

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 15 A 60, 15 A 18, secondary 47 H 07

1. Introduction

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

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

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:

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

Furthermore, the spectral norm is well known:

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

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 $p \neq 2$.

J. Albrecht: Techn. Univ. Clausthal, Inst. Math., Erzstr. 1, D-38678 Clausthal-Zellerfeld

P. P. Klein: Techn. Univ. Clausthal, Rech.-Zentrum, Erzstr. 51, D-38678 Clausthal-Zellerfeld

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.

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

and

Theorem 1. Assume that the eigenvalue problem

$$T\alpha = \lambda\alpha \tag{3}$$

has an eigenvector α with positive components only, corresponding to a positive eigenvalue λ . Then

$$\|P\|_p = \lambda^{1/q}. \tag{4}$$

Proof. 1.1 In the case $(P\alpha)_i > 0$, for $v \in \mathbb{R}_+^n$,

$$(Pv)_i = \sum_{j=1}^n p_{ij} v_j = \sum_{j=1}^n p_{ij} \alpha_j \frac{v_j}{\alpha_j} = (P\alpha)_i \frac{\sum_{j=1}^n p_{ij} \alpha_j \frac{v_j}{\alpha_j}}{\sum_{j=1}^n p_{ij} \alpha_j}$$

holds and Hölder's inequality for convex functions φ (see [6, 8])

$$\varphi \left(\frac{\sum_j p_j t_j}{\sum_j p_j} \right) \leq \frac{\sum_j p_j \varphi(t_j)}{\sum_j p_j}$$

yields

$$(Pv)_i^p \leq (P\alpha)_i^p \frac{\sum_{j=1}^n p_{ij} \alpha_j \left(\frac{v_j}{\alpha_j}\right)^p}{\sum_{j=1}^n p_{ij} \alpha_j} = (P\alpha)_i^{p-1} \sum_{j=1}^n \frac{p_{ij}}{\alpha_j^{p-1}} v_j^p.$$

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

2. Hence it follows that

$$\sum_{i=1}^m (Pv)_i^p \leq \sum_{j=1}^n \frac{\sum_{i=1}^m p_{ij} (P\alpha)_i^{p-1}}{\alpha_j^{p-1}} v_j^p = \sum_{j=1}^n \frac{(T\alpha)_j^{p-1}}{\alpha_j^{p-1}} v_j^p = \lambda^{p-1} \sum_{j=1}^n v_j^p$$

and

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

If $v = \alpha$, then equality holds ■

The theorem is illustrated by the following

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

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

The assumption that the eigenvalue problem $T\alpha = \lambda\alpha$ has an eigenvector α with positive components only, corresponding to a positive eigenvalue λ , will be shown to be fulfilled if $P^T P$ is irreducible.

3. $P^T P$ irreducible

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

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

holds.

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

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 $j \in \{1, \dots, n\}$ the equivalence

$$(T^\nu v)_j = (T^\nu w)_j \iff ((P^T P)^\nu v)_j = ((P^T P)^\nu w)_j \tag{5}$$

holds.

Proof. 1. $\nu = 1$: Let $j \in \{1, \dots, n\}$ be fixed. Then $(Tv)_j = (Tw)_j$ is equivalent to

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

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

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

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

2. Induction from ν to $\nu + 1$: Let $j \in \{1, \dots, n\}$ 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

$$(T\tilde{v})_j = (T\tilde{w})_j \iff (P^T P\tilde{v})_j = (P^T P\tilde{w})_j$$

i.e. $(T^{\nu+1}v)_j = (T^{\nu+1}w)_j$ is equivalent to

$$\sum_{k=1}^n (P^T P)_{jk} (T^\nu w - T^\nu v)_k = 0. \tag{8}$$

As all terms in (8) are non-negative, $(T^\nu w - T^\nu v)_k = 0$ holds for every $k \in \{1, \dots, n\}$ with $(P^T P)_{jk} > 0$. Since (5) is assumed to be true for ν ,

$$\sum_{k=1}^n (P^T P)_{jk} ((P^T P)^\nu (w - v))_k = 0 \tag{9}$$

follows and thus $((P^T P)^{\nu+1}v)_j = ((P^T P)^{\nu+1}w)_j$ 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 $P^T P$ irreducible. Then:

1. $P^T P$ and T are strongly monotone on K .
2. The eigenvalue problem $T\alpha = \lambda\alpha$ has an eigenvector α with positive components only, corresponding to a positive eigenvalue λ .

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

Since $P^T P$ is irreducible and as its diagonal elements are positive, [2: p. 111/Theorem 2.3] says that $(P^T P)^{n-1}$ consists of positive elements only, i.e. $P^T P$ 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 α with positive components only and a corresponding positive eigenvalue λ ■

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 $P^T P$ 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

$$\underline{\lambda}^{(k)} := \min_{j=1, \dots, n} \frac{\alpha_j^{(k+1)}}{\alpha_j^{(k)}} \quad \text{and} \quad \bar{\lambda}^{(k)} := \max_{j=1, \dots, n} \frac{\alpha_j^{(k+1)}}{\alpha_j^{(k)}} \quad (k \in \mathbb{N}) \tag{11}$$

the eigenvalue inclusion

$$\underline{\lambda}^{(1)} \leq \dots \leq \underline{\lambda}^{(k)} \leq \underline{\lambda}^{(k+1)} \leq \dots \leq \lambda \leq \dots \leq \bar{\lambda}^{(k+1)} \leq \bar{\lambda}^{(k)} \leq \dots \leq \bar{\lambda}^{(1)} \tag{12}$$

is obtained. Furthermore,

$$\lim_{k \rightarrow \infty} \underline{\lambda}^{(k)} = \lambda = \lim_{k \rightarrow \infty} \bar{\lambda}^{(k)} \tag{13}$$

holds.

