

CONVERGENCE THEOREMS FOR THE NELDER-MEAD METHOD

AURÉL GALÁNTAI

Óbuda University

1034 Budapest, Bécsi út 96/b, Hungary

galantai.aurel@nik.uni-obuda.hu

[Received: September 8, 2020, Accepted: October 14, 2020]

Dedicated to Professor Barna Szabó on the occasion of his 85th birthday

Abstract. We develop a matrix form of the Nelder-Mead method and after discussing the concept of convergence we prove a general convergence theorem. The new theorem is demonstrated in low dimensional spaces.

Mathematical Subject Classification: 65K10, 90C56

Keywords: Nelder-Mead simplex method, derivative-free algorithms, convergence

1. INTRODUCTION

We study the Nelder-Mead simplex method [1] for the solution of the unconstrained minimization problem

$$f(x) \rightarrow \min \quad (f : \mathbb{R}^n \rightarrow \mathbb{R}),$$

where f is continuous. Since 1965, the Nelder-Mead algorithm and its later variants have become highly popular in various application areas and derivative-free optimization [2–6]. Although the original paper [1] has a lot of citations (over 31 000 in Google Scholar on August 3, 2020), notably few theoretical results are known on the convergence (see Kelley [7] and Lagarias et al. [8, 9]).

In the paper we develop a matrix form of the Nelder-Mead method, discuss the concept of convergence and its consequences, prove a general convergence theorem under plausible assumptions and demonstrate the convergence of the algorithm for low dimensional spaces. This approach partially answers some of the questions raised by Wright [10] concerning the Nelder-Mead method.

Section 2 contains the description of the algorithm. The next section summarizes the most important earlier results on the convergence. The matrix reformulation of the Nelder-Mead method is given Section 4. The concept of convergence and some of its consequences are developed and discussed in Section 5. The spectra of the occurring matrices is investigated in Section 6. The general convergence theorem is developed in Section 7. Finally, the convergence of the Nelder-Mead method is demonstrated for $n = 1, 2, 3$ in the last section.

2. THE NELDER-MEAD SIMPLEX METHOD

We use the following form of the original method [8]. The vertices of the initial simplex are denoted by $x_1, x_2, \dots, x_{n+1} \in \mathbb{R}^n$. It is assumed that vertices x_1, \dots, x_{n+1} are ordered such that

$$f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1}) \quad (2.1)$$

and this condition is maintained during the iterations of the Nelder-Mead algorithm. Define $x_c = \frac{1}{n} \sum_{i=1}^n x_i$ and $x(\lambda) = (1 + \lambda)x_c - \lambda x_{n+1}$. The related evaluation points are

$$x_r = x(1), \quad x_e = x(2), \quad x_{oc} = x\left(\frac{1}{2}\right), \quad x_{ic} = x\left(-\frac{1}{2}\right).$$

Then one iteration step of the method is the following.

Operation	Nelder-Mead simplex method
0. Ordering	$f(x_1) \leq \dots \leq f(x_{n+1})$
1. Reflect	if $f(x_1) \leq f(x_r) < f(x_n)$, then $x_{n+1} \leftarrow x_r$ and goto 0.
2. Expand	if $f(x_r) < f(x_1)$ and $f(x_e) < f(x_r)$, then $x_{n+1} \leftarrow x_e$ and goto 0. If $f(x_e) \geq f(x_r)$, then $x_{n+1} \leftarrow x_r$ and goto 0.
3. Contract outside	If $f(x_n) \leq f(x_r) < f(x_{n+1})$ and $f(x_{oc}) \leq f(x_r)$, then $x_{n+1} \leftarrow x_{oc}$ and goto 0.
4. Contract inside	If $f(x_r) \geq f(x_{n+1})$ and $f(x_{ic}) < f(x_{n+1})$ then $x_{n+1} \leftarrow x_{ic}$ and goto 0.
5. Shrink	$x_i \leftarrow (x_i + x_1)/2$, $f(x_i)$ (for all i) and goto 0.

There are two rules that apply to reindexing after each iteration. If a nonshrink step occurs, then x_{n+1} is discarded and a new point $v \in \{x_r, x_e, x_{oc}, x_{ic}\}$ is accepted. The following cases are possible:

$$f(v) < f(x_1), \quad f(x_1) \leq f(v) \leq f(x_n), \quad f(v) < f(x_{n+1}).$$

If

$$j = \begin{cases} 1, & \text{if } f(v) < f(x_1) \\ \max_{2 \leq \ell \leq n+1} \{f(x_{\ell-1}) \leq f(v) \leq f(x_\ell)\}, & \text{otherwise} \end{cases}.$$

then the new simplex vertices are

$$x_i^{new} = x_i \quad (1 \leq i \leq j-1), \quad x_j^{new} = v, \quad x_i^{new} = x_{i-1} \quad (i = j+1, \dots, n+1). \quad (2.2)$$

This rule inserts v into the ordering with the highest possible index. If shrinking occurs, then

$$x'_1 = x_1, \quad x'_i = (x_i + x_1)/2 \quad (i = 2, \dots, n+1)$$

plus a reordering takes place. By convention, if $f(x'_1) \leq f(x'_i)$ ($i = 2, \dots, n$), then $x_1^{new} = x_1$.

Lagarias et al. [9] also investigated a restricted version, where expansion steps are not allowed.

We adopt the following notations. The simplex of iteration k is denoted by $S^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_{n+1}^{(k)}]$ with vertices that satisfy the condition

$$f(x_1^{(k)}) \leq f(x_2^{(k)}) \leq \dots \leq f(x_{n+1}^{(k)}) \quad (k \geq 0).$$

The initial simplex is $S^{(0)}$. The reflection, expansion and contraction points of simplex $S^{(k)}$ are denoted by $x_r^{(k)}$, $x_e^{(k)}$, $x_{oc}^{(k)}$ and $x_{ic}^{(k)}$, respectively. The function values at the vertices $x_j^{(k)}$ and the points $x_r^{(k)}$, $x_e^{(k)}$, $x_{oc}^{(k)}$ and $x_{ic}^{(k)}$ are denoted by $f(x_j^{(k)}) = f_j^{(k)}$ ($j = 1, \dots, n+1$), $f_r^{(k)} = f(x_r^{(k)})$, $f_e^{(k)} = f(x_e^{(k)})$, $f_{oc}^{(k)} = f(x_{oc}^{(k)})$ and $f_{ic}^{(k)} = f(x_{ic}^{(k)})$, respectively.

3. A REVIEW OF THE EARLIER CONVERGENCE RESULTS

In 1998 McKinnon [11] constructed a function f , which is strictly convex and has continuous first derivatives for certain parameters. He showed that for this f , the Nelder-Mead simplex algorithm may fail to converge.

Lagarias et al. [8] proved several convergence properties of the simplex method for one and two-variable strictly convex functions by giving a deep insight look of the method. They summarize their main results as follows (see p. 114 of [8]):

1. *In dimension 1, the Nelder-Mead method converges to a minimizer, and convergence is eventually M -step linear.*
2. *In dimension 2, the function values at all simplex vertices in the standard Nelder-Mead algorithm converge to the same value.*
3. *In dimension 2, the simplices in the standard Nelder-Mead algorithm have diameters converging to zero.*

In 1999 Kelley [7, 12] developed a sufficient decrease condition for the average of the object function values (evaluated at the vertices) and proved that if this condition is satisfied during the process, then any accumulation point of the simplices is a critical point of f . For similar results on other variants of the Nelder-Mead algorithm, see Tseng [13], Nazareth and Tseng [14], Pryce et al. [15].

For the restricted Nelder-Mead method, Lagarias et al. [9] significantly improved the results of [8]. Let \mathcal{F} denote the class of twice-continuously differentiable functions $\mathbb{R}^2 \rightarrow \mathbb{R}$ with bounded level sets and everywhere positive definite Hessian. Lagarias et al. proved that if the restricted Nelder-Mead algorithm is applied to a function $f \in \mathcal{F}$, starting from any nondegenerate simplex, then the algorithm converges to the unique minimizer of f .

