newton formulas and mentioned their different stability properties, we have not found in the literature a quantitative analysis to explain this different behavior. This has motivated our study, which we have also extended to other representations and polynomial least squares problems.

Next we will be pointing out the main contributions of this work.

- We have defined a conditioning associated to a representation of a quasi-interpolation operator. We have compared the conditioning of different representations of the same operator in terms of the matrix of change of basis, showing the important role played by the Skeel condition number of a matrix in this study.
- In the case of the Lagrange polynomial interpolation problem, we have shown that the Lagrange representation is optimal and, in addition, its conditioning coincides with the norm of the operator.
- In comparisons between the Lagrange and the Newton formulas, a lower triangular matrix of change of basis  $L$  relates the values at the nodes and the divided differences. This matrix  $L$  is the lower triangular factor of a triangular decomposition of the Vandermonde matrix  $V = LU$ . We have proven that  $V$ ,  $L$  and  $U$  and their inverses can be computed under certain conditions with high relative accuracy.
- We have described the conditioning of the Newton formula, which depends on the order of the nodes.
- In the case of the Newton formula with equidistant nodes in increasing order, we have proven that the maximum conditioning is  $3^n$  and it is attained in the last node. This value is much greater than the conditioning of the Lagrange representation,

which has an asymptotic growth of  $\frac{2^{n+1}}{en \log(n+\gamma)}$ . We have also performed this study for almost equidistant nodes and nodes in geometric progression.

- As a consequence of the previous analysis, we obtain that the Skeel condition number of the lower triangular Pascal matrix is  $3^n$ .
- We have obtained the condition number in  $\infty$ -norm of the lower triangular matrix  $L$  and proven its exponential growth with a ratio  $r$  that depends on the interval length. The lowest asymptotic growth is attained for intervals of length 3 and, in this case, the conditioning has the following asymptotic expression  $\sqrt{\frac{3}{2}}3^n$ .
- We analyze the Leja order for the Newton formula. We show that the formula of Newton with nodes following the Leja order has maximum conditioning in the nodes bounded above by  $2^{n+1} - 1$ . In view of the numerical experiments, this bound seems pessimistic.
- We have introduced and analyzed the Newton formula with nodes following a central order with respect to a center. In the particular case of equidistant nodes, the conditioning is bounded above by  $(1 + \sqrt{2})^{n+1}$ , a value lower than the maximum conditioning  $3^n$ , corresponding to the increasing order.
- We present the Newton formula with nodes following a central order with respect to the evaluation point that, despite its higher computational cost, has certain optimality properties. In the particular case of the equidistant nodes, we obtain bounds of the conditioning very close to the conditioning of the Lagrange formula.
- We present and analyze the Newton formula with nodes the inverse central order. For equidistant nodes, we also obtain bounds close to the conditioning of the Lagrange formula.
- We interpret the Leja and the inverse central orderings in terms of partial pivoting strategies for Gauss and Neville eliminations, respectively.
- Considering our theoretical results and numerical experiments, we can conclude that the Leja order gives good global results but it is difficult to analyze theoretically. In contrast, the central order with respect to the evaluation point and the inverse central order allow us a quantitative theoretical analysis and provide a tight maximum conditioning.
- We study representations with orthogonal bases, often used in practice because they correspond to Fourier sums.

- In the case of classical orthogonal polynomials, we show that the conditioning of the continuous least squares problem grows subexponentially. For Legendre polynomials the growth of the conditioning is of the order  $n^{3/2}$ , for Chebyshev polynomials of the first kind, it is a linear growth and for Chebyshev polynomials of the second kind, it is a quadratic growth. In the case of discrete least squares, we have the same properties.
- If the number of nodes is equal to  $n + 1$ , we recover orthogonal representations of the interpolating polynomial at the zeros of orthogonal polynomials.
- We study the consequences of the positivity of the bases and functionals for obtaining bases with good conditioning. We confirm the optimality of the Lagrange representation with respect to other one with positive functionals and of the Bernstein basis with respect to other totally positive bases.



# Bibliografía

---

- [1] M. ABRAMOWITZ, I. STEGUN: *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Applied Mathematics Series, 55, Washington D.C., NY, 1972.
- [2] P. ALONSO, J. DELGADO, R. GALLEGO, J.M. PEÑA: *Growth factors of pivoting strategies associated to Neville elimination*, J. Comput. Appl. Math. 235 (2011), pp. 1755-1762.
- [3] T. ANDO: *Totally positive matrices*, Linear Algebra Appl. 90 (1987), pp. 165-219.
- [4] S. BANACH, H. STEINHAUS: *Sur le principe de la condensation de singularités*, Fundamenta Mathematicae, 9 (1927), pp. 50-61
- [5] R. BEALS, R. WONG: *Special functions and orthogonal polynomials*, Cambridge University Press, Cambridge, 2016.
- [6] S. N. BERNSTEIN: *Sur la limitation des valeurs d'un polynôme  $P_n(x)$  de degré  $n$  sur tout un segment par ses valeurs en  $n + 1$  pointes du segment*, Bull. Acad. Scie. URSS, Leningrad (1931), pp. 1025-1050.
- [7] C. DE BOOR: *A practical guide to splines*, Revised edition, Applied Mathematical Sciences, 27. Springer-Verlag, New York, 2001.
- [8] C. DE BOOR, G. J. FIX: *Spline approximation by quasiinterpolants*, J. Approximation Theory 8 (1973), pp. 19-45.
- [9] C. DE BOOR, A. PINKUS: *Backward error analysis for totally positive linear systems*, Numer. Math, 27 (1976/77), pp. 485-490.
- [10] L., BRUTMAN: *On the Lebesgue function for polynomial interpolation*, SIAM J. Numer. Anal. 15(4)(1978), pp. 694-704.

