



HAL
open science

An equioscillation theorem for multivariate Chebyshev approximation

Alexandre Goldsztejn

► **To cite this version:**

Alexandre Goldsztejn. An equioscillation theorem for multivariate Chebyshev approximation. 2023.
hal-04224541v1

HAL Id: hal-04224541

<https://hal.science/hal-04224541v1>

Preprint submitted on 2 Oct 2023 (v1), last revised 22 Aug 2024 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An equioscillation theorem for multivariate Chebyshev approximation

Alexandre Goldsztejn

^a*CNRS, Ecole Centrale de Nantes, Nantes, France*

Abstract

The equioscillation condition is extended to multivariate approximation. To this end, it is reformulated as the synchronized oscillations between the error maximizers and the components of a related Haar matrix kernel vector. This new condition gives rise to a multivariate equioscillation theorem where the Haar condition is not assumed and hence the existence and the characterization by equioscillation become independent of uniqueness. This allows the theorem to be applicable to problems with no strong uniqueness or even no uniqueness. A technical additional requirement on the involved Haar matrix and its kernel vector is proved to be sufficient for strong uniqueness. Instances of multivariate problems with strongly unique, unique and nonunique solutions are presented to illustrate the scope of the theorem.

Keywords: Chebyshev approximation problem, multivariate approximation, equioscillation theorem, convex analysis

1. The multivariate equioscillation theorem

Given $f : X \rightarrow \mathbb{R}$ continuous, we consider the Chebyshev approximation problem

$$\min_{a \in \mathbb{R}^n} \max_{x \in X} \left| \sum_{i=1}^n a_i \phi_i(x) - f(x) \right|, \quad (1)$$

where basis function $\phi_i : X \rightarrow \mathbb{R}$ are continuous. The index set X is just assumed to be a compact topological space, so that continuity is well defined and entails the existence of the maximum (typically, we will just consider boxes in \mathbb{R}^n). The basis function vector $\phi : X \rightarrow \mathbb{R}^n$ is defined by $\phi(x) = (\phi_1(x), \dots, \phi_n(x))$. We define $e(a, x) = a^T \phi(x) - f(x)$ and

$$m(a) = \max_{x \in X} |e(a, x)|, \quad (2)$$

so that the Chebyshev approximation problem is $\min_{a \in \mathbb{R}^n} m(a)$. We will suppose that this minimum is not 0, otherwise the problem and the characterization of the its solutions are meaningless.

Given $x_1, \dots, x_k \in X$, we define the Haar matrix

$$H(x_1, \dots, x_k) = (\phi(x_1) \quad \phi(x_2) \quad \cdots \quad \phi(x_k)) \in \mathbb{R}^{n \times k}. \quad (3)$$

When $x_1 < \cdots < x_k$ and $\phi(x) = (1, x, \dots, x^k)$, a Haar matrix is a Vandermonde matrix. The *Haar condition* holds for X and ϕ if and only if for all $x_1, \dots, x_n \in X$ the square Haar matrix $H(x_1, \dots, x_n)$ is nonsingular. When $X = [\underline{x}, \bar{x}] \subseteq \mathbb{R}$ is a compact interval, the *equioscillation condition* is satisfied for $a \in \mathbb{R}^n$ if and only if there exist $x_1 < \cdots < x_{n+1} \in X$ such that $|e(a, x_i)| = m(a)$ for all $i \in \{1, \dots, n+1\}$ and $e(a, x_i) e(a, x_{i+1}) \leq 0$ hold for all $i \in \{1, \dots, n\}$. Values of $x \in X$ that satisfy $|e(a, x_i)| = m(a)$ will be called active indices at $a \in \mathbb{R}^n$. With this settings, the equioscillation theorem characterizes the solutions of the Chebyshev univariate approximation problem: if the Haar condition holds then the Chebyshev approximation problem has one strongly unique¹ solution, which is equivalently characterized by the equioscillation condition.

Both the Haar condition and the equioscillation condition don't extend to multivariate approximation. This is somehow related the existence of several minimizers to multivariate Chebyshev approximation problems, and one way to overcome this situation is to restrict the scope of the approximation problem, so that uniqueness holds. This was successfully done by [1, 2, 3]. Instead, we aim here at weakening both the Haar condition and the equioscillation condition so that they can apply to multivariate Chebyshev approximation problems, including instances that enjoy infinitely many minimizers. While the Haar condition and the equioscillation condition are interleaved in the univariate case, where existence and uniqueness always hold together, they need to be separated in the multivariate case so that existence and the equioscillation characterization can be proved independently of the uniqueness. An additional technical condition that is tested on the minimizer only will be proved sufficient for strong uniqueness.

