A generalisation of de la Vallée-Poussin procedure to multivariate approximations

Nadezda Sukhorukova1 · Julien Ugon2

Received: 11 January 2021 / Accepted: 5 December 2021 / Published online: 7 January 2022
© Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
The theory of Chebyshev approximation has been extensively studied. In most cases, the optimality conditions are based on the notion of alternance or alternating sequence (that is, maximal deviation points with alternating deviation signs). There are a number of approximation methods for polynomial and polynomial spline approximation. Some of them are based on the classical de la Vallée-Poussin procedure. In this paper we demonstrate that under certain assumptions the classical de la Vallée-Poussin procedure, developed for univariate polynomial approximation, can be extended to the case of multivariate approximation. There are two main advantages of our approach. First of all, it provides an elegant geometrical interpretation of the procedure. Second, the corresponding basis functions are not restricted to be monomials and therefore can be extended to a larger family of functions.

Keywords Multivariate polynomial · Chebyshev approximation · de la Vallée-Poussin procedure · Radon theorem

PACS: 58C05 · 41A10 · 41A50

1 Introduction
The theory of Chebyshev approximation for univariate functions was developed in the late nineteenth (Chebyshev) and twentieth century (just to name a few [1–3]). Many papers are dedicated to polynomial and polynomial spline approximations, however, other types of functions (for example, trigonometric polynomials) have
also been used. In most cases, the optimality conditions are based on the notion of alternance (that is, maximal deviation points with alternating deviation signs).

There have been several attempts to extend this theory to the case of multivariate functions. One of them is [4]. The main obstacle in extending these results to the case of multivariate functions is that the notion of alternation depends on the ordering of the real numbers. The geometrical interpretation of our approach is designed to overcome this difficulty.

It has to be pointed out that multivariate monomials (and most other basis functions) do not satisfy Haar’s condition, and so do not form a Chebyshev system. This means that classical results cannot be generalised trivially to multivariate approximation. The absence of Haar’s condition does not mean that the generalisation of de la Vallée-Poussin procedure [5] is not possible. One example of such families of functions is univariate polynomial splines [1, 6].

The main contribution of this paper is the extension of the classical de la Vallée-Poussin procedure (originally developed for univariate polynomial approximation [5]) to the case of multivariate approximation under certain assumptions. The corresponding basis functions are not restricted to be monomials (that is, non-polynomial approximation). This extra generalisation makes the procedure very attractive for a number of applications, including numerical analysis, signal processing [7] and numerical integration [8], where approximation plays a major role.

The most recent developments in the area of Chebyshev (uniform) approximation relate to multivariate polynomial and rational approximation, which is seen as a flexible generalisation of polynomial approximation for approximating nonsmooth and non-Lipschitz functions. One of the most promising approaches [9] is dedicated to “nearly optimal” solutions, whose construction is based on Chebyshev polynomials. This is a very efficient method, but it constructs only approximately optimal solutions, while our method explores the properties of optimal solutions and therefore serves different purposes.

The paper is organised as follows. In Section 2 we demonstrate that the corresponding optimisation problems are convex. Then, in Section 3 we extend the classical de la Vallée-Poussin procedure to the case of multivariate approximation. Section 4 contains numerical examples illustrating the procedure. Finally, Section 5 highlights our future research directions.

## 2 Convexity of the objective function

Let us now formulate the objective function. Suppose that a continuous function \( f(x) \) is to be approximated by a function

\[
L(A, x) = a_0 + \sum_{i=1}^{n} a_i g_i(x),
\]

where \( L(A, x) \) is a modelling function, \( g_i(x) \), \( i = 1, \ldots, n \) are the basis functions and the multipliers \( A = (a_0, a_1, \ldots, a_n) \) are the corresponding coefficients. In the case of polynomial approximation, basis functions are monomials. In this paper, however,
we do not restrict ourselves to polynomials. At a point $x$ the deviation between the
function $f$ (also referred as approximation function) and the approximation is

$$d(A, x) = |f(x) - L(A, x)|.$$  \hspace{1cm} (2)

Then we can define the uniform approximation error over the set $Q$ by

$$\Psi(A) = \sup_{x \in Q} \max \{f(x) - a_0 - \sum_{i=1}^{n} a_i g_i(x), a_0 + \sum_{i=1}^{n} a_i g_i(x) - f(x)\}. \hspace{1cm} (3)$$