Wright [10] raised several open questions concerning the Nelder-Mead method such as

- Why is it sometimes so effective (compared to other direct search methods) in obtaining a rapid improvement in f ?

- One failure mode is known (McKinnon [11]) – but are there other failure modes?
- Why, despite its apparent simplicity, should the Nelder-Mead method be difficult to analyze mathematically?

Our purpose is to analyze and prove the convergence of the method using a matrix formalism. This technique will also shed light on these questions, even if only, partially. The failure modes are the subject of paper [16].

4. THE NELDER-MEAD SIMPLEX METHOD IN MATRIX FORM

Assume that simplex $S^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_{n+1}^{(k)}]$ is such that

$$f(x_1^{(k)}) \leq f(x_2^{(k)}) \leq \dots \leq f(x_{n+1}^{(k)}).$$

If the incoming vertex v is of the form

$$x(\alpha) = \frac{1+\alpha}{n} \sum_{i=1}^n x_i^{(k)} - \alpha x_{n+1}^{(k)}$$

for some $\alpha \in \{1, 2, \frac{1}{2}, -\frac{1}{2}\}$, we can define the transformation matrix

$$T(\alpha) = \begin{bmatrix} I_n & \frac{1+\alpha}{n}e \\ 0 & -\alpha \end{bmatrix} \quad (e = [1, 1, \dots, 1]^T).$$

Since $S^{(k)}T(\alpha) = [x_1^{(k)}, \dots, x_n^{(k)}, x(\alpha)]$, we have to reorder the matrix columns according to the insertion rule (2.2). Define the permutation matrix

$$P_j = [e_1, \dots, e_{j-1}, e_{n+1}, e_j, \dots, e_n] \in \mathbb{R}^{(n+1) \times (n+1)} \quad (j = 1, \dots, n+1).$$

Then $S^{(k)}T(\alpha)P_j$ is the new simplex $S^{(k+1)}$. Particularly, we have the following cases.

1. If the reflection point $x_r^{(k)}$ is the new incoming vertex, then

$$S^{(k+1)} = S^{(k)}T(1)P_j \quad (j = 2, \dots, n).$$

- 2a) If the expansion point $x_e^{(k)}$ is the new incoming vertex, then

$$S^{(k+1)} = S^{(k)}T(2)P_1.$$

- 2b) If the expansion point is the reflection point $x_r^{(r)}$, then

$$S^{(k+1)} = S^{(k)}T(1)P_1.$$

- 3) If the outside contraction point $x_{oc}^{(k)}$ is the new incoming vertex, then

$$S^{(k+1)} = S^{(k)}T\left(\frac{1}{2}\right)P_j \quad (j = 1, \dots, n+1).$$

- 4) If the inside contraction point $x_{ic}^{(k)}$ is the new incoming vertex, then

$$S^{(k+1)} = S^{(k)}T\left(-\frac{1}{2}\right)P_j \quad (j = 1, \dots, n+1).$$

5) In the case of shrinking, the new vertices before reordering are

$$x'_i = \left(x_1^{(k)} + x_i^{(k)} \right) / 2 \quad (i = 1, 2, \dots, n + 1).$$

Hence $[x'_1, x'_2, \dots, x'_{n+1}] = S^{(k)}T_{shr}$, where

$$T_{shr} = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{2} \\ 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{2}I + \frac{1}{2}e_1e^T.$$

Since the new vertices are subject to the ordering

$$f\left(x_1^{(k+1)}\right) \leq \cdots \leq f\left(x_k^{(k+1)}\right) \leq \cdots \leq f\left(x_{n+1}^{(k+1)}\right),$$

the new simplex is defined by

$$S^{(k+1)} = S^{(k)}T_{shr}P,$$

where $P \in \mathcal{P}_{n+1}$, which is the set of all possible permutation matrices of order $n + 1$.

There are only $n + 1$ permutation matrices of the type P_j , while there are $(n + 1)!$ possible permutation matrices of the type \mathcal{P}_{n+1}

Hence in any of the above cases the new simplex $S^{(k+1)}$ is given by

$$S^{(k+1)} = S^{(k)}T_kP^{(k)},$$

where T_k is either $T(\alpha)$ ($\alpha \in \{-\frac{1}{2}, \frac{1}{2}, 1, 2\}$) and $P^{(k)} \in \{P_1, \dots, P_{n+1}\}$ or $T_k = T_{shr}$ and $P^{(k)} \in \mathcal{P}_{n+1}$. The number of different $T_kP^{(i)}$ matrices is at most $3n + 3 + (n + 1)!$ indicating an increasing complexity if n increases.

Observe that matrices $T(\alpha)$, T_{shr} , $T(\alpha)P$ and $T_{shr}P$, for any $P \in \mathcal{P}_{n+1}$, have the property that their column sums are 1. We exploit the following simple results.

Claim 1. (i) If $A \in \mathbb{R}^{n \times n}$ is a matrix whose column sums are 1, then A has an eigenvalue $\lambda = 1$ and a corresponding left eigenvector $x = e^T$. (ii) If $A, B \in \mathbb{R}^{n \times n}$ are two matrices whose column sums are 1, then $C = AB$ also has this property. (iii) If $A \in \mathbb{R}^{n \times n}$ is a matrix whose column sums are 1, then $\|A\| \geq 1$ in any induced matrix norm.

Proof. By definition $e^T A = [\sum_{i=1}^n a_{i1}, \dots, \sum_{i=1}^n a_{in}] = 1 \cdot e^T$. This implies $e^T B = e^T$, $e^T AB = e^T B = e^T$. Since $\rho(A) \geq 1$ and $\|A\| \geq \rho(A)$, (iii) also follows. \square

Particularly, $\|T(\alpha)\|_1 = \|T(\alpha)P\|_1 = |1 + \alpha| + |\alpha|$ and $\|T_{shr}\|_1 = \|T_{shr}P\|_1 = 1$.

A matrix A is called left stochastic if $a_{ij} \geq 0$ for all i, j and the column sums are 1. A matrix is called stochastic if $a_{ij} \geq 0$ for all i, j and both the column sums and the row sums are 1.

Matrix $T(\alpha)$ is left stochastic for $-1 \leq \alpha \leq 0$. The shrinking transformation matrix $T_{shr} = \frac{1}{2}I_{n+1} + \frac{1}{2}e_1e^T$ is a left stochastic matrix, $T_{shr}^k = \frac{1}{2^k}I_{n+1} + \frac{2^k - 1}{2^k}e_1e^T$ and $T_{shr}^k \rightarrow e_1e^T$.

For any α, β , $T(\alpha)T(\beta) = T(-\alpha\beta)$. Hence $T(\alpha)^{-1} = T(\frac{1}{\alpha})$ ($\alpha \neq 0$) and $T(\alpha)^k = T((-1)^{k+1}\alpha^k)$. If $|\alpha| < 1$, then $\lim_{k \rightarrow \infty} T(\alpha)^k = T(0)$. Matrix $T(1)$ is an involution ($T(1)^2 = I_{n+1}$). If $T(1)$ is multiplied by a permutation matrix P , this property may change. For $n = 2$, $T(1)P_2$ (reflection) is a 6-involutory matrix (for k -involutory matrices, see Trench [17]). $T(\alpha)^k$ is unbounded if $|\alpha| > 1$ and uniformly bounded if $|\alpha| \leq 1$.

