ON THE HÖLDER CONTINUITY OF MATRIX FUNCTIONS FOR NORMAL MATRICES

THOMAS P. WIHLER
Mathematics Institute
University of Bern
Sidlerstrasse 5, CH-3012 Bern
Switzerland.
EMail: wihler@math.unibe.ch

Received: 28 October, 2009

Accepted: 08 December, 2009

Communicated by: F. Zhang

2000 AMS Sub. Class.: 15A45, 15A60.

Key words: Matrix functions, stability bounds, Hölder continuity.

Abstract: In this note, we shall investigate the Hölder continuity of matrix functions applied to normal matrices provided that the underlying scalar function is Hölder continuous. Furthermore, a few examples will be given.

Acknowledgements: I would like to thank Lutz Dümbgen for bringing the topic of this note to my attention and for suggestions leading to a more straightforward presentation of the material.

journal of inequalities in pure and applied mathematics
issn: 1443-5756

Contents

journal of inequalities in pure and applied mathematics
issn: 1443-5756

1. Introduction

We consider a scalar function $f : D \to \mathbb{C}$ on a (possibly unbounded) subset D of the complex plane $\mathbb{C}$. In this note, we shall be particularly interested in the case where f is *Hölder continuous with exponent* α on D, that is, there exists a constant $\alpha \in (0, 1]$ such that the quantity

$$[f]_{\alpha,D} := \sup_{\substack{x,y\in D\\ x\neq y}} \frac{|f(x) - f(y)|}{|x - y|^{\alpha}} \tag{1.1}$$

is bounded. We note that Hölder continuous functions are indeed continuous. Moreover, they are *Lipschitz continuous* if $\alpha = 1$; cf., e.g., [4].

Let us extend this concept to functions of matrices. To this end, consider

$$\mathbb{M}^{n\times n}_{\mathrm{normal}}(\mathbb{C}) = \left\{ \boldsymbol{A} \in \mathbb{C}^{n\times n} : \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\mathsf{H}} \right\},$$

the set of all normal matrices with complex entries. Here, for a matrix $\boldsymbol{A} = [a_{ij}]^{n}_{i,j=1}$, we use the notation $\boldsymbol{A}^{\mathsf{H}} = [\overline{a_{ji}}]^{n}_{i,j=1}$ to denote the conjugate transpose of $\boldsymbol{A}$. By the spectral theorem normal matrices are unitarily diagonalizable, i.e., for each $\boldsymbol{X} \in \mathbb{M}^{n\times n}_{\mathrm{normal}}(\mathbb{C})$ there exists a unitary $n \times n$-matrix $\boldsymbol{U}$, $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^{\mathsf{H}} = \mathbf{1} = \operatorname{diag}(1, 1, \ldots, 1)$, such that

$$\boldsymbol{U}^{\mathsf{H}}\boldsymbol{X}\boldsymbol{U} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where the set $\sigma(\boldsymbol{X}) = \{\lambda_i\}^{n}_{i=1}$ is the spectrum of $\boldsymbol{X}$. For any function $f : D \to \mathbb{C}$, with $\sigma(\boldsymbol{X}) \subseteq D$, we can then define a corresponding matrix function "value" by

$$\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{U}\operatorname{diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n))\,\boldsymbol{U}^{\mathsf{H}};$$

see, e.g., [5, 6]. Here, we use the bold face letter $\boldsymbol{f}$ to denote the matrix function corresponding to the associated scalar function f.

journal of inequalities in pure and applied mathematics
issn: 1443-5756

We can now easily widen the definition (1.1) of Hölder continuity for a scalar function $f : D \to \mathbb{C}$ to its associated matrix function $\boldsymbol{f}$ applied to normal matrices: Given a subset $\mathbb{D} \subseteq \mathbb{M}^{n\times n}_{\text{normal}}(\mathbb{C})$, then we say that the matrix function $\boldsymbol{f} : \mathbb{D} \to \mathbb{C}^{n\times n}$ is Hölder continuous with exponent $\alpha \in (0, 1]$ on $\mathbb{D}$ if

