

21.5 The Trace Inequality

Sections 3 and 4 built a dictionary: a unitarily invariant norm is a symmetric gauge applied to the list of singular values, and a symmetric gauge respects weak majorization. To use that dictionary we need a supply of inequalities between singular values, and the cleanest source of them is a single scalar inequality about how two matrices pair against each other. This section proves it. The pairing is the Frobenius one, $\text{tr}(\mathbf{A}^*\mathbf{B})$, and the theorem says that it is largest when the two matrices are *aligned*: when their singular value decompositions can be written with the same unitary factors.

21.5.1 How large can a pairing be?

Chapter 11 §01 made $M_{m \times n}(F)$ an inner product space with $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{B}^*\mathbf{A})$, whose norm is the Frobenius norm $\|\mathbf{A}\|_F$. So $\text{tr}(\mathbf{A}^*\mathbf{B}) = \langle \mathbf{B}, \mathbf{A} \rangle$ is an inner product of two matrices, and Cauchy–Schwarz (Theorem 11.1.5) already bounds it:

$$|\text{tr}(\mathbf{A}^*\mathbf{B})| \leq \|\mathbf{A}\|_F \|\mathbf{B}\|_F = \left(\sum_i \sigma_i(\mathbf{A})^2 \right)^{1/2} \left(\sum_i \sigma_i(\mathbf{B})^2 \right)^{1/2},$$

using $\|\mathbf{A}\|_F^2 = \sum_i \sigma_i(\mathbf{A})^2$. That bound depends only on the two lists of singular values, which is the right kind of answer; but it is not the best one.

Here is the smallest example. Let $\mathbf{A} = \text{diag}(2, 1)$ and $\mathbf{B} = \mathbf{I}_2$. Then $\text{tr}(\mathbf{A}^*\mathbf{B}) = 3$, while Cauchy–Schwarz allows $\sqrt{5} \cdot \sqrt{2} \approx 3.162$. The number 3 is $\sigma_1(\mathbf{A})\sigma_1(\mathbf{B}) + \sigma_2(\mathbf{A})\sigma_2(\mathbf{B})$, the two lists of singular values multiplied **term by term in decreasing order**. That is the true bound, and the pairing $\sum_i \sigma_i \tau_i$ is never larger than the Cauchy–Schwarz product, by Cauchy–Schwarz applied to the two lists.

Two matrices pair as well as their singular values allow, and no better; and the best case is when the two matrices stretch the same directions by the same ranking.

The proof takes the two singular value decompositions apart and finds, sitting between them, a square array of nonnegative weights whose rows and columns each add up to at most 1. Everything then reduces to a fact about such arrays, which is worth isolating first.

21.5.2 Weights that do not add up to more than one

Definition 21.5.1. Doubly Substochastic Matrix

A matrix $\mathbf{D} = (d_{kl}) \in M_p(\mathbb{R})$ is **doubly substochastic** if

- (S1) $d_{kl} \geq 0$ for all k, l ;
- (S2) $\sum_{l=1}^p d_{kl} \leq 1$ for **every** row index k ; and
- (S3) $\sum_{k=1}^p d_{kl} \leq 1$ for **every** column index l .

In words: the entries are nonnegative, and no row and no column adds up to more than one. Replacing both “ ≤ 1 ” by “ $= 1$ ” gives the doubly stochastic matrices of Definition 18.8.4, so every doubly stochastic matrix is doubly substochastic, and in particular so is every permutation matrix. The zero matrix is doubly substochastic, which the stochastic condition excludes; that degenerate case matters, because it is what lets a substochastic matrix throw weight away instead

of merely moving it. A nonexample by minimal change: $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ satisfies (S1) and (S3) but its first row adds to 2, so (S2) fails.

The lemma we need says that such an array of weights cannot help two decreasing nonnegative lists pair better than they do term by term.