The approximation problem is

$$\text{minimise } \Psi(A) \text{ subject to } A \in \mathbb{R}^{n+1}. \hspace{1cm} (4)$$

Since the function $L(A, x)$ is linear in $A$, the approximation error function $\Psi(A)$, as
the supremum of affine functions, is convex. Furthermore, its subdifferential at a
point $A$ is trivially obtained using the gradients of the active affine functions in the
supremum (see [10] for details):

$$\partial \Psi(A) = \text{co}\left\{ \begin{pmatrix} 1 \\ g_1(x) \\ g_2(x) \\ \vdots \\ g_n(x) \end{pmatrix} : x \in E_+(A), - \begin{pmatrix} 1 \\ g_1(x) \\ g_2(x) \\ \vdots \\ g_n(x) \end{pmatrix} : x \in E_-(A) \right\}, \hspace{1cm} (5)$$

where $E_+(A)$ and $E_-(A)$ are respectively the points of maximal positive and negative
deviation (extreme points):

$$E^+(A) = \left\{ x \in Q : f(x) - L(A, x) = \max_y d(A, y) \right\}$$

$$E^-(A) = \left\{ x \in Q : -f(x) + L(A, x) = \max_y d(A, y) \right\}.$$

Note that in the case of multivariate polynomial approximation, $g_i(x), i = 1, \ldots, n$
are monomials.

Define by $G^+$ and $G^-$ the sets

$$G^+(A) = \left\{ (1, g_1(x), \ldots, g_n(x))^T : x \in E^+(A) \right\}$$

$$G^-(A) = \left\{ (1, g_1(x), \ldots, g_n(x))^T : x \in E^-(A) \right\}.$$

The following theorem provides a geometric interpretation for multivariate
approximation optimality conditions [11]. One of the main strengths of this
interpretation is that the optimality condition does not depend on the nature of the
basis functions.
Theorem 1 ([11]) $A^*$ is an optimal solution to problem (4) if and only if the convex hulls of the sets $G^+(A^*)$ and $G^-(A^*)$ intersect.

3 de la Vallée-Poussin procedure for nonsingular basis

3.1 Definitions and existing results

We start with necessary definitions from convex analysis.

Definition 1 The relative interior of a set $S$ (denoted by relint($S$)) is defined as its interior within the affine hull of $S$. That is,

$$\text{relint}(S) = \{x \in S : \exists \varepsilon > 0, B_\varepsilon(x) \cap \text{aff}(S) \subseteq S\},$$

where $B_\varepsilon(x)$ is a ball of radius $\varepsilon$ centred in $x$ and $\text{aff}(S)$ is the affine hull of $S$.

A useful property of relative interiors of convex hulls of finite number of points is formulated in the following lemma (see [12], Remark 2.1.4 and [13]).

Lemma 1 Any relative interior point of a convex combination of a finite number of points can be presented as a convex combination of all these points with strictly positive convex combination coefficients and vice versa.

In the classical de la Vallé-Poussin procedure for the univariate case polynomial approximation, a basis is an arbitrary collection of $n+2$ points, where $n$ is the number of monomials. What do we call basis in multivariate case? Based on necessary and sufficient optimality conditions (Theorem 1) the convex hulls built over positive and negative maximal deviation points should intersect. Is it always possible to partition $n+2$ points into two subsets in such a way that the corresponding convex hulls are intersecting. The answer to this question is “yes”, if $n \geq d$. The following theorem holds.

Theorem 2 (Radon [14]) Any set of $d+2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint sets whose convex hulls intersect.

Definition 2 A point in the intersection of these convex hulls (Theorem 2) is called a Radon point of the set.

In the rest of the paper we assume that $n \geq d$.

It will be demonstrated that it is not possible to extend de la Vallée-Poussin procedure to multivariate approximations without imposing additional assumptions (non-singular basis). It may be possible that some (or all) of these assumptions can be removed if we restrict ourselves to a particular class of basis functions (for example, monomials). This research direction is out of scope of this paper.
**Definition 3** Consider a set $S$ of $n + 2$ points partitioned into two sets, the sets $Y$ of points with positive deviation and $Z$ of points with negative deviation. These points are said to form a basis if the convex hulls of $Y$ and $Z$ intersect. Furthermore, if the relative interiors of the convex hulls intersect and any $(n + 1)$ point subset of this basis form an affine independent system then the basis is said to be non-singular.

