Advances in Computational Mathematics (2023) 49:52
https://doi.org/10.1007/s10444-023-10057-9

On the stability of the representation of finite rank operators

J. M. Carnicer[1] · E. Mainar[1] · J. M. Peña[1]

Received: 9 September 2022 / Accepted: 12 June 2023 / Published online: 10 July 2023

Abstract
The stability of the representation of finite rank operators in terms of a basis is analyzed. A conditioning is introduced as a measure of the stability properties. This conditioning improves some other conditionings because it is closer to the Lebesgue function. Improved bounds for the conditioning of the Fourier sums with respect to an orthogonal basis are obtained, in particular, for Legendre, Chebyshev, and disk polynomials. The Lagrange and Newton formulae for the interpolating polynomial are also considered.

Keywords Conditioning · Lebesgue function · Least squares approximation · Lagrange interpolation

Mathematics Subject Classification (2010) 65F35 · 41A10 · 41A63 · 42C05

1 Introduction

The representation of a continuous function in a finite dimensional space depends on the choice of a basis. A first possibility consists of expanding the function in terms of a Lagrange basis with respect to some set of nodes so that the coefficients with respect to the basis are values of the function. The representation in terms of the Lagrange basis with respect to the Chebyshev sites turns out to be a very stable representation of polynomials (see pp. 12–15 of [2]).

Communicated by: Robert Schaback

✉ E. Mainar
esmemain@unizar.es

J. M. Carnicer
carnicer@unizar.es

J. M. Peña
jmpena@unizar.es

1 Departamento de Matemática Aplicada/IUMA, Universidad de Zaragoza, Zaragoza, Spain

A related problem is the representation of operators of finite rank in terms of a given basis. The Lebesgue function provides a pointwise bound for the error propagation of an operator independently of the formula used for evaluation. However, the choice of a basis for expressing the operator might provide worse stability results for the evaluation than those predicted by the Lebesgue function. Some operators, such as orthogonal projections, have good stability properties if an orthogonal basis is used [5]. In order to measure the stability properties, we introduce a condition number associated with the representation.

The bound provided by the Lebesgue function is lower than the conditioning of any representation of the operator with respect to a basis. It would be an ideal situation to have an evaluation formula for the operator whose condition is exactly the Lebesgue function. However, many common formulae for the operator in terms of an orthogonal basis give rise to a condition number higher than the Lebesgue function. If the conditioning is closer to the Lebesgue function, then the corresponding representation is more stable. In the case that the conditioning coincides with the Lebesgue function, the representation is optimal. The Lagrange formula for the interpolation operator is optimally stable as we will show in Section 5 (see also [3]). In contrast, the Newton formula may present numerical instability [1].

In previous research, a different conditioning for comparing different representations of an operator has been considered [3–6, 8]. However, this conditioning tends to overestimate the instability, especially when dealing with Fourier representations with respect to orthogonal polynomials, as we shall see later. The conditioning proposed in this paper is sharper in the sense that it resembles more closely the Lebesgue function. This conditioning might be harder to compute in some cases, but it can be easily bounded in the case of orthogonal projections, giving rise to sharper and practical bounds.

In Sect. 2, we introduce a conditioning $\kappa(x, B, B^{-1})$ of a basis B, and we show that it is smaller than the condition number $\operatorname{cond}(x, B, B^{-1})$ used in [3]. In fact, Example 1 illustrates that it can be considerably smaller. In Sect. 3, we extend the proposed conditioning to the case of representations of a continuous linear operator on the space of continuous functions on a compact domain with finite rank $\kappa(x, B, \Phi)$. We compare it with the condition number $\operatorname{cond}(x, B, \Phi)$, discussed in [5], and prove that $\kappa(x, B, \Phi) \leq \operatorname{cond}(x, B, \Phi)$. We also show that both conditionings are invariant under reordering or rescaling of the basis. Moreover, both conditionings coincide in the case where the functionals associated with the representation are nonnegative. In Sect. 4, we consider the conditioning of least squares problems. The Christoffel function relates the values of the Fourier sum operator $S_n[f](x)$ with $\|f\|_2$, the norm associated with the scalar product. This relation allows us to provide practical bounds for the conditioning $\kappa(x, P, \Phi)$ of the representation of S_n in terms of an orthogonal basis P. In Theorem 1, we provide bounds for both conditionings, and we can say that the bound for $\kappa(x, P, \Phi)$ is lower than the corresponding bound for $\operatorname{cond}(x, P, \Phi)$. We describe some relevant examples, considering Legendre polynomials, Chebyshev polynomials, and disk polynomials. In the three cases, the bounds for $\kappa(x, P, \Phi)$ considerably improve the bounds for $\operatorname{cond}(x, P, \Phi)$. Section 5 focusses on the conditioning of Lagrange interpolation. The representation of the interpolation operator with respect to the Lagrange basis L and the evaluation functionals X' is optimal, because $\kappa(x, L, X')$

coincides with the Lebesgue function. Moreover, both condition numbers coincide, $\kappa(x, L, X') = \mathrm{cond}(x, L, X')$. In Sect. 6, we consider the conditioning of the Newton representation of the interpolating polynomial. In this case, the conditioning depends on the ordering of the nodes. We characterize the orderings such that both conditionings coincide. As a consequence, if the nodes are increasingly ordered or, more generally, if they follow a central ordering with respect to a center (see [4]), both conditionings coincide. Finally, Sect. 7 considers the discrete case, which can be analyzed as a particular case of a Lagrange interpolation operator.

2 Conditioning of a basis

Let K be a compact domain in $\mathbb{R}^d$. Let $U = \langle b_0, \ldots, b_n \rangle$ be the vector space generated by $(b_0, \ldots, b_n)$ with linearly independent $b_0, \ldots, b_n \in C(K)$. Then, the linear mapping

$$B : c \in \mathbb{R}^{n+1} \mapsto \sum_{i=0}^{n} c_i b_i \in U$$

can be regarded as a basis of U whose inverse B^{-1} is the corresponding coordinate mapping. Let us denote by $\pi_i(u) := (B^{-1}u)_i, i = 0, \ldots, n$, the coordinate projections. Each function $u \in U$ can be written in terms of the basis

$$u(x) = B(\pi_0 u, \ldots, \pi_n u)(x) = \sum_{i=0}^{n} \pi_i(u) b_i(x), \quad x \in K.$$

In order to compute $u(x)$ expressed in terms of a given basis, we evaluate each of the terms $\pi_i(u)b_i(x)$ and sum up all of them. Since the computation of each coefficient $\pi_i(u)$, $i = 0, \ldots, n$, can be affected by an error ε_i, we shall obtain instead

$$\sum_{i=0}^{n} (\pi_i(u) + \varepsilon_i) b_i(x) = u(x) + \sum_{i=0}^{n} \varepsilon_i b_i(x), \quad x \in K.$$

So, we can assume that the computed value is the exact expression of a perturbed function $u(x) + e(x)$, where the perturbation $e(x) = \sum_{i=0}^{n} \varepsilon_i b_i(x)$ belongs to the space U. The sign of the errors in the coefficients is difficult to predict. In the worst of the cases, when we evaluate the function u at a given point x, all summands may have the same nonstrict sign, for instance, $\varepsilon_i b_i(x) \geq 0$ for all $i = 0, \ldots, n$. So, the size of the perturbation can reach the following upper bound

$$|e(x)| \leq \sum_{i=0}^{n} |\varepsilon_i||b_i(x)|, \quad x \in K.$$

We can write

$$|e(x)| \leq \|\varepsilon\|_\infty \sum_{i=0}^{n} |b_i(x)|, \quad x \in K,$$

and bound the size of the perturbation in terms of the norm of the error vector $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_n)$. So, $\sum_{i=0}^{n} |b_i(x)|$ gives a bound for the relative error $|e(x)|/\|\varepsilon\|_\infty$.

The size of the error of a coefficient depends on how the coefficient has been computed. However, the previous bound does not reveal the influence of the error ϵ_i of each coefficient to the error propagation $e(x)$. Moreover, the starting point in some problems is a perturbed function, and we want to measure how the perturbation might affect the evaluation of the function in terms of a given basis. For this purpose, we note that $\varepsilon_i = \pi_i(e)$, $i = 0, \dots, n$, and then $\sum_{i=0}^n |\varepsilon_i||b_i(x)| = \sum_{i=0}^n |\pi_i(e)b_i(x)|$. In order to measure the size of the error, we introduce

$$\|e\|_\infty := \max_{x\in K} |e(x)|.$$

The quantity $\sup_{e\in U, \|e\|_\infty=1} \sum_{i=0}^n |\pi_i(e)b_i(x)|$ can be regarded as a pointwise bound for the error in the computation of a function in U at x with respect to a given basis B, relative to the size of the perturbation $\|e\|_\infty$. This suggests the following definition.

Definition 1 Let B be a basis of a $(n+1)$-dimensional space of functions U. The *conditioning* of B at a point x of the domain K is

$$\kappa(x, B, B^{-1}) := \sup_{e\in U, \|e\|_\infty=1} \sum_{i=0}^n |\pi_i(e)b_i(x)|, \quad x \in K.$$

Since any $e \in U$ with $\|e\|_\infty = 1$ can be expressed in the form

$$e(x) = \frac{\sum_{i=0}^n c_i b_i(x)}{\|\sum_{i=0}^n c_i b_i\|_\infty}, \quad x \in K,$$

for some $c = (c_0, \dots, c_n) \neq 0$, we can write

$$\kappa(x, B, B^{-1}) = \sup_{c\neq 0} \frac{\sum_{i=0}^n |c_i||b_i(x)|}{\|\sum_{i=0}^n c_i b_i\|_\infty}, \quad x \in K.$$

Proposition 1 *Let B be a basis of an $(n+1)$-dimensional space of continous functions U defined on a compact set K and let $\pi_0, \dots, \pi_n$ be the corresponding coordinate projections. Let us define*

$$\mathrm{cond}(x, B, B^{-1}) := \sum_{i=0}^n \|\pi_i\|_\infty |b_i(x)|, \quad x \in K,$$

where $\|\pi_i\|_\infty := \sup_{e\in U, \|e\|_\infty=1} |\pi_i e|$. Then, we have

$$\kappa(x, B, B^{-1}) \leq \mathrm{cond}(x, B, B^{-1}), \quad x \in K.$$

Proof Since for each $e \in U$ we have

$$|\pi_i(e)| \leq \|\pi_i\|_\infty \|e\|_\infty,$$

we can write

$$\kappa(x, B, B^{-1}) = \sup_{e \in U, \|e\|_\infty = 1} \sum_{i=0}^{n} |\pi_i(e) b_i(x)| \leq \sum_{i=0}^{n} \|\pi_i\|_\infty |b_i(x)| = \mathrm{cond}(x, B, B^{-1}).$$

□

The conditioning $\mathrm{cond}(x, B, B^{-1})$ has been used in previous papers [3–5]. We want to show that, in some cases, $\mathrm{cond}(x, B, B^{-1})$ is much bigger than $\kappa(x, B, B^{-1})$ and so, $\mathrm{cond}(x, B, B^{-1})$ overestimates the error propagation of the representation of a function in terms of a basis.

