

**On the Component-wise
Convergence Rate**

Amr El-Bakry and Trond Steihaug

**CRPC-TR98761
August 1998**

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

On the Component-wise Convergence Rate

Amr El-Bakry*
Computational and
Applied Mathematics
Rice University
Houston TX 77005-1892

Trond Steihaug†
Department of Informatics
University of Bergen
Høyteknologisenteret
N-5020 Bergen Norway

DRAFT August 26, 1998

Abstract

In this paper we investigate the convergence rate of a sequence of vectors provided that the convergence rates of the components are known. The result of this investigation is then used to study the m-step convergence rate of sequences.

Key Words. convergence rate - Q-factor - multi-step convergence.

1 Introduction

Convergence and convergence rate of iteration sequences play an essential role in the design and analysis of optimizations methods. Convergence rate has been used as a measure of efficiency and a tool for performance comparison of optimization algorithms. See for example Ortega and Rheinboldt Chapter 9 (Ref. 1). In certain methods, the convergence and convergence

*Research supported by Center for Research on Parallel Computations, Rice University, Houston, Texas.

†Research supported by The Research Council of Norway.

rate of component-wise sequence of the underlying vector sequence played an important role in analyzing the performance of the underlying method. One such example is in studying interior-point methods for complementarity problems in the absence of strict complementarity, see Monteiro and Wright (Ref. 2) and El-Bakry, Tapia, and Zhang (Ref. 3).

A natural question then arises: What can one infer about the convergence rate of a vector sequence provided that the convergence rates of the components' sequences are known? This question was partially answer by El-Bakry, Tapia, and Zhang (Ref. 3). In this paper we attempt to further investigate this question. We further demonstrate that the m -step convergence rate of a given sequence is ought to be seen naturally in a certain “augmented” space.

2 Component-wise Convergence Rates

In this section we prove the main result concerning the convergence rate of a vector sequence provided that the convergence rate of the components' sequences are known. Then we use this result to investigate the m -step convergence rate of sequences.

Definition 1 *Let $\{\alpha^k\}$ be a sequence of real non-negative numbers converging to 0. We say that $\{\alpha^k\}$ converges with Q -order at least $p \in [1, \infty)$ if there exist k_0 and non-negative number q such that*

$$\alpha^{k+1} \leq q (\alpha^k)^p \text{ for } k \geq k_0 \quad (1)$$

Moreover, if

$$\limsup_{k \rightarrow \infty} \frac{\alpha^{k+1}}{(\alpha^k)^p} = q, \quad (2)$$

then we say that $\{\alpha^k\}$ converges with Q -order p .

In the following, we will consider the ℓ_t norm unless otherwise specified. For a vector $a = (\alpha_1, \alpha_2, \dots, \alpha_n)^T \in \mathbb{R}^n$, the ℓ_t norm is defined as

$$\|a\|_t = \left(\sum_{i=1}^n |\alpha_i|^t \right)^{1/t} \text{ and } \|a\|_\infty = \max_{1 \leq i \leq n} \{|\alpha_i|\}.$$

Theorem 1 Consider the sequence $\{a^k\} \subset \mathbb{R}_+^n$, where $a^k = (\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k)^T$. Assume that the sequences $\{\alpha_1^k\}, \{\alpha_2^k\}, \dots, \{\alpha_n^k\}$ converge to zero with Q -orders at least p_1, p_2, \dots, p_n , respectively. Then we have :

1. The sequence $\{a^k\}$ converges to zero with Q -order at least

$$\hat{p} = \min_{1 \leq i \leq n} p_i,$$

in any norm in \mathbb{R}^n . Moreover, there exists K such that $\|a^{k+1}\|_t \leq \hat{q} \|a^k\|_t^{\hat{p}}$ for $k \geq K$, where

$$\hat{q} = \max\{q_i \mid p_i = \hat{p}, i = 1, 2, \dots, n\}.$$

2. If there exists at least one $j \in \{1, \dots, n\}$ such that $\{\alpha_j^k\}$ converges to zero with Q -order \hat{p} , then the sequence $\{a^k\}$ has the exact Q_p -factor $\hat{q} = \max\{q_i \mid p_i = \hat{p}, i = 1, 2, \dots, n\}$.

