Rate of Convergence of the Linear Discrete Polya Algorithm

J. M. Quesada and J. Navas

Departamento de Matemáticas, Universidad de Jaén, Paraje las Lagunillas, Campus Universitario, 23071 Jaén, Spain
E-mail: jquesada@ujaen.es, jnavas@ujaen.es

Communicated by E. W. Cheney
Received May 23, 2000; accepted November 6, 2000; published online March 16, 2001

In this paper we consider the problem of best approximation in \( l^p \), \( 1 < p \leq \infty \). If \( h_p \), \( 1 < p < \infty \), denotes the best \( p \)-approximation of the element \( h \in \mathbb{R}^n \) from a proper affine subspace \( K \) of \( \mathbb{R}^n \), then \( \lim_{p \to 1} h_p = h^* \), where \( h^*_p \) is a best uniform approximation of \( h \) from \( K \), the so-called strict uniform approximation. Our aim is to give a complete description of the rate of convergence of \( |h_p - h^*_p| \) as \( p \to \infty \).

1. INTRODUCTION

For \( x = (x(1), x(2), \ldots, x(n)) \in \mathbb{R}^n \), the \( \ell_p \)-norms, \( 1 \leq p \leq \infty \), are defined by
\[
\|x\|_p = \left( \sum_{j=1}^n |x(j)|^p \right)^{1/p}, \quad 1 \leq p < \infty,
\]
\[
\|x\| := \|x\|_\infty = \max_{1 \leq j \leq n} |x(j)|.
\]

Let \( K \neq \emptyset \) be a closed 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 \( p \)-approximation of \( h \) from \( K \) if
\[
\|h_p - h\|_p \leq \|f - h\|_p, \quad \forall f \in K.
\]
The existence of \( h_p \) is a well known fact. Moreover, if \( K \) is in addition a convex set and \( 1 < p < \infty \), then there is an unique best \( p \)-approximation.

\(^1\)This work was partially supported by Junta de Andalucía, Research Group 0268.

Throughout this paper we assume that $K$ is a proper affine subspace of $\mathbb{R}^n$. In this case, it is known (see for instance [6]) that $h_p$, $1 < p < \infty$, is a best $p$-approximation of $h$ from $K$ if and only if

$$\sum_{j=1}^{n} (h_p(j) - f(j)) |h_p(j) - h(j)|^{p-1} \text{sgn}(h_p(j) - h(j)) = 0, \quad \forall f \in K.$$ (1)

If $p = \infty$ we call $h_\infty$ a best uniform approximation of $h$ 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]. The Polya algorithm is an attempt to define $h_\infty^*$ as the limit of the best $p$-approximation $h_p$ as $p \to \infty$. If $K$ is an affine subspace of $\mathbb{R}^n$, then the Polya algorithm converges to the strict uniform approximation [1, 4, 5].

$$\lim_{p \to \infty} h_p = h_\infty^*.$$ (2)

The strict uniform approximation also verifies the next property. Let $H$ denote the set of the best uniform approximations of $h$ from $K$. For every $g \in H$ we consider the vector $\tau(g)$ whose coordinates are given by $|g(j) - h(j)|$, $j = 1, 2, \ldots, n$, arranged in decreasing order. Then $\tau(h_\infty^*)$ is the only one that gives a minimal lexicographic ordering.

In [2, 4] it is proved that the convergence of $h_p$ to $h_\infty^*$ occurs at a rate no worse than $1/p$. In [4] the authors give a necessary and sufficient condition on $K$ for $h_p$ to coincide with $h_\infty^*$ for $p$ large, and also a necessary and sufficient condition for $p \|h_p - h_\infty^*\| = 0$ as $p \to \infty$. (2)

The aim of this paper is to give a detailed description of the rate of convergence of the Polya algorithm; more precisely, we prove that if (2) holds then there is a number $0 < a < 1$ such that $p \|h_p - h_\infty^*\|/a^p$ is bounded.

2. NOTATION AND PRELIMINARY RESULTS

Without loss of generality we may assume that $h = 0$ and $h_\infty^*(j) \geq 0$, $1 \leq j \leq n$, and that the coordinates of $h_\infty^*$ are in decreasing ordering. Let $1 = p_1 > p_2 > \cdots > p_s \geq 0$ denote all the different values of $h_\infty^*(j)$, $1 \leq j \leq n$, and $\{J_l\}_{l=1}^{s}$ the partition of $J := \{1, 2, \ldots, n\}$ defined by $J_l := \{j \in J : h_\infty^*(j) = p_l\}$, $1 \leq l \leq s$. We henceforth put $r = s$ if $p_s > 0$ and $r = s - 1$ if $p_s = 0$. 

