PERTURBATION BOUNDS FOR THE OPERATOR ABSOLUTE VALUE

Rajendra Bhatia

Indian Statistical Institute,
New Delhi 110 016, India
and
Sonderforschungsbereich 343,
Universität Bielefeld, Bielefeld, Germany

Abstract. It is shown how to estimate the norms of the derivatives (of all orders) of the map that takes an invertible operator to the positive part in its polar decomposition. Using this, perturbation bounds of any order can be obtained for this map.

Perturbation Bounds for the Operator Absolute Value.

Let A be a bounded linear operator on a Hilbert space $\mathcal{H}$. Let $|A| = (A^*A)^{1/2}$ be the **positive part** or the **absolute value** of A.

In this note we show how to derive inequalities of the type

$$(1) \qquad || \, |A| - |B| \, || \leq \sum_{n=1}^{N} a_n \, ||A-B||^n + O(||A-B||^{N+1}),$$

where A is an invertible operator, B an operator close to it, N is any positive integer and a_n are coefficients which can be explicitly determined.

For $N = 2$ this was done in [1] and the approach used in that paper is developed further here.

Let $\mathcal{B}(\mathcal{H})$ denote the space of all linear operators on $\mathcal{H}$ and let $\mathcal{B}_{\mathrm{inv}}(\mathcal{H})$, $\mathcal{B}_s(\mathcal{H})$ and $B_+(\mathcal{H})$ denote its subsets consisting of the invertible, the self–adjoint and the positive operators, respectively. Let $\varphi : \mathcal{B}_{\mathrm{inv}}(\mathcal{H}) \to \mathcal{B}_+(\mathcal{H})$ be the map $\varphi(A) = |A|$. Let $D^n\varphi(A)$ denote the nth order (Frechét) derivative of φ. Let

$$a_n = ||D^n\varphi(A)||.$$

Then by Taylor's Theorem [3, Ch. 8] we have the inequality (1). So, our problem is reduced to estimating $||D^n\varphi(A)||$ for all n.

Now $\varphi = f \circ g$, where $g(A) = A^*A$ and $f(A) = A^{1/2}$ is the positive square root of a positive operator A. We will study f and g separately and then combine the information obtained.

More generally, let f be any function mapping $(0,\infty)$ into itself. This induces a map on $\mathcal{B}_+(\mathcal{H})$ which, for convenience, is again denoted by f. Let $f^{(n)}$ be the (ordinary) n–th derivative of f when it is viewed as a map on $(0,\infty)$, and let $D^nf(A)$ be its n–th order Frechét derivative at A when f is viewed as a map on $\mathcal{B}_+(\mathcal{H})$. If

$$(3) \qquad ||D^nf(A)|| = ||f^{(n)}(A)|| \quad \text{for all } A \in \mathcal{B}_+(\mathcal{H}),$$

we will say that f is in the class $\mathcal{D}_n$. The following Proposition is crucial for our analysis.

Proposition 1. Let f be an operator monotone function. Then $f \in \bigcap_{n=1}^{\infty} \mathcal{D}_n$.

Proof. In [1, Equations (10), (13)] we showed that an operator monotone function satisfies (3) for $n = 1, 2$. The same argument will show that this is the case for all n. First note that if $h(A) = A^{-1}$ then

$$(4) \qquad \begin{aligned} &D^nh(A)\,(B_1, B_2, \ldots, B_n) \\ &= (-1)^n \sum_{\sigma} A^{-1}\, B_{\sigma(1)}A^{-1}\, B_{\sigma(2)}\, A^{-1} \ldots A^{-1}\, B_{\sigma(n)}\, A^{-1}, \end{aligned}$$

where, σ runs over all cyclic permutations on n symbols. This gives

$$||D^n h(A)|| = n!\ ||A^{-1}||^{n+1}. \tag{5}$$

Next, use the fact that if f is operator monotone then it can be expressed as

$$f(t) = \alpha + \beta\ t + \int_0^\infty \left(\frac{\lambda}{\lambda^2+1} - \frac{1}{\lambda+t}\right) d\mu(\lambda), \tag{6}$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$ and μ is a positive measure. From this one obtains using (4), for $n > 2$,

$$||D^n f(A)|| \leq n! \int_0^\infty ||(\lambda + A)^{-1}||^{n+1}\ d\mu(\lambda) = ||f^{(n)}(A)||.$$

We skip the details, as they are essentially the same as in [1]. ∎

Corollary 2. Let $f(A) = A^{1/2}$. Then

$$||Df(A)|| = \frac{1}{2}\ ||A^{-1/2}||,$$

$$||D^n f(A)|| = \frac{1 \cdot 3 \cdot 5 \ldots (2n-3)}{2^n}\ ||A^{-n+1/2}||,\ n \geq 2.$$