Proof: Since all the sequences $\{\alpha_i^k\}, i = 1, \dots, n$ converge to zero, then there exists some positive k_0^0 ,

$$\alpha_i^k < 1, \text{ for } i = 1, \dots, n \text{ and for all } k > k_0^0.$$

From the assumptions that the sequence $\{\alpha_i^k\}$ converges to zero with Q -order at least p_i , we have from (1) that

$$\alpha_i^{k+1} \leq q_i (\alpha_i^k)^{p_i}, \quad \text{for } k \geq k_i^0,$$

for some positive integer $k_i^0 \geq k_0^0$. Consider the sequence $\{\alpha_i^k\}$ for which $p_i > \hat{p}$ and let $K_i \geq k_i^0$ such that

$$q_i (\alpha_i^k)^{p_i - \hat{p}} \leq \hat{q} \text{ for } k \geq K_i. \quad (3)$$

This can be done since $\alpha_i^k \rightarrow 0$ as $k \rightarrow \infty$. Note that (3) holds for all i if we let $K_i = k_i^0$ for those i such that $p_i = \hat{p}$. Let $K = \max_{1 \leq i \leq n} K_i$ and let $k \geq K$.

We first consider the case with finite t .

$$\begin{aligned}
\|a^{k+1}\|_t^t &= \sum_{i=1}^n (\alpha_i^{k+1})^t \leq \sum_{i=1}^n (q_i (\alpha_i^k)^{p_i})^t = \sum_{i=1}^n (q_i (\alpha_i^k)^{p_i - \hat{p}})^t (\alpha_i^k)^{\hat{p}t} \\
&\leq \sum_{i=1}^n \hat{q}^t (\alpha_i^k)^{\hat{p}t} \\
&\leq \hat{q}^t \left[\sum_{i=1}^n (\alpha_i^k)^t \right]^{\hat{p}}
\end{aligned}$$

Hence

$$\|a^{k+1}\|_t \leq \hat{q} \|a^k\|_t^{\hat{p}}.$$

Consider the ℓ_∞ norm. Then

$$\begin{aligned}
\|\alpha^{k+1}\|_\infty &= \max_{i=1, \dots, n} |\alpha_i^{k+1}| = \alpha_j^{k+1} \\
&\leq q_j (\alpha_j^k)^{p_j} = q_j (\alpha_j^k)^{p_j - \hat{p}} (\alpha_j^k)^{\hat{p}} \\
&\leq \hat{q} (\max_{i=1, \dots, n} |\alpha_i^k|)^{\hat{p}} \\
&\leq \hat{q} \|\alpha^k\|_\infty^{\hat{p}}.
\end{aligned}$$

The convergence with Q-orders at least \hat{p} in any norm follows from the equivalence of norms in \mathbb{R}^n .

Now we turn our attention to prove the second part of the theorem. We have just shown that the Q_p -factor of $\{a^k\}$ is less than or equal to \hat{q} . We now show that it is exactly \hat{q} . To do this we need to show that there exists a subsequence of $\{a^k\}$ with Q_p factor equals to \hat{q} . We assume that all components have the same order; otherwise we may use the above procedure to show that components with higher order vanish in the limit. Without loss of generality, consider the case $\hat{p} = 1$. In this case we have

$$\hat{q} = \max_{1 \leq i \leq n} q_i.$$

Let \mathcal{J} be the set of indices such that

$$\limsup_{k \rightarrow \infty} \frac{\alpha_j^{k+1}}{\alpha_j^k} = \hat{q}, \quad j \in \mathcal{J}. \quad (4)$$

To simplify the proof, we assume that \mathcal{J} is a singleton. Let \mathcal{K} be a subsequence and let j be the index such that

$$\lim_{k \in \mathcal{K}} \frac{\alpha_j^{k+1}}{\alpha_j^k} = \hat{q}.$$

To prove the case when \mathcal{J} is not a singleton, we restrict the subsequence \mathcal{K} such that the above relation is satisfied for all $j \in \mathcal{J}$.

We want to show that the Q_1 -factor of $\{a^k\}$ is arbitrarily close to \hat{q} in any ℓ_t norm. We have two cases.