Proof. The monotonicity and the convergence of the sequences $\{\underline{\lambda}^{(k)}\}_{k \in \mathbb{N}}$ and $\{\bar{\lambda}^{(k)}\}_{k \in \mathbb{N}}$ follow from [2, p. 53/Theorem 2.7] as well ■

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

$$P = (P_1, \dots, P_s) \quad \text{with} \quad P_\sigma \in \mathbb{R}_+^{m \times n_\sigma} \quad (\sigma = 1, \dots, s). \tag{14}$$

Correspondingly, a vector $v \in \mathbb{R}_+^n$ is decomposed as

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_s \end{pmatrix} \quad \text{with} \quad v_\sigma \in \mathbb{R}_+^{n_\sigma} \quad (\sigma = 1, \dots, s). \tag{15}$$

The block structure of $P^T P$ implies

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

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),

$$\|Pv\|_p^p = \sum_{\sigma=1}^s \|P_\sigma v_\sigma\|_p^p \quad (v \in \mathbb{R}_+^n) \tag{16}$$

holds.

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

$$P^T P = \text{diag}(P_1^T P_1, \dots, P_s^T P_s) \tag{17}$$

and assume each diagonal submatrix $P_\sigma^T P_\sigma$ ($\sigma = 1, \dots, s$) to be irreducible. Consequently, the eigenvalue problem (3) is split up into subproblems of the same type

$$T_\sigma \alpha_\sigma = \lambda_\sigma \alpha_\sigma \quad (\sigma = 1, \dots, s) \tag{18}$$

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

$$\|P\|_p = \lambda^{1/q} \quad \text{with} \quad \lambda = \max_{\sigma=1, \dots, s} \lambda_\sigma \tag{19}$$

holds.

Proof. For each eigenvalue problem (18) Theorem 2 guarantees the existence of an eigenvector α_σ with positive components only, corresponding to a positive eigenvalue λ_σ . Therefore Theorem 1 ensures

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

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

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

follows, implying

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

Equality holds, if v satisfies

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

■

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

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

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

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 $P^T P$ to have a zero diagonal submatrix $P_{\sigma\sigma}^T P_{\sigma\sigma}$ resulting from a zero column block $P_{\sigma\cdot}$, then $T_{\sigma\cdot}$ is the zero operator with the eigenvalue $\lambda_{\sigma\cdot} = 0$. 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

$$v = Pv + r \tag{20}$$

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 $\|P\|_{p^*} = \min\{\|P\|_p \mid 1 \leq p \leq \infty\}$ is generally not known in advance.

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

$$\left. \begin{aligned} -\left(u_{xx} + u_{yy} + \frac{3}{5-y}u_y\right) &= 1 & \text{in } B = \left(-\frac{1}{2}, \frac{1}{2}\right) \times (-1, 1) \\ u &= 0 & \text{on } \partial B \end{aligned} \right\}$$

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

$$P = \begin{pmatrix} \Theta_{11} & P_{12} \\ P_{21} & \Theta_{22} \end{pmatrix}, \quad \text{and} \quad P^T P = \begin{pmatrix} P_{21}^T P_{21} & \Theta_{12} \\ \Theta_{21} & P_{12}^T P_{12} \end{pmatrix}. \tag{21}$$

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

h	n	$\rho(P)$	p^{**}	$\ P\ _{p^{**}}$	$\ P\ _2$
$\frac{1}{6}$	33	0.91496	2.71	0.94058	0.94608
$\frac{1}{8}$	60	0.95175	2.99	0.97062	0.97689
$\frac{1}{12}$	138	0.97843	3.62	0.99003	0.99690
$\frac{1}{16}$	248	0.98784	4.38	0.99587	1.00289
$\frac{1}{24}$	564	0.99459	6.79	0.99915	1.00632
$\frac{1}{32}$	1008	0.99696	12.6	0.99986	1.00719

Table 1: Discrete minimization of $\|P\|_p$

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

$$\left. \begin{aligned} -\left(\left[\frac{1}{(5-y)^3}u_x\right]_x + \left[\frac{1}{(5-y)^3}u_y\right]_y\right) &= \frac{1}{(5-y)^3} & \text{in } B, \\ u &= 0 & \text{on } \partial B \end{aligned} \right\}$$

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

- [1] Albrecht, J.: *Minimal norms of nonnegative irreducible matrices*. Lin. Alg. Appl. 249 (1996), 255 – 258.
- [2] Bohl, E.: *Monotonie: Lösbarkeit und Numerik bei Operatorgleichungen*. Berlin: Springer Verlag 1974.
- [3] Collatz, L.: *Einschließungssatz für die charakteristischen Zahlen von Matrizen*. Math.Z. 48 (1942), 221 – 226.
- [4] Collatz, L.: *Funktionalanalysis und Numerische Mathematik*. Berlin: Springer Verlag 1964.
- [5] Collatz, L.: *The Numerical Treatment of Differential Equations*. Berlin: Springer Verlag 1966.
- [6] Hölder, L. O.: *Über einen Mittelwertsatz*. Nachr. Ges. Wiss. Göttingen 1889, 38 – 47.
- [7] Krein, M. G. and M. A. Rutman: *Linear operators leaving invariant a cone in a Banach space*. AMS Transl. (Series One) 10 (1962), 307 – 320.
- [8] Polya, G. and G. Szegö: *Problems and Theorems in Analysis* (Vol. 1). New York: Springer Verlag 1978.
- [9] Varga, R. S.: *Matrix Iterative Analysis*. Englewood Cliffs: Prentice Hall Inc. 1962.

Received 02.06.1998