**3.2 de la Vallée-Poussin procedure for multivariate approximations**

**3.2.1 Classical univariate procedure**

The classical univariate de la Vallée-Poussin procedure contains three steps.

1. For any basis (arbitrary $(n + 2)$ points) there exists a unique polynomial, such that the absolute deviation at the basis points is the same and the deviation sign is alternating. This polynomial is also called Chebyshev interpolation polynomial.
2. If there is a point (outside of the current basis), such that the absolute deviation at this point is higher than at the basis points then this point can be included in the basis by removing one of the current basis points and the deviation signs are deviating.
3. The absolute deviation of the new Chebyshev interpolating polynomial is at least as high as the absolute deviation for the original basis.

In the rest of this section we extend the procedure for a non-singular basis.

**3.2.2 Step one extension**

We start with constructing Chebyshev interpolation polynomials. The following theorem holds.

**Theorem 3** Assume that a system of points $y_i, i = 1, \ldots, N_+$ and $z_i, i = 1, \ldots, N_-$ forms a non-singular basis. Then there exists a unique polynomial deviating from $f$ at the points $y_i, i = 1, \ldots, N_+$ and $z_i, i = 1, \ldots, N_-$ by the same value and the deviation signs are opposite for $y_i, i = 1, \ldots, N_+$ and $z_i, i = 1, \ldots, N_-$.

**Proof** Consider the following linear system:

$$
\begin{pmatrix}
1 & g(y_1) & 1 \\
1 & g(y_2) & 1 \\
\vdots & \vdots & \vdots \\
1 & g(y_{N_+}) & 1 \\
1 & g(z_1) & -1 \\
1 & g(z_2) & -1 \\
\vdots & \vdots & \vdots \\
1 & g(z_{N_-}) & -1
\end{pmatrix} \begin{pmatrix}
A \\
\sigma
\end{pmatrix} = \begin{pmatrix}
f(y_1) \\
f(y_2) \\
\vdots \\
f(y_{N_+}) \\
f(z_1) \\
f(z_2) \\
\vdots \\
f(z_{N_-})
\end{pmatrix},
$$

(6)
where \( \mathbf{A} \) represents the parameters of the polynomial, while \( \sigma \) is the deviation. If \( \sigma = 0 \), there exists a polynomial passing through the chosen points (interpolation). Denote the system matrix in (6) by \( \mathbf{M} \). Since the basis is non-singular, that is, the relative interiors of the convex hulls of sets \( Y \) and \( Z \) are intersecting, there exist two sets of strictly positive coefficients

\[
\alpha_1, \ldots, \alpha_{N_+} : \sum_{i=1}^{N_+} \alpha_i = 1
\]

and

\[
\beta_1, \ldots, \alpha_{N_-} : \sum_{i=1}^{N_-} \beta_i = 1,
\]

such that

\[
\sum_{i=1}^{N_+} \alpha_i g(y_i) = \sum_{i=1}^{N_-} \beta_i g(z_i). \tag{7}
\]

Multiply the first row of \( \mathbf{M} \) by the convex coefficient \( \alpha_1 \) from (7). For each remaining row of \( \mathbf{M} \) one can apply the following update:

- multiply by the corresponding convex coefficient and add all the rows that correspond to the vertices with the same deviation sign as the first row;
- multiply by the corresponding convex coefficient and subtract all the rows that correspond to the vertices with the deviation sign opposite to the sign of the first row.

Then

\[
\alpha_l \det(\tilde{\mathbf{M}}) = 2(-1)^{l+2+i} \det(M^+_i), \quad l = 1, \ldots, N_+ \tag{8}
\]

where \( M^+_i \) is obtained from \( \tilde{\mathbf{M}} \) by removing the last column and the \( i-\)th row and \( M^-_j \) is obtained from \( \tilde{\mathbf{M}} \) by removing the last column and the \( (N_+ + j)\)-th row. Also note that

\[
\det(M^+_i) = 2(-1)^{l+2+N_+ + j+1} \det(M^-_j), \quad l = 1, \ldots, N_+. \tag{9}
\]

If now we evaluate the the determinant of \( \mathbf{M} \) directly, then

\[
\det \mathbf{M} = \sum_{i=1}^{N_+} (-1)^{l+2+i} \Delta_i + \sum_{j=N_+ + 1}^{N_+ + N_-} (-1)^{l+2+j+1} \Delta_j. \tag{10}
\]

Based of (8), each component in the right hand side of (10) has the same sign. Therefore, the linear system (6) has a unique solution for any right hand side of the system.
Note that the division into “positive” and “negative” basis points does not mean that the deviation sign is positive for “positive” basis points and negative for “negative” basis points. The actual deviation sign also depends on the sign of $u_1D70E$ from (6).