$$[\boldsymbol{f}]_{\alpha,\mathbb{D}} := \sup_{\substack{\boldsymbol{X},\boldsymbol{Y}\in\mathbb{D}\\ \boldsymbol{X}\neq\boldsymbol{Y}}} \frac{\|\boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}(\boldsymbol{Y})\|_{\mathsf{F}}}{\|\boldsymbol{X} - \boldsymbol{Y}\|^{\alpha}_{\mathsf{F}}} \tag{1.2}$$

is bounded. Here, for a matrix $\boldsymbol{X} = [x_{ij}]^{n}_{i,j=1} \in \mathbb{C}^{n\times n}$ we define $\|\boldsymbol{X}\|_{\mathsf{F}}$ to be the Frobenius norm of $\boldsymbol{X}$ given by

$$\|\boldsymbol{X}\|^2_{\mathsf{F}} = \operatorname{trace}\left(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}\right) = \sum_{i,j=1}^{n} |x_{ij}|^2, \qquad \boldsymbol{X} = (x_{ij})^{n}_{i,j=1} \in \mathbb{M}^{n\times n}(\mathbb{C}).$$

Evidently, for the definition (1.2) to make sense, it is necessary to assume that the scalar function f associated with the matrix function $\boldsymbol{f}$ is well-defined on the spectra of all matrices $\boldsymbol{X} \in \mathbb{D}$, i.e.,

$$\bigcup_{\boldsymbol{X}\in\mathbb{D}} \sigma(\boldsymbol{X}) \subseteq D. \tag{1.3}$$

The goal of this note is to address the following question: Provided that a scalar function f is Hölder continuous, what can be said about the Hölder continuity of the corresponding matrix function $\boldsymbol{f}$? The following theorem provides the answer:

Theorem 1.1. *Let the scalar function $f : D \to \mathbb{C}$ be Hölder continuous with exponent $\alpha \in (0, 1]$, and $\mathbb{D} \subseteq \mathbb{M}^{n\times n}_{\text{normal}}(\mathbb{C})$ satisfy (1.3). Then, the associated matrix function $\boldsymbol{f} : \mathbb{D} \to \mathbb{C}^{n\times n}$ is Hölder continuous with exponent α and*

$$[\boldsymbol{f}]_{\alpha,\mathbb{D}} \leq n^{\frac{1-\alpha}{2}} [f]_{\alpha,D} \tag{1.4}$$

journal of inequalities in pure and applied mathematics
issn: 1443-5756

holds true. In particular, the bound

$$\left\| \boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}(\boldsymbol{Y}) \right\|_{\mathsf{F}} \leq [f]_{\alpha,D} n^{\frac{1-\alpha}{2}} \left\| \boldsymbol{X} - \boldsymbol{Y} \right\|_{\mathsf{F}}^{\alpha}, \tag{1.5}$$

holds for any $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{D}$.

Hölder Continuity of
Matrix Functions
Thomas P. Wihler
vol. 10, iss. 4, art. 91, 2009

journal of inequalities
in pure and applied
mathematics
issn: 1443-5756

2. Proof of Theorem 1.1

We shall check the inequality (1.5). From this (1.4) follows immediately. Consider two matrices $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{D}$. Since they are normal we can find two unitary matrices $\boldsymbol{V}, \boldsymbol{W} \in \mathbb{M}^{n\times n}(\mathbb{C})$ which diagonalize $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, i.e.,

$$\boldsymbol{V}^{\mathsf{H}}\boldsymbol{X}\boldsymbol{V} = \boldsymbol{D}_{\boldsymbol{X}} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$
$$\boldsymbol{W}^{\mathsf{H}}\boldsymbol{Y}\boldsymbol{W} = \boldsymbol{D}_{\boldsymbol{Y}} = \operatorname{diag}(\mu_1, \mu_2, \dots, \mu_n),$$

where $\{\lambda_i\}_{i=1}^n$ and $\{\mu_i\}_{i=1}^n$ are the eigenvalues of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. Now we need to use the fact that the Frobenius norm is unitarily invariant. This means that for any matrix $\boldsymbol{X} \in \mathbb{C}^{n\times n}$ and any two unitary matrices $\boldsymbol{R}, \boldsymbol{U} \in \mathbb{C}^{n\times n}$ there holds

$$\|\boldsymbol{R}\boldsymbol{X}\boldsymbol{U}\|_{\mathsf{F}}^2 = \|\boldsymbol{X}\|_{\mathsf{F}}^2 .$$

