Asymptotic expression of the linear discrete best \( \ell_p \)-approximation

J. Fernández-Ochoa, J. Martínez-Moreno, J.M. Quesada∗

Departamento de Matemáticas, Universidad de Jaén, 23071 Jaén, Spain

Received 4 August 2005; accepted 21 December 2005
Communicated by G. Nürnberger
Available online 10 February 2006

Abstract

Let \( h, 1 < p < \infty \), be the best \( \ell_p \)-approximation of the element \( h \in \mathbb{R}^n \) from a proper affine subspace \( K \) of \( \mathbb{R}^n \), \( h \notin K \), and let \( h^\ast_{\infty} \) denote the strict uniform approximation of \( h \) from \( K \). We prove that there are a vector \( \alpha \in \mathbb{R}^n \setminus \{0\} \) and a real number \( a, 0 \leq a \leq 1 \), such that

\[
h_p = h^\ast_{\infty} + \frac{a^p}{p-1} \alpha + \gamma_p,
\]

for all \( p > 1 \), where \( \gamma_p \in \mathbb{R}^n \) with \( \|\gamma_p\| = o \left( \frac{a^p}{p} \right) \).

Keywords: Strict best approximation; Rate of convergence; Polya algorithm; Asymptotic expansion

1. Introduction

For \( 1 \leq p \leq \infty \), we consider the linear space \( \mathbb{R}^n \) endowed with the usual \( p \)-norm. For convenience we will denote \( \|\cdot\| := \|\cdot\|_{\infty} \). Also we will use the functional notation \( x = (x(1), x(2), \ldots, x(n)) \) to denote the element \( x \in \mathbb{R}^n \).

Let \( K \neq \emptyset \) be a subset of \( \mathbb{R}^n \). For \( h \in \mathbb{R}^n \setminus K \) and \( 1 \leq p \leq \infty \) we say that \( h_p \in K \) is a best \( \ell_p \)-approximation of \( h \) from \( K \) if

\[
\|h_p - h\|_p \leq \|f - h\|_p \quad \text{for all } f \in K.
\]

∗ Partially supported by Junta de Andalucía, Research Groups FQM268, FQM178 and by Ministerio de Ciencia y Tecnología, Project BFM2003-05794.

* Corresponding author. Fax: 34 953212141.
E-mail addresses: jquesada@ujaen.es (J.M. Quesada), jmmoreno@ujaen.es (J. Martínez-Moreno).
Throughout this paper, $K$ will denote a proper affine subspace of $\mathbb{R}^n$ and, without loss of generality, we will assume that $h = 0$ and $0 \notin K$. In this context, the existence of $h_p$ is guaranteed. Moreover, there exists a unique best $\ell_p$-approximation if $1 < p < \infty$. In the case $p = \infty$ we will say that $h_\infty$ is a best uniform approximation of 0 from $K$. In general, the unicity of the best uniform approximation is not guaranteed. However, a unique "strict uniform approximation", $h_\infty^*$, can be defined [3,8]. For convenience we will write $K = h_\infty^* + V$, where $V$ is a proper linear subspace of $\mathbb{R}^n$. It is well known (see for instance [9]) that $h_p$, $1 < p < \infty$, is the best $\ell_p$-approximation of 0 from $K$ if and only if
\[
\sum_{j=1}^n v(j) |h_p(j)|^{p-1} \, \text{sgn}(h_p(j)) = 0 \quad \text{for all } v \in V.
\] (1)
It is also known, [1,4,8], that $\lim_{p \to \infty} h_p = h_\infty^*$. This convergence is called Polya algorithm and occurs at a rate no worse than $1/p$ (see [2,4]). In [6] it is proved that for all $r \in \mathbb{N}$ there are $\alpha \in V$, $1 \leq \ell \leq r$, such that
\[
h_p = h_\infty^* + \frac{\alpha_1}{p-1} + \frac{\alpha_2}{(p-1)^2} + \cdots + \frac{\alpha_r}{(p-1)^r} + \gamma_p^{(r)},
\] (2)
where $\gamma_p^{(r)} \in \mathbb{R}^n$ and $\|\gamma_p^{(r)}\| = O(p^{-r-1})$. In [4] the authors give a necessary and sufficient condition on $K$ for
\[
p \|h_p - h_\infty^*\| \to 0 \quad \text{as } p \to \infty
\] (3)
and in [7] it is proved that if (3) holds then there are real numbers $a$, $L_1$ and $L_2$, with $0 \leq a < 1$ and $L_1, L_2 > 0$, such that
\[
L_1 a^p \leq p \|h_p - h_\infty^*\| \leq L_2 a^p
\] (4)
for all $p \geq 1$. In particular, (4) implies that if (3) holds, then we have an exponential rate of convergence of $h_p$ to $h_\infty^*$ as $p \to \infty$ and so the asymptotic expansion in (2) follows immediately with $\alpha_l = 0$, $1 \leq l \leq r$, for all $r \in \mathbb{N}$. The aim of this paper is to complete the results in [4,6,7] giving an asymptotic expression of $h_p$ in the general case. More precisely, we prove that there does exist a vector $\alpha \in V$, $\alpha \neq 0$, such that
\[
h_p = h_\infty^* + \frac{a^p}{p-1} \alpha + \gamma_p,
\] (5)
where $\gamma_p \in V$ and $\|\gamma_p\| = o(a^p/p)$. 