First, the Haar condition cannot be assumed anymore in the context of multivariate approximation: it is well known that it entails that X is either a circle or a compact interval [4] (up to homeomorphism), hence it

¹Strong uniqueness means that not only $m(a) - m(a^*) > 0$ for $a \neq a^*$, but that $m(a) - m(a^*) \geq c \|a - a^*\|$ for some $c > 0$ and the norm $\|a - a^*\| = \max_{x \in X} \|(a - a^*)^T \phi(x)\|$. The definition is in fact independent of the norm since all norms are equivalent in \mathbb{R}^n .

cannot hold for multivariate approximation with other index sets. We still need a similar but weaker condition, and define the *weak Haar condition* as follows: it holds for X and ϕ if and only if there exist $x_1, \dots, x_n \in X$ such that the square Haar matrix $H(x_1, \dots, x_n)$ is nonsingular. The weak Haar condition is likely to hold in most situations, e.g., $X = [-1, 1] \times [-1, 1]$, $\phi(x) = (1, x_1, x_2)$ and $H((0, 0), (1, 0), (1, 1))$ is nonsingular. The weak Haar condition has several consequences, among which $\text{card}(X) \geq n$, the existence of a minimizer of the Chebyshev approximation problem (cf. Section 3.1), the fact that $\|a\|_{X, \phi} = \max_{x \in X} |a^T \phi(x)|$ is a norm on \mathbb{R}^n (the triangular inequality and absolute homogeneity are trivial, the positive definiteness follows directly from the weak Haar condition), with the consequence that strong uniqueness implies uniqueness.

Second, the equioscillation condition requires an order on indices and does not make any sense for multivariate approximation, so we need to reformulate it in a way that can apply more generally. To this end, we rely on the following observation, which generalizes the well-known fact that Vandermonde matrices with one more column than rows have kernel vectors with oscillating signs. Its proof is given in Appendix A.

Lemma 1. *If $\phi : [x, \bar{x}] \rightarrow \mathbb{R}^n$ is continuous and satisfies the Haar condition on $[x, \bar{x}]$ then all Haar matrices $H(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n \times (n+1)}$, with $x_1 < \dots < x_{n+1} \in X$, have a one dimensional kernel $\text{span}\{u\}$ with $0 \neq u \in \mathbb{R}^{n+1}$ having nonzero components with alternating signs, i.e., $u_i u_{i+1} < 0$ for all $i \in \{1, \dots, n\}$.*

This observation allows reformulating the equioscillation as a synchronized oscillation between the error at active indices and the components of the corresponding Haar matrix kernel vector: $e(a, x_i) u_i \geq 0$ for all $i \in \{1, \dots, n+1\}$. Note that $e(a, x_i) u_i$ could be all negative as well, but the opposite of a kernel vector is also in the kernel, hence we don't lose any generality by requiring $e(a, x_i) u_i \geq 0$ for some kernel vector. Under this form, the equioscillation condition can be used for multivariate Chebyshev approximation problems. Uniqueness is not granted anymore in the context of multivariate Chebyshev approximation, so a technical sufficient condition for uniqueness is added below for the definition of the strong equioscillation condition, which will be proved to be sufficient for strong uniqueness.

Definition 2 (Multivariate equioscillation condition). *The (multivariate) equi-oscillation condition is satisfied at $a \in \mathbb{R}^n$ if and only if there exist $x_1, \dots, x_K \in X$, $K \geq 1$, such that $|e(a, x_i)| = m(a)$, and a kernel vector $0 \neq u \in \mathbb{R}^K$ of $H(x_1, \dots, x_K)$ such that $e(a, x_i) u_i \geq 0$ hold for all $i \in \{1, \dots, K\}$.*

The strong (multivariate) equioscillation condition holds if and only if furthermore $H(x_1, \dots, x_K)$ is full rank, hence $K \geq n + 1$, and u no zero components.

First note that by Lemma 1 the Haar condition and the (univariate) equioscillation condition together imply the strong multivariate equioscillation with $K = n + 1$. We observe several direct consequences of this definition: the active indices $x_1, \dots, x_K \in X$ need not to be ordered anymore, so the definition can be applied in the context of multivariate approximation. In fact, the definition is not affected by any permutation the active indices $x_1, \dots, x_K \in X$: indeed, the columns of the Haar matrix, together with the components of its kernel vectors, permute simultaneously with permutations of the active indices, so the sign synchronization $e(a, x_i) u_i \geq 0$ is not affected by permutation of the active indices. Another observation is that $K < n + 1$ implies that $H(x_1, \dots, x_K)$ not full rank. Indeed in this case $H(x_1, \dots, x_K)$ has more rows than columns and has a non trivial kernel. In particular, if $K = 1$ then $H(x_1)$ has to be a zero matrix (see Example 10 for such a non trivial situation with infinitely many minimizers).