5. THE CONCEPT OF CONVERGENCE AND CONSEQUENCES

Simplex $S^{(k)}$ is given by

$$S^{(k)} = S^{(0)}B_k, \tag{5.1}$$

where

$$B_k = \prod_{i=1}^k T_i P^{(i)} \tag{5.2}$$

and

$$T_i P^{(i)} \in \left\{ T(\alpha)P_k : \alpha \in \left\{ -\frac{1}{2}, \frac{1}{2}, 1, 2 \right\}, k = 1, \dots, n+1 \right\} \tag{5.3}$$

or

$$T_i P^{(i)} \in \{T_{shr}P : P \in \mathcal{P}_{n+1}\}. \tag{5.4}$$

Note that each $T_i P^{(i)}$ is nonsingular and $e^T T_i P^{(i)} = e^T$.

A simplex S is nondegenerate if the matrix

$$M = [x_1 - x_{n+1}, x_2 - x_{n+1}, \dots, x_n - x_{n+1}]$$

is nonsingular. Then S must be affinely independent, which is equivalent to (see, e.g., [18], [19]) that vectors

$$\begin{bmatrix} 1 \\ x_1 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ x_{n+1} \end{bmatrix}$$

are linearly independent. Hence $\text{rank}\left(\begin{bmatrix} e^T \\ S \end{bmatrix}\right) = n + 1$.¹ Assume that the initial simplex $S^{(0)}$ is nondegenerate. Since $e^T B_k = e^T$ and

$$\begin{bmatrix} e^T \\ S^{(k)} \end{bmatrix} = \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B_k$$

is nonsingular, $S^{(k)}$ is also nondegenerate.

For the convergence of the Nelder-Mead algorithm, it is natural to require that the simplex vertices $x_j^{(k)}$ ($j = 1, 2, \dots, n + 1$) should converge to the same vector \hat{x} as $k \rightarrow \infty$. In such a case

$$\lim_{k \rightarrow \infty} S^{(k)} = [\hat{x}, \dots, \hat{x}] = \hat{x}e^T. \tag{5.5}$$

Claim 2. *If $\{B_k\}$ is bounded, then $\{S^{(k)}\}$ converge to some S^∞ if and only if $\{B_k\}$ converge to some B .*

¹It is assumed through the whole paper that the sizes of e and the unit vectors e_i are compatible with the operation and/or partition.

Proof. If $S^{(k)} \rightarrow S^\infty$ (whatever S^∞ is) and $B_k \rightarrow B$, then $S^{(0)}B_k \rightarrow S^{(0)}B = S^\infty$. Assume that $S^{(k)} \rightarrow S^\infty$ and $\{B_k\}$ has no limit point. Since $\{B_k\}$ is bounded it must have at least one accumulation point, say B^* and there is a subsequence $\{B_{i_j}\} \subset \{B_k\}$ such that $B_{i_j} \rightarrow B^*$ and $S^{(i_j)} \rightarrow S^{(0)}B^* = S^\infty$. Assume that there exists a second accumulation point $B^{**} \neq B^*$ and a subsequence $\{B_{k_j}\} \subset \{B_k\}$ such that $B_{k_j} \rightarrow B^{**}$. It follows that

$$\begin{bmatrix} e^T \\ S^{(i_j)} \end{bmatrix} \rightarrow \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B^* = \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B^{**} \leftarrow \begin{bmatrix} e^T \\ S^{(k_j)} \end{bmatrix}.$$

Since $\begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix}$ is nonsingular, we obtain that $B^* = B^{**}$, which is a contradiction. It follows that $\{B_k\}$ converges. □

Hence it is enough to study the convergence of $\{B_k\}$, or more precisely the convergence of the right infinite matrix product

$$B = \prod_{i=1}^{\infty} T_i P^{(i)}. \tag{5.6}$$

Let A be an $n \times n$ matrix. The 1-eigenspace of the matrix A is

$$E(A) = \{x : xA = x\}.$$

Lemma 3. *Assume that $B_k \rightarrow B$ and $T_s P^{(s)}$ occurs infinitely often in the product $\prod_{j=1}^{\infty} T_j P^{(j)}$, then every row of B is in $E(T_s P^{(s)})$.*

Proof. Since $T_s P^{(s)}$ occurs infinitely often in the product $\prod_{i=1}^{\infty} T_i P^{(i)}$, there is a subsequence of $\{B_{i_j}\}$ with rightmost factor TP , say

$$B_{i_1} T_s P^{(s)}, B_{i_2} T_s P^{(s)}, \dots,$$

where the B_{i_j} 's are products of $T_i P^{(i)}$'s. Since $B_{i_j} \rightarrow B$, so does $B_{i_j} T_s P^{(s)}$. Thus $B_{i_j} T_s P^{(s)} \rightarrow B T_s P^{(s)} = B$. □

In fact the rows of B , if not zero vectors, are the left eigenvectors of $T_s P^{(s)}$ belonging to $\lambda = 1$. If several $T_s P^{(s)}$ occur infinitely often in the product $\prod_{i=1}^{\infty} T_i P^{(i)}$, then the rows of B belong to $\cap E(T_s P^{(s)})$, where the intersection is over all matrices $T_s P^{(s)}$ that occur infinitely often in $\prod_{k=1}^{\infty} T_k P^{(k)}$.

Each of the matrices $T_i P^{(i)}$ has at least one left eigenvector (e^T) belonging to $\lambda = 1$. Hence $e^T \in \cap E(T_s P^{(s)})$. If $\cap E(T_s P^{(s)}) = \{\lambda e^T : \lambda \in \mathbb{R}\}$, then B has the form $w e^T$ for some $w \in \mathbb{R}^{n+1}$. However it is not always the case.

Example 4. *Let $n = 2$, $\{d_k\}_{k=0}^{\infty}$ be a strictly monotone decreasing sequence, $d_k > \delta_1 > \delta_2$ for all k , and define $f_1^{(k)} = \delta_2$, $f_2^{(k)} = \delta_1$, $f_3^{(k)} = d_{k+1}$, $f_r^{(k)} = d_{k+1}$ and $f_{ic}^{(k)} = d_{k+2}$. Then $x_1^{(k+1)} = x_1^{(k)}$, $x_2^{(k+1)} = x_2^{(k)}$, $x_3^{(k+1)} = x_3^{(k)}$, $f(x_1^{(k+1)}) = \delta_2$,*

$f(x_2^{(k+1)}) = \delta_1$ and $f(x_3^{(k+1)}) = d_{k+2}$. Hence a repeated inside contraction occurs,

$$B_k = \left[T \left(-\frac{1}{2} \right) \right]^k \rightarrow B = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \neq we^T.$$

Here $x_1^{(k)} \rightarrow x_1^{(0)}$, $x_2^{(k)} \rightarrow x_2^{(0)}$, $x_3^{(k)} \rightarrow \frac{1}{2}(x_1^{(0)} + x_2^{(0)})$, and the simplex diameters do not converge to 0. Note that B is not a rank one matrix.

Assume now that $B_k \rightarrow B = we^T$. Then $S^{(k)} = S^{(0)}B_k \rightarrow S^{(0)}we^T = \widehat{w}e^T$, $\text{diam}(S^{(k)}) = \max_{i,j} \|S^{(k)}(e_i - e_j)\|$ and

$$\text{diam}(S^{(k)}) \leq \max_{i,j} \|S^{(0)}\| \|B_k(e_i - e_j)\| \rightarrow 0. \quad (5.7)$$

Since $B(e_i - e_j) = 0$, $B_k(e_i - e_j) = (B_k - B)(e_i - e_j)$, we have the speed estimate

$$\text{diam}(S^{(k)}) \leq \sqrt{2} \|S^{(0)}\| \|B_k - B\|. \quad (5.8)$$

Note again that in Example 4, where B is of rank 2, the simplex diameters do not converge to 0.

If B_k converges to a rank one matrix we^T , then all simplex vertices $x_i^{(k)}$ ($i = 1, \dots, n+1$) converge to the same limit $\widehat{x} = S^{(0)}w$ implying that $f_i^{(k)} \rightarrow f(S^{(0)}w)$ ($i = 1, \dots, n+1$) and $\text{diam}(S^{(k)}) \rightarrow 0$. In such a case the results of Lagarias et al. [8] mentioned as 2. and 3. in Section 3, are direct consequences.