Proposition 3. Let g be the map on $\mathcal{B}(\mathcal{H})$ defined as $g(A) = A^*A$. Then

$$\begin{aligned} ||Dg(A)|| &= 2\ ||A||, \\ ||D^2 g(A)|| &= 2, \\ ||D^n g(A)|| &= 0 \ \text{ for } \ n \geq 3. \end{aligned}$$

Proof. The first two equalities were derived in [1] from the relations

$$\begin{aligned} Dg(A)(B) &= A^*B + B^*A, \\ D^2 g(A)(B_1, B_2) &= B_1^*, B_2 + B_2^* B_1. \end{aligned}$$

Since $D^2 g(A)$ is a constant map we have $D^n g(A) = 0$ for $n \geq 3$. ∎

Our next task is to combine the information provided by the above two propositions. For this we first need expressions for the n–th Frechét derivative of a composite map $\varphi = f \circ g$. We will write these down in a general set up. Let X, Y, Z be Banach spaces and let g be a smooth map from X to Y and f a smooth map from Y

to Z . Let $\varphi = f \circ g$. If $X = Y = Z = \mathbb{R}$ we have the following formulae for the derivatives $\varphi^{(n)}$.

$$\varphi^{(1)}\ (x) = f^{(1)}\ (g(x))\ g^{(1)}\ (x),$$
$$\varphi^{(2)}\ (x) = f^{(2)}\ (g(x))\ [g^{(1)}(x)]^2 + f^{(1)}\ (g(x))\ g^{(2)}\ (x),$$
$$\varphi^{(3)}\ (x) = f^{(3)}\ (g(x))\ [g^{(1)}(x)]^3 + 3\ f^{(2)}\ (g(x))\ g^{(1)}(x)\ g^{(2)}\ (x) + f^{(1)}\ (g(x))\ g^{(3)}(x),$$
$$\begin{aligned}\varphi^{(4)}(x) = {} & f^{(4)}(g(x))\ [g^{(1)}(x)]^4 + 6\ f^{(3)}(g(x))\ [g^{(1)}(x)]^2\ g^{(2)}(x) + 3\ f^{(2)}(g(x))\ [g^{(2)}(x)]^2 \\ & + 4\ f^{(2)}(g(x))\ g^{(1)}(x)\ g^{(3)}(x) + f^{(1)}\ (g(x))\ g^{(4)}(x),\end{aligned}$$

etc.

When X, Y, Z are general Banach spaces analogoues of these formulae are more complicated. Recall that $D^{(n)}g(x)$ is a symmetric n–linear map from $X \times \ldots \times X$ to Y , etc.

To write our expressions for $D^n \varphi$ compactly let us adopt the following convention. A summation $\sum\limits_{\sigma}$ will indicate summation over permutations σ on n symbols. Since the higher Frechét derivatives are symmetric in their variables several summands in the sum $\sum\limits_{\sigma}$ will be identically equal. If we retaim only one representative from each of these identically equal terms and sum them up, the resulting sum will be written as $\sum_{\sigma}^{*}$. Thus, for example, we have, for the first two derivatives of $\varphi = f \circ g$ the expressions

$$\begin{aligned} D\ \varphi(x) &= Df(g(x))\ Dg(x) \qquad \text{(chain rule)}\ , \\ D^2\ \varphi(x) &= D^2 f(g(x))\ (Dg(x)(x_1),\ Dg(x)(x_2)) \\ &\qquad + Df(g(x))\ (D^2 g(x)\ (x_1, x_2)). \end{aligned}$$

With our notation we could also write

$$\begin{aligned} D^2\ \varphi(x) = {} & \sum\nolimits_{\sigma}^{*}\ D^2 f(g(x))\ (Dg(x)\ (x_{\sigma(1)}),\ Dg(x)\ (x_{\sigma(2)})) \\ & + \sum\nolimits_{\sigma}^{*}\ Df(g(x))\ (D^2 g(x)\ (x_{\sigma(1)}, x_{\sigma(2)})). \end{aligned}$$

Since the second derivative is a symmetric bilinear map, from each of the sum $\sum_{\sigma}$ only one of the two summands is retained when we go to $\sum_{\sigma}^{*}$. Of course, in this case there is no advantage in going to this notation. However, for higher derivatives it is helpful to use this notation and write

$$\begin{aligned} & D^3 \varphi(x)\ (x_1, x_2, x_3) \\ & = D^3 f(g(x))\ (Dg(x)\ (x_1),\ Dg(x)\ (x_2),\ Dg(x)\ (x_3)) \\ & + \sum\nolimits_{\sigma}^{*}\ D^2 f(g(x))\ (Dg(x)\ (x_{\sigma(1)}),\ D^2 g(x)\ (x_{\sigma(2)}, x_{\sigma(3)})) \\ & + Df(g(x))\ D^3 g(x)\ (x_1, x_2, x_3). \end{aligned}$$

