On the Matrix Norm
Subordinate to the Hölder Norm

J. Albrecht and P. P. Klein

Dedicated to Prof. L. von Wolfersdorf on the occasion of his retirement

Abstract. For non-negative matrices $P$ the matrix norm subordinate to the Hölder norm of index $p$ with $p \in (1, \infty)$ is determined by an eigenvalue problem $T\alpha = \lambda\alpha$, where $T$ is a homogeneous, strongly monotone operator.

Keywords: Hölder vector norms, subordinate matrix norms, non-negative matrices

AMS subject classification: Primary 15A60, 15A18, secondary 47H07

1. Introduction

Assume $v \in \mathbb{R}^n$ and $M \in \mathbb{R}^{m \times n}$. For the Hölder vector norm

$$||v||_p = \left\{ \begin{array}{ll}
\left[ \sum_{i=1}^{n} |v_i|^p \right]^{1/p} & \text{for } 1 \leq p < \infty \\
\max_{i=1, \ldots, n} |v_i| & \text{for } p = \infty
\end{array} \right.$$ 

the subordinate matrix norm

$$||M||_p = \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{||Mv||_p}{||v||_p} \quad (1 \leq p \leq \infty)$$

can be easily calculated in the limiting cases:

$$||M||_1 = \max_{j=1, \ldots, n} \sum_{i=1}^{m} |m_{ij}| \quad \text{and} \quad ||M||_{\infty} = \max_{i=1, \ldots, m} \sum_{j=1}^{n} |m_{ij}|.$$ 

Furthermore, the spectral norm is well known:

$$||M||_2 = \left[ \rho(M^T M) \right]^{1/2}.$$

Beyond that in the special case of non-negative matrices $P \in \mathbb{R}_+^{m \times n}$ for all $p \in (1, \infty)$ the matrix norm $||P||_p$ can be determined by an eigenvalue problem, which is nonlinear for $p \neq 2$. 


ISSN 0232-2064 / $2.50 \quad \circledast$ Heldermann Verlag Berlin
2. The eigenvalue problem

Let \( P \in \mathbb{R}^{m \times n}_+ \), \( p \in (1, \infty) \) and \((p - 1)(q - 1) = 1\). Because of \(|Pv| \leq P|v|\) for \( v \in \mathbb{R}^n\),

\[
\|P\|_p = \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\|Pv\|_p}{\|v\|_p} \tag{1}
\]

holds. Discussing this maximum problem leads to

Definition 1.

\[
T : \mathbb{R}^n_+ \to \mathbb{R}^n_+, \quad (Tv)_j = \left[ \sum_{i=1}^m p_{ij}(Pv)_i^{p-1} \right]^{q-1} \quad (j = 1, \ldots, n) \tag{2}
\]

and

Theorem 1. Assume that the eigenvalue problem

\[
Ta = \lambda a \tag{3}
\]

has an eigenvector \( a \) with positive components only, corresponding to a positive eigenvalue \( \lambda \). Then

\[
\|P\|_p = \lambda^{1/q} \tag{4}
\]

Proof. 1.1 In the case \((Pa)_i > 0\), for \( v \in \mathbb{R}^n_+\),

\[
(Pv)_i = \sum_{j=1}^n p_{ij}v_j = \sum_{j=1}^n p_{ij} \frac{v_j}{\alpha_j} = (Pa)_i \sum_{j=1}^n p_{ij} \frac{v_j}{\alpha_j} \tag{5}
\]

holds and Hölder’s inequality for convex functions \( \varphi \) (see [6, 8])

\[
\varphi \left( \sum_j p_j \frac{v_j}{\alpha_j} \right) \leq \sum_j p_j \varphi \left( \frac{v_j}{\alpha_j} \right) \tag{6}
\]

yields

\[
(Pv)_i^p \leq (Pa)_i^p \sum_{j=1}^n p_{ij} \frac{v_j}{\alpha_j} \sum_{j=1}^n p_{ij} \frac{v_j}{\alpha_j} = (Pa)_i^{p-1} \sum_{j=1}^n \frac{p_{ij}}{\alpha_j^{p-1}} v_j^p. \tag{7}
\]

1.2 In the case \((Pa)_i = 0\), because of \( \alpha_j > 0 \) \((j = 1, \ldots, n)\), \( p_{ij} = 0 \) \((j = 1, \ldots, n)\) holds and therefore \((Pv)_i = 0\) is valid for all \( v \in \mathbb{R}^n_+\).