The next example indicates that the boundedness assumption on $\{B_k\}$ is also needed.

Example 5. Let $n = 2$, $\{d_k\}_{k=1}^\infty$ be a strictly monotone decreasing sequence, and define

$$f_1^{(k)} = d_{3+k}, \quad f_2^{(k)} = d_{2+k}, \quad f_3^{(k)} = d_{1+k}$$

and

$$f_1^{(k)} = d_{3+k} > f_r^{(k)} = \frac{1}{2}(d_{3+k} + d_{4+k}) > d_{4+k} = f_e^{(k)}.$$

This guarantees the selection of $x_e^{(k)}$ as the incoming vertex for each iteration. The sequences $\{f(x_1^{(k)})\}$, $\{f(x_2^{(k)})\}$ and $\{f(x_3^{(k)})\}$ are strictly monotone decreasing, while $B_k = [T(2)P_1]^k$ is unbounded. For $d_k \rightarrow d$, the function values are converging to d , while there is no convergence for the simplex vertices. A similar example can be given for $x_r^{(k)}$ if it is selected as an expansion point.

For a given $S^{(0)}$ and f , the sequence $S^{(k)}$ (B_k) is uniquely defined. In fact, $S^{(k+1)}$ is determined by $S^{(k)}$ and the relative value distribution of f at the vertices of $S^{(k)}$ and trial points $x_r^{(k)}$, $x_e^{(k)}$, $x_{oc}^{(k)}$ and $x_{ic}^{(k)}$. Hence we study the convergence of the matrix product B_k and the convergence of the simplex vertices as a consequence. The selection of the initial simplex $S^{(0)}$ may also influence the convergence of the Nelder-Mead algorithm but it is difficult to consider it within this approach (for experimental observations on the initial simplex $S^{(0)}$, see [20], [21]).

6. EIGENVALUES OF THE TRANSFORMATION MATRICES

The eigenvalues of the matrices $T_i P^{(i)}$ play a key role in the convergence of the infinite matrix product $\prod_{i=1}^{\infty} T_i P^{(i)}$. For $P^{(i)} = I_{n+1}$, we have the following simple result.

Lemma 6. (i) Matrix $T(\alpha)$ has the eigenvalues $\lambda_i = 1$ ($i = 1, \dots, n$) and $\lambda_{n+1} = -\alpha$ and the diagonal Jordan form

$$T(\alpha) = \begin{bmatrix} I_n & -\frac{1}{n}e \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -\alpha \end{bmatrix} \begin{bmatrix} I_n & \frac{1}{n}e \\ 0 & 1 \end{bmatrix}.$$

(ii) The eigenvalues of T_{shr} are $\lambda_1 = 1$ and $\lambda_i = \frac{1}{2}$ for $i = 2, \dots, n + 1$. The corresponding eigenvectors are $x_1 = e_1$, $x_i = -e_1 + e_i$. Furthermore T_{shr} has a diagonal Jordan normal form

$$T_{shr} = X \Lambda X^{-1}, \tag{6.1}$$

where $\Lambda = \text{diag}(\lambda_i) = \frac{1}{2} (I_{n+1} + e_1 e_1^T)$ and $X = I_{n+1} + e_1 (e_1^T - e^T)$.

Lemma 7. The matrix $T(\alpha) P_k$ ($1 \leq k \leq n$) has $k - 1$ eigenvalues $\lambda = 1$. The remaining $n - k + 2$ eigenvalues are the zeros of the polynomial

$$p_{n+2-k}(\lambda) = \lambda^{n+2-k} - c \sum_{i=1}^{n+1-k} \lambda^i + \alpha, \tag{6.2}$$

where $c = \frac{1+\alpha}{n}$. If $k = 1$, then $p_{n+1}(\lambda)$ has at least one eigenvalue $\lambda = 1$. If $\alpha = 1$, $p_{n+1}(\lambda)$ has at least two eigenvalues $\lambda = 1$. If $\alpha = 2$, $p_{n+1}(\lambda)$ has an eigenvalue in the interval $(1, 2)$. For $k \geq 2$, there are exactly $k - 1$ eigenvalues $\lambda = 1$. If $2 \leq k \leq n$ and $\alpha < 0$, $p_{n-k+2}(\lambda)$ has all roots in the open unit disk.

Proof. For $1 \leq k \leq n$,

$$T(\alpha) P_k = \begin{bmatrix} I_{k-1} & c e e_1^T \\ 0 & A_{n+2-k} \end{bmatrix},$$

where

$$A_{n+2-k} = \begin{bmatrix} c & 1 & 0 & \cdots & 0 \\ c & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ c & 0 & \cdots & 0 & 1 \\ -\alpha & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Since A_{n+2-k} is a companion matrix (for this form, see, e.g. [22]), its characteristic polynomial is

$$p_{n+2-k}(\lambda) = \lambda^{n+2-k} - c \sum_{i=1}^{n+1-k} \lambda^i + \alpha \quad (1 \leq k \leq n)$$

and the characteristic polynomial of $T(\alpha) P_k$ is

$$\det(T(\alpha) P_k - \lambda I_{n+1}) = (1 - \lambda)^{k-1} p_{n+2-k}(\lambda).$$

Note that $p_{n+2-k}(1) = \frac{k-1}{n}(\alpha + 1)$. If $k = 1$, $p_{n+1}(1) = 0$, that is $\lambda_1 = 1$ (the column sums of $T(\alpha)P_1$ are 1!). Since $p'_{n+1}(\lambda) = \frac{1-\alpha}{2}(n+1)$, there is a second zero $\lambda_2 = 1$ if $\alpha = 1$ (expansion by reflection). For $\alpha = 2$, $p'_{n+1}(1) < 0$ and $p_{n+1}(2) = 2^{n+1}(1 - \frac{3}{n}) + 2 + \frac{6}{n} > 0$ ($n \geq 2$). Hence $p_{n+1}(\lambda)$ has a zero in the interval $(1, 2)$ for $\alpha = 2$ and $n \geq 2$. If $k \geq 2$, $p_{n+2-k}(1) \neq 0$. If $\alpha < 0$ and $2 \leq k \leq n$, the roots of $p_{n+2-k}(\lambda)$ are inside the unit disk since for $|\lambda| \geq 1$,

$$\begin{aligned} |p_{n+2-k}(\lambda)| &= |\lambda|^{n+2-k} \left| 1 - c \sum_{i=1}^{n+1-k} \frac{1}{\lambda^i} + \alpha \frac{1}{\lambda^{n+2-k}} \right| \\ &\geq |\lambda|^{n+2-k} \left(1 - c \sum_{i=1}^{n+1-k} \frac{1}{|\lambda|^i} - |\alpha| \frac{1}{|\lambda|^{n+2-k}} \right) \\ &\geq |\lambda|^{n+2-k} (1 - (n+1-k)c - |nc - 1|) \\ &\geq |\lambda|^{n+2-k} (k-1)c > 0. \end{aligned}$$

□

For $\alpha > 0$, there is no estimate on the location of the roots of p_{n+2-k} relative to the open unit disk. However the Schur-Cohn test may help to decide if the roots of $p_{n+2-k}(\lambda)$ are in the open unit disk.

For the eigenvalues of $T_{shr}P$ ($P \in \mathcal{P}_{n+1}$), we cite the following result

Theorem 8. (Langville and Meyer [23, 24]). *If the spectrum of the stochastic matrix P is $\{1, \lambda_2, \dots, \lambda_n\}$, then the spectrum of*

$$W = \alpha P + (1 - \alpha)ev^T$$

is $\{1, \alpha\lambda_2, \alpha\lambda_3, \dots, \alpha\lambda_n\}$, where v^T is a probability vector².

