Best Constant in Norm Bounds on Commutators

Koenraad M.R. Audenaert
k.audenaert@imperial.ac.uk

September 18, 2007

1 Problem

Find the best constant c in the inequality

$$|| \, [X, Y] \, ||_2 \leq c \, ||X||_2 \, ||Y||_2, \tag{1}$$

where X and Y are arbitrary complex $d \times d$ matrices, $[X, Y]$ is their commutator $XY - YX$, and $||A||_2 = (\mathrm{Tr}[A^* A])^{1/2}$ is the Frobenius norm.

2 Background

Commutators come up almost everywhere, in diverse areas of mathematical physics, including quantum information theory. The actual answer to the above problem, the value of this best constant, is likely to be irrelevant for most applications of such bounds. Nevertheless it is quite an intruiging problem, since it is one of the simplest problems one can ask about commutators, and yet it has not been answered yet. The hope is that the tools used to solve it will provide additional handles on the structure of commutators, and may in this way be helpful to the solution of other problems involving them.

3 Partial Solution

It is well-known and easy to see that the inequality holds for $c = 2$. This is a straightforward application of the triangle inequality and submultiplicativity of unitarily invariant norms. In [2], A. Böttcher and D. Wenzel conjectured that, actually, $c = \sqrt{2}$ is the best constant (they only consider the real case, but the complex case most likely gives the same answer). One can certainly not do better, since $c = \sqrt{2}$ is achieved by two anti-commuting Pauli matrices. Say $X = \sigma_x$ and $Y = \sigma_z$, then $[X, Y] = -2i\sigma_y$, so that $||[X, Y]||_2 = 2||\sigma_y||_2 = 2\sqrt{2}$, while $||X||_2 = ||Y||_2 = \sqrt{2}$.

One of the matrices can be eliminated from the problem by regarding the matrices as vectors in a Hilbert space. Let $\mathrm{Vec}(X)$ be the operation that "stacks the columns of X on top of each other", i.e. in terms of the standard matrix and vector basis elements, $\mathrm{Vec}(e^{ij}) = e^i \otimes e^j$. Then one easily sees that

$$\mathrm{Vec}([X,Y]) = (X \otimes \mathbb{1} - \mathbb{1} \otimes X^T)\,\mathrm{Vec}(Y).$$

Since the Frobenius norm has the special property that $||X||_2$ is equal to the Euclidian 2-norm $||.||_2$ of $\mathrm{Vec}(X)$, the norm inequality turns into

$$||(X \otimes \mathbb{1} - \mathbb{1} \otimes X^T)\,\mathrm{Vec}(Y)||_2 \le c||X||_2\,||\,\mathrm{Vec}(Y)||_2.$$

The matrix Y can now be eliminated by noting that

$$\max_y \frac{||Ay||_2}{||y||_2} = ||A||,$$

where the maximisation is over all vectors and $||A||$ denotes the operator norm of A (largest singular value). Hence, an equivalent form of the problem is finding the best constant c in

$$||X \otimes \mathbb{1} - \mathbb{1} \otimes X^T|| \le c\,||X||_2.$$

Known results are listed below:

1. When X is normal, the best c is indeed $\sqrt{2}$. Normality of X implies that X is unitarily diagonalisable. Thus there is a unitary U and diagonal Λ such that $UXU^* = \Lambda$. By unitary invariance of the operator norm and Frobenius norm, the inequality can then be converted to the diagonal case, which is an easy exercise to solve. Two proofs are presented in [2]. When X is not normal, one could try to exploit the singular value decomposition $X = U\Sigma V^*$, but the problem is that one cannot get rid of both U and V in the inequality.
2. When X and Y are real, $c \le \sqrt{3}$; proven in [2].
3. When X and Y are real 2×2 matrices, or when one of the matrices is rank 1, $c = \sqrt{2}$; again, see[2].
4. When $X = Y^*$, $c = \sqrt{2}$; this case is due to Spiros Michalakis (unpublished).

A related problem is to see what happens when the right-hand side is based on other norms. It turns out that this gives rise to "boring" results. Given a UI norm $|||.|||$, one can define the corresponding Q-norm $||.||_Q$ by

$$||X||_Q := |||X^*X|||^{1/2}. \tag{2}$$

Theorem 1 *For general complex matrices X and Y, and for any UI norm $|||.|||$ with corresponding Q-norm $||.||_Q$, the following inequality holds:*

$$|||[X,Y]||| \le 2||X||_Q||Y||_Q. \tag{3}$$

Proof. Follows from the triangle inequality and Hölder's inequality:

$$|||XY||| \leq |||\,|X|^2\,|||^{1/2}\;|||\,|Y|^2\,|||^{1/2}$$

see [1] (IV.42). QED

Despite the siicity of this derivation, the ensuing bound is sharp, as equality is obtained for X and Y two anticommuting Pauli operators. Take $X = \sigma_x$ and $Y = \sigma_z$, then $[X, Y] = -2i\sigma_y$, so that $|||[X, Y]||| = 2|||\mathbb{1}|||$ and $||X||_Q = ||Y||_Q = |||\mathbb{1}|||^{1/2}$.

For example, when applied to Schatten p-norms, this gives the bound

$$||[X, Y]||_p \leq 2||X||_{2p}||Y||_{2p}. \tag{4}$$

As a closing remark, one can replace X by $X + Z$ for any Z that commutes with Y, since this transformation does not change the commutator. It might, for example, make sense to replace X and Y by their traceless versions $X \mapsto X - \mathrm{Tr}(X)\mathbb{1}/d$.

References

[1] R. Bhatia, *Matrix Analysis*, Springer, Heidelberg (1997).

[2] A. Böttcher and D. Wenzel, "How big can the commutator of two matrices be and how big is it typically?", Lin. Alg. Appl. **403** (2005) 216–228.