2. Hence it follows that

\[
\sum_{i=1}^m (Pv)_i^p \leq \sum_{i=1}^m \sum_{j=1}^n \frac{(Pa)_i^{p-1}}{\alpha_j^{p-1}} v_j^p = \sum_{i=1}^m \frac{(T\alpha)_i^{p-1}}{\alpha_j^{p-1}} v_j^p = \lambda^{p-1} \sum_{j=1}^n v_j^p
\]

and

\[
\|Pv\|_p \leq \lambda^{1/q} \|v\|_p. \tag{8}
\]

If \( v = a \), then equality holds.

The theorem is illustrated by the following

Example (\( f \in \mathbb{R}^m_+ \), \( g \in \mathbb{R}^n_+ \)).

\[
P = fg^T : \quad \alpha = (g_i^{p-1})_{i=1}^n, \quad \|P\|_p = \|f\|_p \|g\|_q.
\]

The assumption that the eigenvalue problem \( Ta = \lambda a \) has an eigenvector \( a \) with positive components only, corresponding to a positive eigenvalue \( \lambda \), will be shown to be fulfilled if \( P^TP \) is irreducible.
3. $P^TP$ irreducible

In a real linear space $X$ let the cone $K$ define the partial ordering $\leq$. Eigenvalue problems with operators $T : K \to K$ having the properties

1. $T$ is monotone on $K$, i.e. $u, v \in K$ with $u \leq v$ implies $Tu \leq Tv$

2. $T$ is homogeneous on $K$, i.e. $T(cv) = cTv$ for $c \geq 0$ and $v \in K$

3. $T$ is completely continuous on $K$

have been investigated by Krein and Rutman [7] and by Bohl [2]. The results in [2] necessitate another assumption, namely that $T$ is strongly monotone on $K$. In the case $X = \mathbb{R}^n$, $K = \mathbb{R}^+_n$ this means the following.

**Definition 2.** An operator $T$ being monotone on $\mathbb{R}^+_n$ is called strongly monotone on $\mathbb{R}^+_n$, if for all $v, w \in \mathbb{R}^+_n$ with $v \leq w$ and $v \neq w$ there exists a number $\mu \in \mathbb{N}$ such that

$$(T^\mu v)_j < (T^\mu w)_j \quad (j = 1, \ldots, n)$$

holds.

By the following lemma the strong monotonicity of the operator $T$ defined in (2) can be concluded from the strong monotonicity of $P^TP$.

**Lemma 1.** Assume $P \in \mathbb{R}^{m \times n}_+$ and $p \in (1, \infty)$. For arbitrary vectors $v, w \in \mathbb{R}^+_n$ with $v \leq w$, all $\nu \in \mathbb{N}$ and each fixed $j \in \{1, \ldots, n\}$ the equivalence

$$(T^\nu v)_j = (T^\nu w)_j \iff ((P^T P)^\nu v)_j = ((P^T P)^\nu w)_j.$$  (5)

holds.

**Proof.** 1. $\nu = 1$: Let $j \in \{1, \ldots, n\}$ be fixed. Then $(Tv)_j = (T^1 w)_j$ is equivalent to

$$\sum_{i=1}^{m} p_{ij} ((Pw)_i^{p-1} - (Pv)_i^{p-1}) = 0.$$  (6)

As $v \leq w$ implies $Pv \leq Pw$ and $(Pv)_i^{p-1} \leq (Pw)_i^{p-1}$ ($i = 1, \ldots, m$), all terms of (6) are non-negative. For every $i \in \{1, \ldots, m\}$ with $p_{ij} > 0$ equation (6) requires that $(Pw)_i^{p-1} - (Pv)_i^{p-1} = 0$, yielding $(P(w - v))_i = 0$. Therefore

$$\sum_{i=1}^{m} p_{ij}(P(w - v))_i = 0.$$  (7)

follows and thus $(P^T P v)_j = (P^T P w)_j$ holds. Analogously (6) can be deduced from (7).