The notion of Chebyshev interpolation can be extended to multivariate polynomials and to non-polynomial modelling functions.

The additional requirement for a basis to be non-singular may be removed by

– restricting to some particular types of basis functions (for example, polynomials);
– allowing the system (6) to have more than one solution.

These will be included in our future research directions.

3.2.3 Step two extension

Our next step is to demonstrate

**Theorem 4** Consider two intersecting sets $Y$ and $Z$ such that the points in $Y$ all have the same deviation and opposite deviation to all the points in $Z(g(\tilde{y}) = -g(\tilde{z}), \forall \tilde{y} \in Y, \tilde{z} \in Z)$. Assume now that $g(y) = g(\tilde{y}), \forall y \in Y$, and that the set

$$\mathcal{K} = \text{relint}(\{Y \cup y\}) \cap \text{relint}(Z) \neq \emptyset.$$ 

There exists a point in the combined collection of vertices of $Y$ and $Z$, that can be removed while $y$ is included in $Y$, such that the updated sets $\tilde{Y}$ and $\tilde{Z}$ intersect.

**Proof** Since $\text{relint}(Y) \cap \text{relint}(Z) \neq \emptyset$, there exist strictly positive coefficients

$$\alpha_i, \ i = 1, \ldots, N_+ \ \text{and} \ \beta_j, \ j = 1, \ldots, N_-,$$

such that $\sum_{i=1}^{N_+} \alpha_i = 1$ and $\sum_{j=1}^{N_-} \beta_j = 1$ and $\sum_{i=1}^{N_+} \alpha_i y_i = \sum_{j=1}^{N_-} \beta_j z_j$.

Since $\mathcal{K} \neq \emptyset$ there exist strictly positive coefficients

$$\alpha, \ \tilde{\alpha}_i, \ i = 1, \ldots, N_+$$

such that $\alpha + \sum_{i=1}^{N_+} \tilde{\alpha}_i = 1$ and $\tilde{\beta}_j, \ j = 1, \ldots, N_-$, such that $\sum_{j=1}^{N_-} \tilde{\beta}_j = 1$ and $\alpha y + \sum_{i=1}^{N_+} \tilde{\alpha}_i y_i = \sum_{j=1}^{N_-} \tilde{\beta}_j z_j$.

Find

$$\gamma = \min \left\{ \min_{i=1, \ldots, N_+} \frac{\tilde{\alpha}_i}{\alpha_i}, \min_{j=1, \ldots, N_-} \frac{\tilde{\beta}_j}{\beta_j} \right\}. \quad (11)$$

First, assume that $\gamma = \frac{\tilde{\alpha}_1}{\alpha_1}$. Note that $\alpha_1 \neq 0$, then (7) can be written as

$$y_1 = \frac{1}{\alpha_1} \left( \sum_{j=1}^{N_-} \beta_j g(z_j) - \sum_{i=2}^{N_+} \alpha_i g(y_i) \right).$$
Then, the convex hull with the new point \( y \) is

\[
\alpha g(y) + \frac{\tilde{\alpha}_i}{\alpha_1} \left( \sum_{j=1}^{N} \beta_j g(z_j) - \sum_{i=2}^{N} \alpha_i g(y_i) \right) + \sum_{i=2}^{N} \alpha_i \tilde{\alpha}_i g(y_i) = \sum_{j=1}^{N} \tilde{\beta}_j g(z_j)
\]

(12)

and finally

\[
\alpha g(y) + \sum_{i=2}^{N} (\tilde{\alpha}_i - \frac{\tilde{\alpha}_i}{\alpha_1}) g(y_i) = \sum_{j=1}^{N} (\tilde{\beta}_j - \frac{\tilde{\alpha}_1}{\alpha_1}) g(z_j).
\]

(13)

Since \( \alpha_i > 0, i = 1, \ldots, N_+ \) and the definition of \( \gamma \), one can obtain that for any \( i = 1, \ldots, N_+ \)

\[
\tilde{\alpha}_i - \frac{\tilde{\alpha}_i}{\alpha_1} \geq \tilde{\alpha}_i - \frac{\tilde{\alpha}_i}{\alpha_1} = 0.
\]

(14)