Lemma 21.5.1. Substochastic Pairing

Let $x_1 \geq \cdots \geq x_p \geq 0$ and $y_1 \geq \cdots \geq y_p \geq 0$ be real numbers, and let $\mathbf{D} \in M_p(\mathbb{R})$ be doubly substochastic. Then

$$\sum_{k=1}^p \sum_{l=1}^p d_{kl} x_k y_l \leq \sum_{k=1}^p x_k y_k.$$

Idea. Decreasing nonnegative lists are nonnegative combinations of the “staircase” lists $(1, \dots, 1, 0, \dots, 0)$: write $x_k = \sum_{i \geq k} (x_i - x_{i+1})$. Summation by parts in both variables at once turns each side into a nonnegative combination of the same coefficients, and all that is left to compare is what multiplies them: on the left a partial sum of $\mathbf{D}$ over a top-left rectangle, on the right the number $\min(i, j)$. The two substochastic conditions say exactly that an $i \times j$ rectangle of $\mathbf{D}$ holds at most $\min(i, j)$.

Proof. Put $x_{p+1} = y_{p+1} = 0$ and, for $1 \leq i \leq p$, write $\Delta x_i = x_i - x_{i+1}$ and $\Delta y_i = y_i - y_{i+1}$. These are ≥ 0 , since the lists are decreasing and $x_p, y_p \geq 0$, and telescoping gives

$$x_k = \sum_{i=k}^p \Delta x_i, \quad y_l = \sum_{j=l}^p \Delta y_j \quad (1 \leq k, l \leq p).$$

Substituting both and exchanging the order of the four finite sums,

$$\sum_{k,l} d_{kl} x_k y_l = \sum_{i=1}^p \sum_{j=1}^p \Delta x_i \Delta y_j S_{ij}, \quad S_{ij} := \sum_{k \leq i} \sum_{l \leq j} d_{kl},$$

because the pair (k, l) contributes to (i, j) exactly when $k \leq i$ and $l \leq j$. The same substitution with $\mathbf{D}$ replaced by the identity matrix gives

$$\sum_{k=1}^p x_k y_k = \sum_{i=1}^p \sum_{j=1}^p \Delta x_i \Delta y_j \min(i, j),$$

since the number of $k \leq \min(i, j)$ is $\min(i, j)$.

It remains to compare S_{ij} with $\min(i, j)$. Summing (S2) over the i rows $k \leq i$ and discarding the nonnegative entries with $l > j$ gives $S_{ij} \leq i$; summing (S3) over the j columns $l \leq j$ gives $S_{ij} \leq j$. Hence $S_{ij} \leq \min(i, j)$ for all i, j . Since every Δx_i and every Δy_j is ≥ 0 , replacing S_{ij} by $\min(i, j)$ can only increase the double sum. This proves the lemma. ■

Taking $\mathbf{D}$ to be the permutation matrix of $\pi \in S_p$ gives $\sum_k x_k y_{\pi(k)} \leq \sum_k x_k y_k$: two decreasing lists pair best when they are paired in order. That is the **rearrangement inequality**, and it is the special case that carries the meaning of the lemma. The general case allows the weights

to be spread out and partly thrown away, and says that neither helps.

Quick check. Where does the proof use $x_p \geq 0$ and $y_p \geq 0$, rather than just the decrease? Find a doubly substochastic $\mathbf{D}$ and a decreasing pair with a negative entry for which the conclusion fails.

Solution. It is used to make $\Delta x_p = x_p$ and $\Delta y_p = y_p$ nonnegative, which is what allows S_{ij} to be replaced by the larger $\min(i, j)$ in the last step. Without it the conclusion fails: take $p = 2$, $\mathbf{x} = (1, -1)$, $\mathbf{y} = (1, 1)$, and $\mathbf{D} = \text{diag}(1, 0)$, which is doubly substochastic. The left side is 1 and the right side is $1 - 1 = 0$, so the asserted inequality $1 \leq 0$ is false. The proof breaks at the step named: here $\Delta x_2 = x_2 = -1 < 0$ and $\Delta y_2 = 1$, so replacing $S_{22} = 1$ by $\min(2, 2) = 2$ lowers the double sum instead of raising it. Throwing weight away helps when a pairing is negative, and that is exactly what nonnegativity rules out.