Therefore, it follows that

$$\begin{aligned}
\|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}}^2 &= \left\|\boldsymbol{V}\boldsymbol{D}_{\boldsymbol{X}}\boldsymbol{V}^{\mathsf{H}} - \boldsymbol{W}\boldsymbol{D}_{\boldsymbol{Y}}\boldsymbol{W}^{\mathsf{H}}\right\|_{\mathsf{F}}^2 \\
&= \left\|\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{D}_{\boldsymbol{X}}\boldsymbol{V}^{\mathsf{H}}\boldsymbol{V} - \boldsymbol{W}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{D}_{\boldsymbol{Y}}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right\|_{\mathsf{F}}^2 \\
&= \left\|\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{D}_{\boldsymbol{X}} - \boldsymbol{D}_{\boldsymbol{Y}}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right\|_{\mathsf{F}}^2 \\
&= \sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{D}_{\boldsymbol{X}} - \boldsymbol{D}_{\boldsymbol{Y}}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2 \\
&= \sum_{i,j=1}^{n} \left|\sum_{k=1}^{n} \left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,k} (\boldsymbol{D}_{\boldsymbol{X}})_{k,j} - (\boldsymbol{D}_{\boldsymbol{Y}})_{i,k}\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{k,j}\right|^2 \\
&= \sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2 |\lambda_j - \mu_i|^2 .
\end{aligned} \tag{2.1}$$

Hölder Continuity of Matrix Functions
Thomas P. Wihler
vol. 10, iss. 4, art. 91, 2009

journal of inequalities in pure and applied mathematics
issn: 1443-5756

In the same way, noting that

$$\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{V} f(\boldsymbol{D}_{\boldsymbol{X}}) \boldsymbol{V}^{\mathsf{H}}, \qquad \boldsymbol{f}(\boldsymbol{Y}) = \boldsymbol{W} f(\boldsymbol{D}_{\boldsymbol{Y}}) \boldsymbol{W}^{\mathsf{H}},$$

we obtain

$$\|\boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}(\boldsymbol{Y})\|_{\mathsf{F}}^2 = \sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2 |f(\lambda_j) - f(\mu_i)|^2 .$$

Employing the Hölder continuity of f, i.e.,

$$|f(x) - f(y)| \leq [f]_{\alpha,D} |x-y|^{\alpha}, \qquad x, y \in D,$$

it follows that

$$\|\boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}(\boldsymbol{Y})\|_{\mathsf{F}}^2 \leq [f]_{\alpha,D}^2 \sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2 |\lambda_j - \mu_i|^{2\alpha} . \tag{2.2}$$

For $\alpha = 1$ the bound (1.5) results directly from (2.1) and (2.2). If $0 < \alpha < 1$, we apply Hölder's inequality. That is, for arbitrary numbers $s_i, t_i \in \mathbb{C}$, $i = 1, 2, \ldots$, there holds

$$\sum_{i \geq 1} |s_i t_i| \leq \left(\sum_{i\geq 1} |s_i|^{\frac{1}{\alpha}}\right)^{\alpha} \left(\sum_{i \geq 1} |t_i|^{\frac{1}{1-\alpha}}\right)^{1-\alpha} .$$