110 QUESADA AND NAVAS
We can write $K = h^*_m + V$, where $V$ is a proper linear subspace of $\mathbb{R}^n$. Note that it is possible to choose a basis $B = \{ v_1, v_2, ..., v_m \}$ of $V$ and a partition $\{I_k\}_{k=1}^s$ of $I := \{1, 2, ..., m\}$ such that for all $i \in I_k$, $1 \leq k \leq s$,

\begin{align*}
(p1) & \quad v_i(j) = 0, \quad \forall j \in J_i, \quad 1 \leq l < k, \\
(p2) & \quad v_i(j) \neq 0 \text{ for some } j \in J_k.
\end{align*}

Denote $n_l = \text{card}(J_l)$ and $m_k = \text{card}(I_k)$. We have $m_k < n_k$; otherwise we can take a linear combination of the vectors $v_i$, $i \in I_k$, in such way that the definition of $h^*_m$ is contradicted. In this partition it is possible that $I_k = \emptyset$ for some $k$, $1 \leq k \leq s$. However, for simplicity of notation, we suppose that $I_k \neq \emptyset$ for $1 \leq k \leq r$, this involves no loss of generality. In order to get our main theorem, we use the following result [4].

**Theorem 1.** In the above conditions, $p \| h - h^*_m \| \to 0$ as $p \to \infty$ if and only if

$$
\sum_{j \in J_k} v_i(j) = 0, \quad \forall i \in I_k, \quad 1 \leq k \leq r.
$$

Moreover, $h = h^*_m$ for $p$ large if and only if

$$
\sum_{j \in J_k} v_i(j) = 0,
$$

for all $i \in I_k$, $1 \leq k \leq r$, and for all $1 \leq l \leq r$.

The following notation will be used throughout the paper. For $k, l = 1, 2, ..., r$, we consider the matrices $M_{kl} = (v_i(j))_{i \in I_k, j \in J_l}$, and we put $I_0 = \bigcup_{k=1}^s I_k$, $m_0 = \text{card}(I_0)$, $J_0 = \bigcup_{l=1}^r J_l$, $n_0 = \text{card}(J_0)$. Finally, if $A$ is a matrix then we denote by $A^T$ the transpose matrix of $A$ and by $\| A \|$ the row-sum norm of $A$.

**Lemma 1.** Let $\{ x_p \}$ be a sequence of vectors in $\mathbb{R}^m \setminus \{0\}$ such that $p \| x_p \| \to 0$ as $p \to \infty$. Then, for a fixed vector $b \in \mathbb{R}^m$ and for all $\beta > 0$,

$$
\left( \beta + \sum_{j=1}^m b(j) x_p(j) \right)^p = \beta^p + \beta^{p-1} p \sum_{j=1}^m b(j) x_p(j) + \beta^{p-2} R(p),
$$

where $R(p) = o(p \| x_p \|)$.

**Proof.** The proof follows immediately from the application of Taylor’s formula to the function $\varphi(x) = (1 + x)^p$ at $x = 0$. \qed
3. RATE OF CONVERGENCE

Theorem 2. Let $K$ be a proper affine subspace of $\mathbb{R}^n$, $0 \notin K$. For $1 < p < \infty$, let $h_p$ denote the best $p$-approximation of $0$ from $K$ and let $h^*_p$ be the strict uniform approximation. Suppose that $p \|h_p - h^*_p\| \to 0$ as $p \to \infty$. Then there are $L_1, L_2 > 0$ such that

$$L_1 a \leq p \|h_p - h^*_p\| \leq L_2 a^p,$$

(3)

where

$$a = \max_{1 \leq i, l \leq r} \left\{ \frac{\rho_l}{\rho_k} \sum_{j \in I_k} v_i(j) \neq 0 \text{ for some } i \in I_k \right\}$$

(4)

and $a$ is assumed to be $0$ if $\sum_{j \in I_k} v_i(j) = 0$ for all $i \in I_k$, $1 \leq k, l \leq r$.