- [11] L. BOS, S. DE MARCHI, A. SOMMARIVA, M. VIANELLO: *Computing multivariate Fekete and Leja points by numerical linear algebra*, SIAM J. Numer. Anal., 48 (2010), pp. 1984-1999.
- [12] J. M. CARNICER, Y. KHIAR, J. M. PEÑA: *Matrix analysis of the Newton interpolation formula*, Proceedings of the XXIV Congress on Differential Equations and Applications/XIV Congress on Applied Mathematics, J. M. Díaz Moreno et al. (Eds.), Universidad de Cádiz, Cádiz, 2015, pp. 441-446.
- [13] J. M. CARNICER, Y. KHIAR, J. M. PEÑA: *A matrix approach to the Newton formula and divided differences*, Trends in differential equations and applications, SEMA SIMAI Springer Ser., Springer, 8 (2016), pp. 107-123.
- [14] J. M. CARNICER, Y. KHIAR, J. M. PEÑA: *Factorization of Vandermonde matrices and the Newton formula*, Thirteenth International Conference Zaragoza-Pau on Mathematics and its Applications, Monogr. Mat. García Galdeano, 40, Prensas Univ. Zaragoza, Zaragoza, 2016, pp. 53-60.
- [15] J. M. CARNICER, Y. KHIAR, J. M. PEÑA: *Optimal stability of the Lagrange formula and conditioning of the Newton formula*, aparecerá en J. Approx. Theory, DOI: 10.1016/j.jat.2017.07.005.
- [16] J. M. CARNICER, J. M. PEÑA: *Shape preserving representations and optimality of the Bernstein basis*, Advances in Computational Mathematics, 1 (1993), pp. 173-196.
- [17] J. M. CARNICER, J. M. PEÑA: *Totally positive bases for shape preserving curve design and optimality of B-splines*, Comput. Aided Geom. Design, 11 (1994), pp. 633-654.
- [18] J. M. CARNICER, J. M. PEÑA: *Total positivity and optimal bases*. En: *Totally positivity and its applications* (M. Gasca and C. A. Micchelli, Eds.), 1996, pp. 133-155.
- [19] J. M. CARNICER, T. N. T. GOODMAN, J. M. PEÑA: *Roundoff errors for polynomial evaluation by a family of formulae*, Computing, 82 (2008), pp. 199-215.
- [20] E. W. CHENEY: *Introduction to Approximation Theory*, Second Edition, AMS Chelsea Publishing, Providence RI, 1982.

- [21] E. W. CHENEY: *Analisis for Applied Mathematics*, Springer-Verlag, New York, 2001.
- [22] E. W. CHENEY, W. LIGHT: *A Course in Approximation Theory*, Brooks/Cole Publishing Company, Pacific Groove CA, 2000.
- [23] C. W. CRYER: *Some properties of Totally Positive Matrices*, Linear Algebra Appl., 15 (1976), pp. 1-25.
- [24] P. J. DAVIS: *Interpolation and Approximation*, Dover Publications, Mineola NY, 2014.
- [25] J. DEMMEL, I. DUMITRIU, O. HOLTZ: *Accurate and efficient expression evaluation and linear algebra*, Acta Numer., 17 (2008), pp. 87-145.
- [26] J. DEMMEL, M. GU, S. EISENSTAT, I. SLAPNIŃAR, K. VESELIĆ, Z. DRMAČ: *Computing the singular value decomposition with high relative accuracy*, Linear Algebra Appl., 299 (1999), pp. 21-80.
- [27] J. DEMMEL, J., P. KOEV: *The accurate and efficient solution of a totally positive generalized Vandermonde linear system*, SIAM J. Matrix Anal. Appl., 27 (2005) pp.142-152.
- [28] P. ERDŐS: *Problems and results on the theory of interpolation. II*, Acta Math. Acad. Sci. Hungar., 12 (1961), pp. 235-244
- [29] FABER: *Über die interpolatorische Darstellung stetiger Funktionen*, DMV, 23 (1914), pp. 192-210.
- [30] S. M. FALLAT, C.R. JOHNSON: *Totally Nonnegative Matrices*, Princeton University Press, Princeton and Oxford, 2011.
- [31] R. T. FAROUKI, V.T. RAJAN: *On the numerical condition of algebraic curves and surfaces 1. Implicit equations*, Computer Aided Geometric Design, 5 (1988), pp. 215-252.
- [32] M. FOURIER: *Mémoire sur la propagation de la chaleur dans les corps solides*, Nouveau Bulletin des Sciencies par la Société philomathique de Paris, 1 (1807), pp. 112-116.
- [33] F. P. GANTMACHER, M. G. GREIN: *Oscillation matrices and kernels and small vibrations of mechanical systems. Revised edition*, AMS Chelsea publishing, Providence, 2002.