21.5.3 The trace inequality

Theorem 21.5.2. Von Neumann's Trace Inequality

Let $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$ with $F = \mathbb{R}$ or $\mathbb{C}$, and let $p = \min(m, n)$. Then

$$|\text{tr}(\mathbf{A}^* \mathbf{B})| \leq \sum_{i=1}^p \sigma_i(\mathbf{A}) \sigma_i(\mathbf{B}).$$

The bound is attained whenever $\mathbf{A}$ and $\mathbf{B}$ have **aligned** singular value decompositions, that is, whenever there are unitary $\mathbf{U} \in M_m(F)$ and $\mathbf{V} \in M_n(F)$ with $\mathbf{A} = \mathbf{U} \Sigma_{\mathbf{A}} \mathbf{V}^*$ and $\mathbf{B} = \mathbf{U} \Sigma_{\mathbf{B}} \mathbf{V}^*$, where $\Sigma_{\mathbf{A}}, \Sigma_{\mathbf{B}} \in M_{m \times n}(F)$ carry $\sigma_i(\mathbf{A})$ and $\sigma_i(\mathbf{B})$ on the main diagonal and zeros elsewhere.

Idea. Write both singular value decompositions and push the four unitary factors into two: one, $\mathbf{M}$, sits between the two diagonal matrices, the other, $\mathbf{N}$, closes the trace up. Multiplying out, the trace becomes $\sum_{k,l} \sigma_k(\mathbf{A}) \sigma_l(\mathbf{B}) m_{kl} n_{lk}$: every pairing of a singular value of $\mathbf{A}$ with one of $\mathbf{B}$ occurs, weighted by one entry of each unitary. The weights $|m_{kl} n_{lk}|$ are exactly where unitarity is spent — each row of $\mathbf{M}$ and each column of $\mathbf{N}$ is a unit vector, so Cauchy–Schwarz makes the weight array doubly substochastic — and Lemma 21.5.1 finishes.

Proof. Fix singular value decompositions $\mathbf{A} = \mathbf{U}_1 \Sigma_{\mathbf{A}} \mathbf{V}_1^*$ and $\mathbf{B} = \mathbf{U}_2 \Sigma_{\mathbf{B}} \mathbf{V}_2^*$ (Theorem 13.9.1), with $\mathbf{U}_1, \mathbf{U}_2 \in M_m(F)$ and $\mathbf{V}_1, \mathbf{V}_2 \in M_n(F)$ unitary and $\Sigma_{\mathbf{A}}, \Sigma_{\mathbf{B}}$ as in the statement. Put

$$\mathbf{M} := \mathbf{U}_1^* \mathbf{U}_2 \in M_m(F), \quad \mathbf{N} := \mathbf{V}_2^* \mathbf{V}_1 \in M_n(F),$$

both unitary, being products of unitary matrices. Then $\mathbf{A}^* \mathbf{B} = \mathbf{V}_1 \Sigma_{\mathbf{A}}^* \mathbf{M} \Sigma_{\mathbf{B}} \mathbf{V}_2^*$, so by Theorem 0.9.6 (3), moving the factor $\mathbf{V}_1$ to the back,

$$\text{tr}(\mathbf{A}^* \mathbf{B}) = \text{tr}(\Sigma_{\mathbf{A}}^* \mathbf{M} \Sigma_{\mathbf{B}} \mathbf{N}).$$

Now compute that trace entry by entry. The matrix $\Sigma_{\mathbf{A}}^* \in M_{n \times m}(F)$ has (k, a) entry

$\sigma_k(\mathbf{A})\delta_{ka}$, which vanishes unless $k = a \leq p$, and $\Sigma_{\mathbf{B}}$ has (b, l) entry $\sigma_b(\mathbf{B})\delta_{bl}$, which vanishes unless $b = l \leq p$. Hence

$$\operatorname{tr}(\mathbf{A}^*\mathbf{B}) = \sum_{k=1}^p \sum_{l=1}^p \sigma_k(\mathbf{A}) \sigma_l(\mathbf{B}) m_{kl} n_{lk}, \quad (21.5.1)$$

all indices being in range because $p \leq m$ and $p \leq n$. By the triangle inequality (Theorem 0.6.4),

$$|\operatorname{tr}(\mathbf{A}^*\mathbf{B})| \leq \sum_{k=1}^p \sum_{l=1}^p \sigma_k(\mathbf{A}) \sigma_l(\mathbf{B}) d_{kl}, \quad d_{kl} := |m_{kl}| |n_{lk}|.$$

Claim. $\mathbf{D} = (d_{kl})_{k,l \leq p}$ is doubly substochastic.