Corollary 9. *Since the eigenvalues of W and W^T coincide, we have the same result for the transposed matrix*

$$W^T = \alpha P^T + (1 - \alpha)ve^T$$

as well.

Corollary 10. *The spectrum of $T_{shr}P = \frac{1}{2}P + \frac{1}{2}e_1e^T$ is $\{1, \frac{1}{2}\lambda_2, \frac{1}{2}\lambda_3, \dots, \frac{1}{2}\lambda_{n+1}\}$. Since the eigenvalues of a permutation matrix are on the unit circle $|\lambda| = 1$, we have $|\frac{1}{2}\lambda_i| = \frac{1}{2}$ for $i = 2, \dots, n+1$.*

7. GENERAL CONVERGENCE RESULTS FOR THE NELDER-MEAD METHOD

For every $T(\alpha)P_k$, the spectral radius is bigger than or equal to 1. Since $e^T \in E(T_iP^{(i)})$ for all $T_iP^{(i)}$ occurring here, we first block triangularize them by a common similarity transformation (for left infinite matrix products, see Theorem 6.10 of Hartfiel [25]). We show that for

$$F = \begin{bmatrix} 1 & -e^T \\ 0 & I_n \end{bmatrix}, \quad F^{-1} = \begin{bmatrix} 1 & e^T \\ 0 & I_n \end{bmatrix} \tag{7.1}$$

²Vector v is such that $v_i \geq 1$ and $e^T v = 1$.

and for all possible $T_i P^{(i)}$,

$$F^{-1}T_i P^{(i)}F = \begin{bmatrix} 1 & 0^T \\ b_i & C_i \end{bmatrix}, \tag{7.2}$$

where $b_i \in \mathbb{R}^n$ and $C_i \in \mathbb{R}^{n \times n}$ depends on $T_i P^{(i)}$.

Lemma 11. For every $T(\alpha) P_k$ ($\alpha \in \{-\frac{1}{2}, \frac{1}{2}, 1, 2\}$, $P_k \in \{P_1, \dots, P_{n+1}\}$), the matrix $F^{-1}T(\alpha) P_k F$ has the form (7.2).

Proof. For $k > 1$ we can write

$$T(\alpha) P_k = \begin{bmatrix} 1 & ce_k^T \\ 0 & W \end{bmatrix} \quad (W \in \mathbb{R}^{n \times n}),$$

and so

$$F^{-1}T(\alpha) P_k F = \begin{bmatrix} 1 & -e^T + ce_k^T + e^T W \\ 0 & W \end{bmatrix}.$$

Since $e^T W e_{k-1} = (n-1)c - \alpha = 1 - c$, $e^T W = [1, \dots, 1 - c, 1, \dots, 1]$, we obtain the form

$$F^{-1}T(\alpha) P_k F = \begin{bmatrix} 1 & 0 \\ 0 & W \end{bmatrix}.$$

For $k = 1$, we can write

$$T(\alpha) P_1 = \begin{bmatrix} c & e_1^T \\ z & W \end{bmatrix} \quad (W \in \mathbb{R}^{n \times n})$$

with $z = [c, \dots, c, -\alpha]^T$. Hence

$$F^{-1}T(\alpha) P_1 F = \begin{bmatrix} c + e^T z & -ce^T + e_1^T - e^T z e^T + e^T W \\ z & -ze^T + W \end{bmatrix}.$$

Since $e^T W = [0, 1, \dots, 1]$, $e^T z = 1 - c$, $c + e^T z = 1$,

$$-ce^T + e_1^T - e^T z e^T + e^T W = -ce^T + e_1^T - (1 - c)e^T + e^T W = 0.$$

The final result is

$$F^{-1}T(\alpha) P_1 F = \begin{bmatrix} 1 & 0 \\ z & -ze^T + W \end{bmatrix}.$$

□

Remark 12. For $k > 1$, $b = 0$, and for $k = 1$, $\|b\|_2 = \left(\frac{(1+\alpha)^2}{n} + \alpha^2\right)^{\frac{1}{2}}$.

Lemma 13. For every $T_{shr} P$ ($P \in \mathcal{P}_{n+1}$), the matrix $F^{-1}T_{shr} P F$ has the form (7.2).

Proof. Note that $T_{shr} P = \frac{1}{2}P + \frac{1}{2}e_1 e^T$ and $P = [e_{i_1}, \dots, e_{i_{n+1}}]$. If $i_1 = 1$, then

$$T_{shr} P = \begin{bmatrix} 1 & \frac{1}{2}e^T \\ 0 & W_1 \end{bmatrix}$$

where W_1 is an $n \times n$ permutation matrix multiplied by $\frac{1}{2}$. Hence $e^T W_1 = \frac{1}{2} e^T$ and

$$F^{-1} T_{shr} P F = \begin{bmatrix} 1 & -\frac{1}{2} e^T + e^T W_1 \\ 0 & W_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix}.$$

If $i_1 > 1$ and $T_{shr} P e_j = e_1$, then

$$T_{shr} P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} e^T + \frac{1}{2} e_{j-1}^T \\ \frac{1}{2} e_{i_1-1} & W_2 \end{bmatrix},$$

where $W_2 e_{j-1} = 0$, $e^T W_2 e_i = \frac{1}{2}$ ($i \neq j-1$) and $e^T W_2 = \frac{1}{2} (e^T - e_{j-1}^T)$. Since $e^T e_{i_1-1} = 1$,

$$\begin{aligned} F^{-1} T_{shr} P F &= \begin{bmatrix} \frac{1}{2} + \frac{1}{2} e^T e_{i_1-1} & \frac{1}{2} e_{j-1}^T - \frac{1}{2} e^T e_{i_1-1} e^T + e^T W_2 \\ \frac{1}{2} e_{i_1-1} & -\frac{1}{2} e_{i_1-1} e^T + W_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ \frac{1}{2} e_{i_1-1} & -\frac{1}{2} e_{i_1-1} e^T + W_2 \end{bmatrix} \end{aligned}$$

□

Remark 14. If $i_1 = 1$, then the first column entries are 0 except for entry $(1, 1)$. If $i_1 \geq 2$, then entry $(i_1, 1)$ is $\frac{1}{2}$, while the remaining entries are 0 ($\ell \neq 1, i_1$). Hence $\|b\|_2 \leq \frac{1}{2}$. The entries of submatrix C are only 0, $\frac{1}{2}$ and $-\frac{1}{2}$. In column j , there can be at most two nonzero elements. Theorem 8 and Corollaries 9 and 10 imply that $\rho(C) = \frac{1}{2}$. Note that $\|C\|_1 \leq 1$.

We also need the following simple results.

Assume that for $i \geq 1$,

$$A_i = \begin{bmatrix} 1 & 0 \\ b_i & C_i \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)} \quad (C_i \in \mathbb{R}^{n \times n}). \quad (7.3)$$

It is easy to see that

$$L_k = \prod_{j=1}^k A_j = \begin{bmatrix} 1 & 0 \\ \sum_{i=1}^k \left(\prod_{j=1}^{i-1} C_j \right) b_i & \prod_{j=1}^k C_j \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ x_k & \prod_{j=1}^k C_j \end{bmatrix}. \quad (7.4)$$

Lemma 15. Assume that $\left\| \prod_{j=1}^k C_j \right\| \leq c_k$, $\sum_{k=1}^{\infty} c_k$ is convergent ($< \infty$) and $\|b_k\| \leq \gamma$ for all k . Then L_k converges and

$$\lim_{k \rightarrow \infty} L_k = \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} \quad (7.5)$$

for some \tilde{x} .