Similarly, for any \( j = 1, \ldots, N_- \),

\[
\tilde{\beta}_j - \frac{\tilde{\alpha}_1}{\alpha_1} \beta_j \geq 0.
\]

Note that

\[
\sum_{j=1}^{N_-} (\tilde{\beta}_j - \frac{\tilde{\alpha}_1}{\alpha_1} \beta_j) = 1 - \frac{\tilde{\alpha}_1}{\alpha_1}
\]

(15)

and

\[
\alpha + \sum_{i=2}^{N} \tilde{\alpha}_i - \frac{\tilde{\alpha}_1}{\alpha_1} \sum_{i=2}^{N} \alpha_i = \alpha + (1 - \alpha \tilde{\alpha}_1) - \frac{\tilde{\alpha}_1}{\alpha_1} = 1 - \frac{\tilde{\alpha}_1}{\alpha_1} = 1 - \gamma.
\]

(16)

Since \( \alpha \) is strictly positive, \( \gamma < 1 \). Therefore, the new point can be included instead of \( y_1 \) and the convex hulls of the updated sets are intersecting (and so their relevant interiors).

Second, assume that \( \gamma = \frac{\tilde{\beta}_1}{\beta_1} \). Note that \( \beta_1 \neq 0 \), otherwise \( y \) can be included instead of \( z_1 \).

Similarly to part 1, obtain

\[
\alpha g(y) + \sum_{i=1}^{N_+} (\tilde{\alpha}_i - \frac{\tilde{\alpha}_1}{\beta_1}) g(y_i) = \sum_{j=2}^{N_-} (\tilde{\beta}_j - \frac{\tilde{\alpha}_1}{\beta_1} \beta_j) g(z_j).
\]

(17)

Since

\[
\alpha + 1 - \alpha - \frac{\tilde{\beta}_1}{\beta_1} = 1 - \tilde{\beta}_1 - \frac{\tilde{\alpha}_1}{\beta_1} (1 - \beta_1) = 1 - \frac{\tilde{\beta}_1}{\beta_1} > 0,
\]

the convex hulls of the updated sets are intersecting.

Note that for the extension of this step we only need the assumption that the relative interiors are intersecting, moreover, if this is the case, the new basis preserves this property.
3.2.4 Step three extension

The final step is to show that the proposed exchange rule leads to a modelling function whose deviation at the new basis is strictly higher than the deviation at the points of the original basis.

**Theorem 5** Assume that a point with a higher absolute deviation is included in the basis instead of one of the points of the original basis (which is also non-singular). The absolute deviation of the Chebyshev interpolation modelling function that corresponds to the new basis is higher than the one of the Chebyshev interpolation modelling function on the original basis.

**Proof** Denote by

\[ \mathcal{Y} = \{ y_i, \ i = 1, \ldots, N_+ \} \]

and

\[ \mathcal{Z} = \{ z_j, \ j = 1, \ldots, N_- \} \]

respectively. Assume that \( \mathring{\mathcal{Y}} = \mathcal{Y} \cup \{ y \} \setminus \{ y_1 \} \) and \( \mathring{\mathcal{Z}} = \mathcal{Z} \) (when the a point from the set \( \mathcal{Z} \) is removed instead, the proof is similar).

Since the convex hulls of positive and negative deviation points are intersecting, there exist nonnegative convex coefficients

- \( \alpha_1, \ldots, \alpha_{N_+} : \sum_{i=1}^{N_+} \alpha_i = 1 \) and \( \beta_1, \ldots, \beta_{N_-} : \sum_{j=1}^{N_-} \beta_j = 1 \) (original basis);
- \( \alpha, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_{N_+} : \alpha + \sum_{i=2}^{N_+} \tilde{\alpha}_i = 1 \) and \( \beta_1, \ldots, \beta_{N_-} : \sum_{j=1}^{N_-} \beta_j = 1 \) (new basis),

such that on the original basis

\[
\sum_{i=1}^{N_+} \alpha_i y_i - \sum_{j=1}^{N_-} \beta_j z_j = 0 \tag{18}
\]

and on the new basis

\[
\alpha y + \sum_{i=2}^{N_+} \tilde{\alpha}_i y_i - \sum_{j=1}^{N_-} \tilde{\beta}_j z_j = 0 \tag{19}
\]