Example 1 Let us consider the basis $b_0(x) = 1, b_1(x) = x$ of the space P_1 of polynomials of degree less than or equal to 1 on the domain $[-1, 1]$. Then, the corresponding basis mapping is $B(c_0, c_1) = c_0 + c_1 x$. Let us show that the coordinate projections have unit norm. Since $\pi_0(p) = (p(-1) + p(1))/2$, we have that

$$|\pi_0(p)| = \frac{1}{2}|p(-1) + p(1)| \leq \|p\|_\infty, \quad \forall p \in P_1.$$

Taking $p(x) = b_0(x)$, we see that $|\pi_0(p)| = 1 = \|p\|_\infty$ and conclude that $\|\pi_0\|_\infty = 1$. On the other hand, we can write $\pi_1(p) = (p(1) - p(-1))/2$ and deduce that

$$|\pi_1(p)| = \frac{1}{2}|p(1) - p(-1)| \leq \|p\|_\infty, \quad \forall p \in P_1.$$

Taking $p(x) = b_1(x)$, we see that $|\pi_1(p)| = 1 = \|p\|_\infty$ and we deduce that $\|\pi_1\|_\infty = 1$. Then, in this case

$$\mathrm{cond}(x, B, B^{-1}) = 1 + |x|.$$

On the other hand, we can write

$$\begin{aligned}
\kappa(x, B, B^{-1}) &= \sup_{\|e\|_\infty = 1} (|\pi_0(e)| + |\pi_1(e)||x|) \\
&= \sup_{\|e\|_\infty = 1} \frac{1}{2}(|e(1) + e(-1)| + |e(1) - e(-1)||x|).
\end{aligned}$$

Denoting $a_0 := |e(1) + e(-1)|/2 \geq 0$ and $a_1 := |e(1) - e(-1)|/2 \geq 0$, we find that

$$a_0 + a_1 = \max(|e(-1)|, |e(1)|) \leq 1$$

and

$$a_0 + a_1|x| \leq a_0 + a_1 \leq 1.$$

Therefore, $\kappa(x, B, B^{-1}) \leq 1$ and, since the value 1 is attained for $e(x) = 1$, we conclude that

$$\kappa(x, B, B^{-1}) = 1.$$

So, we find that $\kappa(x, B, B^{-1}) < \operatorname{cond}(x, B, B^{-1})$ if $x \neq 0$. In particular, for $x = 1$, we have

$$\kappa(1, B, B^{-1}) = 1, \quad \operatorname{cond}(1, B, B^{-1}) = 2.$$

3 Conditioning of an operator

The norm of a continuous linear operator on the space of continuous functions on a compact domain $K \subset \mathbb{R}^d$

$$T : C(K) \to C(K),$$

is also called the Lebesgue constant. We can introduce a Lebesgue function as

$$\lambda(x; T) := \sup_{\|e\|_\infty = 1} |T[e](x)|.$$

The following result proves that the supremum value of the Lebesgue function coincides with the Lebesgue constant.

Proposition 2 *Let $T : C(K) \to C(K)$ be a continuous linear operator on the space of continuous functions defined on the compact domain K. Then, we have*

$$\|T\|_\infty = \sup_{x \in K} \lambda(x; T).$$

Proof From

$$\lambda(x; T) = \sup_{\|e\|_\infty = 1} |T[e](x)| \leq \sup_{\|e\|_\infty = 1} \|T[e]\|_\infty = \|T\|_\infty,$$

we deduce that

$$\sup_{x \in K} \lambda(x; T) \leq \|T\|_\infty.$$

On the other hand, for each $e \in C(K)$ with $\|e\|_\infty = 1$, we have that

$$|T[e](x)| \leq \sup_{\|f\|_\infty = 1} |T[f](x)| = \lambda(x; T).$$

Then, we deduce that

$$\|T[e]\|_\infty \leq \sup_{x \in K} \lambda(x; T), \quad e \in C(K), \quad \|e\|_\infty = 1,$$

and

$$\|T\|_\infty = \sup_{\|e\|_\infty = 1} \|T[e]\|_\infty \leq \sup_{x \in K} \lambda(x; T).$$

$\square$

If T has finite rank, $\dim U = n + 1$, with $U = T[C(K)]$, we can obtain the representation of T with respect to a basis $B : \mathbb{R}^{n+1} \to U$ of the form

$$T[f](x) = \sum_{i=0}^{n} \pi_i T[f] b_i(x).$$

Defining $\Phi := B^{-1} \circ T$, we have that $\Phi f = (\phi_0 f, \ldots, \phi_n f)$, where $\phi_i : C(K) \to \mathbb{R}$ is the linear functional obtained applying the i-th coordinate map to $T[f]$

$$\phi_i f = \pi_i T[f], \quad f \in C(K).$$

In this way, the relation $T = B \circ \Phi$ can be understood as a way of representing the operator by choosing a basis $(b_0, \ldots, b_n)$ and a system of functionals $(\phi_0, \ldots, \phi_n)$.

Definition 2 Let U be a finite dimensional subspace of $C(K)$ with $\dim U = n + 1$ and let $B : \mathbb{R}^{n+1} \to U$ be a basis mapping for U. Let $\Phi : C(K) \to \mathbb{R}^{n+1}$ whose components $\phi_i : C(K) \to \mathbb{R}$, $i = 1, \ldots, n$, are continuous linear functionals defined on $C(K)$. We define

$$\kappa(x, B, \Phi) := \sup_{e \in C(K), \|e\|_\infty = 1} \sum_{i=0}^{n} |\phi_i(e) b_i(x)|.$$

The conditioning $\kappa(x, B, \Phi)$ can be regarded as a pointwise bound for the error in the computation of the operator $T = B \circ \Phi$ expressed in terms of the basis B, relative to the size of any perturbation $e \in C(K)$.

Proposition 3 *Let $T : C(K) \to C(K)$ be a continuous linear operator of finite rank $U = T[C(K)]$ with $\dim(U) = n + 1$. Let $B : U \to \mathbb{R}^{n+1}$ be a basis mapping and let $\Phi = B^{-1} \circ T$. Then, we have*

$$|T[\tilde{f}](x) - T[f](x)| \leq \kappa(x, B, \Phi) \|\tilde{f} - f\|_\infty, \quad x \in K, \quad f, \tilde{f} \in C(K).$$

We also have the following inequality relating $\kappa(x, B, \Phi)$ and the Lebesgue function $\lambda(x; T)$

$$\lambda(x; T) \leq \kappa(x, B, \Phi), \quad x \in K,$$

and so,

$$\|T\|_\infty \leq \max_{x \in K} \kappa(x, B, \Phi).$$

Proof Let $\phi_i : C(K) \to \mathbb{R}$ be the components of Φ, $i = 0, \ldots, n$, and let $e = \tilde{f} - f$ be the perturbation function. Since $T[e] = \sum_{i=0}^{n} \phi_i(e) b_i$, we can write

$$|T[f + e](x) - T[f](x)| = |T[e](x)| \leq \sum_{i=0}^{n} |\phi_i(e) b_i(x)| \leq \kappa(x, B, \Phi) \|e\|_\infty. \quad (1)$$

By the above inequality, we have that

$$\lambda(x;T) = \sup_{\|e\|_\infty=1} |T[e](x)| \le \kappa(x, B, \Phi)$$

and

$$\|T\|_\infty = \sup_{\|e\|_\infty=1} \|T[e]\|_\infty = \sup_{\|e\|_\infty=1} \max_{x\in K} |T[e](x)| \le \max_{x\in K} \kappa(x, B, \Phi).$$

□

Note that, in contrast to $\kappa(x, B, \Phi)$, the Lebesgue function $\lambda(x;T)$ depends only on the finite rank operator and not on the choice of the basis. Different bases B may lead to different conditionings $\kappa(x, B, \Phi)$. If $\kappa(x, B, \Phi)$ is close to $\lambda(x;T)$, the basis B provides a quasi-optimal conditioned representation of the operator at the point x.

In [5], the following measure for the conditioning was introduced:

$$\operatorname{cond}(x, B, \Phi) := \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)|.$$

Let us show that $\kappa(x, B, \Phi)$ provides a measure of the conditioning sharper than $\operatorname{cond}(x, B, \Phi)$.

Proposition 4 *Let U be a finite dimensional subspace of $C(K)$ with $\dim U = n+1$ and let $B : \mathbb{R}^{n+1} \to U$ be a basis mapping for U. Let $\Phi : C(K) \to \mathbb{R}^{n+1}$ whose components $\phi_i : C(K) \to \mathbb{R}$, $i = 0, 1, \ldots, n$, are continuous linear functionals defined on $C(K)$. Then, we have the following inequality:*

$$\kappa(x, B, \Phi) \le \operatorname{cond}(x, B, \Phi), \quad x \in K.$$

Proof Taking into account that $|\phi_i(e)| \le \|\phi_i\|_\infty \|e\|_\infty$, we have

$$\kappa(x, B, \Phi) = \sup_{\|e\|_\infty=1} \sum_{i=0}^{n} |\phi_i(e) b_i(x)| \le \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)| = \operatorname{cond}(x, B, \Phi).$$

□

For a given operator $T : C(K) \to C(K)$, with $U := T[C(K)]$, we can compare different representations with respect to different bases. If B and $\bar{B}$ are two basis mappings of U, we can write

$$T[f](x) = \sum_{i=0}^{n} \phi_i(f) b_i(x) = \sum_{i=0}^{n} \bar{\phi}_i(f) \bar{b}_i(x),$$

where $\Phi = B^{-1} \circ T$ and $\Phi' = \bar{B}^{-1} \circ T$. We have already mentioned that the conditioning depends on the choice of the basis. A reordering of the elements of a basis corresponds to a reordering of the associated functionals. In the same way, a rescaling of the basis $\bar{b}_i(x) = k_i b_i(x)$, implies a rescaling of the functions $\bar{\phi}_i = \frac{1}{k_i}\phi_i$. Let us show now that the conditionings $\mathrm{cond}(x, B, \Phi)$ and $\kappa(x, B, \Phi)$ are invariant under reordering or rescaling of the basis.