The connection of the multivariate equioscillation condition to the optimality of the Chebyshev approximation problem can be seen by introducing the matrix

$$G(x_1, \dots, x_d) = (\epsilon(x_1)\phi(x_1) \cdots \epsilon(x_d)\phi(x_d)). \quad (4)$$

The capitale G is for "G"radient, since one can see that the columns of G are gradients of $|e(a, x_i)|$, provided that $|e(a, x_i)| > 0$. Furthermore observe that by definition we have

$$G(x_1, \dots, x_d) \lambda = H(x_1, \dots, x_d) u, \text{ with } u_i = \epsilon(x_i) \lambda_i. \quad (5)$$

Therefore, the multivariate equioscillation condition corresponds to the non negativity of a kernel vector of G . We see will the later that the columns of G are actually subgradients of $m(a)$ so the multivariate equioscillation condition is directly related to 0 belonging to the convex hull of some subgradients, that is to the optimality condition for unconstrained convex optimization.

These conditions lead to the following generalization of the equioscillation theorem to multivariate Chebyshev approximation problems, where X is not restricted anymore to be any interval or circle.

Theorem 3 (Multivariate equioscillation theorem). *Let $\phi : X \rightarrow \mathbb{R}^n$ be continuous. If the weak Haar condition holds for ϕ and X then there exists an optimal solution to the Chebyshev approximation problem. Furthermore,*

$a \in \mathbb{R}^n$ is a minimizer of this problem if and only if it satisfies the multivariate equioscillation condition (MEC). Finally, if the strong MEC holds then the minimizer is strongly unique, hence unique.

The usual equioscillation theorem for univariate approximation is a simple consequence of Theorem 3: on the one hand by Lemma 1 the equioscillation implies the multivariate equioscillation, on the other hand the Haar condition implies both the weak Haar condition and, together with Lemma 1, the strong multivariate equioscillation.

The rest of the paper is organized as follows: some basic definitions and properties of convex functions are presented in Section 2. Section 3 presents the proof of the multivariate equioscillation theorem. Finally, several instances of multivariate problems are presented in Section 4 to illustrate the scope of the theorem.

2. Standard convex properties of $m(a)$

The function $m : \mathbb{R}^n \rightarrow \mathbb{R}$ in (2) is called a pointwise supremum, and its properties are well known in the context of nonsmooth convex optimization. First note that the function $|e(a, x)|$ is convex, and the pointwise supremum of convex functions is convex, hence $m(a)$ is convex. Convex functions defined in \mathbb{R}^n being continuous, so is $m(a)$. The subdifferential of $m(a)$ is of central importance here. It is denoted by $\partial m(a)$ and is made of all subgradients: $u \in \mathbb{R}^n$ is a subgradient at \bar{a} if it gives rise to an affine under-estimator, i.e., $m(a) \geq m(\bar{a}) + u^T(a - \bar{a})$. The optimality condition for unconstrained convex optimization then reads a^* is a minimizer of $m(a)$ if and only if $0 \in \partial m(a^*)$.

When $|e(a, x)|$ is defined in $\mathbb{R}^n \times X$ with values in \mathbb{R} , and X is compact, the pointwise supremum is actually an unconstrained maximum² and its subdifferential enjoys a simple explicit expression: $\partial m(a) = \text{conv}\{\partial_a |e(a, x)| : x \in \text{act}(a)\}$ with $\text{act}(a) = \{x \in X : |e(a, x)| = m(a)\}$ is the set of active indices. Let us illustrate this formula on a simple finite pointwise maximum example.

Example 4. Let $U \in \mathbb{R}^{2 \times 3}$ with columns $u_1 = (-2, 1)^T$, $u_2 = (1, 1)^T$ and $u_3 = (1, -3)^T$. Define the pointwise maximum $m(a) = \max\{u_1^T a, u_2^T a, u_3^T a\} = \|U^T a\|_\infty$, which is a piecewise linear function. Figure 1 shows the level sets

²In the convex analysis literature, the wording pointwise supremum applied to infinite possibly compact index sets, and the working pointwise maximum is restricted to finite index sets.