Proof. If $\sum_{k=1}^{\infty} c_k$ is convergent, then $c_k \rightarrow 0$. Hence $\prod_{j=1}^k C_j \rightarrow 0$ as $k \rightarrow \infty$. Since $s_k = \sum_{j=1}^k c_j$ is convergent, for any $\varepsilon > 0$ there is a number $k_0 = k_0(\varepsilon)$ such that for

$m > k \geq k_0$, $|s_m - s_k| < \varepsilon$. Thus for $m > k \geq k_0$, we obtain

$$\|x_m - x_k\| \leq \sum_{i=k+1}^m \left\| \prod_{j=1}^{i-1} C_j \right\| \|b_i\| \leq \gamma \sum_{i=k+1}^m c_{i-1} \leq \gamma\varepsilon.$$

Hence $x_k \rightarrow \tilde{x}$ for some \tilde{x} . □

Remark 16. If $\|C_j\| \leq q < 1$ for $j \geq 1$, then $\left\| \prod_{j=1}^k C_j \right\| \leq q^k$ and the series $\sum_{i=1}^{\infty} q^i$ is convergent.

Remark 17. For the matrices C_i of (7.2), either $\rho(C_i) < 1$ or $\rho(C_i) \geq 1$. Since in any induced matrix norm $\|C_i\| \geq \rho(C_i)$, we cannot expect convergence improvement from the latter matrices, unless their effect is compensated.

Assume now that for some indices j , $1 \leq \|C_j\| \leq Q$, while for other indices i , $\|C_i\| \leq q < 1$. Denote by $t_1(k)$ the number of those C_i 's that satisfies $\|C_i\| \leq q < 1$ ($1 \leq i \leq k$), and denote by $t_2(k)$ those C_i 's that satisfy $1 \leq \|C_i\| \leq Q$ ($1 \leq i \leq k$). Clearly, $0 \leq t_i(k) \leq k$ and $t_1(k) + t_2(k) = k$. Then

$$\left\| \prod_{j=1}^k C_j \right\| \leq q^{t_1(k)} Q^{t_2(k)}.$$

Assume that $\kappa \in \mathbb{N}$ is such that $\frac{1}{q^{\kappa-1}} \leq Q \leq \frac{1}{q^\kappa}$. Then $\left\| \prod_{j=1}^k C_j \right\| \leq q^{t_1(k) - \kappa t_2(k)}$. If $\varphi(k) := t_1(k) - \kappa t_2(k)$ is a monotone increasing sequence converging to infinity, then $\prod_{j=1}^k C_j \rightarrow 0$ as $k \rightarrow \infty$. Note that $\varphi(k) \leq k$. The root test of infinite series guarantees that if for some $k_0 > 0$,

$$q^{\frac{\varphi(k)}{k}} \leq r < 1 \quad (k \geq k_0),$$

then $\sum_{i=1}^{\infty} q^{\varphi(i)}$ is convergent. This condition is certainly satisfied if $\frac{\varphi(k)}{k} \geq \mu$, where $0 < \mu < 1$ is a fixed number. Observe that in such a case $k \geq t_1(k) \geq \mu k + \kappa t_2(k)$ and $t_2(k) \leq \frac{1-\mu}{\kappa} k$. If $Q = 1$, then $\kappa = 1$.

We can also give an estimate for the speed of convergence. For $\prod_{i=1}^k C_j \rightarrow 0$, we have the estimate $\left\| \prod_{j=1}^k C_j \right\| \leq q^{\varphi(k)} \leq q^{\mu k}$. For the speed of the convergence $\sum_{i=1}^k \left(\prod_{j=1}^{i-1} C_j \right) b_i \rightarrow \tilde{x}$, we have the estimate

$$\|\tilde{x} - x_k\| = \left\| \sum_{i=k+1}^{\infty} \left(\prod_{j=1}^{i-1} C_j \right) b_i \right\| \leq \gamma \sum_{i=k}^{\infty} q^{\varphi(i)} \leq \gamma \sum_{i=k}^{\infty} q^{\mu i} \leq \frac{\gamma q^{\mu k}}{1 - q^\mu}.$$

We have just proved the following

Corollary 18. Assume that $\|C_k\| \leq Q$ and $\|b_k\| \leq \gamma$ for all k . Denote by $t_1(k)$ the number of those C_i 's that satisfies $\|C_i\| \leq q < 1$ ($1 \leq i \leq k$), and denote by $t_2(k)$ those C_i 's that satisfies $1 \leq \|C_i\| \leq Q$ ($1 \leq i \leq k$). Define $\kappa \in \mathbb{N}$ by the inequality

$\frac{1}{q^{\kappa-1}} \leq Q \leq \frac{1}{q^{\kappa}}$. If for some $k_0 \in \mathbb{N}$ and $\mu \in (0, 1)$, $t_1(k) \geq \mu k + \kappa t_2(k)$ holds ($k \geq k_0$), then L_k converges and

$$\lim_{k \rightarrow \infty} L_k = \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} = \tilde{L} \tag{7.6}$$

for some \tilde{X} . For the speed of the convergence we have the estimate

$$\|L_k - \tilde{L}\| \leq \Gamma q^{\mu k}, \tag{7.7}$$

where $\Gamma > 0$ is a suitable constant depending on μ and γ .

Remark 19. If $t_2(k) = 0$ for $k \geq 1$, then $t_1(k) = k$ and $\mu = 1$. If there are only a finite number of cases, when $1 \leq \|C_i\| \leq Q$ holds, we can omit it in the sense that if after iteration, say k_0 , only $\|C_i\| \leq q < 1$ occurs, then $t_2(k) \leq k_0$, $t_1(k) \geq k - k_0$ and we can set $\mu = 1$ in the above Corollary 18.

Given n , we have a finite set of matrices $T_i P^{(i)}$, say, \mathcal{T} that may occur in the infinite product (5.6) if the Nelder-Mead method is applied to some function f . For each $T_i P^{(i)}$, we have the representation

$$T_i P^{(i)} = F \begin{bmatrix} 1 & 0^T \\ b_i & C_i \end{bmatrix} F^{-1}. \tag{7.8}$$

Hence $B_k = \prod_{i=1}^k T_i P^{(i)} = F L_k F^{-1}$ and B_k is convergent if and only if

$$L_k = \prod_{i=1}^k \begin{bmatrix} 1 & 0^T \\ b_i & C_i \end{bmatrix}$$

is convergent. For some of the C_i 's, $\rho(C_i) \leq q \leq 1$, while for the others $\rho(C_i) \geq 1$. Define the sets

$$\mathcal{W}_1 = \left\{ T \left(\frac{1}{2} \right) P_j, T \left(-\frac{1}{2} \right) P_j : j = 1, 2 \right\} \cup \{ T_{shr} P : P \in \mathcal{P}_{n+1} \}$$

and

$$\begin{aligned} \mathcal{W}_2 = \{ & T(2) P_1, T(1) P_j : j = 1, \dots, n \} \\ & \cup \left\{ T \left(\frac{1}{2} \right) P_j, T \left(-\frac{1}{2} \right) P_j : j = 3, \dots, n+1 \right\}. \end{aligned}$$

If $T_i P^{(i)} \in \mathcal{W}_2$, the matrix has a second eigenvalue $|\lambda_2| \geq 1$, and for the corresponding C_i , $\rho(C_i) \geq 1$ holds. Although Lemma 7 does not guarantee that for matrices $T_i P^{(i)} \in \{ T \left(\frac{1}{2} \right) P_j, T \left(-\frac{1}{2} \right) P_j : j = 1, 2 \}$, the second largest eigenvalue $|\lambda_2| < 1$ or $\rho(C_i) < 1$, a thorough computer check shows that it is the case, at least up to $n = 20$. However $|\lambda_2|$'s are approaching 1 as n increases. Hence we assume the following:

- (A) If $T_i P^{(i)} \in \mathcal{W}_1$, then $\rho(C_i) < 1$, and if $T_i P^{(i)} \in \mathcal{W}_2$, then $\rho(C_i) \geq 1$.
- (B) There is a matrix norm $\|A\|_w$ (induced by a vector norm $\|x\|_w$) such that if $T_i P^{(i)} \in \mathcal{W}_1$, then $\|C_i\|_w < 1$.