Proposition 5 *Let $\phi_i : C(K) \to \mathbb{R}$, $i = 0, \ldots, n$, be a sequence of linear functionals and B be a basis mapping of a finite dimensional space $U \subset C(K)$. Let $k_i \neq 0$, $i = 0, \ldots, n$, and $\sigma : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, n\}$ be any permutation. Let $\bar{b}_i(x) = k_i b_{\sigma(i)}(x)$ and $\bar{\phi}_i(f) = \phi_{\sigma(i)}(f)/k_i$. Then*

$$\mathrm{cond}(x, B, \Phi) = \mathrm{cond}(x, \bar{B}, \bar{\Phi}), \quad \kappa(x, B, \Phi) = \kappa(x, \bar{B}, \bar{\Phi}), \quad x \in K.$$

Proof Clearly,

$$\sum_{i=0}^{n} \|\bar{\phi}_i\|_\infty |\bar{b}_i(x)| = \sum_{i=0}^{n} \frac{\|\phi_{\sigma(i)}\|_\infty}{|k_i|} |k_i||b_{\sigma(i)}(x)| = \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)|,$$

which implies that $\mathrm{cond}(x, B, \Phi) = \mathrm{cond}(x, \bar{B}, \bar{\Phi})$. On the other hand,

$$\sum_{i=0}^{n} |\bar{\phi}_i(e)||\bar{b}_i(x)| = \sum_{i=0}^{n} \frac{|\phi_{\sigma(i)}(e)|}{|k_i|} |k_i||b_{\sigma(i)}(x)| = \sum_{i=0}^{n} |\phi_i(e)||b_i(x)|,$$

and taking the supremum when $\|e\|_\infty = 1$, we deduce that $\kappa(x, B, \Phi) = \kappa(x, \bar{B}, \bar{\Phi})$. □

In some cases, we can deduce the equality of both conditionings. We say that a functional $\phi : C(K) \to \mathbb{R}$ is *nonnegative* if $\phi(f) \geq 0$ for any nonnegative continuous function $f \in C(K)$, $f \geq 0$.

Lemma 1 *If $\phi : C(K) \to \mathbb{R}$ is a nonnegative continuous linear functional, then $\|\phi\|_\infty = \phi(1)$.*

Proof Clearly,

$$\phi(1) = |\phi(1)| \leq \sup_{\|e\|_\infty = 1} |\phi(e)| \leq \|\phi\|_\infty.$$

On the other hand, we have for any $e \in C(K)$ with $\|e\|_\infty = 1$ that both $1 + e(x)$ and $1 - e(x)$ are nonnegative functions. By the nonnegativity of ϕ, $\phi(1) + \phi(e)$ and $\phi(1) - \phi(e)$ are nonnegative values and hence $|\phi(e)| \leq \phi(1)$. Therefore

$$\|\phi\|_\infty = \sup_{\|e\|_\infty = 1} |\phi(e)| \leq \phi(1).$$

□

Now, we deduce the equality of both conditionings if the corresponding functionals are nonnegative.

Proposition 6 *Let $\phi_i : C(K) \to \mathbb{R}$, $i = 0, \ldots, n$, be a sequence of nonnegative linear functionals and B be a basis mapping of a finite dimensional space $U \subset C(K)$. Then*

$$\kappa(x, B, \Phi) = \operatorname{cond}(x, B, \Phi), \quad x \in K.$$

Proof By Lemma 1, we have that $\phi_i(1) = \|\phi_i\|_\infty$. So, we deduce that

$$\kappa(x, B, \Phi) \geq \sum_{i=0}^{n} \phi_i(1)|b_i(x)| = \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)| = \operatorname{cond}(x, B, \Phi)$$

and, using Proposition 4, the result follows. □

In some cases, the basis B of the space U can be chosen such that all functions b_i attain its maximum absolute value at the same point. Legendre and Chebyshev polynomials form relevant bases of the space of polynomials of degree not greater than n and attain its maximum absolute value on $[-1, 1]$ at $x = 1$.

Proposition 7 *Let U be a finite dimensional subspace of $C(K)$ with $\dim U = n + 1$ and let $B : \mathbb{R}^{n+1} \to U$ be a basis mapping for U such that there exists $x^0 \in K$ such that all basis functions attain its maximum at the same point $x^0 \in K$,*

$$\max_{x \in K} |b_i(x)| = |b_i(x^0)|$$

and $\Phi : C(K) \to \mathbb{R}^{n+1}$ whose components $\phi_i : C(K) \to \mathbb{R}$, $i = 0, 1, \ldots, n$, are continuous linear functionals defined on $C(K)$. Then, the maximum condition is attained at x^0,

$$\max_{x \in K} \operatorname{cond}(x, B, \Phi) = \operatorname{cond}(x^0, B, \Phi), \quad \max_{x \in K} \kappa(x, B, \Phi) = \kappa(x^0, B, \Phi).$$

Proof Clearly, we have

$$\operatorname{cond}(x, B, \Phi) = \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)| \leq \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x^0)| = \operatorname{cond}(x^0, B, \Phi)$$

and

$$\kappa(x, B, \Phi) = \sup_{\|e\|_\infty = 1} \sum_{i=0}^{n} |\phi_i(e)||b_i(x)| \leq \sup_{\|e\|_\infty = 1} \sum_{i=0}^{n} |\phi_i(e)||b_i(x^0)| = \kappa(x^0, B, \Phi).$$

□

4 Conditioning of least squares problems

In this section, orthogonal projections arising in least squares problems will be considered. We first deduce some general properties of the conditioning of projectors.

Definition 3 Let $T : C(K) \to C(K)$ be a linear operator. We say that T is a *projector* if $T[u] = u$ for all $u \in U$, where $U := T[C(K)]$.

If T is a projector, then $T \circ B = B$ for any basis mapping of U. So, if $\Phi = B^{-1} \circ T$, we have

$$\Phi \circ B = B^{-1} \circ T \circ B = I,$$

where I is the identity map from $\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$. In other words, the system of functionals $(\phi_0, \dots, \phi_n)$ is dual to the basis $(b_0, \dots, b_n)$ in the sense that

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

where δ_{ij} is the Kronecker symbol. This implies that the restriction of Φ to the space U coincides with the coordinate mapping B^{-1}, that is,

$$\Phi u = B^{-1} u, \quad u \in U.$$

The following proposition shows that the condition of the representation of a projector is always greater than or equal to the corresponding condition of the basis.

Proposition 8 *Let $T : C(K) \to C(K)$ be a projector on a finite dimensional space U and let B be a basis mapping of U. If $\Phi = B^{-1} \circ T$, then*

$$\operatorname{cond}(x, B, B^{-1}) \le \operatorname{cond}(x, B, \Phi), \quad x \in K,$$

and

$$\kappa(x, B, B^{-1}) \le \kappa(x, B, \Phi), \quad x \in K.$$

Proof Let $\pi_0, \dots, \pi_n$ be the coordinate mappings with respect to B. Since T is a projector, then $\phi_i u = \pi_i u$ for all $u \in U$, $i = 0, \dots, n$. Then

$$\|\phi_i\|_\infty = \sup_{e \in C(K), \|e\|_\infty = 1} |\phi_i e| \ge \sup_{e \in U, \|e\|_\infty = 1} |\phi_i e| = \sup_{e \in U, \|e\|_\infty = 1} |\pi_i e| = \|\pi_i\|_\infty$$

and we conclude that

$$\operatorname{cond}(x, B, \Phi) = \sum_{i=0}^{n} \|\phi_i\|_\infty |b_i(x)| \ge \sum_{i=0}^{n} \|\pi_i\|_\infty |b_i(x)| = \operatorname{cond}(x, B, B^{-1}).$$

In the same way, we deduce that

$$\begin{aligned}\kappa(x, B, \Phi) &= \sup_{e \in C(K), \|e\|_\infty = 1} \sum_{i=0}^{n} |\phi_i(e) b_i(x)| \ge \sup_{e \in U, \|e\|_\infty = 1} \sum_{i=0}^{n} |\phi_i(e) b_i(x)| \\ &= \sup_{e \in U, \|e\|_\infty = 1} \sum_{i=0}^{n} |\pi_i(e) b_i(x)| = \kappa(x, B, B^{-1}).\end{aligned}$$

□

Let K be a compact set of $\mathbb{R}^d$ and μ be a nonnegative regular Borel measure with $0 < \mu(K) < \infty$. Let us define the semidefinite symmetric bilinear form

$$\langle f, g \rangle = \int_K f(x) g(x) d\mu(x), \tag{2}$$

and $\|f\|_2 := \langle f, f \rangle^{1/2}$. The best approximation of $f \in C(K)$ in a finite dimensional subspace U is a function minimizing $\|f - u\|_2$, $u \in U$. If the bilinear form is positive definite on the finite dimensional subspace U, then the best approximation exists, it is unique and it is characterized by the property that the error is orthogonal to the space U. If $P = (p_0, \ldots, p_n)$ is an orthogonal basis of U, then the solution of the least squares problem can be given as the n-th Fourier sum

$$S_n[f] = \sum_{i=0}^{n} \frac{\langle f, p_i \rangle}{\|p_i\|_2^2} p_i, \quad f \in C(K).$$

Introducing the Christoffel-Darboux kernel

$$K_n(x, y) := \sum_{i=0}^{n} \frac{p_i(x) p_i(y)}{\|p_i\|_2^2},$$

we can express the n-th Fourier sum in the form

$$S_n[f](x) = \int_K K_n(x, y) f(y) d\mu(y).$$

Using the Cauchy-Schwarz inequality

$$|S_n[f](x)|^2 \le \int_K K_n(x, y)^2 d\mu(y) \int_K f(y)^2 d\mu(y) = K_n(x, x) \|f\|_2^2,$$

we deduce that the values of the Fourier sum at x and the norm $\|f\|_2$ for $f \in C(K)$ can be related by the Christoffel function $1/K_n(x, x)$ (see Theorem 3.6.6 of [7]). Since the measure of K is finite, we have that

$$\|f\|_2^2 \le \|f\|_\infty^2 \mu(K)$$

and the following upper bound for the Lebesgue function in terms of the Christoffel function follows

$$|\lambda(x; S_n)|^2 \le \|1\|_2 \sqrt{K_n(x,x)}.$$

The next result shows that the same bound can be deduced for the conditioning $\kappa(x, P, \Phi)$ for the Fourier sum expressed in terms of an orthogonal basis P.