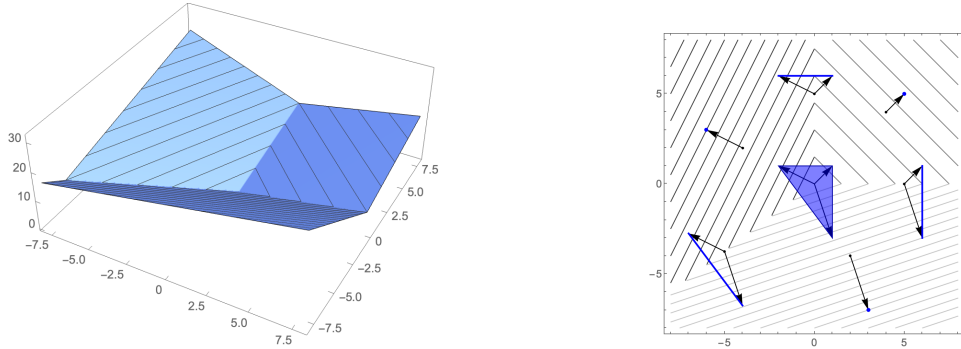


Figure 1: Pointwise maximum of Example 4, with the subdifferential evaluated at several points.

of $m(a)$. From darker to lighter level sets correspond to areas where $u_1^T a$, $u_2^T a$ or $u_3^T a$ is active. Dashed lines are the place where two linear functions are equal, and the three linear functions are equal at the origin. The subdifferential are represented in blue: when one linear constraint is active, the pointwise maximum is differentiable and the subdifferential contains only the gradient. On dashed lines, where two linear constraints are active, the subdifferential is the convex hull of the two corresponding gradients, hence a segment. Finally, at the origin where the three functions are active, the subdifferential is the convex hull of the three gradients, hence a triangle. From this analysis, we see that the origin is the only point where the subdifferential contains zero, hence the only minimizer of $m(a)$.

In the typical situation where $\min_{a \in \mathbb{R}^n} m(a) > 0$, we have that $m(a) > 0$ and $|e(a, x)|$ differentiable for all $a \in \mathbb{R}^n$ and $x \in \text{act}(a)$. Therefore, $\partial e(a, x) = \{\epsilon(x) \phi(x)\}$ with

$$\epsilon(x) = \text{sign}(e(a, x)), \quad (6)$$

whose value is inside $\{-1, 1\}$ for $x \in \text{act}(a)$ since in that case $|e(a, x)| > 0$ by assumption, and finally

$$\partial m(a) = \text{conv}\{\epsilon(x) \phi(x) : x \in \text{act}(a)\}. \quad (7)$$

Also note that $\epsilon(x) \phi(x)$ for $x \in \text{act}(a)$ are subgradients at a .

This section is ended with the following two simple lemmas. The first provides a sufficient condition for the uniqueness of the minimizer of some pointwise maximum of linear functions. It applies directly to the pointwise maximum function of Example 4. In the sequel, the maximal component of the vector a is denoted by $\max u$.

Lemma 5. Let $U = (u_1 u_2 \cdots u_d) \in \mathbb{R}^{n \times (d)}$, $d \geq n + 1$, be full rank and suppose that $\lambda \in \ker(U)$ with $\lambda_i > 0$ for all $i \in \{1, \dots, d\} = I$. Then the piecewise linear function $l(a) = \max(U^T a) = \max\{u_1^T a, \dots, u_d^T a\}$ has the origin as unique minimizer.

Proof. By way of contradiction, suppose $a \neq 0$ and $l(a) \leq 0$, i.e., $a^T u_i \leq 0$ for all $i \in I$ and in matrix form $U^T a \leq 0$. Now $U^T a \neq 0$ (because U^T is full rank and $a \neq 0$) therefore exists $i^* \in I$ such that $a^T u_{i^*} < 0$. Finally, $a^T u_{i^*} = -\sum_{i \in I \setminus \{i^*\}} \frac{\lambda_i}{\lambda_{i^*}} a^T u_i \geq 0$, since $\frac{\lambda_i}{\lambda_{i^*}} > 0$ and $a^T u_i \leq 0$, a contradiction. \square

The second lemma is a matrix formulation of the property that the null vector belongs to a the convex hull of some given vectors.

Lemma 6. Let $U = (u_1 u_2 \cdots u_m) \in \mathbb{R}^{n \times m}$. Then $0 \in \text{conv}\{u_1, \dots, u_m\}$ if and only if $0 \neq \lambda \in \ker(U)$ with $\lambda_i \geq 0$.