Proof. Each $d_{kl} \geq 0$, which is (S1). Fix $k \leq p$. By Cauchy–Schwarz (Theorem 11.1.5) in $\mathbb{R}^p$, and then by extending both sums to their full ranges $l \leq m$ and $l \leq n$, which only adds nonnegative terms,

$$\sum_{l=1}^p |m_{kl}| |n_{lk}| \leq \left(\sum_{l=1}^m |m_{kl}|^2 \right)^{1/2} \left(\sum_{l=1}^n |n_{lk}|^2 \right)^{1/2} = 1,$$

because row k of the unitary $\mathbf{M}$ and column k of the unitary $\mathbf{N}$ are unit vectors. This is (S2), and (S3) is the same computation with the roles of the two indices exchanged: column l of $\mathbf{M}$ and row l of $\mathbf{N}$ are unit vectors as well. ■

The lists $\sigma_1(\mathbf{A}) \geq \cdots \geq \sigma_p(\mathbf{A}) \geq 0$ and $\sigma_1(\mathbf{B}) \geq \cdots \geq \sigma_p(\mathbf{B}) \geq 0$ are decreasing and nonnegative (Definition 13.9.1), so Lemma 21.5.1 applies and gives

$$\sum_{k,l} \sigma_k(\mathbf{A}) \sigma_l(\mathbf{B}) d_{kl} \leq \sum_{k=1}^p \sigma_k(\mathbf{A}) \sigma_k(\mathbf{B}),$$

which is the stated bound.

For the last sentence, suppose $\mathbf{A} = \mathbf{U}\Sigma_{\mathbf{A}}\mathbf{V}^*$ and $\mathbf{B} = \mathbf{U}\Sigma_{\mathbf{B}}\mathbf{V}^*$ with the same $\mathbf{U}$ and $\mathbf{V}$. Then $\mathbf{A}^*\mathbf{B} = \mathbf{V}\Sigma_{\mathbf{A}}^*\mathbf{U}^*\mathbf{U}\Sigma_{\mathbf{B}}\mathbf{V}^* = \mathbf{V}\Sigma_{\mathbf{A}}^*\Sigma_{\mathbf{B}}\mathbf{V}^*$, and $\Sigma_{\mathbf{A}}^*\Sigma_{\mathbf{B}}$ is the $n \times n$ diagonal matrix with entries $\sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$ for $i \leq p$ and zeros after. Taking traces and using Theorem 0.9.6 (3) again, $\operatorname{tr}(\mathbf{A}^*\mathbf{B}) = \sum_{i \leq p} \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$. This proves the theorem. ■

Three things about the proof are worth keeping. The weight matrix $\mathbf{D}$ is where all the information about the two sets of singular directions has gone; the only thing about it that survives is a pair of inequalities. Unitarity is spent exactly once, on the unit rows and columns inside the claim, and nothing else about $\mathbf{M}$ and $\mathbf{N}$ is used. And the equality case shows the bound is not merely an upper bound to be improved later: it is attained for every prescribed pair of singular value lists.

Warning. The singular values are paired **in matching order**, and no other pairing gives an upper bound. For $\mathbf{A} = \operatorname{diag}(2, 1)$ and $\mathbf{B} = \operatorname{diag}(20, 15)$ the reversed pairing is $2 \cdot 15 + 1 \cdot 20 = 50$, while $\operatorname{tr}(\mathbf{A}^*\mathbf{B}) = 40 + 15 = 55$. What is true for every permutation π is the **lower** bound

$\sum_i \sigma_i(\mathbf{A})\sigma_{\pi(i)}(\mathbf{B}) \leq \sum_i \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$, which is the rearrangement inequality above.

Example. A strict case and an equality case

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 12 & -12 \\ 16 & 9 \end{pmatrix}, \quad \mathbf{B}' = \begin{pmatrix} 20 & 0 \\ 0 & 15 \end{pmatrix}.$$

Check that $\mathbf{B}$ and $\mathbf{B}'$ have the same singular values, and compare the two sides of [Theorem 21.5.2](#) for the pairs $(\mathbf{A}, \mathbf{B})$ and $(\mathbf{A}, \mathbf{B}')$.