First assume that $\hat{q} > 0$. Let $\varepsilon > 0$ be such that

$$\varepsilon < \hat{q} - \max_{i \neq j} q_i.$$

First we show that α_i^k / α_j^k can be made arbitrarily small. Let $\varepsilon > \varepsilon_0 > 0$ and consider $k \in \mathcal{K}$ sufficiently large so that

$$\alpha_j^{k+1} \geq (\hat{q} - \varepsilon_0) \alpha_j^k.$$

Let

$$0 < \delta \leq 1 - \left(\frac{\hat{q} - \varepsilon}{\hat{q} - \varepsilon_0} \right)^t$$

and choose $k \in \mathcal{K}$ sufficiently large and $i \neq j$ so that

$$\frac{\alpha_i^k}{\alpha_j^k} \leq \left(\frac{q_i}{\hat{q} - \varepsilon_0} \right)^k \frac{\alpha_i^0}{\alpha_j^0} \leq \left(\frac{\delta}{n-1} \right)^{\frac{1}{t}}. \quad (5)$$

Consider

$$\begin{aligned} \|a^{k+1}\|_t^t &\geq (\alpha_j^{k+1})^t \geq (\hat{q} - \varepsilon_0)^t (\alpha_j^k)^t \\ &= (\hat{q} - \varepsilon_0)^t \left(\|a^k\|_t^t - \sum_{i \neq j} (\alpha_i^k)^t \right) \\ &\geq (\hat{q} - \varepsilon_0)^t (\|a^k\|_t^t - \delta (\alpha_j^k)^t) \\ &\geq (\hat{q} - \varepsilon_0)^t (1 - \delta) \|a^k\|_t^t \\ &\geq (\hat{q} - \varepsilon)^t \|a^k\|_t^t \end{aligned}$$

and we have shown that there exists a subsequence \mathcal{K} so that for k sufficiently large $k \in \mathcal{K}$ that

$$\frac{\|a^{k+1}\|_t}{\|a^k\|_t} \geq \hat{q} - \varepsilon.$$

If we consider the ℓ_∞ norm, then we see from (5) that

$$\|a^k\|_\infty = \alpha_j^k$$

for all k sufficiently large and $k \in \mathcal{K}$. Then

$$\|a^{k+1}\|_\infty \geq \alpha_j^{k+1} \geq (\hat{q} - \varepsilon_0)\alpha_j^k = (\hat{q} - \varepsilon_0)\|a^k\|_\infty \geq (\hat{q} - \varepsilon)\|a^k\|_\infty$$

Now assume that $\hat{q} = 0$. Suppose that the Q_1 factor of $\{\alpha^k\}$ is not zero. Then there exists a subsequence \mathcal{K} such that

$$\lim_{k \in \mathcal{K}} \frac{\|\alpha^{k+1}\|_t^t}{\|\alpha^k\|_t^t} = \bar{\varepsilon} > 0,$$

This implies that there exists $\tilde{\varepsilon} < \bar{\varepsilon}$ such that

$$\sum_{i=1}^n (\alpha_i^{k+1})^t > \tilde{\varepsilon} \sum_{i=1}^n (\alpha_i^k)^t. \quad (6)$$

Since $q_i = 0$ for all $i = 1, \dots, n$, then there exist sequences $\{c_i^k\}, i = 1, \dots, n$ converging to zero such that

$$\sum_{i=1}^n (\alpha_i^{k+1})^t \leq \sum_{i=1}^n c_i^k (\alpha_i^k)^t. \quad (7)$$

Combining (6) and (7), we obtain $\sum c_i^k (\alpha_i^k)^t > \tilde{\varepsilon} \sum_{i=1}^n (\alpha_i^k)^t$, which implies that $\sum (c_i^k - \tilde{\varepsilon})(\alpha_i^k)^t > 0$. Since $c_i^k \rightarrow 0$ for all $i = 1, \dots, n$, the last inequality leads to a contradiction and completes the proof. \square

The above theorem implies that both the convergence rate and Q-factor in the vector is determined by the slowest converging component.