Proof. The only if part is a direct application of the convex hull definition. Now suppose that $0 \neq \lambda \in \ker(U)$ with $\lambda_i \geq 0$ and define $\mu = \frac{1}{\|\lambda\|_1} \lambda$. Then $\mu \in \ker(U)$, i.e., $\sum \mu_i u_i = 0$, $\mu_i \geq 0$ and $\sum \mu_i = \frac{\|\lambda\|_1}{\|\lambda\|_1} = 1$ hence $0 \in \text{conv}\{u_1, \dots, u_m\}$. \square

3. Proof of the multivariate equioscillation theorem

First, if $\min_{a \in \mathbb{R}^n} m(a) = 0$ then by the weak Haar condition there is a unique solution ($a^T \phi(x_i) = f(x_i)$ with n linearly independent vectors $\phi(x_i)$ uniquely defines a), which is trivially identified with a vacuous multivariate equioscillation condition. We now suppose that $\min_{a \in \mathbb{R}^n} m(a) > 0$, hence the subdifferential formula (7) holds true.

3.1. Existence of a minimizer

By the weak Haar condition, there exists n distinct points $x_1, \dots, x_n \in X$ such that $H(x_1, \dots, x_n)$ is nonsingular. We now build consecutive lower bounds for $m(a)$. First note that $m(a) \geq \max_{x \in X} (|a^T \phi(x)| - |f(x)|) \geq (\max_{x \in X} |a^T \phi(x)|) - M$ where M is an upper bound of the continuous function f inside the compact X . Finally performing the maximum over $\{x_1, \dots, x_n\} \subseteq X$ we obtain the lower bound $(\max_{x \in \{x_1, \dots, x_n\}} |a^T \phi(x)|) - M$. The pointwise maximum in this last lower bound is $\|H(x_1, \dots, x_n)^T a\|_\infty$, which is coercive since the Haar matrix is nonsingular (indeed $\sqrt{n} \|H^T a\|_\infty \geq \|H^T a\|_2 \geq \lambda_{\min}(HH^T) \|a\|_2$ while $\lambda_{\min}(HH^T) > 0$ because H is nonsingular). Continuous coercive functions defined in \mathbb{R}^n have a minimizer.

3.2. Necessity of the multivariate equioscillation condition at the minimizer

Consider a minimizer $a^* \in \mathbb{R}^n$, hence $0 \in \partial m(a^*) = \text{conv } A$, where is A given in (7). By Carathéodory theorem, 0 is the convex hull of $d \leq n + 1$ vectors of A , i.e., by Lemma 6 we have

$$G(x_1, \dots, x_d)\lambda = 0 \text{ with } \lambda \neq 0, \lambda_i \geq 0 \text{ and } x_1, \dots, x_d \in \text{act}(a^*). \quad (8)$$

Therefore from (5) we obtain $H(x_1, \dots, x_{n+1})u = 0$ with $u_i = \epsilon(x_i)\lambda_i$. Finally $u_i \epsilon(x_i) = \epsilon(x_i)^2 \lambda_i \geq 0$ proving that the multivariate equioscillation condition holds.

3.3. Sufficiency of the multivariate equioscillation condition for a minimizer

Now, we prove that the multivariate equioscillation condition (MEC) at a^* implies $0 \in \partial m(a^*)$. The MEC at a^* means there exists $x_1, \dots, x_d \in \text{act}(a^*)$ and a kernel vector $u \neq 0$ of $H(x_1, \dots, x_d)$ such that $\epsilon(x_i)u_i \geq 0$. Again from (5) we have $G(x_1, \dots, x_d)\lambda = 0$ with $u_i = \epsilon(x_i)\lambda_i$, so that $\lambda \neq 0$ and by the MEC $0 \leq \epsilon(x_i)u_i = \epsilon(x_i)^2 \lambda_i = \lambda_i$. Finally, Lemma 6 proves that 0 is in the convex hull of the columns of $G(x_1, \dots, x_d)$ which are subgradients of m at a^* . Therefore $0 \in \partial m(a^*)$ and a^* is a minimizer.

3.4. Uniqueness

Consider a minimizer $a^* \in \mathbb{R}^n$, so $0 \in \partial m(a^*)$. We restart at (8) but with the additional assumption that the strong multivariate equioscillation condition holds: $H(x_1, \dots, x_d)$ is full rank, hence so is $G(x_1, \dots, x_d)$ and $d \geq n+1$, and u and λ have at least $n+1$ nonzero components. Since the columns of $G(x_1, \dots, x_d)$ are subgradients at a^* we have a piecewise affine lower bound $m(a) \geq m(a^*) + l(a - a^*)$ with $l(a) = \max(G(x_1, \dots, x_d)^T(a - a^*))$. Finally all hypothesis of Lemma 5 hold, and it shows that this piecewise affine lower bound has a unique minimizer a^* , therefore $a \neq a^*$ implies $m(a) > m(a^*)$ proving the uniqueness of the minimizer.