2. Induction from $\nu$ to $\nu + 1$: Let $j \in \{1, \ldots, n\}$ be fixed. As $T$ is monotone, $v \leq w$ implies $T^\nu v \leq T^\nu w$. Define $\tilde{v} = T^\nu v$ and $\tilde{w} = T^\nu w$. Using (5) with $\nu = 1$ leads to

$$(Tv)_j = (T\tilde{w})_j \iff (P^T P \tilde{v})_j = (P^T P \tilde{w})_j.$$
\[(T^{\nu+1}v)_j = (T^{\nu+1}w)_j \text{ is equivalent to} \]
\[
\sum_{k=1}^{n} (P^TP)_{jk} (T^{\nu}w - T^{\nu}v)_k = 0. \tag{8}
\]

As all terms in (8) are non-negative, \((T^{\nu}w - T^{\nu}v)_k = 0\) holds for every \(k \in \{1, \ldots, n\}\) with \((P^TP)_{jk} > 0\). Since (5) is assumed to be true for \(\nu\),
\[
\sum_{k=1}^{n} (P^TP)_{jk} ((P^TP)^\nu(w - v))_k = 0 \tag{9}
\]
follows and thus \(((P^TP)^{\nu+1}v)_j = ((P^TP)^{\nu+1}w)_j\) is obtained. In the same way (8) can be concluded from (9).

**Theorem 2.** Assume \(P \in \mathbb{R}_{++}^{m \times n}, p \in (1, \infty)\) and \(P^TP\) irreducible. Then:

1. \(P^TP\) and \(T\) are strongly monotone on \(K\).

2. The eigenvalue problem \(T\alpha = \lambda\alpha\) has an eigenvector \(\alpha\) with positive components only, corresponding to a positive eigenvalue \(\lambda\).

**Proof.**

1. All diagonal elements of \(P^TP\) are positive. Assuming the contrary, namely that \((P^TP)_{jj} = 0\) for at least one \(j \in \{1, \ldots, n\}\), all elements of the \(j\)-th column of \(P\) would be zero. This would imply \(P^TP\) to be reducible, in contradiction to the assumption.

Since \(P^TP\) is irreducible and as its diagonal elements are positive, [2: p. 111/Theorem 2.3] says that \((P^TP)^{n-1}\) consists of positive elements only, i.e. \(P^TP\) is strongly monotone. Using Lemma 1 for \(\nu = n - 1\) proves that \(T\) is strongly monotone as well.

2. As the operator \(T\) is completely continuous and strongly monotone, by [2: p. 53/Theorem 2.7] with \(S = T\), \(T\) has an eigenvector \(\alpha\) with positive components only and a corresponding positive eigenvalue \(\lambda\).

**Example.** Doubly stochastic matrices, e.g.

\[
P = \frac{1}{15} \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix} : \quad \alpha = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \|P\|_p = 1.
\]

**Theorem 3.** Assume \(P \in \mathbb{R}_{++}^{m \times n}, p \in (1, \infty)\) and \(P^TP\) irreducible. Starting from \(\alpha^{(1)} \in \mathbb{R}_+^n\) having positive components only, the iterates \(\alpha^{(k+1)}\) defined by

\[
\alpha^{(k+1)} := T\alpha^{(k)} \quad (k \in \mathbb{N}) \tag{10}
\]

have the same property. With

\[
\lambda^{(k)} := \min_{j=1, \ldots, n} \frac{\alpha^{(k+1)}_j}{\alpha^{(k)}_j} \quad \text{and} \quad \overline{\lambda}^{(k)} := \max_{j=1, \ldots, n} \frac{\alpha^{(k+1)}_j}{\alpha^{(k)}_j} \quad (k \in \mathbb{N}) \tag{11}
\]
On the Matrix Norm Subordinate to the Hölder Norm

The eigenvalue inclusion

\[ \lambda^{(1)} \leq \ldots \leq \lambda^{(k)} \leq \lambda^{(k+1)} \leq \ldots \leq \lambda \leq \ldots \leq \lambda^{(k+1)} \leq \lambda^{(k)} \leq \ldots \leq \lambda^{(1)} \]  

is obtained. Furthermore,

\[ \lim_{k \to \infty} \lambda^{(k)} = \lambda = \lim_{k \to \infty} \lambda^{(k)} \]  

holds.

Proof. The monotonicity and the convergence of the sequences \( \{\lambda^{(k)}\}_{k \in \mathbb{N}} \) and \( \{\lambda^{(k)}\}_{k \in \mathbb{N}} \) follow from [2, p. 53/Theorem 2.7] as well \footnote{Remark. For \( p = 2 \) Theorem 3 reduces to the inclusion theorem of Collatz [3] for non-negative irreducible matrices applied to \( P^T P \).

4. \( P^T P \) reducible

Allowing \( P^T P \) to be reducible, it may be assumed that \( P^T P \) already has the normal block diagonal form of symmetric reducible matrices [9]. Otherwise the columns of \( P \) have to be permuted appropriately, which implies the same permutations for the rows of \( P^T \) and thus results in the normal form of \( P^T P \). Permuting the columns of \( P \) has no effect on \( \|P\|_p \).