Since the sets \mathcal{W}_1 and \mathcal{W}_2 are finite, there are numbers $0 < q < 1 \leq Q$ such that for $T_i P^{(i)} \in \mathcal{W}_1$, $\|C_i\|_w \leq q < 1$, and for $T_i P^{(i)} \in \mathcal{W}_2$, $1 \leq \|C_i\|_w \leq Q$. Also we can assume that for every $T_i P^{(i)} \in \mathcal{W}_1 \cup \mathcal{W}_2$, $\|b_i\|_w \leq \gamma$.

Theorem 20. *Assume (A) and (B) are satisfied and the initial simplex $S^{(0)}$ is non-degenerate. Let $t_1(k)$ be the number of Nelder-Mead steps (operations $T_i P^{(i)}$) that belong to \mathcal{W}_1 , and $t_2(k)$ be the number of those steps (operations $T_i P^{(i)}$) that belong to \mathcal{W}_2 during the first k iterations of the Nelder-Mead method. Also assume that $\kappa \in \mathbb{N}$ is such that $\frac{1}{q^{\kappa-1}} \leq Q \leq \frac{1}{q^\kappa}$. If for some $k_0 \in \mathbb{N}$ and $\mu \in (0, 1)$, $t_1(k) \geq \mu k + \kappa t_2(k)$ holds ($k \geq k_0$), then the Nelder-Mead algorithm converges in the sense that for all simplex vertices $x_j^{(k)} \rightarrow \hat{x}$ ($j = 1, \dots, n+1$) holds as $k \rightarrow \infty$ with a convergence speed proportional to $q^{\mu k}$. If f is continuous at \hat{x} , then $f(x_j^{(k)}) \rightarrow f(\hat{x})$ ($j = 1, \dots, n+1$) holds as well.*

Proof. Under the assumptions

$$L_k = \prod_{i=1}^k \begin{bmatrix} 1 & 0^T \\ b_i & C_i \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} = \tilde{L}$$

with the speed

$$\|L_k - \tilde{L}\|_w \leq \Gamma q^{\mu k}.$$

Hence

$$B_k \rightarrow F \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} F^{-1} = \begin{bmatrix} 1 - e^T \tilde{x} \\ \tilde{x} \end{bmatrix} e^T = \hat{x} e^T = B \tag{7.9}$$

with the speed

$$\|B_k - B\|_w \leq \Gamma \text{cond}(F) q^{\mu k}. \tag{7.10}$$

□

Corollary 21. *diam($S^{(k)}$) $\rightarrow 0$ ($k \rightarrow \infty$) with a speed of $O(q^{\mu k})$.*

For higher dimension, we can expect slower convergence, since q approaches 1.

8. THE CONVERGENCE OF THE NELDER-MEAD METHOD IN LOW DIMENSIONS

Here we demonstrate the use of Theorem 20 for $n = 1, 2, 3$. The cases show an increasing technical complexity and also the growth of $\max\{\rho(C_i) : T_i P^{(i)} \in \mathcal{W}_1\}$.

8.1. The Nelder-Mead method in one dimension. For $n = 1$,

$$\mathcal{W}_1 = \left\{ T \begin{pmatrix} 1 \\ 2 \end{pmatrix} P_j, T \begin{pmatrix} -1 \\ 2 \end{pmatrix} P_j : j = 1, 2 \right\}$$

and

$$\mathcal{W}_2 = \{T(1) P_2, T(2) P_2, T(1) P_1\}.$$

Lemma 11 implies

$$F^{-1} B_k F = \prod_{i=1}^k \begin{bmatrix} 1 & 0 \\ b_i & c_i \end{bmatrix}.$$

Note that for $T_i P^{(i)} \in \mathcal{W}_1$, $|c_i| = \frac{1}{2} = q$, and for $T_i P^{(i)} \in \mathcal{W}_2$, $|c_i| \geq 1$ and $1 \leq |c_i| \leq 2 = Q$. Here the norm $\|\cdot\|_w = |\cdot|$. Hence Theorem 20 implies the convergence with $\kappa = 1$.

8.2. The Nelder-Mead method in two dimensions. For $n = 2$, the six permutation matrices of set \mathcal{P}_3 are

$$P_1 = [e_3, e_1, e_2], \quad P_2 = [e_1, e_3, e_2], \quad P_3 = [e_1, e_2, e_3],$$

$$P_4 = [e_2, e_1, e_3], \quad P_5 = [e_2, e_3, e_1], \quad P_6 = [e_3, e_2, e_1].$$

Define

$$\mathcal{W}_1 = \left\{ T \left(\frac{1}{2} \right) P_j, T \left(-\frac{1}{2} \right) P_j : j = 1, 2 \right\} \cup \{ T_{shr} P_j : j = 1, \dots, 6 \}$$

and

$$\mathcal{W}_2 = \{ T(1) P_2, T(2) P_1, T(1) P_1 \} \cup \left\{ T \left(\frac{1}{2} \right) P_3, T \left(-\frac{1}{2} \right) P_3 \right\}.$$

The similarity transformation

$$F^{-1} T P F = \begin{bmatrix} 1 & 0^T \\ b & C \end{bmatrix} \quad (8.1)$$

on the elements of $\mathcal{W}_1 \cup \mathcal{W}_2$ yields matrices C_i for which $\rho(C_i) < 1$, while $\|C_i\|_2 > 1$.

Since all such C_i 's have diagonal Jordan forms, it would be an ideal situation if for some norm $\|\cdot\|_w$, every $\|C_i\|_w$ would be close to $\rho(C_i)$. Such thing is possible, if the matrices C_i are simultaneously similar to diagonal matrices (see, e.g. [26–28]). However, this requires that all matrices C_i must be pairwise commuting, which is not the case here. Instead of this we look for a 2×2 matrix S such that $\|C_i\|_w = \|S^{-1} C_i S\|_2$ (induced by the vector norm $\|x\|_w = \|S^{-1} x\|_2$) is as close to $\rho(C_i)$ as possible. So we try to solve the optimization problem

$$\min_S \max \left\{ \|S^{-1} C_i S\|_2 : T_i P^{(i)} \in \left\{ T \left(\frac{1}{2} \right) P_j, T \left(-\frac{1}{2} \right) P_j : j = 1, 2 \right\} \right\}. \quad (8.2)$$

For the matrix

$$S = \begin{bmatrix} 2 & -\frac{4}{5} \\ 0 & \frac{8}{5} \end{bmatrix} \quad (\text{cond}(S) \approx 1.6404),$$

we have the following numerical results.

Case	Operation	C_i	$\rho(C_i)$	$\ C_i\ _2$	$\ S^{-1}C_iS\ _2$
1	$F^{-1}T(1)P_2F$	C_1	1	1.6180	1.3765
2a	$F^{-1}T(2)P_1F$	C_2	1.6861	3.1787	2.7043
2b	$F^{-1}T(1)P_1F$	C_3	1	1.6180	1.3765
3a	$F^{-1}T\left(\frac{1}{2}\right)P_1F$	C_4	0.8431	0.9056	0.8438
3b	$F^{-1}T\left(\frac{1}{2}\right)P_2F$	C_5	0.7071	1.2892	0.8438
3c	$F^{-1}T\left(\frac{1}{2}\right)P_3F$	C_6	1	1.2892	1.0
4a	$F^{-1}T\left(-\frac{1}{2}\right)P_1F$	C_7	0.7071	0.9056	0.8438
4b	$F^{-1}T\left(-\frac{1}{2}\right)P_2F$	C_8	0.8431	1.0399	0.8438
4c	$F^{-1}T\left(-\frac{1}{2}\right)P_3F$	C_9	1	1.0399	1.0
5a	$F^{-1}T_{shr}P_1F$	C_{10}	$\frac{1}{2}$	0.8090	0.6882
5b	$F^{-1}T_{shr}P_2F$	C_{11}	$\frac{1}{2}$	0.5	0.6882
5c	$F^{-1}T_{shr}P_3F$	C_{12}	$\frac{1}{2}$	0.5	0.5
5d	$F^{-1}T_{shr}P_4F$	C_{13}	$\frac{1}{2}$	0.8090	0.5
5e	$F^{-1}T_{shr}P_5F$	C_{14}	$\frac{1}{2}$	0.8090	0.6882
5f	$F^{-1}T_{shr}P_6F$	C_{15}	$\frac{1}{2}$	0.8090	0.6882

Here we can select the values $q = 0.85$ ($q \approx \rho(C_4), \rho(C_8)$), $Q = 2.71$ and $\kappa = 7$. Note that for $T_iP^{(i)} \in \mathcal{W}_1$, $\|C_i\|_w \leq q < 1$, and for $T_iP^{(i)} \in \mathcal{W}_2$, $1 \leq \|C_i\|_w \leq Q$.