3.5. Strong uniqueness

We proved so far that $m(a) - m(a^*) \geq l(a - a^*)$, which is not as strong as strong uniqueness, but we prove it entails strong uniqueness. The maximum is homogenous for positive scalar hence we have the lower bound $l(a - a^*) = \|a - a^*\|_1 l\left(\frac{a - a^*}{\|a - a^*\|_1}\right)$ which is in turn less than $\alpha \|a - a^*\|_1$ with $\alpha = \min_{\|a - a^*\|_1 = 1} l(a - a^*) > 0$. Now, $\|a\|_{X,\phi} = \max_{x \in X} |a^T \phi(x)|$ is a norm on \mathbb{R}^n (the triangular inequality and absolute homogeneity are trivial, the positive definiteness follows from Haar condition) which is therefore equivalent to $\|a\|_1$, hence $\|a\|_1 \geq \beta \|a\|_{X,\phi}$ for some $\beta > 0$. We finally obtain the lower bound $\alpha\beta \|a - a^*\|_{X,\phi}$, that is strong uniqueness.

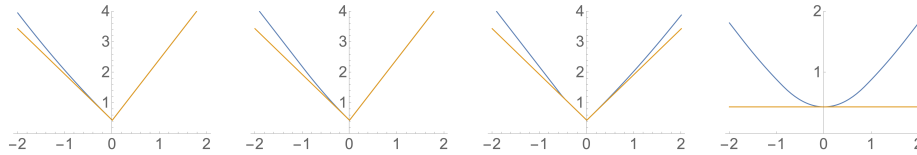


Figure 2: Four directional restrictions of $m(a)$ in blue and its piecewise linear lower bound $l(a)$ in orange, showing a unique non-strongly unique minimizer.

4. Some multivariate problem instances

The first example shows an instance with a strongly unique solution.

Example 7. Let $X = [0, 1] \times [0, 1]$, $f(x) = x_1^2 + x_2^2$ and $\phi(x) = (1, x_1, x_2)$, so that we approximate f by affine functions $a^T \phi(x) = a_1 + a_2 x_1 + a_3 x_2$. The weak Haar condition is satisfied since $H((0, 0), (1, 0), (1, 1))$ is non-singular, so there is an optimal solution. Let us test the approximation $a^* = (-\frac{1}{4}, 1, 1)$ using the multivariate equioscillation theorem. The error absolute value $|e(a^*, x)|$ has five maximizers $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$ and $(\frac{1}{2}, \frac{1}{2})$, whose error values are respectively $(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, \frac{1}{4})$. The Haar matrix for these active indices is

$$H((0, 0), (0, 1), (1, 0), (1, 1), (\frac{1}{2}, \frac{1}{2})) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & \frac{1}{2} \\ 0 & 1 & 0 & 1 & \frac{1}{2} \end{pmatrix}. \quad (9)$$

We can extract the kernel vector $u = (-1, -3, -3, -1, 8)$, which satisfies the multivariate equioscillation condition. Therefore the sincoscillation theorem proves that $-\frac{1}{4} + x_1 + x_2$ is one optimal solution of this Chebyshev approximation problem. Furthermore the Haar matrix is full rank and the kernel vector has at least 4 nonzero components, hence the strong multivariate equioscillation condition holds and this solution is strongly unique.

The second example shows an example where the strong multivariate equioscillation condition does not hold, but where further numerical investigations let us conjecture there is a unique solution. This is related to the singular Chebyshev approximation problems that were investigated in [5, 6, 7, 8].

Example 8. Let $X = [0, 1] \times [0, 1]$, $f(x) = x_1^2 + 2x_2^2 - \frac{1}{2}x_1 x_2$ and $\phi(x) = (1, x_1, x_2)$, so that again we approximate f by affine functions $a^T \phi(x) = a_1 + a_2 x_1 + a_3 x_2$. Again the weak Haar condition holds and implies the existence

of a minimizer. Let us test the approximation $a^* = (-\frac{3}{16}, \frac{3}{4}, \frac{7}{4})$ using the multivariate equioscillation theorem. The error absolute value $|e(a^*, x)|$ has three maximizers $(0, 1)$, $(1, 0)$ and $(\frac{1}{2}, \frac{1}{2})$, whose error values are respectively $(-\frac{7}{16}, -\frac{7}{16}, \frac{7}{16})$. The Haar matrix for these maximizers is

$$H((0, 1), (1, 0), (\frac{1}{2}, \frac{1}{2})) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \end{pmatrix}. \quad (10)$$