$$
\begin{aligned}
&D^4\varphi(x)\ (x_1,x_2,x_3,x_4)\\
&= D^4 f(g(x))\ (Dg(x)\ (x_1),\ Dg(x)\ (x_2),\ Dg(x)\ (x_3),\ Dg(x)\ (x_4))\\
&+\sum\nolimits_\sigma^* \ D^3 f(g(x))\ (Dg(x)\ (x_{\sigma(1)}),\ Dg(x)\ (x_{\sigma(2)}),\ D^2 g(x)\ (x_{\sigma(3)},x_{\sigma(4)}))\\
&+,\sum\nolimits_\sigma^* \ D^2 f(g(x))\ (D^2 g(x)\ (x_{\sigma(1)},x_{\sigma(2)}),\ D^2 g(x)\ (x_{\sigma(3)},x_{\sigma(4)}))\\
&+\sum\nolimits_\sigma^* \ D^2 f(g(x))\ (D^2 g(x)\ (x_{\sigma(1)}),\ D^3 g(x)\ (x_{\sigma(2)},\ x_{\sigma(3)},x_{\sigma(4)}))\\
&+ Df(g(x))\ D^4 g(x)\ (x_1,x_2,x_3,x_4).
\end{aligned}
$$

The reader may check that in the three starred sums occuring here, the summation involves six, three and four summands, respectively, and that when $X = Y = Z = \mathbb{R}$ this reduces to the expression for $\varphi^{(4)}(x)$ written earlier.

Now return to the special situation $g(A) = A^*A,\ f(A) = A^{1/2},\ \varphi(A) = g(f(A)) = |A|$. Then using the above expressions for $D^{(n)}\varphi$ and the results of Corollary 2 and Proposition 3 one obtains the following bounds

$$
\begin{aligned}
||D\varphi(A)|| &\leq ||A^{-1}||\ ||A||,\\
||D^2\varphi(A)|| &\leq ||A^{-1}||^3\ ||A||^2 + ||A^{-1}||,\\
||D^3\varphi(A)|| &\leq 3\ ||A^{-1}||^5\ ||A||^3 + 3\ ||A^{-1}||^3\ ||A||,\\
||D^4\varphi(A)|| &\leq 15\ ||A^{-1}||^7\ ||A||^4 + 18\ ||A^{-1}||^5\ ||A^2|| + 3\ ||A^{-1}||^3.
\end{aligned}
\tag{7}
$$

The first two inequalities in (7) were derived in [1].

Bounds for derivatives of all orders can be calculated using this procedure. A simple rule which can be skimmed from the above analysis is the following. For the composite function $\varphi(x) = f(g(x))$ of a real variable write down the expression for its derivative $\varphi^{(n)}(x)$. This will be a sum of terms each of which is a product of factors $f^{(1)}(g(x)),\ f^{(2)}(g(x)),\ldots, f^{(n)}(g(x))$ and $g^{(1)}(x),\ g^{(2)}(x),\ldots, g^{(n)}(x)$. In this expression replace $f^{(1)}(g(x))$ by $||A^{-1}||$ and for $n \geq 2$ replace $f^{(n)}(g(x))$ by $1\cdot 3\cdot\ldots\cdot 5\ (2n-3)\ ||A^{-1}||^{2n-1}$; replace $g^{(1)}(x)$ by $||A||$, $g^{(2)}(x)$ by 1 and for $n \geq 3$ replace $g^{(n)}(x)$ by 0. The resulting expression will be a bound for the norm $||D^n\varphi(A)||$, where $\varphi(A) = |A|$.

The reader can check that this rule is a consequence of the above analysis, that the inequalities (7) conform to this and that this gives, for instance,

$$
||D^5\varphi(A)|| \leq 105\ ||A^{-1}||^9\ ||A||^5 + 150\ ||A^{-1}||^7\ ||A||^3 + 45\ ||A^{-1}||^5\ ||A||.
$$

We can thus obtain perturbation bounds like (1) to any desired order.

It seems a difficult problem to characterize the classes $\mathcal{D}_n$ of functions that satisfy the relation (3). When $n = 1$ this is already quite intricate [2].

Acknowledgement This work was done at the University of Toronto with the support of NSERC (Canada) and at the Universität Bielefeld with the support of Sonderforschungsbereich 343.

References

[1] R. Bhatia, *First and second order perturbation bounds for the operator absolute value*, Linear Algebra and Appl., (??? 1994 ??).

[2] R. Bhatia and K.B. Sinha, *Variation of real powers of positive operators*, Indiana Univ. Math. J., (?? 1994 ??).

[3] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York (1969).