2.1 Multi-Step Convergence Rate

Assume that the sequence $\{x^k\} \subset \mathbb{R}^n$ converges to x^* . We say that the sequence $\{x^k\}$ has an m -step Q-order at least p if for some positive integer k_0 we have

$$\|x^{k+m} - x^*\| \leq Q_* \|x^k - x^*\|^p, \quad (8)$$

for $k \geq k_0$. The main result of this section is the following. If we consider a certain sequence, composed of elements of $\{x^k\}$, then this sequence has a 1-step Q-order at least p .

Choose a positive integer j_0 such that $m(j_0 - 1) + 1 \geq k_0$ and define the sequence $\{y^j\}_{j=j_0}^\infty$ in $\mathbb{R}^{n'}$, where $n' = m \cdot n$, by stacking every m consequent elements of the sequence $\{x^k\}$ together. Thus

$$y^j = \begin{pmatrix} x^{m(j-1)+1} \\ \vdots \\ x^{mj} \end{pmatrix}, \quad j \geq j_0. \quad (9)$$

Furthermore, define the vector

$$y^* = \begin{pmatrix} x^* \\ \vdots \\ x^* \end{pmatrix} \in \mathbb{R}^{n'}. \quad (10)$$

If $\{x^k\} \subset \mathbb{R}^n$ converges to x^* then the sequence $\{y^j\}$ converges to y^* . We now investigate the convergence rate of the sequence $\{y^j\}$. If the norm used in (8) is the ℓ_1 norm define $\alpha_i^j = \|x^{m(j-1)+i} - x^*\|_1$ for $i = 1, \dots, m$ and the m -vector $a^j = (\alpha_1^j, \dots, \alpha_m^j)^T$. It follows that

$$\|y^{j+1} - y^*\|_1 = \|a^{j+1}\|_1 \leq Q_* \|a^j\|_1^p = \|y^j - y^*\|_*^p$$

using Theorem 1. We will now show that if we have m -step Q-order at least p in some norm in \mathbb{R}^n then the sequence $\{y^j\}$ converges to y^* with Q-order at least p in any $\mathbb{R}^{n'}$ -norm.

Theorem 2 *Assume that the sequence $\{x^k\} \subset \mathbb{R}^n$ converges to x^* with m -step Q-order at least p . Then the sequence $\{y^j\} \subset \mathbb{R}^{n'}$ defined by (9), converges to y^* with Q-order at least p in any $\mathbb{R}^{n'}$ -norm.*

Proof : Let $\alpha_i^j = \|x^{m(j-1)+i} - x^*\|$ for $i = 1, \dots, m$. Then from the assumption that the sequence $\{x^k\}$ has an m -step Q-order at least p

$$\alpha_i^{j+1} \leq Q_* (\alpha_i^j)^p.$$

Define the m -vector

$$a^j = (\alpha_1^j, \dots, \alpha_m^j)^T.$$

From Theorem 1 we know that the sequence $\{a^j\} \subset \mathbb{R}^m$ converges to zero with Q-order at least p and in ant ℓ_t norm we have

$$\|a^{j+1}\|_t \leq Q_* \|a^j\|_t.$$

For $z \in \mathbb{R}^{n'}$ and the norms $\|\cdot\|_t, \|\cdot\|$, we can define a norm in $\mathbb{R}^{n'}$ in the following manner. Partition z into m -projections z_i , $i = 1, \dots, m$ each in \mathbb{R}^n . Define $a(z) \in \mathbb{R}^m$ by

$$a(z) = (\|z_1\|, \dots, \|z_m\|)^T.$$

Then the function $\|z\|_* = \|a(z)\|_t$ defines a norm in $\mathbb{R}^{n'}$. Thus

$$\|y^{j+1} - y^*\|_* \leq Q_* \|y^j - y^*\|_*^p.$$

Convergence with Q-order at least p in any norm follows from the equivalence of norms. \square

References

1. ORTEGA, J. M., and RHEINBOLDT, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, New York 1970.
2. MONTEIRO, E. C., and WRIGHT, S., *Local Convergence of Interior-Point Algorithms for Degenerate Monotone LCP Problems*, Computational Optimization and Applications, Vol. 3, pp. 131–155, 1994.
3. EL-BAKRY, A. S., TAPIA, R. A., and ZHANG, Y., *On the Convergence Rate of Newton Interior-Point Methods in the Absence of Strict Complementarity*, Computational Optimization and Applications, Vol. 6, pp. 157–167, 1996.