We can extract the kernel vector $u = (-1, -1, 2)$, which satisfies the multivariate equioscillation condition. Therefore $-\frac{3}{16} + \frac{3}{4}x_1 + \frac{7}{4}x_2$ is one optimal solution of the Chebyshev approximation problem. The Haar matrix is not full rank, so the multivariate equioscillation theorem does not allow proving the uniqueness. Let us investigate in more details this situation. The subgradient matrix is

$$G((0, 1), (1, 0), (\frac{1}{2}, \frac{1}{2})) = \begin{pmatrix} -1 & -1 & 1 \\ 0 & -1 & \frac{1}{2} \\ -1 & 0 & \frac{1}{2} \end{pmatrix}, \quad (11)$$

whose kernel vector $\lambda = (1, 1, 2)$ is positive as expected. The three columns of G are three subgradient, which give rise to a piecewise linear lower bound $l(a) = \max(G^T(a - a^*))$ as defined in Subsection 3.4. But in the present case, $l(a)$ does not allow to prove uniqueness because it is constant in the direction $v = (-1, 1, 1)$, which is a kernel vector of G^T . Figure 2 shows four directional restrictions of both $m(a)$ and its lower bound $l(a)$. The first three directions are columns of G , hence steepest ascent directions of the lower bound. The fourth direction is the kernel vector v of G^T , so as expected $l(a)$ is flat in this direction. However, we can see that a second order curvature make the minimizer unique.

The following example shows a case with infinitely many solutions to the Chebyshev approximation problem.

Example 9. Again $X = [-1, 1] \times [-1, 1]$ and $\phi(x) = (x_1, x_2)$. We now build a function that will have infinitely many best approximations. Let $\bar{f}(t) = 2t^2 - 1$ be the degree 2 Chebyshev polynomial, with 0 as best Chebyshev approximation in $[-1, 1]$, and $\hat{f}(t) = \exp(-10t)$. Finally, we define $f(x) = \bar{f}(\frac{x_1+x_2}{2}) \hat{f}(\text{abs}(x_1 - x_2))$ so that f is the degree 2 Chebyshev polynomial on the line $x_1 = x_2$ and the quickly converge to zero with the distance to this line. Its graph is shown in the left graphic of Figure 3. Let us test $a^* = (0, 0, 0)$ using the multivariate equioscillation theorem. The error absolute

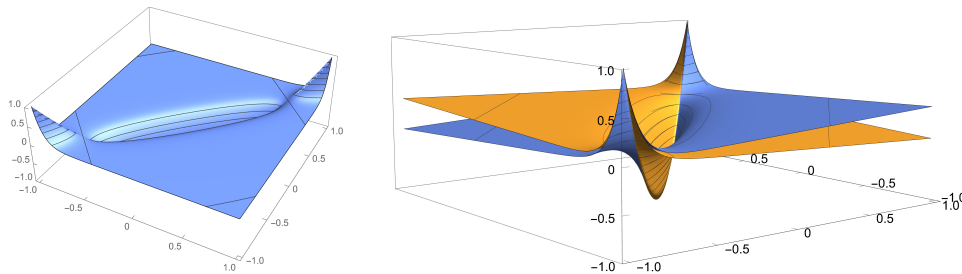


Figure 3: Four directional restrictions of $m(a)$ in blue and its piecewise linear lower bound $l(a)$ in orange, showing a unique non-strongly unique minimizer.

value $|e(a^*, x)|$ has three maximizers $(-1, -1)$, $(0, 0)$ and $(1, 1)$, whose error values are respectively $(1, -1, 1)$. The Haar matrix for these maximizers is

$$H((-1, -1), (0, 0), (1, 1)) = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}. \quad (12)$$

We can extract the kernel vector $u = (1, -2, 1)$, which satisfies the multivariate equioscillation condition. Therefore 0 is one optimal solution of the Chebyshev approximation problem. Now, neighbor affine approximations satisfying $a_2 + a_3 = 0$, so that its value is zero when $x_1 = x_2$, will show the same maximizers for the error absolute value (this is due to the non-differentiability of f on the line $x_1 = x_2$, as see on the right graphics of Figure 3, which fixes the local maximizer independently to small linear perturbations³) and same Haar matrix, hence will be optimal as well.