Proof. Let $B = \{v_1, v_2, \ldots, v_m\}$ and $I_k, 1 \leq k \leq s$, as above. If $h_p = h^*_p$ for $p$ large then, by Theorem 1, $a = 0$ and (3) holds. Therefore we assume $h_p \neq h^*_p$ for $p$ large. This condition is just equivalent to

$$\max_{1 \leq k, l \leq r} \left| \sum_{j \in I_k} v_i(j) \right| > 0.\quad (5)$$

Putting $f = h_p + v_i, i \in I_0$, in (1) we have, for $p$ large,

$$\sum_{j \in I_0} v_i(j) h_p(j)^{p-1} + \sum_{j \in I_{r+1}} v_i(j) |h_p(j)|^{p-1} \text{sgn}(h_p(j)) = 0, \quad \forall i \in I_0.\quad (6)$$

This non-linear system can be written as

$$MH_p^T + NK_p^T = 0,\quad (7)$$

where $M$ and $N$ are the matrices defined by

$$M = (v_i(j)), (i, j) \in I_0 \times I_r, \quad N = (v_i(j)), (i, j) \in I_0 \times I_{r+1}$$

and $H_p, K_p$ denote the vectors in $\mathbb{R}^{n_0}$ and $\mathbb{R}^{n_{r+1}}$, respectively, whose components are given by

$$H_p(j) = h_p(j)^{p-1}, \quad 1 \leq j \leq n_0$$

$$K_p(j) = |h_p(n_0 + j)|^{p-1} \text{sgn}(h_p(n_0 + j)), \quad 1 \leq j \leq n_{r+1}.$$ 

If $r = s$ then $NK_p^T$ is assumed to be the null vector in $\mathbb{R}^{n_0}$. 

112 QUESADA AND NAVAS
Let \( \hat{\lambda}_p = (\hat{\lambda}_p(1), \hat{\lambda}_p(2), \ldots, \hat{\lambda}_p(m)) \) be the vector in \( \mathbb{R}^m \) such that

\[
h_p = h_{\infty}^* + \sum_{k=1}^{m} \hat{\lambda}_p(k) v_k.
\]

Note that the condition \( p \| h_p - h_{\infty}^* \| \to 0 \) as \( p \to \infty \) is equivalent to \( p \| \hat{\lambda}_p \| \to 0 \) as \( p \to \infty \).

If \( j \in J_p \), \( 1 \leq l \leq r \), then from Lemma 1,

\[
h_p(j)^{p-1} = \left( h_{\infty}^*(j) + \sum_{k=1}^{m} \hat{\lambda}_p(k) v_k(j) \right)^{p-1}
\]

\[
= \rho \delta_p^{p-1} + \rho \delta_p^{p-2} (p-1) \sum_{k=1}^{m} \hat{\lambda}_p(k) v_k(j) + \rho \delta_p^{p-3} R_p(j),
\]

with \( R_p(j) = o(p \| A_p \|) \), where \( A_p := (\hat{\lambda}_p(1), \ldots, \hat{\lambda}_p(m)) \). Thus we can express the vector \( H_p^T \) like

\[
H_p^T = A_p^T \delta_p^{p-1} \theta^T + (p-1) A_p^T \delta_p^{p-2} M^T A_p^T + A_p^T \delta_p^{p-3} R_p^T,
\]

where \( \theta := (1, 1, \ldots, 1) \in \mathbb{R}^n \), \( \delta_p := (R_p(1), \ldots, R_p(n_0)) \) and \( \delta_p := (\delta_p(j)_{(i),j})_{(n_0 \times n_0)} \) is the diagonal matrix of order \( n_0 \times n_0 \) such that \( \delta_p(j)_{(i),j} = \rho \) if \( j \in J_p \), \( 1 \leq l \leq r \).

Substituting in (7) we obtain the system

\[
M(\delta_p^{p-1} \theta^T + (p-1) \delta_p^{p-2} M^T \theta^T + \delta_p^{p-3} R_p^T) + N K_p^T = 0. \tag{8}
\]

Let \( A_{n_0} = (\delta_p(j)_{(i),j})_{n_0 \times n_0} \) be the diagonal matrix of order \( m_0 \times m_0 \) such that \( \delta_p(j)_{(i),j} = \rho \) if \( i \in L_k \), \( 1 \leq k \leq r \). Multiplying (8) by \( A_{n_0}^{p+2} := (A_{n_0}^{-1})^{p-2} \) we have

\[
(p-1) A_{n_0}^{p+2} M A_{n_0}^{p-2} M^T A_{n_0}^T
\]

\[
= -A_{n_0}^{p+2} M A_{n_0}^{p-1} \theta^T - A_{n_0}^{p+2} M A_{n_0}^{p-2} R_p^T - A_{n_0}^{p+2} N K_p^T. \tag{9}
\]