Theorem 1 *Let K be a compact set of $\mathbb{R}^d$ and μ be a nonnegative finite Borel measure. Let $P = (p_0, \ldots, p_n)$ be an orthogonal basis of a space U with respect to the bilinear form $\langle f, g\rangle := \int_K f(x)g(x)d\mu(x)$ and*

$$\phi_i f := \frac{\langle f, p_i\rangle}{\|p_i\|_2^2}, \quad i = 0, \ldots, n.$$

Let

$$K_n(x, y) := \sum_{i=0}^{n} \frac{p_i(x)p_i(y)}{\|p_i\|_2^2},$$

be the Christoffel-Darboux kernel associated to the basis P. Then, the Lebesgue function of the Fourier sum operator $S_n[f](x) := \sum_{i=0}^n \phi_i f p_i(x)$ is given by

$$\lambda(x; S_n) = \int_K |K_n(x, y)| d\mu(y),$$

and we have the following bounds for the conditionings

$$\operatorname{cond}(x, P, \Phi) = \sum_{i=0}^{n} \frac{\|p_i\|_1}{\|p_i\|_2^2} |p_i(x)| \le \|1\|_2 \sum_{i=0}^{n} \frac{|p_i(x)|}{\|p_i\|_2}, \tag{3}$$

$$\kappa(x, P, \Phi) \le \|1\|_2 \sqrt{K_n(x,x)} = \|1\|_2 \Big(\sum_{i=0}^{n} \frac{|p_i(x)|^2}{\|p_i\|_2^2} \Big)^{1/2}. \tag{4}$$

Proof Let

$$K_n(x, y) := \sum_{i=0}^{n} \frac{p_i(x)p_i(y)}{\|p_i\|_2^2},$$

be the Christoffel-Darboux kernel associated to the orthogonal basis P. The Lebesgue function of the projector S_n can be written in the form

$$\lambda(x; S_n) := \sup_{\|e\|_\infty = 1} \Big| \int_K K_n(x, y) e(y) d\mu(y) \Big|.$$

Choosing for each $x \in K$ a sequence $e_n(y)$ of functions in $C(K)$ with $\|e_n\|_\infty = 1$ converging to $\mathrm{sign}(K(x, y))$, where

$$\mathrm{sign}(x) := \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0, \end{cases}$$

we deduce that

$$\lambda(x; S_n) = \int_K |K_n(x, y)| d\mu(y).$$

By the Riesz representation Theorem (see Theorem 6.19 of Chapter 6 of [9]),

$$\|\phi_i\|_\infty = \frac{\int_K |p_i(y)| d\mu(y)}{\|p_i\|_2^2} = \frac{\|p_i\|_1}{\|p_i\|_2^2},$$

where $\|p_i\|_1 := \int_K |p_i(x)| d\mu(x)$. Then, we obtain (see Proposition 1 of [5])

$$\mathrm{cond}(x, P, \Phi) = \sum_{i=0}^n \frac{\|p_i\|_1}{\|p_i\|_2^2} |p_i(x)|.$$

Using the Cauchy-Schwarz inequality $\int_K |p_i(x)| d\mu(x) \leq \|p_i\|_2 \|1\|_2$, the bound (3) follows. Now, let us bound

$$\kappa(x, P, \Phi) = \sup_{\|e\|_\infty = 1} \sum_{i=0}^n \frac{|\langle e, p_i \rangle|}{\|p_i\|_2^2} |p_i(x)|.$$

From the Cauchy-Schwarz inequality, we obtain

$$\sum_{i=0}^n \frac{|\langle e, p_i \rangle|}{\|p_i\|_2^2} |p_i(x)| \leq \Big(\sum_{i=0}^n \frac{|\langle e, p_i \rangle|^2}{\|p_i\|_2^2} \Big)^{1/2} \Big(\sum_{i=0}^n \frac{|p_i(x)|^2}{\|p_i\|_2^2} \Big)^{1/2}.$$

Using Bessel inequality and the fact that $|e(x)| \leq 1$, for all x, we deduce that

$$\sum_{i=0}^n \frac{|\langle e, p_i \rangle|^2}{\|p_i\|_2^2} \leq \|e\|_2^2 \leq \mu(K) = \|1\|_2^2.$$

Then, the bound (4) follows. □

We remark that the bound in (4) is lower than the bound in (3) because

$$\sum_{i=0}^n \frac{|p_i(x)|^2}{\|p_i\|_2^2} \leq \Big(\sum_{i=0}^n \frac{|p_i(x)|}{\|p_i\|_2} \Big)^2.$$

Now, let us analyze some common Fourier approximations with respect to different scalar products, giving rise to classical orthogonal polynomials.

In the case that μ is the Lebesgue measure on $[-1, 1]$, we can take the basis of Legendre polynomials $P = (P_0, \dots, P_n)$. Taking into account that

$$\|P_i\|_2^2 = \frac{2}{2i+1}, \quad \|P_i\|_1 \le \sqrt{2}\|P_i\|_2 = \frac{2}{\sqrt{2i+1}},$$

we have that

$$\mathrm{cond}(x, P, \Phi) = \sum_{i=0}^{n} \frac{\|P_i\|_1}{\|P_i\|_2^2} |P_i(x)| \le \sum_{i=0}^{n} \sqrt{2i+1}|P_i(x)|.$$

From the fact that the Legendre polynomials attain its maximum value $P_n(1) = 1$ at $x = 1$ (see Section 7.2 of [10]), we deduce from Proposition 7 that

$$\max_{x\in[-1,1]} \mathrm{cond}(x, P, \Phi) = \mathrm{cond}(1, P, \Phi) \le \sum_{i=0}^{n} \sqrt{2i+1}$$

(see Proposition 2 of [5]). Using Proposition 7, we deduce that $\kappa(x, P, \Phi)$ attains its maximum value at $x = 1$, and from (4), we get

$$\max_{x\in[-1,1]} \kappa(x, P, \Phi) = \kappa(1, P, \Phi) \le \sqrt{2}\Big(\sum_{i=0}^{n} \frac{2i+1}{2}\Big)^{1/2} = n + 1.$$

We remark that

$$\frac{2\sqrt{2}}{3}\sqrt{(n+1)^3} \le \sum_{i=0}^{n} \sqrt{2i+1} \le \frac{1 + (4n+5)\sqrt{2n+1}}{6},$$

which implies that the bound for $\mathrm{cond}(x, P, \Phi)$ is higher than the bound for $\kappa(x, P, \Phi)$.

Example 2 Let us consider the Fourier sum that associates to each function its best approximation in the least squares sense in the space of polynomials of degree not greater than 1

$$S_1[f](x) := \frac{1}{2}\int_{-1}^{1} f(t)dt + \frac{3x}{2}\int_{-1}^{1} tf(t)dt = \int_{-1}^{1} \frac{1+3xt}{2} f(t)dt.$$

Let us first compute the Lebesgue function of S_1. By Theorem 1,

$$\lambda(x; S_1) = \sup_{\|e\|_\infty=1} \Big| \int_{-1}^{1} \frac{1+3xt}{2} e(t)dt \Big| = \frac{1}{2}\int_{-1}^{1} |1+3xt|dt.$$

Using the change of variables $\tau = -t$, we deduce

$$\lambda(-x; S_1) = \frac{1}{2}\int_{-1}^{1} |1 - 3xt|dt = \frac{1}{2}\int_{-1}^{1} |1 + 3x\tau|d\tau = \lambda(x; S_1).$$

So, the Lebesgue function is an even function.

If $|x| \le 1/3$, then $1 + 3xt$ is nonnegative for each $t \in [-1, 1]$, and we obtain

$$\lambda(x; S_1) = \frac{1}{2}\int_{-1}^{1} \big|1 + 3xt\big|dt = \frac{1}{2}\int_{-1}^{1} (1 + 3xt)dt = 1, \quad |x| \le 1/3.$$

Otherwise, if $|x| > 1/3$, then $1 + 3xt$ changes its sign at $t = \frac{-1}{3x}$, and we deduce that

$$\begin{aligned}
\frac{1}{2}\int_{-1}^{1} \big|1 + 3xt\big|dt &= \frac{1}{2}\Big(\Big|\int_{-1}^{\frac{-1}{3x}} (1 + 3xt)dt\Big| + \Big|\int_{\frac{-1}{3x}}^{1} (1 + 3xt)dt\Big|\Big)\\
&= \frac{1}{2}\Big(\Big|1 - \frac{1}{3x} - \frac{3x}{2}\Big(1 - \frac{1}{9x^2}\Big)\Big| + \Big|1 + \frac{1}{3x} + \frac{3x}{2}\Big(1 - \frac{1}{9x^2}\Big)\Big|\Big)\\
&= \frac{1}{2}\Big(\Big|1 - \frac{3x}{2} - \frac{1}{6x}\Big| + \Big|1 + \frac{3x}{2} + \frac{1}{6x}\Big|\Big).
\end{aligned}$$

Using the identity $\max(|a|, |b|) = |a + b|/2 + |a - b|/2$, we obtain

$$\lambda(x; S_1) = \max\Big(1, \frac{3|x|}{2} + \frac{1}{6|x|}\Big) = \frac{1}{2}\Big(3|x| + \frac{1}{3|x|}\Big), \quad |x| \ge 1/3.$$

The continuous linear functionals for the representation of S_1 with respect to the basis of Legendre polynomials $P_0(x) = 1$, $P_1(x) = x$, are

$$\phi_0[f] = \frac{1}{2}\int_{-1}^{1} f(t)dt, \quad \phi_1[f] = \frac{3}{2}\int_{-1}^{1} tf(t)dt.$$

The corresponding norms are $\|\phi_0\|_\infty = 1$, $\|\phi_1\|_\infty = 3/2$. So, a direct computation gives

$$\mathrm{cond}(x, P, \Phi) = 1 + \frac{3}{2}|x|.$$

Now, we compute

$$\kappa(x, P, \Phi) = \sup_{\|e\|_\infty = 1} \frac{1}{2}\Big|\int_{-1}^{1} e(t)dt\Big| + \frac{3|x|}{2}\Big|\int_{-1}^{1} te(t)dt\Big|.$$