Systems (18) is equivalent to

\[
\begin{bmatrix}
\alpha, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_{N_+}, \tilde{\beta}_1, \ldots, \tilde{\beta}_{N_-}
\end{bmatrix}
\begin{bmatrix}
y \\
y_2 \\
\vdots \\
z_1 \\
z_{N_-}
\end{bmatrix}
= 0. \tag{20}
\]
Then

\[
\begin{bmatrix}
\alpha, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_{N_+}, \tilde{\beta}_1, \ldots, \tilde{\beta}_{N_-}
\end{bmatrix}
\begin{bmatrix}
1 & y \\
1 & y_2 \\
\vdots & \vdots \\
1 & y_{N_+} \\
1 & z_1 \\
\vdots & \vdots \\
1 & z_{N_-}
\end{bmatrix} = 0
\]  
(21)

for any \( A \in \mathbb{R}^{n+1} \). Let \( A_o \) and \( A_{new} \) be parameter coefficients of the Chebyshev interpolation modelling functions that correspond to the original and new basis respectively. Then

\[
\alpha P_n(A_o, y) + \sum_{i=2}^{N_+} \tilde{\alpha}_i P_n(A_o, y_i) - \sum_{j=1}^{N_-} \tilde{\beta}_j P_n(A_o, z_j) = 0
\]  
(22)

and

\[
\alpha P_n(A_{new}, y) + \sum_{i=2}^{N_+} \tilde{\alpha}_i P_n(A_{new}, y_i) - \sum_{j=1}^{N_-} \tilde{\beta}_j P_n(A_{new}, z_j) = 0.
\]  
(23)

Assume that

\[
f(y_1) - P_n(A_{new}, y_1) = \sigma_{new} > 0.
\]  
(24)

Then

\[
\sigma_{new} + P_n(A_{new}, y) = f(y),
\]  
(25)

\[
\sigma_{new} + P_n(A_{new}, y_i) = f(y_i), \quad i = 2, \ldots, N_+,
\]  
(26)

and

\[
-\sigma_{new} + P_n(A_{new}, z_j) = f(z_j), \quad j = 2, \ldots, N_-.
\]  
(27)

Denote

\[
\Sigma = \sum_{j=1}^{N_-} \tilde{\beta}_j (f(z_j) - P_n(A_o, z_j)).
\]

Due to (22)-(23)
Therefore, \( \sigma_{\text{new}} > \sigma_o \).

Therefore, the notion of basis and de la Vallée-Poussin procedure is extended to multidimensional functions. Also, it has been extended to any basis functions (not only traditional polynomials). If the newly obtained basis is non-singular, one can make another de la Vallée-Poussin procedure step.

4 Numerical examples

In this section we demonstrate how the extension of de la Vallée-Poussin procedure works in multivariate settings. The goal is to approximate \( f(x, y) = (x + y)^2 \) by a lower degree multivariate polynomial \( p(x, y) = ax + by + c \), where \( a, b \) and \( c \) are the decision variables, \(-2 \leq x \leq 2\) and \(-2 \leq y \leq 2\). The choice of the initial basis is essential. Currently, there is no procedure for selecting an efficient initial basis. The following two examples demonstrate that the initial basis choice can change drastically the number of required iterations. Case 1 corresponds to a much better choice of the initial basis compared to Case 2.

4.1 Initial basis choice: Case 1

We start with the basis points located at
\[
(1, 1), (1, -1), (-1, 1), (-1, -1).
\]
Iterations 1: Points (1, 1) and (−1, −1) are assigned to $\mathcal{Y}$ (red) and points (−1, 1) and (1, −1) to $\mathcal{Z}$ (blue). The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ are intersecting. The coefficients of the Chebyshev interpolation polynomial and the deviation at the basis points can be obtained by solving the system (6). In this example, the Chebyshev interpolation polynomial is $p_1(x, y) = 2$ and the absolute deviation at the basis points is $\sigma = 2$.

The maximal deviation of polynomial $p_1$ over the whole domain ($−2 \leq x \leq 2$ and $−2 \leq y \leq 2$) is 14, achieved, for example, at (2,2). This point is included into the new basis (set $\mathcal{Y}$), while point (1,1) will be removed. The minimal deviation is $−2$, and is attained at basis points. The absolute deviation is higher for positive deviation.