According to the number and the sizes of the diagonal submatrices of \( P^T P \), the matrix \( P \in \mathbb{R}^{m \times n} \) is split up into column blocks

\[ P = (P_1, \ldots, P_s) \quad \text{with} \quad P_\sigma \in \mathbb{R}^{m \times n_\sigma} \quad (\sigma = 1, \ldots, s). \]  

Correspondingly, a vector \( v \in \mathbb{R}^n \) is decomposed as

\[ v = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \quad \text{with} \quad v_\sigma \in \mathbb{R}^{n_\sigma} \quad (\sigma = 1, \ldots, s). \]  

The block structure of \( P^T P \) implies

\[ P_\rho^T P_\sigma = \Theta_\rho \sigma \in \mathbb{R}_{+}^{n_\rho \times n_\sigma} \quad (\rho \neq \sigma; \ \rho, \sigma = 1, \ldots, s) \]

which means that each non-zero row of \( P \) has non-zero elements exactly in one column block of \( P \). Therefore, taking notice of (15),

\[ \|Pv\|_p^p = \sum_{\sigma = 1}^{s} \|P_\sigma v_\sigma\|_p^p \]  

(\( v \in \mathbb{R}_+^n \))

holds.
Theorem 4. Assume $P \in \mathbb{R}_+^{m \times n}$ and $p \in (1, \infty)$. Let $P^TP$ be reducible such that

$$P^TP = \text{diag}(P_1^TP_1, \ldots, P_s^TP_s)$$

(17)

and assume each diagonal submatrix $P_{\sigma}^T P_{\sigma}$ ($\sigma = 1, \ldots, s$) to be irreducible. Consequently, the eigenvalue problem (2) is split up into subproblems of the same type

$$T_{\sigma} \alpha_{\sigma} = \lambda_{\sigma} \alpha_{\sigma} \quad (\sigma = 1, \ldots, s)$$

(18)

where each $T_{\sigma} : \mathbb{R}_+^{k_{\sigma}} \rightarrow \mathbb{R}_+^{k_{\sigma}}$ results from (2) with $P_{\sigma}$ instead of $P$. Then

$$\|P\|_p = \lambda^{1/p} \quad \text{with} \quad \lambda = \max_{\sigma = 1, \ldots, s} \lambda_{\sigma}$$

(19)

holds.

Proof. For each eigenvalue problem (18) Theorem 2 guarantees the existence of an eigenvector $\alpha_{\sigma}$ with positive components only, corresponding to a positive eigenvalue $\lambda_{\sigma}$. Therefore Theorem 1 ensures

$$\|P_{\sigma} v_\sigma\|_p \leq \lambda_{\sigma}^{p-1} \|v_\sigma\|_p \quad (v_\sigma \in \mathbb{R}_+^{k_{\sigma}}, \sigma = 1, \ldots, s)$$

with equality, if $v_\sigma = \alpha_{\sigma}$ ($\sigma = 1, \ldots, s$). For $v \in \mathbb{R}_+^n$, using (16),

$$\|Pv\|_p = \sum_{\sigma = 1}^s \|P_{\sigma} v_\sigma\|_p \leq \sum_{\sigma = 1}^s \lambda_{\sigma}^{p-1} \|v_\sigma\|_p \leq \lambda^{p-1} \sum_{\sigma = 1}^s \|v_\sigma\|_p = \lambda^{p-1} \|v\|_p$$

follows, implying

$$\|Pv\|_p \leq \lambda^{1/p} \|v\|_p.$$

Equality holds, if $v$ satisfies

$$v_\sigma = \begin{cases} \alpha_{\sigma} & \text{for } \lambda_{\sigma} = \lambda \\ \theta_{\sigma} & \text{for } \lambda_{\sigma} < \lambda \end{cases} \quad (\sigma = 1, \ldots, s)$$

Remark. Since permuting the rows of $P$ leaves $P^TP$ as well as $\|P\|_p$ unchanged, additional splittings of $P \in \mathbb{R}_+^{m \times n}$ into row blocks can be obtained such that

$$P = (P_{\rho \sigma}) \quad \text{with} \quad P_{\rho \sigma} \in \mathbb{R}_+^{m_{\rho} \times n_{\sigma}} \quad (\rho, \sigma = 1, \ldots, s)$$

and, with $\pi$ denoting any permutation of $\{1, \ldots, s\}$, each column block $P_{\sigma}$ has exactly one non-zero subblock $P_{\pi(\sigma) \sigma}$ ($\sigma = 1, \ldots, s$).