We use the identity $|a| + |b| = \max(|a - b|, |a + b|)$, to obtain

$$\frac{1}{2}\Big|\int_{-1}^{1} e(t)dt\Big| + \frac{3|x|}{2}\Big|\int_{-1}^{1} te(t)dt\Big| = \frac{1}{2}\max_{s\in\{-1,1\}}\Big|\int_{-1}^{1}(s + 3|x|t)e(t)dt\Big| \leq \frac{\|e\|_\infty}{2}\max_{s\in\{-1,1\}}\int_{-1}^{1}\big|s + 3|x|t\big|dt.$$

By the symmetry of the functions $t \in [-1, 1] \mapsto |1 + 3|x|t|$, $t \in [-1, 1] \mapsto |-1 + 3|x|t|$, we see that both have the same integral and obtain

$$\kappa(x, P, \Phi) \leq \frac{1}{2}\int_{-1}^{1}\big|1 + 3|x|t\big|dt = \lambda(|x|; S_1) = \lambda(x; S_1).$$

From Proposition 3, we deduce that

$$\kappa(x, B, \Phi) = \lambda(x; T) = \frac{1}{2}\int_{-1}^{1}\big|1 + 3xt\big|dt.$$

So, we have that

$$\lambda(x; S_1) = \kappa(x, P, \Phi) < \mathrm{cond}(x, P, \Phi).$$

Thus, we can say that the basis of Legendre polynomials has the optimal conditioning $\kappa(x, B, \Phi) = \lambda(x; T)$. However, the conditioning $\mathrm{cond}(x, P, \Phi)$ does not show the good behavior of the basis. By Proposition 7, the maximum value of both conditionings is attained at $x = 1$

$$\|T\|_\infty = \kappa(1, P, \Phi) = \frac{5}{3} < \frac{5}{2} = \mathrm{cond}(1, P, \Phi).$$

Chebyshev polynomials $T = (T_0, \ldots, T_n)$ are orthogonal with respect to $d\mu(x) = dx/\sqrt{1 - x^2}$ on the interval $[-1, 1]$. Taking into account that $\|T_0\|^2 = 2\pi$, $\|T_n\|_2^2 = \pi$, for $n \geq 1$, we deduce from (4) that

$$\kappa(x, T, \tau) \leq \sqrt{2\pi}\Big(\frac{1}{2\pi} + \frac{1}{\pi}\sum_{i=1}^{n} T_i(x)^2\Big)^{1/2} = \Big(1 + 2\sum_{i=1}^{n} T_i(x)^2\Big)^{1/2}.$$

The norm of the functionals $\tau_i f = \langle f, T_i\rangle/\|T_i\|^2$ can be computed from the fact that $\|T_0\|_1 = \pi$ and $\|T_n\|_1 = 2$, for $n \geq 1$,

$$\|\tau_0\|_\infty = 1, \quad \|\tau_n\|_\infty = \frac{4}{\pi}, \quad n \geq 1,$$

giving rise to

$$\mathrm{cond}(x, T, \tau) = \sum_{i=0}^{n} \|\tau_i\|_\infty |T_i(x)| = 1 + \frac{4}{\pi} \sum_{i=1}^{n} |T_i(x)|,$$

So, we find again that the bound for $\kappa(x, T, \tau)$ is lower than $\mathrm{cond}(x, T, \tau)$. Since Chebyshev polynomials attain its maximum absolute value at $x = 1$, we deduce from Proposition 7, that the maximum conditionings $\kappa(x, T, \tau)$ and $\mathrm{cond}(x, T, \tau)$ are attained at $x = 1$

$$\max_{x\in[-1,1]} \mathrm{cond}(x, T, \tau) = \mathrm{cond}(1, T, \tau) = 1 + \frac{4n}{\pi}$$

and

$$\max_{x\in[-1,1]} \kappa(x, T, \tau) = \kappa(1, T, \tau) \leq \sqrt{2n + 1}.$$

Disk polynomials (cf. section 2.6 of [7]) can be used for approximation of functions in the disk $\mathbb{D} = \{(r\cos\theta, r\sin\theta)|0 \leq r \leq 1; \theta \in [-\pi, \pi]\}$. An orthogonal basis of Gegegenbauer-like orthogonal polynomials with respect to the measure

$$d(\mu, \theta) = \frac{\alpha + 1}{\pi}(1 - r^2)^\alpha dr d\theta$$

is given by

$$Z^\alpha_{j,m}(r, \theta) := \begin{cases} R^\alpha_{j,|m|}(r)\cos(|m|\theta), & \text{if } m \geq 0, \\ R^\alpha_{j,|m|}(r)\sin(|m|\theta), & \text{if } m < 0, \end{cases} \quad 2j + |m| \leq n, \tag{5}$$

where

$$R^\alpha_{j,|m|}(r) := \frac{P_j^{(\alpha,|m|)}(2r^2 - 1)}{P_j^{(\alpha,|m|)}(1)} r^{|m|}$$

and $P_j^{(\alpha,m)}$ denotes the usual Jacobi polynomial of degree j in $[-1, 1]$. In Theorem 3 of [6], the conditioning of the basis has been computed as

$$\begin{aligned} \mathrm{cond}(r, \theta, Z^\alpha, \Phi^\alpha) = &\sum_{j=0}^{\lfloor n/2 \rfloor} \frac{H^\alpha_{j,j}}{h^\alpha_{j,j}} |R^\alpha_{j,0}(r)| \\ &+ \frac{4}{\pi} \sum_{m=1}^{n} \sum_{j=0}^{\lfloor (n-m)/2 \rfloor} \frac{H^\alpha_{j+m,j}}{h^\alpha_{j+m,j}} |R^\alpha_{j,m}(r)|(|\cos m\theta| + |\sin m\theta|), \end{aligned} \tag{6}$$

where

$$H^{\alpha}_{j+m,j} := 2(\alpha+1)\int_0^1 |R^{\alpha}_{j,m}(r)| r(1-r^2)^{\alpha} dr,$$
$$h^{\alpha}_{j+m,j} := \frac{2(\alpha+1)}{P_j^{(\alpha,m)}(1)^2}\int_0^1 P_j^{(\alpha,m)}(2r^2-1)^2 r^{2m+1}(1-r^2)^{\alpha} dr.$$

The integral defining $h^{\alpha}_{j+m,j}$ can be expressed in terms of the square of the norm of the usual Jacobi polynomials (see formula (8) of [6]) giving rise to

$$h^{\alpha}_{j+m,m} = \frac{(j+m)!j!}{(\alpha+2j+m+1)(\alpha+2)_{j+m-1}(\alpha+1)_j}, \quad j+m>0,$$

where $(t)_j := t(t+1)\cdots(t+j-1)$ denotes the usual Pochhammer symbol. For $j=m=0$, we have $h^{\alpha}_{0,0}=1$. First, we provide the values of the norm of the basis (see the proof of Theorem 3 of [6])

$$\|Z^{\alpha}_{j,0}\|^2_{\alpha} = h^{\alpha}_{j,j}, \quad \|Z^{\alpha}_{j,m}\|^2_{\alpha} = \frac{1}{2}h^{\alpha}_{j+m,j}.$$

We now compute a bound for $\kappa(r,\theta,Z^{\alpha},\Phi^{\alpha})$ using formula (4).

$$\kappa(r,\theta,Z^{\alpha},\Phi^{\alpha})^2 \le \sum_{j=0}^{\lfloor n/2\rfloor}\frac{Z^{\alpha}_{j,0}(r,\theta)^2}{h^{\alpha}_{j,j}} + 2\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}\frac{Z^{\alpha}_{j,m}(r,\theta)^2+Z^{\alpha}_{j,-m}(r,\theta)^2}{h^{\alpha}_{j+m,j}}$$
$$= \sum_{j=0}^{\lfloor n/2\rfloor}\frac{R^{\alpha}_{j,0}(r)^2}{h^{\alpha}_{j,j}} + 2\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}\frac{R^{\alpha}_{j,m}(r)^2}{h^{\alpha}_{j+m,j}}.$$

Since the polynomials $Z_{j,m}$ attain their maximum value at the boundary $r=1$, we get that $R^{\alpha}_{j,m}(r) \le R^{\alpha}_{j,m}(1)=1$ and deduce the following bound

$$\max_{r\in[0,1],\theta\in[0,2\pi]}\kappa(r,\theta,Z^{\alpha},\Phi^{\alpha}) \le \left(\sum_{j=0}^{\lfloor n/2\rfloor}\frac{1}{h^{\alpha}_{j,j}} + 2\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}\frac{1}{h^{\alpha}_{j+m,j}}\right)^{1/2}.$$

Since $h^{\alpha}_{j+m,j}$ is a decreasing function of α, $\alpha\in[0,\infty)$, we deduce that the least bound is attained for the Zernike polynomials corresponding to $\alpha=0$

$$\max_{r\in[0,1],\theta\in[0,2\pi]}\kappa(r,\theta,Z^{0},\Phi^{0}) \le \left(\sum_{j=0}^{\lfloor n/2\rfloor}(2j+1) + 2\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}(2j+m+1)\right)^{1/2}.$$

Then, we can write

$$
\begin{aligned}
&\sum_{j=0}^{\lfloor n/2\rfloor}(2j+1)+2\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}(2j+m+1)\\
&=\sum_{j=0}^{\lfloor n/2\rfloor}(2j+1)+\sum_{l=1}^{n}2\lfloor\frac{l+1}{2}\rfloor(l+1)=\sum_{l=0}^{n}(l+1)^2=\frac{(n+1)(n+2)(2n+3)}{6},
\end{aligned}
$$

and obtain the bound

$$
\max_{r\in[0,1],\theta\in[0,2\pi]}\kappa(r,\theta,Z^0,\Phi^0)\le\sqrt{\frac{(n+1)(n+2)(2n+3)}{6}}.
$$

In Proposition 2 of [6] it was shown that

$$
\max_{r\in[0,1],\theta\in[0,2\pi]}\operatorname{cond}(r,\theta,Z^\alpha,\Phi^\alpha)\le\sum_{j=0}^{\lfloor n/2\rfloor}\frac{1}{\sqrt{h^\alpha_{j,j}}}+\frac{4\sqrt{2}}{\pi}\sum_{m=1}^{n}\sum_{j=0}^{\lfloor (n-m)/2\rfloor}\frac{1}{\sqrt{h^\alpha_{j+m,j}}}.
$$

With an analogous reasoning, we get that the bound is a decreasing function of $\alpha\in[0,\infty)$ and the least bound is obtained for $\alpha=0$

$$
\max_{r\in[0,1],\theta\in[0,2\pi]}\operatorname{cond}(r,\theta,Z^0,\Phi^0)\le\frac{4\sqrt{2}}{5\pi}(n+5/2)(n+3/2)^{3/2}.
$$

Clearly, the bound for $\kappa(r,\theta,Z^0,\Phi^0)$ is lower than the bound for $\operatorname{cond}(r,\theta,Z^0,\Phi^0)$.