In the present situation this yields

$$\begin{aligned} &\|\boldsymbol{f}(\boldsymbol{X}) - \boldsymbol{f}(\boldsymbol{Y})\|^2 \\ &\quad \leq [f]_{\alpha,D}^2 \sum_{i,j=1}^{n} \left(|\lambda_j - \mu_i| \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|\right)^{2\alpha} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^{2-2\alpha} \\ &\quad \leq [f]_{\alpha,D}^2 \left(\sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2 |\lambda_j - \mu_i|^2\right)^{\alpha} \left(\sum_{i,j=1}^{n} \left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^2\right)^{1-\alpha} . \end{aligned}$$

journal of inequalities in pure and applied mathematics
issn: 1443-5756

Therefore, using the identity (2.1), there holds

$$\|\boldsymbol{f}(\boldsymbol{X})-\boldsymbol{f}(\boldsymbol{Y})\|_{\mathsf{F}} \leq [f]_{\alpha,D}\,\|\boldsymbol{X}-\boldsymbol{Y}\|_{\mathsf{F}}^{\alpha}\left(\sum_{i,j=1}^{n}\left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^{2}\right)^{\frac{1-\alpha}{2}}.$$

Then, recalling again that $\|\cdot\|_{\mathsf{F}}$ is unitarily invariant, yields

$$\left(\sum_{i,j=1}^{n}\left|\left(\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right)_{i,j}\right|^{2}\right)^{\frac{1-\alpha}{2}} = \left\|\boldsymbol{W}^{\mathsf{H}}\boldsymbol{V}\right\|_{\mathsf{F}}^{1-\alpha} = \|\mathbf{1}\|_{\mathsf{F}}^{1-\alpha} = n^{\frac{1-\alpha}{2}},$$

This implies the estimate (1.5).

journal of inequalities
in pure and applied
mathematics
issn: 1443-5756

3. Applications

We shall look at a few examples which fit in the framework of the previous analysis. Here, we consider the special case that all matrices are *real* and *symmetric*. In particular, they are normal and have only real eigenvalues.

Let us first study some functions $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}$ is an interval, which are continuously differentiable with bounded derivative on D. Then, by the mean value theorem, we have

$$[f]_{1,D} = \sup_{\substack{x,y\in D\\ x\neq y}} \left| \frac{f(x)-f(y)}{x-y} \right| = \sup_{\xi\in D} |f'(\xi)| < \infty,$$

i.e., such functions are Lipschitz continuous.

Trigonometric Functions:

Let $m \in \mathbb{N}$. Then, the functions $t \mapsto \sin^m(t)$ and $t \mapsto \cos^m(t)$ are Lipschitz continuous on $\mathbb{R}$, with constant

$$\begin{aligned} L_m := [\sin^m]_{1,\mathbb{R}} = [\cos^m]_{1,\mathbb{R}} &= \sup_{t\in\mathbb{R}} \left| \frac{\mathrm{d}}{\mathrm{d}t} \sin^m(t) \right| \\ &= \sup_{t\in\mathbb{R}} \left| \frac{\mathrm{d}}{\mathrm{d}t} \cos^m(t) \right| \\ &= \sqrt{m} \left(\frac{\sqrt{m-1}}{\sqrt{m}} \right)^{m-1}. \end{aligned}$$

Hölder Continuity of Matrix Functions
Thomas P. Wihler
vol. 10, iss. 4, art. 91, 2009

journal of inequalities in pure and applied mathematics
issn: 1443-5756

Thence, we immediately obtain the bounds

$$\|\sin^m(\boldsymbol{X}) - \sin^m(\boldsymbol{Y})\|_{\mathsf{F}} \le \sqrt{m}\left(\frac{\sqrt{m-1}}{\sqrt{m}}\right)^{m-1} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}}$$

$$\|\cos^m(\boldsymbol{X}) - \cos^m(\boldsymbol{Y})\|_{\mathsf{F}} \le \sqrt{m}\left(\frac{\sqrt{m-1}}{\sqrt{m}}\right)^{m-1} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}}$$

for any real symmetric $n \times n$-matrices $\boldsymbol{X}, \boldsymbol{Y}$. We note that

$$\lim_{m\to\infty}\left(\frac{\sqrt{m-1}}{\sqrt{m}}\right)^{m-1} = e^{-\frac{1}{2}},$$

and hence $L_m \sim \sqrt{m}$ with $m \to \infty$.