Iterations 2: The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ intersect. The Chebyshev interpolation polynomial is $p_2(x, y) = 4 + 2x + 2y$ and the absolute deviation at the basis points is $\sigma = 4$. The maximal deviation of polynomial $p_2$ over the whole domain is 20, achieved, for example, at (−2,−2). This point is included into the new basis (set $\mathcal{Y}$), while point (−1,−1) is removed. The minimal deviation is $−5$, the absolute
deviation is higher for positive deviation. Note that our results indicate that one can include the minimal deviation point into the new basis, but we include maximal deviation points (similar to the classical univariate version).

Iterations 3: The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ intersect. The Chebyshev interpolation polynomial is $p_3(x, y) = 8$ and the absolute deviation at the basis points is $\sigma = 8$. The maximal deviation of polynomial $p_3$ over the whole domain is 8 and the minimal deviation is -8. Based on Theorem 1, this polynomial is an optimal approximation.

4.2 Initial basis choice: Case 2

We start with the basis points located at

$$(1, -1), (0, 1), (0, 0), (-1, -1).$$
Iteration 1: Points \((1, -1), (-1, -1)\) and \((0, 1)\) are assigned to \(\mathcal{Y}\) and point \((0, 0)\) to \(\mathcal{Z}\). The relative interiors of the convex hulls of \(\mathcal{Y}\) and \(\mathcal{Z}\) are intersecting.

The Chebyshev interpolation polynomial is \(p_1(x, y) = 0.75 - 2x - 0.5y\) and the absolute deviation at the basis points is \(\sigma = 0.75\).

The maximal deviation of polynomial \(p_1\) over the whole domain \((-2 \leq x \leq 2\) and \(-2 \leq y \leq 2\)) is 20.25 achieved at \((2, 2)\). If we include this point the new basis (set \(\mathcal{Y}'\)), while point \((0,1)\) is removed, sets \(\mathcal{Y}\) and \(\mathcal{Z}\) will have a non-empty intersection, but their relative interiors are not intersecting. Therefore, the basis update has to be done differently. Based on our rules, we need to include a point with the absolute deviation higher than at the current basis points. One possibility is to include point \((1.5, 2)\), whose deviation is 15.5, and remove point \((0, 1)\) in \(\mathcal{Y}\). We suggest, however, to include a point with the highest negative deviation: point \((-2, 2)\), whose deviation is -3.75, to set \(\mathcal{Z}\) and remove point \((1, -1)\) from \(\mathcal{Y}\).
So, now the number of point in each set (\( \mathcal{Y} \) and \( \mathcal{Z} \)) are the same.

**Iteration 2:** The relative interiors of the convex hulls of \( \mathcal{Y} \) and \( \mathcal{Z} \) are intersecting. The Chebyshev interpolation polynomial is \( p_2(x, y) = 1 - x - y \) and the absolute deviation at the basis points is \( \sigma = 1 \). The maximal deviation of polynomial \( p_2 \) over the whole domain is 19, achieved at (2,2). If this point is included into \( \mathcal{Y} \), while point (0, 1) is excluded, the relative interiors of the convex hulls of \( \mathcal{Y} \) and \( \mathcal{Z} \) are not intersecting and therefore, similar to Iteration 1, we suggest to include the point of the highest negative deviation: point (1, -1.5), whose deviation is -1.25 into \( \mathcal{Z} \), and remove point (0, 0).

**Iteration 3:** The relative interiors of the convex hulls of \( \mathcal{Y} \) and \( \mathcal{Z} \) are intersecting. The Chebyshev interpolation polynomial is \( p_3(x, y) = 0.9079 - 1.5026x - 0.9737y \) and the absolute deviation at the basis points is \( \sigma = 1.0656 \). The maximal deviation of polynomial \( p_3 \) over the whole domain is 19.1447, achieved at point (2, 2) and the minimal deviation is -1.3026. We include this point (2, 2) into set \( \mathcal{Y} \) and exclude point (0, 1).
Iteration 4: The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ are intersecting. The Chebyshev interpolation polynomial is $p_4(x, y) = 4.3946 + 2.1923x + 1.8077y$ and the absolute deviation at the basis points is $\sigma = 3.6154$. The maximal deviation of polynomial $p_4$ over the whole domain is 19.6154, achieved at point $(-2, -2)$ and the minimal deviation is -5.9615. We include point $(-2, -2)$ into set $\mathcal{Y}$ and exclude point $(-1, -1)$.