5 Conditioning of Lagrange interpolation

An interesting case of a projector is the Lagrange interpolation operator. Given a sequence of distinct nodes $X=(x_0,\dots,x_n)$ with $x_0,\dots x_n\in K$ and a subspace $U\subset C(K)$ with $\dim U=n+1$ such that the Lagrange interpolation problem

$$
u(x_i)=f(x_i),\quad i=0,\dots,n,
$$

has a unique solution in U, we can define the operator

$$
T:C(K)\to C(K),
$$

which associates to each $f\in C(K)$ its unique interpolant in U at the sequence of nodes. Let $l_0,\dots,l_n\in U$ be the fundamental solution associated to the sequence of nodes, that is

$$
l_j(x_i)=\delta_{ij},
$$

where δ_{ij} is the Kronecker symbol. Then, we can express the operator in terms of the Lagrange basis

$$L : c \in \mathbb{R}^{n+1} \to \sum_{i=0}^{n} c_i l_i \in U,$$

as follows

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

The functionals $X' := L^{-1} \circ T$ associated to the Lagrange representation of the interpolant $L[f]$ are the evaluation functionals at the nodes

$$x_i' f := f(x_i), \quad i = 0, \ldots, n.$$

Theorem 2 *Let U be a subspace of $C(K)$ with $\dim U = n+1$ and let $X = (x_0, \ldots, x_n)$ be a sequence of nodes such that the Lagrange interpolation problem has a unique solution in U. Let T be the Lagrange interpolation operator, L be the Lagrange basis and let $X' = (x_0', \ldots, x_n')$ be the evaluation functionals at the nodes,*

$$x_i' f := f(x_i), \quad i = 0, \ldots, n.$$

Then

$$\lambda(x; T) = \kappa(x, L, X') = \mathrm{cond}(x, L, X') = \sum_{i=0}^{n} |l_i(x)| \quad x \in K. \tag{7}$$

If, in addition, the constant functions belong to U, we also have that

$$\kappa(x, L, L^{-1}) = \mathrm{cond}(x, L, L^{-1}) = \sum_{i=0}^{n} |l_i(x)| \geq 1, \quad x \in K.$$

Proof By propositions 3 and 4, we have that

$$\lambda(x; T) \leq \kappa(x, L, X') \leq \mathrm{cond}(x, L, X'), \quad x \in K.$$

The evaluation functionals have unit norm, and we can write

$$\mathrm{cond}(x, L, X') = \sum_{i=0}^{n} \|x_i'\|_\infty |l_i(x)| = \sum_{i=0}^{n} |l_i(x)|.$$

Given $\xi \in K$, let $u \in U$ be the solution of the interpolation problem

$$u(x_i) = \mathrm{sign} l_i(\xi), \quad i = 0, \ldots, n.$$

Let us define

$$e(x) := \begin{cases} 1, & \text{if } u(x) > 1, \\ u(x), & \text{if } |u(x)| \le 1, \\ -1, & \text{if } u(x) < -1. \end{cases}$$

Then, we have that $\|e\|_\infty = 1$ and $e \in C(K)$. For this function, we can write

$$T[e](x) = \sum_{i=0}^{n} \operatorname{sign}(l_i(\xi)) l_i(x).$$

In particular,

$$T[e](\xi) = \sum_{i=0}^{n} \operatorname{sign}(l_i(\xi)) l_i(\xi) = \sum_{i=0}^{n} |l_i(\xi)|.$$

So, the value $|T[e](\xi)| = \sum_{i=0}^{n} |l_i(\xi)|$ is attained and we deduce that

$$\lambda(\xi; T) \ge \sum_{i=0}^{n} |l_i(\xi)|$$

for each $\xi \in K$. So, we have shown (7).

Since the interpolation operator is a projection, we deduce from Proposition 8 that $\operatorname{cond}(x, L, L^{-1}) \le \operatorname{cond}(x, L, X')$ and $\kappa(x, L, L^{-1}) \le \kappa(x, L, X')$. Let $\pi_0, \ldots, \pi_n$ be the coordinate projections corresponding to the Lagrange basis. If $1 \in U$, by Proposition 1, we have that

$$\begin{aligned} \operatorname{cond}(x, L, L^{-1}) &\ge \kappa(x, L, L^{-1}) = \sup_{e \in U, \|e\|_\infty = 1} \sum_{i=0}^{n} |\pi_i(e) l_i(x)| \ge \sum_{i=0}^{n} |\pi_i(1) l_i(x)| \\ &= \sum_{i=0}^{n} |l_i(x)| = \operatorname{cond}(x, L, X') \ge \operatorname{cond}(x, L, L^{-1}). \end{aligned}$$

Finally, if $1 \in U$, we also have that $\sum_{i=0}^{n} l_i(x) = 1$ and, by the triangular inequality, $\sum_{i=0}^{n} |l_i(x)| \ge 1$ for all $x \in K$. □

So, we have shown that the Lagrange representation of the interpolation $L[f] = \sum_{i=0}^{n} f(x_i) l_i$ has optimal conditioning and that both $\kappa(x, L, X')$ and $\operatorname{cond}(x, L, X')$ coincide for this representation.

We observe that the Lagrange representation is a particular case of a representation with respect to a set of nonnegative functionals and that the equality of $\kappa(x, L, X')$ and $\operatorname{cond}(x, L, X')$ could also be obtained by direct application of Proposition 6.

Let us compute the conditionings of the representation of the interpolant $L[f]$ with respect to any other basis. Let us denote

$$\operatorname{skeel}(A) := \left\| |A^{-1}||A| \right\|_\infty,$$

the Skeel condition number, where $|A|$ stands for the matrix whose entries are the absolute values of the entries of A.

Theorem 3 *Let U be a subspace of $C(K)$ with $\dim U = n+1$ and let $X = (x_0, \dots, x_n)$ be a sequence of nodes such that the Lagrange interpolation problem has a unique solution in U. Let T be the Lagrange interpolation operator and $B : \mathbb{R}^{n+1} \to U$ be a basis mapping. Let $\Phi = B^{-1} \circ T$ be the corresponding set of functionals of the representation of T with respect to the basis B. Let $b_0(x), \dots, b_n(x)$ be the basis functions associated to B and let*

$$M(B, X) := (b_j(x_i))_{i,j=0,\dots,n} \in \mathbb{R}^{(n+1)\times(n+1)}$$

be the collocation matrix of the basis B at the set of nodes X. Then, we have

$$\operatorname{cond}(x, B, \Phi) = (|b_0(x)|, \dots, |b_n(x)|)|M(B, X)^{-1}|(1, 1, \dots, 1)^T, \tag{8}$$

$$\max_{k\in\{0,\dots,n\}} \operatorname{cond}(x_k, B, B^{-1}) = \left\| |M(B, X)|\, |M(B, X)^{-1}| \right\|_\infty = \operatorname{skeel}(M(B, X)^{-1}), \tag{9}$$

$$\kappa(x, B, \Phi) = \max_{\epsilon_0,\dots,\epsilon_n\in\{-1,1\}} (|b_0(x)|, \dots, |b_n(x)|)|M(B, X)^{-1}\epsilon|, \tag{10}$$

and

$$\max_{k=0,\dots,n} \kappa(x_k, B, \Phi) = \max_{\epsilon_0,\dots,\epsilon_n\in\{-1,1\}} \left\| |M(B, X)||M(B, X)^{-1}\epsilon| \right\|_\infty. \tag{11}$$

Proof The matrix $M(B, X)$ is the matrix of change of basis between $(b_0, \dots, b_n)$ and the Lagrange basis $(l_0, \dots, l_n)$

$$(b_0, \dots, b_n) = (l_0, \dots, l_n)M(B, X).$$

Let us observe that the basis mapping corresponding to $(b_0, \dots, b_n)$ is given by

$$Bc = \sum_{j=0}^{n} c_j b_j = \sum_{i=0}^{n} \sum_{j=0}^{n} b_j(x_i)c_j l_i = L\, M(B, X)c, \quad i = 0, \dots, n.$$

The corresponding functionals $\Phi = B^{-1} \circ T$ can be expressed in terms of the inverse of the matrix $M(B, X)$

$$\Phi(u) = \begin{pmatrix} \phi_0(u) \\ \vdots \\ \phi_n(u) \end{pmatrix} = B^{-1}(T[u]) = M(B, X)^{-1}X'(u) = M(B, X)^{-1} \begin{pmatrix} u(x_0) \\ \vdots \\ u(x_n) \end{pmatrix}.$$

Using the notation $M(B,X)^{-1} = (m_{ij})_{i,j=0,\dots,n}$, we can write

$$\phi_i(u) = \sum_{j=0}^{n} m_{ij} u(x_j).$$

Proposition 2 of [3] can be immediately generalized to a multivariate setting to derive

$$\|\phi_i\| = \sum_{j=0}^{n} |m_{ij}|.$$

Then, we deduce that

$$\begin{pmatrix} \|\phi_0\|_\infty \\ \vdots \\ \|\phi_n\|_\infty \end{pmatrix} = |M(B,X)^{-1}| \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

And we obtain a formula for $\operatorname{cond}(x, B, \Phi)$

$$\operatorname{cond}(x, B, \Phi) = \sum_{i=0}^{n} \|\phi_i\|_\infty \, |b_i(x)| = (|b_0(x)|, \dots, |b_n(x)|) \, |M(B,X)^{-1}| \, (1, 1, \dots, 1)^T$$

and (8) holds. In particular, we can write

$$\begin{pmatrix} \operatorname{cond}(x_0, B, B^{-1}) \\ \vdots \\ \operatorname{cond}(x_n, B, B^{-1}) \end{pmatrix} = |M(B,X)| \, |M(B,X)^{-1}| \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

and (9) follows.