- [34] M. GASCA, J. M. PEÑA: *Total positivity and Neville elimination*, Linear Algebra Appl., 165 (1992), pp. 25–44.
- [35] M. GASCA, J. M. PEÑA: *Scaled pivoting in Gauss and Neville elimination for totally positive systems*, Applied Numerical Mathematics, 13 (1993), pp. 354-355.
- [36] N. J. HIGHAM: *Stability analysis of algorithms for solving confluent Vandermonde-like systems*, SIAM J. Matrix Anal. Appl., 1 (1990), pp. 23-41.
- [37] N. J. HIGHAM: *Accuracy and Stability of Numerical Algorithms (Second Edition)*, SIAM, Philadelphia, PA, 2002.
- [38] S. KARLIN: *Total positivity*, volumen 1, Stanford University Press, Stanford, CA, USA, 1968.
- [39] P. KOEV: *Accurate computations with totally nonnegative matrices*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 731-751.
- [40] H. LEBESGUE: *Sur les intégrales singulières*, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys., 1 (1909), ppg. 25-117.
- [41] F. LEJA: *Sur certaines suites liées aux ensembles plans et leur application à la représentation conforme*, Ann. Polon. Math., 4 (1957), pp. 8–13.
- [42] T. LYCHE, J. M. PEÑA: *Optimally stable multivariate bases*, Advances in Computational Mathematics, 20 (2004), pp. 149-159.
- [43] A. MARCO, J.J. MARTÍNEZ: *A fast and accurate algorithm for solving Bernstein-Vandermonde linear systems*, Linear Algebra Appl. 422 (2007), pp. 616–628.
- [44] H. MINK: *Nonnegative matrices*, Wiley Interscience, New York, 1988.
- [45] V. F. NIKOLAEV: *Concerning the approximation of continuous functions by polynomials*, Doklady Akad. Nauk SSSR, 61 (1948), pp. 201-204.
- [46] J. M. PEÑA: *Pivoting strategies leading to small bounds of the errors for certain linear systems*, IMA Journal of Numerical Analysis, 16 (1996), pp. 141-153.
- [47] J. M. PEÑA: *On the Skeel condition number, growth factor and pivoting strategies for Gaussian elimination*, Proceedings of the SIAM Conference on Applied Linear Algebra, Williamsburg, VA, 2003.

- [48] J. M. PEÑA: *B-splines and optimal stability*, Math. Comp. 66 (1997), pp. 1555-1560.
- [49] G. M. PHILLIPS: *Interpolation and Approximation by Polynomials*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 14. Springer-Verlag, New York, 2003.
- [50] A. PINKUS: *Totally positive matrices*, Cambridge Tracts in Mathematics, 181. Cambridge University Press, Cambridge, 2010.
- [51] L. REICHEL: *Newton interpolation at Leja points*, BIT 30, 2 (1990), pp. 332-346.
- [52] T. J. RIVLIN: *Chebyshev polynomials. From approximation theory to algebra and number theory*, Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1990.
- [53] H. ROBBINS: *A Remark on Stirling's Formula*, The American Mathematical Monthly, 62 (1955), pp. 26-29.
- [54] W. RUDIN: *Real and complex analysis*, Mac Graw-Hill, London, 1970.
- [55] C. RUNGE: *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten*, Zeitschrift für Mathematik und Physik, 46 (1901), pp. 224-243.
- [56] E. B. SAFF, V. TOTIK: *Logarithmic potentials with external fields*, Grundlehren der mathematischen Wissenschaften, Volume 316, Springer, Berlin, 1994.
- [57] E. B. SAFF: *Logarithmic Potential Theory with Applications to Approximation Theory*, Surveys in Approximation Theory, 5 (2010), pp. 165-200.
- [58] A. SCHÖNHAGE: *Fehlerfortpflanzung bei Interpolation*, Numer. Math, 3 (1961), pp. 62-71.
- [59] J. F. STEFFENSEN: *Interpolation*, Second Edition, Chelsey Publishing Company, New York, 1965.
- [60] G. SZEGÖ: *Orthogonal Polynomials*, Fourth Edition, American Mathematical Society, Colloquium Publications, Providence RI, 1975.
- [61] R. TAYLOR, V. TOTIK: *Lebesgue constants for Leja points*, IMA J. Numer. Anal., 30 (2010), pp. 462-486.
- [62] L. N. TREFETHEN, J. A. C. WEIDEMAN: *Two results on polynomial interpolation in equally spaced points*, Journal of Approximation Theory, 65 (1991), pp. 247-260.