Observe that the multiplication by \( A_{n_0}^{p+2} \) is equivalent to divide by \( p^{p-2} \) each of the equations in (6) obtained for \( i \in L_k \). This operation is justified because \( v_i(j) = 0 \) for all \( j \in J_i \) if \( j < k \). Next we study each of the terms in the former system. An easy computation shows that

\[
A(p) := A_{n_0}^{p+2} M A_{n_0}^{p-2} M^T = \begin{pmatrix}
A_{11}(p) & A_{12}(p) & \cdots & A_{1m}(p) \\
A_{21}(p) & A_{22}(p) & \cdots & A_{2m}(p) \\
\vdots & \vdots & \ddots & \vdots \\
A_{m1}(p) & A_{m2}(p) & \cdots & A_{mm}(p)
\end{pmatrix},
\]
where $A_{ij}(p)$, $i, j = 1, 2, \ldots, r$, is the matrix of order $m_i \times m_j$ given by

$$
A_{ij}(p) = \sum_{k=1}^{r} \left( \frac{p_k}{p_k} \right)^{p-2} M_{ik} M_{jk}^T.
$$

Thus,

$$
A_{ij} := \lim_{p \to \infty} A_{ij}(p) = M_{ii} M_{ji}^T.
$$

Since $M_{ji}$ is the null matrix if $j > i$, we conclude that

$$
A := \lim_{p \to \infty} A(p) = \begin{pmatrix}
M_{11} M_{11}^T & 0 & \cdots & 0 \\
M_{22} M_{12}^T & M_{22} M_{22}^T & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
M_{rr} M_{tr}^T & M_{rr} M_{2r}^T & \cdots & M_{rr} M_{rr}^T
\end{pmatrix}
$$

is a triangular matrix by blocks and so

$$
\det(A) = \prod_{i=1}^{r} \det(M_{ii} M_{ii}^T) \neq 0.
$$

In particular we have proved that the matrix $A(p)$ is non singular for $p$ large.

Analogously, denoting by $B_p = -A_{ik}^* p^{-2} M A_{ik}^{* -1} Y^T$ it is easy to check that

$$
B_p(i) = -\rho_k \sum_{l=1}^{r} \left( \frac{p_l}{p_k} \right)^{p-1} \sum_{j \not\in J_k} v_i(j), \quad i \in I_k, \quad 1 \leq k \leq r.
$$

Let $a$ defined by (4). From (5) and Theorem 1, we have $0 < a < 1$. Moreover, $b := \lim_{p \to \infty} \|B_p\|/\|p\| > 0$.

Similarly, writing $C_p = -A_{ik}^* p^{-2} M A_{ik}^{* -1} R_p^T$ we obtain

$$
C_p(i) = -\frac{1}{\rho_k} \sum_{l=1}^{r} \left( \frac{p_l}{p_k} \right)^{p-3} \sum_{j \not\in J_k} v_i(j) R_p(j), \quad i \in I_k, \quad 1 \leq k \leq r.
$$

Since $v_i(j) = 0$ for all $(i, j) \in I_k \times J_k$, $k \leq l$, and $R_p(j) = o(p \|A_p\|)$, it follows that

$$
\lim_{p \to \infty} \frac{\|C_p\|}{\|A_p\|} = 0.
$$

Finally, denoting $D_p = -A_{ik}^* p^{-2} N K_p^T$ we get immediately $\lim_{p \to \infty} \|D_p\|/\|p\| = 0$. 
With this new notation the system (9) can be written as

\[(p - 1) A(p) A_p^T = B_p + C_p + D_p,\]

and so

\[(p - 1) \|A_p\| = \|A(p)^{-1} (B_p + C_p + D_p)\| \leq \|A(p)^{-1}\| \|B_p\| + \|A(p)^{-1}\| \|C_p\| + \|A(p)^{-1}\| \|D_p\|.\]

Therefore,

\[(p - 1) \|A_p\| \left(1 - \frac{\|A(p)^{-1}\| \|C_p\|}{(p - 1) \|A_p\|}\right) \leq \|A(p)^{-1}\| \|B_p\| + \|A(p)^{-1}\| \|C_p\| + \|A(p)^{-1}\| \|D_p\|.\] (10)

Dividing (10) by \(a^p\) and taking limits as \(p \to \infty\) we have

\[\lim_{p \to \infty} \frac{p \|A_p\|}{a^p} \leq \|A^{-1}\| b.\]

