Note #1. A trace inequality on 2×2 block matrices

Author: Suvrit Sra
Date: Jun. 6, 2011
Revised: Nov. 15, 2011
To appear in: ILAS Image #47

Question 1.1 (ILAS Image #46, Problem 46-5)**.** Let A and B be $m \times n$ strictly contractive complex matrices. Show that

$$|\operatorname{tr}(I - B^*A)^{-1}|^2 \leq \operatorname{tr}(I - A^*A)^{-1} \operatorname{tr}(I - B^*B)^{-1}. \tag{1.1}$$

——— * ———

To ease readability, we use lower-case letters for matrices. Proving (1.1) reduces to showing

$$\begin{pmatrix} \operatorname{tr}(e - a^*a)^{-1} & \operatorname{tr}(e - a^*b)^{-1} \\ \operatorname{tr}(e - b^*a)^{-1} & \operatorname{tr}(e - b^*b)^{-1} \end{pmatrix} \geq 0.$$

Since a is strictly contractive, we may write

$$(e - a^*a)^{-1} = \sum\nolimits_{k \geq 0} (a^*a)^k;$$

similarly we can expand $(e - b^*b)^{-1}$ and $(e - a^*b)^{-1}$, etc.

Since sums of positive definite matrices are positive definite, it suffices to prove instead

$$\begin{pmatrix} \operatorname{tr}(a^*a)^k & \operatorname{tr}(a^*b)^k \\ \operatorname{tr}(b^*a)^k & \operatorname{tr}(b^*b)^k \end{pmatrix} \geq 0,$$

Equivalently, we must prove that

$$|\operatorname{tr}(a^*b)^k|^2 \leq \operatorname{tr}(a^*a)^k \operatorname{tr}(b^*b)^k. \tag{1.2}$$

Using the notation $|a| = (a^*a)^{1/2}$, recall by Weyl's majorant relation, $|\operatorname{tr}\, x| \leq \operatorname{tr}|x|$. Thus, we have

$$|\operatorname{tr}(a^*b)^k|^2 \leq (\operatorname{tr}|a^*b|^k)^2 = \|\, |a^*b|^{k/2} \,\|_F^4, \tag{1.3}$$

where as usual we have the *Frobenius norm* $\|x\|_F^2 := \operatorname{tr}|x|^2$. Now, by the known inequality [*Horn and Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991, p.212.*]

$$\|\, |xy|^r \,\|^2 \leq \|\, |x|^{2r} \,\|\|\, |y|^{2r} \,\|, \quad r > 0, \tag{1.4}$$

for any unitarily invariant norm. Using inequality (1.4) with (1.3), we get

$$|\operatorname{tr}(a^*b)^k|^2 \leq \|\, |a^*b|^{k/2} \,\|_F^4 \leq \|\, |a|^k \,\|_F^2 \|\, |b|^k \,\|_F^2 = (\operatorname{tr}|a|^{2k})(\operatorname{tr}|b|^{2k}). \quad \square$$

Note: Alternatively, one can prove the stronger result that the original matrix, without the trace terms is positive definite. Then, apply the fact that the trace function is completely positive.