Iteration 5: The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ are intersecting. The Chebyshev interpolation polynomial is $p_6(x, y) = 8.0769 + 0.0385x - 0.0385y$ and the absolute deviation at the basis points is $\sigma = 7.931$. The maximal deviation of polynomial $p_6$ over the whole domain is 7.9231, achieved at point $(-2, -2)$ and the minimal deviation is -8.2308 at point $(2, -2)$. We include point $(2, -2)$ into set $\mathcal{Z}$ and exclude point $(-1, -1.5)$. 
Iteration 6: The relative interiors of the convex hulls of $\mathcal{Y}$ and $\mathcal{Z}$ are intersecting. The Chebyshev interpolation polynomial is $p_7(x,y) = 8$ and the absolute deviation at the basis points is $\sigma = 8$. The maximal deviation of polynomial $p_7$ over the whole domain is $8$ and the minimal deviation is $-8$. Based on Theorem 1, this polynomial is optimal [15].

5 Conclusions and further research directions

In this paper we extended the classical de la Vallée-Poussin procedure to the case of multivariate approximation under the assumption that the corresponding basis is non-singular. Since the family of multivariate polynomials does not satisfy Haar’s condition, this assumption can not be omitted.
In our future research we will extend the results to the case when the basis is singular. In order to do this, we need to remove two assumptions.

1. Any $(n + 1)$ point subset of the basis $(n + 2)$ points form an affine independent system.
2. Relative interiors of the convex hulls of positive and negative maximal deviation points (restricted to basis) are intersecting.

The first assumption may not be removed for an arbitrary type of basis function. However, it may be possible to remove this assumption for some special types of functions (for example, polynomials). The removal of the second assumption may lead to dimension reduction. These will be included in our future research directions.

Acknowledgements
This research was supported by the Australian Research Council (ARC), Solving hard Chebyshev approximation problems through nonsmooth analysis (Discovery Project DP180100602). This paper was inspired by the discussions during a recent MATRIX program “Approximation and Optimisation” that took place in July 2016. We are thankful to the MATRIX organisers, support team and participants for a terrific research atmosphere and productive discussions.

References

1. Nürenberger, G.: Approximation by spline functions, Springer-Verlag, 1989
2. Rice, J.: Characterization of Chebyshev approximation by splines, SIAM Journal on Numerical Analysis 4 (1967), no. 4, 557–567, (1967)
3. Schumaker, L.: Uniform approximation by Chebyshev spline functions. II: free knots, SIAM Journal of Numerical Analysis 5, 647–656, (1968)
4. Rice, John: Tchebycheff approximation in several variables. Trans. Amer. Math. Soc. 109, 444–466 (1963)
5. Charles Jean de la Vallée Poussin: Sur la méthode de l’approximation minimum. Annales de la Société Scientifique de Bruxelles 35, 1–16 (1911)
6. Sukhorukova, Nadezda: Uniform approximation by the highest defect continuous polynomial splines: necessary and sufficient optimality conditions and their generalisations. Journal of Optimisation Theory and Applications 147(2), 378–394 (2010)
7. Peiris, V., Sharon, N., Sukhorukova, N. and Ugon, J.: Generalised rational approximation and its application to improve deep learning classifiers, Applied Mathematics and Computation 389, 125560 (2021)
8. Blair, J.M., Edwards, C.A., Johnson, J.H.: Rational Chebyshev Approximations for the Inverse of the Error Function. Mathematics of Computation 30(136), 827–830 (1976)
9. Nakatsukasa, Yuji, Sete, Olivier, Trefethen, Lloyd N.: The aaa algorithm for rational approximation. SIAM Journal on Scientific Computing 40(3), A1494–A1522 (2018)
10. Zalinescu, C.: Convex analysis in general vector spaces, World Scientific, (2002)
11. Sukhorukova, N., Ugon, J. and Yost, D.: Chebyshev multivariate polynomial approximation: alternance interpretation, in 2016 MATRIX Annals , David R. Wood, Jan de Gier, Cheryl E. Praeger and Terence Tao (eds.), Springer International Publishing, Switzerland, : 177–182, (2018)
12. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I. Springer-Verlag, Berlin, Heidelberg, New York (1993)
13. Bertsekas, D.: Convex Analysis and Optimization Chapter 1 Solutions, Athena Scientific (2008)
14. Radon, J.: Mengen konvexer körper, die einen gemeinsamen punkt enthalten. Mathematische Annalen 83(1–2), 113–115 (1921)

15. Sukhorukova, Nadezda: Vallée Poussin theorem and Remez algorithm in the case of generalised degree polynomial spline approximation. Pacific Journal of Optimization 6(2), 103–114 (2010)

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