In similar way,

\[\|B_p\| \leq (p - 1) \|A(p)\| \|A_p\| \left(1 + \frac{\|C_p\|}{(p - 1) \|A(p)\| \|A_p\|}\right) + \|D_p\|\]

and then

\[\lim_{p \to \infty} \frac{p \|A_p\|}{a^p} \geq \frac{b}{\|A\|}.\]

Finally, we conclude that

\[\frac{b}{\|A\|} \leq \lim_{p \to \infty} \frac{p \|A_p\|}{a^p} \leq b \|A^{-1}\|.\] (11)

If \(r = s\) or \(J_s = \emptyset\) then the proof is complete. In the other case we put

\[h_p = h^*_w + \sum_{i \in I} \lambda_p(i) v_i,\]

\[= h^*_w + \sum_{i \in I_p} \sigma_p(i) v_i + \sum_{i \in I_s} \lambda_s(i) v_i = h^*_w + u_p + w_p.\]
By (11), we have actually proved that \( p \| u_p \|/a^p \) is bounded. Our purpose is to prove that \( p \| w_p \|/a^p \) is also. Obviously, we need only consider the case \( w_p \neq 0 \) for \( p \) large. Taking \( f = h_p + w_p \) in (1) we obtain

\[
\sum_{j \in J_s} w_p(j) |u_p(j) + w_p(j)|^{p-1} \text{sgn}(u_p(j) + w_p(j)) = 0. \tag{12}
\]

If \( u_p(j) = 0 \) for all \( j \in J_s \), then by (12) \( w_p(j) = 0 \) for all \( j \in J_s \) and hence \( w_p = 0 \). Therefore, we can assume that for all \( p \geq 1 \), \( u_p(j) \neq 0 \) for some \( j \in J_s \) and \( w_p \neq 0 \). Under these conditions, let \( \beta = \inf \| u_p \|/\| w_p \| \). To conclude the proof we will prove that \( \beta > 0 \). Suppose \( \beta = 0 \). Then there exists a subsequence, \( p_k \to \infty \), such that \( \| u_p \|/\| w_p \| \to 0 \). Let \( J_s^{(1)} \) be the set of indices in \( J_s \) such that

\[
\lim_{k \to \infty} |w_p(j)|/\| w_p \| \neq 0.
\]

Note that \( J_s^{(1)} \neq \emptyset \). Multiplying (12) by \( 2^{p_k-1}/\| w_p \|^{p_k} \) we obtain, for \( k \) large,

\[
0 = \sum_{j \in J_s^{(1)}} |w_p(j)| |2u_p(j)| |2w_p(j)|^{p_k-1} \text{sgn}(u_p(j) + w_p(j))
+ \sum_{j \notin J_s^{(1)}} |w_p(j)| |2u_p(j)| |2w_p(j)|^{p_k-1} \text{sgn}(u_p(j) + w_p(j))
\geq \sum_{j \notin J_s^{(1)}} |w_p(j)| |2u_p(j)| |2w_p(j)|^{p_k-1} \text{sgn}(u_p(j) + w_p(j))
- \sum_{j \notin J_s^{(1)}} |w_p(j)| |2u_p(j)| |2w_p(j)|^{p_k-1} \text{sgn}(u_p(j) + w_p(j))
\]

Taking limits as \( k \to \infty \) we get a contradiction.

The following corollary summarizes the results in Theorems 1 and 2.

**Corollary 1.** Let \( K \) be a proper affine subspace of \( \mathbb{R}^n \), \( 0 \not\in K \). Write \( K = h_p^* + \gamma^* \), where \( \gamma^* \) is a proper linear subspace of \( \mathbb{R}^n \), and let \( B = \{ v_1, v_2, \ldots, v_m \} \) a basis of \( \gamma^* \) and \( \{ I_k \}_{k=1}^\infty \) a partition of \( \{ 1, 2, \ldots, m \} \) verifying (p1) and (p2). Then there exist \( L_1, L_2 > 0 \) such that

\[
L_1 a^p \leq p \| h_p - h_p^* \| \leq L_2 a^p, \tag{13}
\]
where

\[
a = \max_{1 \leq i, k \leq r} \left\{ \frac{p_i}{p_k} : \sum_{j \neq j_i} v_i(j) \neq 0 \text{ for some } i \in I_k \right\}
\]

and \(a\) is assumed to be 0 if \(\sum_{j \neq j_i} v_i(j) = 0\) for all \(i \in I_k\), \(1 \leq k \leq r\).