Gaussian Function:

For fixed $m > 0$, the Gaussian function $f : t \mapsto \exp(-mt^2)$ is Lipschitz continuous on $\mathbb{R}$ with constant $[f]_{1,\mathbb{R}} = \sqrt{2m}\exp(-\frac{1}{2})$. Consequently, we have for the matrix exponential that

$$\left\|\mathbf{exp}(-m\boldsymbol{X}^2) - \mathbf{exp}(-m\boldsymbol{Y}^2)\right\|_{\mathsf{F}} \le \sqrt{2m}e^{-\frac{1}{2}} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}},$$

for any real symmetric $n \times n$-matrices $\boldsymbol{X}, \boldsymbol{Y}$.

We shall now consider some functions which are less smooth than in the previous examples. In particular, they are not differentiable at 0.

Absolute Value Function:

Due to the triangle inequality

$$\big|\,|x| - |y|\,\big| \le |x - y|, \qquad x, y \in \mathbb{R},$$

journal of inequalities in pure and applied mathematics
issn: 1443-5756

the absolute value function $f : t \mapsto |t|$ is Lipschitz continuous with constant $[f]_{1,\mathbb{R}} = 1$, and hence

$$\||\boldsymbol{X}| - |\boldsymbol{Y}|\|_{\mathsf{F}} \leq \|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}}\,, \tag{3.1}$$

for any real symmetric $n \times n$-matrices $\boldsymbol{X}, \boldsymbol{Y}$. We note that, for general matrices, there is an additional factor of $\sqrt{2}$ on the right hand side of (3.1), whereas for symmetric matrices the factor 1 is optimal; see [1] and the references therein.

p-th Root of Positive Semi-Definite Matrices:

Finally, let us consider the p-th root ($p > 1$) of a real symmetric positive semi-definite matrix. The spectrum of such matrices belongs to the non-negative real axes $D = \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Here, we notice that the function $f : t \mapsto t^{\frac{1}{p}}$ is Hölder continuous on D with exponent $\alpha = \frac{1}{p}$ and $[f]_{\frac{1}{p},D} = 1$. Hence, Theorem 1.1 applies. In particular, the inequality

$$\left\|\boldsymbol{X}^{\frac{1}{p}} - \boldsymbol{Y}^{\frac{1}{p}}\right\|_{\mathsf{F}}^{p} \leq n^{\frac{p-1}{2}} \|\boldsymbol{X} - \boldsymbol{Y}\|_{\mathsf{F}} \tag{3.2}$$

holds for any real symmetric positive-semidefinite $n \times n$-matrices $\boldsymbol{X}, \boldsymbol{Y}$. We note that the estimate (3.2) is sharp. Indeed, there holds equality if $\boldsymbol{X}$ is chosen to be the identity matrix, and $\boldsymbol{Y}$ is the zero matrix.

We remark that an alternative proof of (3.2) has already been given in [2, Chapter X] in the context of operator monotone functions. Furthermore, closely related results on the Lipschitz continuity of matrix functions and the Hölder continuity of the p-th matrix root can be found in, e.g., [2, Chapter VII] and [3], respectively.

Hölder Continuity of Matrix Functions
Thomas P. Wihler
vol. 10, iss. 4, art. 91, 2009

journal of inequalities in pure and applied mathematics
issn: 1443-5756

References

[1] H. ARAKI AND S. YAMAGAMI, An inequality for Hilbert-Schmidt norm, *Comm. Math. Phys.*, **81**(1) (1981), 89–96.

[2] R. BHATIA, *Matrix Analysis*, Volume 169 of *Graduate Texts in Mathematics*, Springer-Verlag New York, 2007.

[3] Z. CHEN AND Z. HUAN, On the continuity of the mth root of a continuous nonnegative definite matrix-valued function, *J. Math. Anal. Appl.*, **209** (1997), 60–66.

[4] D. GILBARG AND N.S. TRUDINGER, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag Berlin, 2001 (Reprint of the 1998 edition).

[5] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, 3rd edition, 1996.

[6] N.J. HIGHAM, *Functions of Matrices*, Society for Industrial Mathematics, 2008.

journal of inequalities in pure and applied mathematics
issn: 1443-5756