The following example shows an untypical case where there is only one active index and the multivariate equioscillation theorem succeeds. It is presented to show that $K = 1$ actually makes sense in the multivariate equioscillation definition.

Example 10. Let $X = [-1, 1] \times [-1, 1]$, $f(x) = 2 - x_1^2 - x_2^2$ and $\phi(x) = (x_1, x_2)$, so that we approximate f by linear functions $a^T \phi(x) = a_1 x_1 + a_2 x_2$. The weak Haar condition is satisfied since $H((1, 0), (0, 1))$ is nonsingular. Let us test the approximation $a = (0, 0)$ using the multivariate equioscillation

³For example in 1D, the minimizer of differentiable function x^2 changes with linear small perturbations $x^2 + \epsilon x$, but the minimizer of the nondifferentiable function $|x|$ is not affected by small linear perturbations $|x| + \epsilon x$.

theorem. The error is 0 on the four corners of the box, and has a unique maximizer at $(0,0)$ with error 2. The Haar matrix is

$$H((0,0)) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (13)$$

with kernel vector 1. multivariate equioscillation is satisfied, therefore 0 is one best linear approximation of $f(x)$ in X . In fact, this seemingly strange situation is quite normal from the point of view of convex analysis: the column of the Haar matrix is a subgradient, and having a null subgradient entails being a minimizer.

Aknowlegments

The author's work is funded the CNRS project émergence AMEGO.

Appendix A. Proof of Lemma 1

It is well known if $\phi : X \rightarrow \mathbb{R}^n$ is continuous and satisfies the Haar condition on X then the determinant of all square Haar matrices $H(x_1, \dots, x_n)$, with $x_1 < \dots < x_n \in X$, have the same sign, see [9, proof of lemma page 74]. Now, the Haar matrix $H(x_1, \dots, x_{n+1})$ is full rank (because Haar condition entails $\det H(x_1, \dots, x_n) \neq 0$) so its kernel is dimension 1. Let us consider one kernel element $0 \neq u \in \mathbb{R}^{n+1}$, i.e., $\sum_{i=1}^{n+1} u_i \phi(x_i) = 0$. For an arbitrary $i \in \{1, \dots, n+1\}$, we have $u_i \neq 0$ otherwise $\sum_{j \neq i} u_j \phi(x_j) = 0$ and $H_i := H(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$ would be singular contradicting Haar condition. For $i \in \{1, \dots, n\}$ we have $\phi(x_i) = -\sum_{j \neq i} \frac{u_j}{u_i} \phi(x_j)$ and replacing the column $\phi(x_i)$ in H_{i+1} by this sum shows that $\det H_{i+1}$ is equal to

$$\left| \phi(x_1) \cdots \phi(x_{i-1}) \left(-\sum_{j \neq i} \frac{u_j}{u_i} \phi(x_j) \right) \phi(x_{i+2}) \cdots \phi(x_{n+1}) \right| = -\frac{u_{i+1}}{u_i} \det H_i, \quad (\text{A.1})$$

the last equality obtained using the multilinearity and alternativity of the determinant. The determinants $\det H_i$ and $\det H_{i+1}$ having the same sign, we conclude that u_i and u_{i+1} have opposite signs.

References

- [1] S. E. Weinstein, Approximations of functions of several variables: Product chebychev approximations, I, Journal of Approximation Theory 2 (1969) 433–447.

- [2] S. E. Weinstein, Product approximations of functions of several variables, *SIAM Journal on Numerical Analysis* 8 (1971) 178–189.
- [3] M. Henry, S. Weinstein, Best rational product approximations of functions. II, *Journal of Approximation Theory* 12 (1974) 6–22.
- [4] J. C. Mairhuber, On Haar’s theorem concerning Chebychev approximation problems having unique solutions, *Proc. Amer. Math. Soc.* 7 (1956) 609–615.
- [5] M. Osborne, G. Watson, A note on singular minimax approximation problems, *Journal of Mathematical Analysis and Applications* 25 (1969) 692–700.
- [6] G. A. Watson, A multiple exchange algorithm for multivariate chebyshev approximation, *SIAM Journal on Numerical Analysis* 12 (1975) 46–52.
- [7] R. Reemtsen, Modifications of the first remez algorithm, *SIAM Journal on Numerical Analysis* 27 (1990) 507–518.
- [8] R. Reemtsen, Discretization methods for the solution of semi-infinite programming problems, *Journal of Optimization Theory and Applications* 71 (1991) 85–103.
- [9] E. W. Cheney, *Introduction to Approximation Theory*, AMS Chelsea Publishing, 1982.