Hence Theorem 20 implies convergence with $\kappa = 7$.

The applied optimization technique (8.2) is somewhat similar to the joint approximate diagonalization method (see, e.g., [29]).

8.3. The Nelder-Mead method in three dimensions. For $n = 3$, \mathcal{W}_1 has 28 elements, while \mathcal{W}_2 has 8. It can be checked that for $T_iP^{(i)} \in \mathcal{W}_1$, $\rho(C_i) \leq 0.9275$ and $\|C_i\|_2 \leq 1.2622$. A numerical solution of the optimization problem (8.2) gives the following matrix

$$S = \begin{bmatrix} -0.6012 & 1.5707 & 0.3968 \\ 1.4938 & -0.0616 & 0.4419 \\ -0.6500 & -0.7949 & 0.7620 \end{bmatrix},$$

for which $\|C_i\|_w = \|S^{-1}C_iS\|_2 \leq 0.9293$ ($T_iP^{(i)} \in \mathcal{W}_1$) and $\|C_i\|_w \leq 2.7729$ ($T_iP^{(i)} \in \mathcal{W}_2$). Selecting $q = 0.93$ and $Q = 2.78$, we have $\kappa = 15$, and Theorem 20 holds for $n = 3$, as well.

If we exclude the expansion steps, then we can set $Q = 1.3725$, resulting in $\kappa = 5$ and a faster convergence.

REFERENCES

1. J. A. Nelder and R. Mead. "A simplex method for function minimization," *Computer Journal*, **7**, (1965), pp. 308–313. DOI: 10.1093/comjnl/7.4.308.
2. F. H. Walters, S. L. Morgan, L. R. Parker, and S.N. Deming. *Sequential Simplex Optimization*. CRC Press LLC, 1991.

3. R. M. Lewis, V. Torczon, and M. W. Trosset. “Direct search methods: then and now.” *Journal of Computational and Applied Mathematics*, **124**, (2000), pp. 191–207. DOI: 10.1016/S0377-0427(00)00423-4.
4. T. G. Kolda, R. M. Lewis, and V. Torczon. “Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods.” *SIAM Review*, **45**(3), (2003), pp. 385–482. DOI: 10.1137/S003614450242889.
5. A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to Derivative-Free Optimizations*. MPS-SIAM Series on Optimization. SIAM, 2009. DOI: 10.1137/1.9780898718768.
6. C. Audet and W. Hare. *Derivative-free and Blackbox Optimization*. Springer, 2017. DOI: 10.1007/978-3-319-68913-5.
7. C. T. Kelley. “Detection and remediation of stagnation in the Nelder-Mead algorithm using an sufficient decrease condition.” *SIAM Journal of Optimization*, **10**, (1999), pp. 33–45. DOI: 10.1137/S1052623497315203.
8. J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. “Convergence properties of the Nelder-Mead simplex method in low dimensions.” *SIAM Journal on Optimization*, **9**(1), (1998), pp. 112–147. DOI: 10.1137/S1052623496303470.
9. J. C. Lagarias, B. Poonen, and M. H. Wright. “Convergence of the restricted Nelder-Mead algorithm in two dimensions.” *SIAM Journal on Optimization*, **22**(2), (2012), pp. 501–532. DOI: 10.1137/110830150.
10. M. H. Wright. “Nelder, Mead, and the other simplex method,” *Documenta Mathematica, Extra Volume: Optimization Stories*, (2012), pp. 271–2766.
11. K. I. M. McKinnon. “Convergence of the Nelder-Mead simplex method to a nonstationary point,” *SIAM Journal on Optimization*, **9**(1), (1998), pp. 148–158. DOI: 10.1137/S1052623496303482.
12. C. T. Kelley. *Iterative Methods for Optimization*. Frontiers in Applied Mathematics. SIAM, 1999. DOI: 10.1137/1.9781611970920.
13. P. Tseng. “Fortified-descent simplicial search method: A general approach,” *SIAM Journal on Optimization*, **10**(1), (1999), pp. 269–288. DOI: 10.1137/S1052623495282857.
14. L. Nazareth and P. Tseng. “Gilding the Lily: A variant for the Nelder-Mead algorithm based on golden-section search,” *Computational Optimization and Applications*, **22**, (2002), pp. 133–144. DOI: 10.1023/A:1014842520519.
15. C. J. Price, I. D. Coope, and D. Byatt. “A convergent variant of the Nelder-Mead algorithm,” *Journal of Optimization Theory and Applications*, **113**(1), (2002), pp. 5–19. DOI: 10.1023/A:1014849028575.
16. A. Galántai. “A convergence analysis of the Nelder-Mead simplex method.” *Acta Polytechnica Hungarica*, (2020). Submitted.
17. W. F. Trench. “Characterization and properties of matrices with k -involutory symmetries,” *Linear Algebra and its Applications*, **429**, (2008), pp. 2278–2290. DOI: 10.1016/j.laa.2008.07.002.
18. J. G. Hocking and G. S. Young. *Topology*. Addison–Wesley, 1961.
19. J. Matousek. *Using the Borsuk-Ulam Theorem*. 2nd edition. Springer, 2008. DOI: 10.1007/978-3-540-76649-0.

20. J. M. Parkinson and D. Hutchinson. “An investigation into the efficiency of variants on the simplex method.” *Numerical Method for Non-linear Optimization*. Ed. by F. A. Lootsma. Academic Press, 1972, pp. 115–135.
21. S. Wessing. “Proper initialization is crucial for the Nelder-Mead simplex search,” *Optimization Letters*, **13**, (2019), pp. 847–856. DOI: 10.1007/s11590-018-1284-4.
22. R. A. Horn and C. R. Johnson. *Matrix Analysis*. 2nd edition. Cambridge University Press, 2013. DOI: 10.1017/9781139020411.
23. A. N. Langville and C. D. Meyer. “Deeper inside PageRank,” *Internet Mathematics Journal*, **1**(3), (2005), pp. 335–380. DOI: 10.1080/15427951.2004.10129091.
24. A. N. Langville and C. D. Meyer. *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006.
25. D. J. Hartfiel. *Nonhomogeneous Matrix Products*. World Scientific, 2001. DOI: 10.1142/4707.
26. L. Mirsky. *An Introduction to Linear Algebra*. Oxford University Press, 1955.
27. S. Friedland. *Matrices: Algebra, Analysis, and Applications*. World Scientific Publishing, 2015. DOI: 10.1142/9567.
28. K. C. O’Meara, J. Clark, and C. L. Vinsonhaler. *Advanced Topics in Linear Algebra: Weaving Matrix Problems through the Weyr Form*. Oxford University Press, 2011.
29. X. Shi. *Blind Signal Processing: Theory and Practice*. Springer, 2011. DOI: 10.1007/978-3-642-11347-5.