Example ($f \in \mathbb{R}_+^{m-1}, g \in \mathbb{R}_+^{n-1}$).

$$P = \begin{pmatrix} \Theta & f \\ g^T & 0 \end{pmatrix} \quad \|P\|_p = \max \{\|f\|_p, \|g\|_q\}.$$

Theorem 4 is supplemented by the following.

Remark. Allowing $P^TP$ to have a zero diagonal submatrix $P_{\sigma}^T P_{\sigma}$ resulting from a zero column block $P_{\rho \sigma}$, then $T_{\sigma}$ is the zero operator with the eigenvalue $\lambda_{\sigma} = 0$. This leaves the result of Theorem 4 unchanged.
5. Numerical example

Applying discretization methods to boundary value problems with partial differential equations, often leads to linear systems

\[ v = P v + r \]  \hspace{1cm} (20)

with non-negative matrices \( P \). If \( P \) is symmetric, \( \rho(P) = \|P\|_2 \leq \|P\|_p \) for \( 1 \leq p \leq \infty \) holds. In case \( P \) is non-symmetric, however, \( p^* \) with \( \|P\|_p^* = \min\{\|P\|_p | 1 \leq p \leq \infty\} \) is generally not known in advance.

Applying the finite difference method to the boundary value problem [5]

\[-(u_{xx} + u_{yy} + \frac{3}{5-y} u_y) = 1 \quad \text{in} \quad B = (-\frac{1}{2}, \frac{1}{2}) \times (-1,1) \]

\[ u = 0 \quad \text{on} \quad \partial B \]

red-black ordering of the unknowns generates linear systems (20) with \( P \) non-symmetric, non-negative and \( P^T P \) reducible:

\[ P = \begin{pmatrix} \Theta_{11} & P_{12} \\ P_{21} & \Theta_{22} \end{pmatrix}, \quad \text{and} \quad P^T P = \begin{pmatrix} P_{21}^T P_{21} & \Theta_{12} \\ \Theta_{21} & P_{12}^T P_{12} \end{pmatrix}. \]  \hspace{1cm} (21)

For different mesh widths \( h \) the following results were obtained by discretely minimizing \( \|P\|_p \) with respect to \( p \) in a finite interval:

<table>
<thead>
<tr>
<th>( h )</th>
<th>( n )</th>
<th>( \rho(P) )</th>
<th>( p^{**} )</th>
<th>( |P|_{p^{**}} )</th>
<th>( |P|_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{6} )</td>
<td>33</td>
<td>0.91496</td>
<td>2.71</td>
<td>0.94058</td>
<td>0.94608</td>
</tr>
<tr>
<td>( \frac{1}{8} )</td>
<td>60</td>
<td>0.95175</td>
<td>2.99</td>
<td>0.97062</td>
<td>0.97689</td>
</tr>
<tr>
<td>( \frac{1}{12} )</td>
<td>138</td>
<td>0.97843</td>
<td>3.62</td>
<td>0.99003</td>
<td>0.99690</td>
</tr>
<tr>
<td>( \frac{1}{16} )</td>
<td>248</td>
<td>0.98784</td>
<td>4.38</td>
<td>0.99587</td>
<td>1.00289</td>
</tr>
<tr>
<td>( \frac{1}{24} )</td>
<td>564</td>
<td>0.99469</td>
<td>6.79</td>
<td>0.99915</td>
<td>1.00632</td>
</tr>
<tr>
<td>( \frac{1}{32} )</td>
<td>1008</td>
<td>0.99696</td>
<td>12.6</td>
<td>0.99986</td>
<td>1.00719</td>
</tr>
</tbody>
</table>

Table 1: Discrete minimization of \( \|P\|_p \)

Rewriting the boundary value problem in self-adjoint form [5]

\[-\left( \frac{1}{(5-y)^3} u_x \right)_x + \left( \frac{1}{(5-y)^3} u_y \right)_y = \frac{1}{(5-y)^3} \quad \text{in} \quad B, \]

\[ u = 0 \quad \text{on} \quad \partial B \]

and applying the finite difference method with red-black ordering of the unknowns again, linear systems (20), (21) are obtained, where \( P \) now is symmetric and non-negative. The spectral radii \( \rho(P) \) in this case are slightly above those given in Table 1.
References


Received 02.06.1998