For the computation of $\kappa(x, B, \Phi)$, we take into account that

$$\kappa(x, B, \Phi) = \max_{\|e\|=1} \sum_{i=0}^{n} |\phi_i(e)||b_i(x)| = \max_{\|e\|=1} \sum_{i=0}^{n} |e(x_i)||b_i(x)|.$$

Denoting $\epsilon_i = e(x_i)$, $i = 0, \dots, n$ and $\epsilon = (\epsilon_0, \dots, \epsilon_n)^T$, we can write $\Phi(e) = M(B,X)^{-1}\epsilon$ and

$$\kappa(x, B, \Phi) = \max_{\|\epsilon\|_\infty = 1} (|b_0(x)|, \dots, |b_n(x)|) |M(B,X)^{-1}\epsilon|.$$

Taking into account that for each x

$$F(\varepsilon_0, \dots, \varepsilon_n) := (|b_0(x)|, \dots, |b_n(x)|)|M(B,X)^{-1}\epsilon| = \sum_{i=0}^{n} \left| b_i(x) \sum_{j=0}^{i} m_{ij} \varepsilon_j \right|$$

is a convex function in each of the variables $\varepsilon_0, \ldots, \varepsilon_n$ and that a convex function defined on an interval attains its maximum value at the ends, we deduce that the maximum is attained only for $\varepsilon_0, \ldots, \varepsilon_n \in \{-1, 1\}$ and then

$$\kappa(x, B, \Phi) = \max_{\epsilon_0,\ldots,\epsilon_n \in \{-1,1\}} (|b_0(x)|, \ldots, |b_n(x)|)|M(B, X)^{-1}\epsilon|,$$

that is, (10) holds. In particular, we have

$$\kappa(x_k, B, \Phi) = \max_{\epsilon_0,\ldots,\epsilon_n \in \{-1,1\}} (|b_0(x_k)|, \ldots, |b_n(x_k)|)|M(B, X)^{-1}\epsilon|$$

and (11) follows. □

6 Conditioning of the Newton interpolation formula

Let us now compute the conditioning of the Newton representation of the Lagrange interpolation formula

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

where

$$d_i f = [x_0, \ldots, x_i] f,$$

is the i-th order divided difference of f at the nodes $x_0, \ldots, x_i$ and

$$\omega_i(x) = \prod_{j=0}^{i-1} (x - x_j), \quad i = 1, \ldots, n,$$

with the convention that ω_0 is the constant polynomial $\omega_0(x) = 1$. Since

$$d_i f = \sum_{k=0}^{i} \frac{f(x_k)}{\omega'_{i+1}(x_k)},$$

we deduce (see Proposition 2 of [3]) that

$$\|d_i\|_\infty = \sum_{k=0}^{i} \frac{1}{|\omega'_{i+1}(x_k)|}$$

and

$$\operatorname{cond}(x, \omega, d) = \sum_{i=0}^{n} |\omega_i(x)| \sum_{k=0}^{i} \frac{1}{|\omega'_{i+1}(x_k)|}.$$

Using formula (10), the conditioning $\kappa(x,\omega,d)$ can be computed in the following way

$$\kappa(x,\omega,d)=\max_{\varepsilon_0,\dots,\varepsilon_n\in\{-1,1\}}\sum_{i=0}^{n}\Big|\omega_i(x)\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\Big|.$$

Definition 4 Let $x_0,\dots,x_n$ be a sequence of distinct nodes. We say that x_n leaves the other nodes at one side if either $x_k<x_n$, $k=0,\dots,n-1$, or $x_k>x_n$, $k=0,\dots,n-1$, that is, there exist $\sigma\in\{-1,1\}$ such that

$$\operatorname{sign}(x_k-x_n)=\sigma,\quad k=0,\dots,n-1.$$

Theorem 4 *Let $x_0,\dots,x_n$ be a sequence of distinct nodes. Then, we have that*

$$\kappa(x,\omega,d)=\operatorname{cond}(x,\omega,d)$$

if and only if each node x_i leaves the previous nodes $x_0,\dots,x_{i-1}$ at one side, that is, there exist $\sigma_1,\dots,\sigma_n$ such that

$$\operatorname{sign}(x_k-x_i)=\sigma_i,\quad k=0,\dots,i-1.$$

Proof Let us define $s_0:=1$ and

$$s_i:=\sigma_1\cdots\sigma_i,\quad i=1,\dots,n.$$

Then, we have that

$$\begin{aligned}\operatorname{sign}(\omega'_{i+1}(x_k))&=\prod_{j\in\{0,\dots,i\}\setminus\{k\}}\operatorname{sign}(x_k-x_j)=\prod_{j\in\{0,\dots,k-1\}}(-\sigma_k)\prod_{j\in\{k+1,\dots,i\}}\sigma_j\\&=(-\sigma_k)^k\sigma_{k+1}\cdots\sigma_i=(-\sigma_k)^k s_k s_i,\quad k\le i.\end{aligned}$$

Choosing $\varepsilon_k:=(-\sigma_k)^k s_k$, $k=0,\dots,n$, we have that

$$\operatorname{sign}(\omega'_{i+1}(x_k))=s_i\varepsilon_k$$

and

$$\|d_i\|_\infty=s_i\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)},\quad i=0,\dots,n.$$

So,

$$\kappa(x,\omega,d)\ge\sum_{i=0}^{n}\Big|\omega_i(x)\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\Big|=\sum_{i=0}^{n}\|d_i\|_\infty\,|\omega_i(x)|=\operatorname{cond}(x,\omega,d).$$

From Proposition 4, we deduce that $\kappa(x,\omega,d)=\mathrm{cond}(x,\omega,d)$.

Conversely, let us assume that $\kappa(x,\omega,d)=\mathrm{cond}(x,\omega,d)$ for a given $x\notin\{x_0,\dots,x_n\}$. Let $\varepsilon_0,\dots,\varepsilon_n\in\{-1,1\}$ be such that

$$\kappa(x,\omega,d)=\sum_{i=0}^{n}\left|\omega_i(x)\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\right|.$$

Since

$$|\omega_i(x)|\left|\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\right|\le|\omega_i(x)|\sum_{k=0}^{i}\frac{1}{|\omega'_{i+1}(x_k)|},$$

we have that

$$\kappa(x,\omega,d)=\sum_{i=0}^{n}|\omega_i(x)|\left|\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\right|\le\sum_{i=0}^{n}|\omega_i(x)|\sum_{k=0}^{i}\frac{1}{|\omega'_{i+1}(x_k)|}=\mathrm{cond}(x,\omega,d).$$

Since $\omega_i(x)\neq 0$, equality $\kappa(x,\omega,d)=\mathrm{cond}(x,\omega,d)$ holds if and only if

$$\left|\sum_{k=0}^{i}\frac{\varepsilon_k}{\omega'_{i+1}(x_k)}\right|=\sum_{k=0}^{i}\frac{1}{|\omega'_{i+1}(x_k)|},\quad i=0,\dots,n,$$

which implies that $\epsilon_k\omega'_{i+1}(x_k)$, $k=0,\dots,i$, have the same strict sign s_i for any $i=0,\dots,n$. So we have that

$$\mathrm{sign}(\omega'_{i+1}(x_k))=s_i\epsilon_k,\quad k=0,\dots,i,\quad i=0,\dots,n.$$

Now, we use the recurrence

$$\omega'_{i+1}(x_k)=(x_k-x_i)\omega'_i(x_k),\quad i<k,$$

to deduce that

$$s_i\varepsilon_k=\mathrm{sign}(x_k-x_i)s_{i-1}\varepsilon_k.$$

Defining $\sigma_i:=s_is_{i-1}$, we conclude that

$$\mathrm{sign}(x_k-x_i)=\sigma_i,\quad k=0,\dots,i-1,$$

that is, each x_i leaves the previous nodes $x_0,\dots,x_{i-1}$ to the left ($\sigma_i=-1$) or to the right ($\sigma_i=1$). □

Definition 5 We say that the sequence $(x_0,\dots,x_n)$ follows a central order with respect to a center c if the sequence of distances of the nodes to the center is monotonically increasing, that is,

$$|x_0-c|\le|x_1-c|\le\cdots\le|x_n-c|.$$

If the nodes form a monotonic sequence, we can consider that they follow a central order with respect to the first node x_0. In fact, if $x_0 < \cdots < x_n$, then the distances $|x_i - x_0| = x_i - x_0$, $i = 0, \ldots, n$, form an increasing sequence. If $x_0 > \cdots > x_n$, then the distances $|x_i - x_0| = x_0 - x_i$, $i = 0, \ldots, n$, also form an increasing sequence.

Corollary 1 *Let $x_0, \ldots, x_n$ be a sequence of distinct nodes following a central order with respect to a center c. Then,*

$$\kappa(x, \omega, d) = \mathrm{cond}(x, \omega, d).$$

Proof In order to apply the characterization of Theorem 4, let us show that each x_i, $i \in \{1, \ldots, n\}$, leaves the previous nodes at one side. Let $\sigma_i := \mathrm{sign}(c - x_i) \in \{-1, 1\}$. If $\sigma_i(c - x_j) > 0$ for some $j \in \{0, \ldots, i-1\}$, then c leaves x_i and x_j at the same side and, since both nodes are distinct, we have that $|x_i - c| > |x_j - c|$ and

$$\sigma_i(x_j - x_i) = \sigma_i(c - x_i) - \sigma_i(c - x_j) = |x_i - c| - |x_j - c| > 0.$$

Otherwise, if $\sigma_i(c - x_j) < 0$ for some $j \in \{0, \ldots, i-1\}$, then

$$\sigma_i(x_j - x_i) = \sigma_i(c - x_i) - \sigma_i(c - x_j) = |x_i - c| + |x_j - c| > 0.$$

Therefore,

$$\mathrm{sign}(x_j - x_i) = \mathrm{sign}(c - x_i) = \sigma_i, \quad j = 0, \ldots, i-1,$$

that is, x_i leaves all previous nodes at the same side as c. □

In [4], some nice properties of the central ordering were described. In particular, for equidistant nodes, the central ordering with respect to the center of the interval provides lower bounds for the conditioning of the Newton formula than the corresponding bounds for increasing nodes.

An ordering of the nodes giving rise to conditionings that are relatively close to the Lebesgue function is the central ordering with respect to the evaluation point (see Section 4 of [4]). Using Corollary 1, we also deduce that

$$\kappa(x, \omega, d) = \mathrm{cond}(x, \omega, d),$$

for evaluation of the Newton formula using nodes following a central ordering with respect to x, that is,

$$|x_n - x| \geq \cdots \geq |x_1 - x| \geq |x_0 - x|.$$

Let us illustrate with an example that $\kappa(x, \omega, d)$ can be lower than $\mathrm{cond}(x, \omega, d)$.