In the case $0 < a < 1$, taking into account (4), we immediately deduce that $p \|h_p - h_\infty^*\|/a^p$ is bounded. However, it is not a trivial question to show that the limit $p(h_p - h_\infty^*)/a^p$ exists as $p \to \infty$. This justifies the present paper.

On the other hand, since there is trivially an expression of the form (5) for some $\alpha$ and $\gamma_p$ in $V$, the only part requiring proof is the error estimate for $\gamma_p$. Also observe that (5) is a particular case of (2) for $a = 1$. However, in the case $0 \leq a < 1$ expression (5) is specially interesting because (2) does not give any information about $h_p$.

2. Notation and preliminary results

Without loss of generality, we will assume that $\|h_\infty^*\| = 1$, $h_\infty^*(j) \geq 0$, $1 \leq j \leq n$, and that the coordinates of $h_\infty^*$ are in decreasing ordering. Let $1 = d_1 > d_2 > \cdots > d_s \geq 0$ denote all the
different values of \( h^*_\infty(j) \), \( 1 \leq j \leq n \), and let \( \{ J_l \}_{l=1}^s \) be the partition of \( J := \{ 1, 2, \ldots, n \} \) defined by \( J_l := \{ j \in J : h^*_\infty(j) = d_l \} \), \( 1 \leq l \leq s \).

If \( J' \subseteq J \) we will denote by \( \| \cdot \|_{J'} \) the restriction of the norm \( \| \cdot \| \) to the set of indices on \( J' \).

Note that it is possible to choose a basis \( B = \{ v_1, v_2, \ldots, v_m \} \) of \( V \) and a partition \( \{ I_k \}_{k=1}^s \) of \( I := \{ 1, 2, \ldots, m \} \) such that for all \( i \in I_k, 1 \leq k \leq s \),

\[(p1) \quad v_i(j) = 0, \quad \forall \ j \in J_l, \ 1 \leq l < k, \]
\[(p2) \quad v_i(j) \neq 0 \text{ for some } j \in J_k. \]

Note that \( I_k \) can be empty for some \( k, 1 \leq k \leq s \). However, as we will notice later, the case \( d_s > 0 \) or \( I_s = \emptyset \) simplify the proof of the results in this paper. For this reason, and to consider the more general situation, we will assume that \( d_s = 0 \) and \( I_s \neq \emptyset \). We will use the following results.

**Theorem 1 (Quesada [7, Corollary 1]).** Let

\[
a = \max_{1 \leq l, k \leq s-1} \left\{ d_l/d_k : \sum_{j \in J_l} v_i(j) \neq 0 \text{ for some } i \in I_k \right\}, \tag{6}
\]

where \( a \) is assumed to be 0 if \( \sum_{j \in J_l} v_i(j) = 0 \) for all \( i \in I_k, 1 \leq k, l \leq s - 1 \). Then there are \( L_1, L_2 > 0 \) such that

\[
L_1 a^p \leq p \| h_p - h^*_\infty \| \leq L_2 a^p \quad \text{for all } p \geq 1. \tag{7}
\]

**Lemma 2.** If \( \{ x_p \} \) is a sequence of real numbers such that \( x_p \to 0 \) as \( p \to \infty \), then

\[
\left( 1 + \frac{x_p}{p} \right)^p = 1 + x_p + R_p,
\]

where \( R_p = O(x_p^2) \).

**Proof.** The proof follows immediately from the application of the Taylor’s formula to the function \( \varphi(z) = (1 + z/p)^p \) at \( z = 0 \). \( \square \)

**3. Asymptotic expression of the best \( \ell_p \)-approximations**

Since \( h_p \to h^*_\infty \) as \( p \to \infty \), then \( h_p(j) > 0 \) for all \( j \in J_l, 1 \leq l \leq s - 1 \), and \( p \) large enough. So, without loss of generality, we will assume that \( h_p(j) > 0 \) for all \( j \in J_l, 1 \leq j \leq s - 1 \).