**Proof.** It is sufficient to note that \(a = 1\) if and only if \(j \notin Jl v_i(j) \neq 0\) for some \(i \in I_l\). In this case, \(p \|h_p - h_m^n\|\) does not converge to 0 as \(p \to \infty\). This condition is equivalent to \(h_p\) converging to \(h_m^n\) with rate exactly \(1/p\).

**A Numerical Example.** Consider \(h_m^n = (1, 1, 1, 1, 1, 1, 1, 1, 0, 0)\) and \(v = (1, -1, 0, 0, 1, -1, 0, 1, 0, 1, 2, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)\) and

\[v' = (1, -1, 0, 0, 1, -1, 0, 1, 0, 1, -1, 1, 2, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0).
\]

First, we construct the matrix

\[
\begin{bmatrix}
J_1 & J_2 & J_3 & J_4 \\
1 & 1 & 1 & 1 & 1/2 & 1/2 & 1/3 & 1/3 & 0 & 0 \\
I_1 & 1 & -1 & 0 & 0 & 1 & -1 & 2 & -1 & 0 & 1 & 1 & v_1 \\
I_2 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 2 & 0 & v_2 \\
I_3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & v_3 \\
I_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & v_5
\end{bmatrix}
\]

Note that in this example \(I_2 = \emptyset\). Since the sum of coordinates in the diagonals blocks, \((I_k J_k), k = 1, 2, 3,\) is zero, we deduce that \(p \|h_p - h_m^n\| \to 0\) as \(p \to \infty\). Next we construct the matrix \(Q = (q_{kl})_{k,l=1,2,3}, q_{kl} = \max_{i \in I_k} |\sum_{j \neq j_i} v_i(j)|\) (if \(k = 2\) we put \(q_{kl} = 0\)), and the matrix \(R = (p_i/p_k)_{k,l=1,2,3},\)

\[
\begin{bmatrix}
J_1 & J_2 & J_3 & J_4 \\
I_1 & 0 & 1 & 2 & 1/2 & 1/3 \\
I_2 & 0 & 0 & 0 & 2 & 1 & 2/3 \\
I_3 & 0 & 0 & 0 & 3 & 2/3 & 1
\end{bmatrix}
\]

By definition we obtain \(a = 1/2\) and we conclude that the rate of convergence of \(h_p\) to \(h_m^n\) is \(1/(p^{2^p})\).
Remarks. It is possible to obtain additional information for the rate of convergence for each of the blocks of coordinates \( h_p(j), j \in J_l, 1 \leq l \leq r \), by means of an inductive procedure. More precisely, considering only the equations in (6) for \( i \in I_1 \) we have

\[
\sum_{j \in J_1} v_i(j) h_p(j) - \sum_{j \in J_1} v_i(j) h_p(j) + \sum_{j \in J_1} v_i(j) |h_p(j)|^{-1} \text{sgn}(h_p(j)). \tag{14}
\]

For \( j \in J_1 \),

\[
h_p(j)^{p-1} = \left( 1 + \sum_{k \in I_1} \lambda_p(k) v_k(j) \right)^{p-1}.
\]

\[
= 1 + (p - 1) \sum_{k \in I_1} \lambda_p(k) v_k(j) + R_p(j),
\]

with \( R_p(j) = o(p \|A_p^{(1)}\|) \), where \( A_p^{(1)} = (\lambda_p(k))_{k \in I_1} \). Substituting in (14) and denoting by \( F_p(i) \) the second member in (14) we obtain the system

\[
(p-1) M_{11} M_{11}^T (A_p^{(1)})^T + M_{11} (R_p^{(1)})^T = F_p^T,
\]

where \( R_p^{(1)} = (R_p(j))_{j \in J_1} \). Since \( M_{11} M_{11}^T \) is non-singular and \( \|F_p^p\|/p_p^p \) is bounded, we conclude that the \( p \lambda_p(i), i \in I_1 \), converge to 0 at a rate no worse than \( p_p^p \). Taking into account this information and applying the same procedure to the equations in (6) for \( i \in I_2 \) we deduce that the rate of convergence of \( p \lambda_p(i) \), \( i \in I_2 \), is at least \( (p_1/p_2)^p \). Now, we can reiterate this argument for the others blocks of coordinates. Finally, this first estimation of the rate of convergence can be used to obtain an estimation more precise. The basic idea is to apply the same technique to the equations in (6) for \( i \in \bigcup_{l=1}^r I_1 \), \( k = 1, 2, \ldots, r \). Note that this inductive procedure supposes, in fact, another strategy to prove Theorem 1.

REFERENCES