Example 3 Let us consider the set of nodes $x_0 = 0$, $x_1 = 2$, $x_2 = 1$ on the interval $[0, 2]$. Then, the conditioning of the Newton formula

$$p(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{2}x + \frac{f(x_2) - 2f(x_1) + f(x_0)}{2}x(x-2),$$

can be computed as follows:

$$\kappa(x,\omega,d) = \sup_{\epsilon_0,\epsilon_1,\epsilon_2\in\{-1,1\}} |\epsilon_0| + \frac{|\epsilon_1-\epsilon_0|}{2}x + \frac{|\epsilon_2-2\epsilon_1+\epsilon_0|}{2}x(2-x), \quad x\in[0,2].$$

We can easily deduce that the supremum can be achieved either for $(\epsilon_0,\epsilon_1,\epsilon_2) = (1,1,-1)$ or $(\epsilon_0,\epsilon_1,\epsilon_2) = (1,-1,1)$ and we have that

$$\kappa(x,\omega,d) = \max(1+2x(2-x), 1+x+x(2-x)) = \begin{cases} 1+2x(2-x), & x\in[0,1], \\ 1+3x-x^2, & x\in[1,2]. \end{cases}$$

On the other hand,

$$\operatorname{cond}(x,\omega,d) = 1+x+2x(2-x).$$

Evaluating at $x=1$, we have

$$3 = \kappa(1,\omega,d) < \operatorname{cond}(1,\omega,d) = 4.$$

The maximum value of κ is attained at $x = 3/2$, and we have

$$3.25 = \frac{13}{4} = \kappa(3/2) < \operatorname{cond}(3/2) = 4.$$

On the other hand, the maximum value of cond is attained at $x = 5/4$ and

$$3.1875 = \frac{51}{16} = \kappa(5/4) < \operatorname{cond}(5/4) = \frac{33}{8} = 4.125.$$

7 Conditioning in the discrete case

The discrete case corresponds to $K = \{x_0,\dots,x_n\}$, where $x_0,\dots,x_n$ are distinct points in $\mathbb{R}^d$. Each real function f defined on K can be completely described by the vector $(f(x_0),\dots,f(x_n)) \in \mathbb{R}^{n+1}$. The functions $l_j \in \mathbb{R}^K$ defined by $l_j(x_i) = \delta_{ij}$, $j = 0,\dots,n$, form a basis of $C(K) = \mathbb{R}^K$ corresponding to the basis mapping

$$L : (c_0,\dots,c_n) \in \mathbb{R}^{n+1} \to \sum_{j=0}^{n} c_j l_j \in \mathbb{R}^K.$$

Since

$$f(x) = \sum_{j=0}^{n} f(x_j) l_j(x),$$

the coordinate projectors π_i are the evaluation functionals

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

Considering the whole set K as a set of nodes, the functions $l_0, \dots, l_n$ can be regarded as the Lagrange basis with respect to the Lagrange interpolation problem, find $u \in \mathbb{R}^K$ such that

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

This problem has a unique solution and the interpolation operator $T : \mathbb{R}^K \to \mathbb{R}^K$ is the identity mapping because the interpolation space coincides with the whole set $C(K)$. Since T is the identity mapping, the evaluation functionals are just the coordinate projections $B^{-1} = \pi = (\pi_0, \dots, \pi_n)$. The discrete case can be analyzed using the tools of Sect. 5. Any basis B of $\mathbb{R}^K$ can be identified with the corresponding collocation matrix $M(B, K) = (b_j(x_i))_{i,j=0,\dots,n}$. By Theorem 2, the maximum conditioning in this case can be expressed with the formulae

$$\max_{k=0,\dots,n} (\mathrm{cond}(x_k, B, B^{-1})) = \mathrm{skeel}(M(B, K)^{-1}),$$

$$\max_{k=0,\dots,n} \kappa(x_k, B, B^{-1}) = \max_{\epsilon_0,\dots,\epsilon_n \in \{-1,1\}} \left\| \, |M(B, K)| \, |M(B, K)^{-1}\epsilon| \, \right\|_\infty .$$

Any linear operator $T : \mathbb{R}^K \to \mathbb{R}^K$ is described by the effect on the basis L

$$T[l_j] = \sum_{i=0}^n a_{ij} l_i$$

and, since $f = \sum_{j=0}^n f(x_j) l_j$, we can write

$$T[f] = \sum_{j=0}^n \sum_{i=0}^n a_{ij} f(x_j) l_i .$$

and we have

$$T[f](x_i) = \sum_{j=0}^n a_{ij} f(x_j).$$

The coordinate mapping π transforms each function f into a vector πf in such a way that

$$\|f\|_\infty = \max_{i=0,\dots,n} |f(x_i)| = \|\pi f\|_\infty .$$

Defining $A = (a_{ij})_{i,j=0,\dots,n} \in \mathbb{R}^{(n+1)\times(n+1)}$, we can express the norm of the operator T in terms of the norm of the matrix A. From $\pi T[f] = A\pi f$, we obtain

$$\|T\|_\infty = \max_{\|\epsilon\|_\infty = 1} \|A\epsilon\|_\infty = \|A\|_\infty = \max_{i=0,\dots,n} \sum_{j=0}^n |a_{ij}|.$$

Conversely, since any square matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$ defines the mapping $x \in \mathbb{R}^{n+1} \mapsto Ax \in \mathbb{R}^{n+1}$ and the space $\mathbb{R}^{n+1}$ can be regarded as the space of real functions on the set $K = \{0, 1, \ldots, n\}$, we can view any matrix as an operator. Let us show that the different conditionings coincide with the Lebesgue function and give rise to a corresponding vector, whose maximum entry is the infinity norm of the matrix.

Proposition 9 *Let $A \in \mathbb{R}^{(n+1)\times(n+1)}$ and let $T : \mathbb{R}^{\{0,1,\ldots,n\}} \to \mathbb{R}^{\{0,1,\ldots,n\}}$ be the linear operator defined by*

$$T[f] = \sum_{j=0}^{n} a_{ij} f(j) l_i, \quad f \in \mathbb{R}^{\{0,1,\ldots,n\}},$$

where l_i is the function defined on $\{0, 1, \ldots, n\}$ whose values are given by $l_i(j) = \delta_{ij}$, $j = 0, \ldots, n$, for each $i = 0, \ldots, n$. Let $\pi = (\pi_0, \ldots, \pi_n)$ be the evaluation functionals $\pi_i f = f(i)$, $i = 0, \ldots, n$, and $\Phi = \pi \circ T$. Then, the Lebesgue function is given by

$$\lambda(i; T) = \sum_{j=0}^{n} |a_{ij}|, \quad i = 0, \ldots, n,$$

and

$$\kappa(x, T, \Phi) = \mathrm{cond}(x, T, \Phi) = \lambda(x; T), \quad x \in \{0, 1, \ldots, n\}.$$

Proof For each error function e, we obtain an error vector

$$\varepsilon = \pi e = (e(0), \ldots, e(n))^T$$

such that $\pi T[e] = A\varepsilon$. So, the values of the Lebesgue function at $x = i$ are given by

$$\lambda(i; T) = \max_{\|e\|_\infty=1} |T[e](i)| = \max_{\|\epsilon\|_\infty=1} |\sum_{j=0}^{n} a_{ij}\epsilon_j| = \sum_{j=0}^{n} |a_{ij}|.$$

Collecting all the values of the Lebesgue function, we can form a vector

$$\pi\lambda(\cdot; T) = |A|(1, \ldots, 1)^T.$$

From the definition of Φ, we have that

$$\phi_i e = \pi_i(T[e]) = T[e](i) = \sum_{j=0}^{n} a_{ij} e(j)$$

and we deduce that

$$\|\phi_i\|_\infty = \max_{\|e\|_\infty=1} |T[e](i)| = \max_{\|\epsilon\|_\infty=1} |\sum_{j=0}^{n} a_{ij}\epsilon_j| = \sum_{j=0}^{n} |a_{ij}|.$$

So,

$$\mathrm{cond}(x, L, \Phi) = \sum_{i=0}^{n} \|\phi_i\|_\infty |l_i(x)| = \sum_{i=0}^{n} \lambda(i; T) l_i(x) = \lambda(x; T), \quad x \in \{0, 1, \ldots, n\}.$$

We also have

$$\kappa(i, L, \Phi) = \sup_{\|e\|_\infty = 1} \sum_{j=0}^{n} |\phi_j(e)| l_j(i) = \sup_{\|e\|_\infty = 1} |\phi_i(e)| = \|\phi_i\|_\infty = \sum_{j=0}^{n} |a_{ij}|.$$

Therefore, the result follows. □

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was partially supported by Spanish research grants PGC2018-096321-B-I00, RED2022-134176-T (MCI/AEI) and by Gobierno de Aragón (E41_23R).

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Berrut, J.P., Trefethen, L.N.: Barycentric Lagrange interpolation. SIAM Rev. **46**, 501–517 (2004)
2. de Boor, C.: A practical guide to splines, Revised Springer Verlag, New York (2001)
3. Carnicer, J.M., Khiar, Y., Peña, J.M.: Optimal stability of the Lagrange formula and conditioning of the Newton formula. J. Approx. Theory **238**, 52–66 (2019)
4. Carnicer, J.M., Khiar, Y., Peña, J.M.: Central orderings for the Newton interpolation formula. BIT **59**(2), 371–386 (2019)
5. Carnicer, J. M.; Khiar, Y.; Peña, J. M. Conditioning of polynomial Fourier sums. Calcolo 56, no. 3, Paper No. 24 (2019)
6. Carnicer, J.M., Mainar, E., Peña, J.M.: Stability properties of disk polynomials. Numer. Algorithms **87**(1), 119–135 (2021)
7. Dunkl, C.F., Xu, Y.: Orthogonal polynomials of several variables, Second Edition, Encyclopedia of Mathematics and its applications, 155. Cambridge University Press, Cambridge (2014)
8. Lam, D. H.; Cuong, L. N.; Van Manh, P.; Van Minh, N., On the conditioning of the Newton formula for Lagrange interpolation. J. Math. Anal. Appl. 505, no. 1, Paper No. 125473, 14 pp (2022)

9. Rudin, W.: Real and complex analysis. Mac Graw-Hill, London (1970)
10. Szegő, G.: Orthogonal polynomials, Colloquium Publ, vol. 23. American Mathematical Society, Providence, Rhode Island (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.