**Theorem 3.** Let \( K \) be a proper affine subspace of \( \mathbb{R}^n \), \( 0 \notin K \). For \( 1 < p < \infty \), let \( h_p \) denote the best \( \ell_p \)-approximation of 0 from \( K \) and let \( h^*_\infty \) be the strict uniform approximation. Let \( a \) be the real number defined in (6). Then there is a vector \( x \in \mathbb{R}^n \), \( x \neq 0 \), such that

\[
h_p = h^*_\infty + \frac{a^p}{p-1} x + \gamma_p, \tag{8}
\]

where \( \gamma_p \in \mathbb{R}^n \) and \( \| \gamma_p \| = o(a^p/p) \).

**Proof.** Write \( K = h^*_\infty + \mathcal{V} \), where \( \mathcal{V} \) is a proper linear subspace of \( \mathbb{R}^n \), and consider the basis \( B = \{ v_1, v_2, \ldots, v_m \} \) defined as above. By the conditions (p1) and (p2) and the definition of \( a \) in (6), we have \( 0 \leq a \leq 1 \). We will consider three cases.
(a) If \( a = 0 \), then by (7), \( h_p = h^*_\infty \) for all \( p \geq 1 \) and (8) follows trivially for all \( x \in \mathbb{R}^n \) and \( \gamma_p = 0 \in \mathbb{R}^n \).

(b) If \( a = 1 \), then (8) is a particular case of (2), with \( \gamma_p = O(1/p^2) \). So, to conclude the proof, we only need to prove that \( x \neq 0 \). Indeed, since \( a = 1 \), there exist \( k \in \{1, 2, \ldots, s-1\} \) and \( i \in I_k \) such that \( \sum_{j \in I_k} v_i(j) \neq 0 \). Applying (1) with \( v = v_i \) we have

\[
\sum_{j \in J} v_i(j)|h_p(j)|^{p-1} \text{sgn}(h_p(j)) = 0.
\]

Since \( v_i \in I_k \), taking into account (p1) and (8) the above equation can be written as

\[
\sum_{j \in I_k} v_i(j) \left( d_k + \frac{\gamma_p(j)}{p-1} \right)^{p-1} + \sum_{l>k \in J} \sum_{j \in I_l} v_i(j) |h_p(j)|^{p-1} \text{sgn}(h_p(j)) = 0.
\]

Dividing by \( d_k^{p-1} \) and letting \( p \to \infty \) we obtain \( \sum_{j \in I_k} v_i(j) e^{\gamma(j)/d_k} = 0 \) and hence \( \gamma(j) \neq 0 \) for some \( j \in I_k \).

(c) If \( 0 < a < 1 \), then \( a = d_{l_0}/d_{l_0} \) for some \( 1 \leq k_0 < l_0 < s \) and \( \sum_{j \in I_{l_0}} v_i(j) \neq 0 \) for some \( i \in I_{l_0} \). On the other hand, by Theorem 1, \( p \|h_p - h^*_\infty\|/a^p \) is bounded. So, we can take a subsequence \( p_k \to \infty \) such that \( (p_k - 1)(h_{p_k} - h^*_\infty)/a^{p_k} \) converges. Define

\[
x := \lim_{k \to \infty} (p_k - 1)(h_{p_k} - h^*_\infty)/a^{p_k} \in \mathcal{V}.
\]

By (7) \( x \neq 0 \). Then we can write

\[
h_p = h^*_\infty + \frac{a^p}{p-1} x + \gamma_p,
\]

where \( \gamma_p := h_p - h^*_\infty - ax^p/(p-1) \in \mathcal{V} \). Note that \( p \|\gamma_p\|/a^p \) is also bounded and \( p_k \|\gamma_{p_k}\|/a^{p_k} \to 0 \) as \( k \to \infty \). Now we prove that \( \gamma_p = o(a^{p_k}/p) \). Indeed, suppose to the contrary that there exists a subsequence \( p_k' \to \infty \) such that \( (p_k' - 1)\gamma_{p_k'}/a^{p_k'} \to u \neq 0 \). Since \( u \in \mathcal{V} \), applying (1) with \( v = u \) we have

\[
\sum_{j \in J} u(j)|h_p(j)|^{p-1} \text{sgn}(h_p(j)) = 0.
\]

Let \( r_0 = \min \{l \in \{1, 2, \ldots, s\} : u(j) \neq 0 \text{ for some } j \in J_l \} \). Note that \( u \in \text{span}\{v_i : i \in I_{k_0} \} \). Now we consider two cases:

(c.1) If \( 1 \leq r_0 \leq s - 1 \), then dividing (10) by \( d^{p-1}_{r_0} \) and keeping in mind (9) we obtain

\[
\sum_{j \in J_{r_0}} u(j) \left( 1 + \frac{\gamma(j)}{d_{r_0}} \right)^{p-1} + \sum_{l=r_0+1}^s \left( \frac{d_l}{d_{r_0}} \right)^{p-1} \sum_{j \in J_l} u(j) \left( 1 + \frac{\gamma(j)}{d_l} \right)^{p-1} + \sum_{j \in J_s} u(j) \left| \frac{h_p(j)}{d_{r_0}} \right|^{p-1} \text{sgn}(h_p(j)) = 0.
\]
Now, applying Lemma 2 we get

\[
\sum_{j \in J_{r_0}} u(j) \left( 1 + \frac{a(j)}{d_{r_0}} a^p + (p - 1) \frac{\gamma_p(j)}{d_{r_0}} + R_p(j) \right) \\
+ \sum_{l = r_0 + 1}^{s-1} \left( \frac{d_l}{d_{r_0}} \right)^{p-1} \sum_{j \in J_l} u(j) \left( 1 + \frac{a(j)}{d_l} a^p + (p - 1) \frac{\gamma_p(j)}{d_l} + R_p(j) \right) \\
+ \sum_{j \in J_s} u(j) \left| \frac{h_p(j)}{d_{r_0}} \right|^{p-1} \operatorname{sgn}(h_p(j)) = 0,
\]

where \( R_p(j) = \mathcal{O}(a^{2p}) \) for all \( j \in J_1, r_0 \leq l \leq s - 1 \).

If \( r_0 \leq l \leq s - 1 \) and \( \frac{a_l}{d_{r_0}} > a \), then \( \frac{a_l}{d_l} > a \) for all \( r \geq r_0 \). Hence, from the definition of \( a \), \( \sum_{j \in J_l} v_i(j) = 0 \) for all \( i \in I_r \) with \( r \geq r_0 \) and hence \( \sum_{j \in J_l} u(j) = 0 \). So, dividing by \( a^p \) and rearranging terms we can write the equality above as

\[
\frac{1}{d_{r_0}} \sum_{j \in J_{r_0}} u(j) \frac{a(j)}{d_{r_0}} + \frac{1}{a} \sum_{j \in J_{r_0}} u(j) + \frac{p - 1}{a^p d_{r_0}} \sum_{j \in J_{r_0}} u(j) \frac{\gamma_p(j)}{a^p} + \tilde{R}_p = 0. \tag{11}
\]

where \( \tilde{R}_p = \mathcal{O}(a^p) \) and \( l_0 \) is the possible index in \( \{r_0 + 1, \ldots, s - 1\} \) such that \( \frac{a_l}{d_{r_0}} = a \).

Particularizing (11) for \( p = p_k \) and taking limits as \( k \to \infty \), we have

\[
\frac{1}{d_{r_0}} \sum_{j \in J_{r_0}} u(j) \frac{a(j)}{d_{r_0}} + \frac{1}{a} \sum_{j \in J_{r_0}} u(j) = 0.
\]

In similar way, letting \( k \to \infty \) in (11) with \( p = p_k' \) and taking into account the equality above, we obtain

\[
\sum_{j \in J_{r_0}} u(j)^2 = 0.
\]

A contradiction.

(c.2) If \( r_0 = s \), then multiplying (10) by \( (p - 1)^{p-1}/a^{p(p-1)} \) we get

\[
\sum_{j \in \hat{J}_s} u(j) \left| \frac{a(j)}{a^p} \right|^{p-1} \operatorname{sgn} \left( \frac{a(j)}{a^p} \right) + \sum_{j \in J_0^s} u(j) \left| \frac{\gamma_p(j)}{a^p} \right|^{p-1} \operatorname{sgn}(\gamma_p(j)) = 0, \tag{12}
\]

where \( \hat{J}_s := \{ j \in J_s : u(j) \frac{a(j)}{a^p} \neq 0 \} \) and \( J_0^s = J_s \setminus \hat{J}_s \).

Note that \( J_s \neq \emptyset \). Otherwise, particularizing (12) for \( p = p_k' \) we get a contradiction for \( k \) large enough because \( \operatorname{sgn}(\gamma_p(j)) = \operatorname{sgn}(u(j)) \) if \( u(j) \neq 0 \).
Let $\hat{J}_s := \{ j \in \hat{J} : |x(j)| = \|x\|_{\hat{J}_s} \}$. Dividing (12) for $\|x\|_{\hat{J}_s}^{p-1}$, we have
\[
\sum_{j \in \hat{J}_s} u(j) \left| 1 + \frac{(p - 1)\gamma_p(j)}{a^p x(j)} \right|^{p-1} \text{sgn} \left( x(j) + \frac{(p - 1)\gamma_p(j)}{a^p} \right)
+ \sum_{j \in \hat{J} \setminus \hat{J}_s} u(j) \left( \frac{x(j)}{\|x\|_{\hat{J}_s}} \right)^{p-1} \left| 1 + \frac{(p - 1)\gamma_p(j)}{a^p x(j)} \right|^{p-1} \text{sgn} \left( x(j) + \frac{(p - 1)\gamma_p(j)}{a^p} \right)
+ \sum_{j \in J^0} u(j) \left( \frac{(p - 1)\gamma_p(j)}{a^p \|x\|_{\hat{J}_s}} \right)^{p-1} \text{sgn}(\gamma_p(j)) = 0.
\] (13)

Since $(p_k - 1)\gamma_{p_k}(j)/a^{p_k} \to 0$ as $k \to \infty$, there exists a real number $\beta$, $0 < \beta < 1$, such that for $k$ large enough
\[
\min_{j \in \hat{J}_s} \left| 1 + \frac{(p_k - 1)\gamma_{p_k}(j)}{a^{p_k} x(j)} \right| > \beta > \max_{j \in \hat{J} \setminus \hat{J}_s} \frac{x(j)}{\|x\|_{\hat{J}_s}} \left| 1 + \frac{(p_k - 1)\gamma_{p_k}(j)}{a^{p_k} x(j)} \right|.
\]

Now, taking into account that
\[
- \sum_{j \in \hat{J}_s \setminus \hat{J}_s^2} \frac{|u(j)|}{\beta^{p_k-1}} \left| 1 + \frac{(p_k - 1)\gamma_{p_k}(j)}{a^{p_k} x(j)} \right|^{p_k-1} < - \sum_{j \in \hat{J}_s} |u(j)|
\]
and
\[
\lim_{k \to \infty} \frac{|u(j)|}{\beta^{p_k-1}} \left( \frac{x(j)}{\|x\|_{\hat{J}_s}} \right)^{p_k-1} \left| 1 + \frac{(p_k - 1)\gamma_{p_k}(j)}{a^{p_k} x(j)} \right|^{p_k-1} = 0
\]
for all $j \in \hat{J}_s \setminus \hat{J}_s^2$, we deduce from (13) that there exists $j_0 \in \hat{J}_s^2$ such that $u(j_0)x(j_0) > 0$. But, if $j \in \hat{J}_s$ and $u(j)x(j) > 0$, then
\[
\lim_{k \to \infty} \left| 1 + \frac{(p_k' - 1)\gamma_{p_k'}(j)}{a^{p_k'} x(j)} \right| = 1 + \frac{|u(j)|}{x(j)} > 1
\]
and $\text{sgn} \left( x(j) + (p_k' - 1)\gamma_{p_k'}(j)/a^{p_k'} \right) = \text{sgn}(u(j))$ for $k$ large enough.

On the other hand, if $j \in \hat{J}_s$ with $u(j)x(j) < 0$ and $|u(j)| \leq |x(j)|$ then
\[
\lim_{k \to \infty} \left| 1 + \frac{(p_k' - 1)\gamma_{p_k'}(j)}{a^{p_k'} x(j)} \right| = \left| 1 + \frac{|u(j)|}{x(j)} \right| < 1.
\]

Finally, if $j \in \hat{J}_s$ with $u(j)x(j) < 0$ and $|u(j)| > |x(j)|$ then
\[
\text{sgn} \left( x(j) + (p_k' - 1)\gamma_{p_k'}(j)/a^{p_k'} \right) = \text{sgn}(u(j)).
\]

So, taking limits in (13) as $k \to \infty$, with $p = p_k'$, we get a contradiction. \[ \square \]

**Remark 4.** Recently, in [5] it is proved that estimation (4) of the order of convergence of the Polya algorithm also holds if $K$ is a finite affine subspace of $\ell_1(\mathbb{N})$. A slight modification of the techniques used in this paper shows that Theorem 3 is also valid in this new context.
References