Solution. For $\mathbf{B}$,

$$\mathbf{B}^\top \mathbf{B} = \begin{pmatrix} 144 + 256 & -144 + 144 \\ -144 + 144 & 144 + 81 \end{pmatrix} = \begin{pmatrix} 400 & 0 \\ 0 & 225 \end{pmatrix},$$

so $\sigma(\mathbf{B}) = (20, 15)$ by Definition 13.9.1, and $\sigma(\mathbf{B}') = (20, 15)$ as well since $\mathbf{B}'$ is diagonal with positive decreasing entries. Also $\sigma(\mathbf{A}) = (2, 1)$. So both pairs have the same right-hand side,

$$\sigma_1(\mathbf{A})\sigma_1(\mathbf{B}) + \sigma_2(\mathbf{A})\sigma_2(\mathbf{B}) = 2 \cdot 20 + 1 \cdot 15 = 55.$$

On the left, $\text{tr}(\mathbf{A}^\top \mathbf{B}') = 2 \cdot 20 + 1 \cdot 15 = 55$: equality. And $\text{tr}(\mathbf{A}^\top \mathbf{B}) = 2 \cdot 12 + 1 \cdot 9 = 33 < 55$: strict.

The reason is visible in the decompositions. Both $\mathbf{A}$ and $\mathbf{B}'$ are diagonal with decreasing nonnegative entries, so $\mathbf{A} = \mathbf{I}\mathbf{A}\mathbf{I}^*$ and $\mathbf{B}' = \mathbf{I}\mathbf{B}'\mathbf{I}^*$ are aligned. For $\mathbf{B}$, on the other hand, $\mathbf{B} = \mathbf{U}_2 \text{diag}(20, 15)$ with

$$\mathbf{U}_2 = \frac{1}{5} \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix},$$

a rotation, so in the notation of the proof $\mathbf{M} = \mathbf{U}_1^* \mathbf{U}_2 = \mathbf{U}_2$ and $\mathbf{N} = \mathbf{I}$. The weight matrix is $\mathbf{D} = \text{diag}(\frac{3}{5}, \frac{3}{5})$, doubly substochastic and far from a permutation, and [Eq. 21.5.1](#) reads $2 \cdot 20 \cdot \frac{3}{5} + 1 \cdot 15 \cdot \frac{3}{5} = 24 + 9 = 33$. The rotation has thrown away two fifths of each weight, and that is exactly the loss.

21.5.4 The dual of a unitarily invariant norm

The trace inequality is a statement about a pairing, so it is really a statement about duality. Recall the dual norm of Definition 18.4.4: it measures a vector by the largest value of the pairing against vectors the given norm calls short. On matrices, the pairing to use is the Frobenius one, and the result is exactly as clean as one could hope: dualizing a unitarily invariant norm dualizes its symmetric gauge and nothing else.

Corollary 21.5.3. Duality of Unitarily Invariant Norms

Let $\|\cdot\|$ be a unitarily invariant norm on $M_{m \times n}(F)$ (Definition 21.4.1) with symmetric gauge Φ on $\mathbb{R}^p$, $p = \min(m, n)$, so that $\|\mathbf{A}\| = \Phi(\sigma(\mathbf{A}))$ (Lemma 21.4.1). Define

$$\|\mathbf{B}\|^* := \max\{|\text{tr}(\mathbf{B}^* \mathbf{A})| : \mathbf{A} \in M_{m \times n}(F), \|\mathbf{A}\| \leq 1\}.$$

Then $\|\mathbf{B}\|^* = \Phi^*(\sigma(\mathbf{B}))$, where Φ^* is the dual gauge of Definition 21.3.3. In particular $\|\cdot\|^*$ is again a unitarily invariant norm, the trace norm $\|\cdot\|_{(p)}$ of Definition 21.4.2 and the spectral norm $\|\cdot\|_2$ are dual to each other, and the Frobenius norm is self-dual.

Idea. Both inequalities are one line each once the trace inequality is available. For $\leq$, bound the pairing by $\sum_i \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$ and then by $\Phi(\sigma(\mathbf{A}))\Phi^*(\sigma(\mathbf{B}))$, which is what a dual norm is for. For $\geq$, build the competitor $\mathbf{A}$ out of $\mathbf{B}$: take the vector where $\Phi^*(\sigma(\mathbf{B}))$ is attained and hang it on $\mathbf{B}$'s own singular vectors, which makes the pair aligned and turns the trace inequality into an equality.

Proof. Write $\sigma(\mathbf{B}) \in \mathbb{R}^p$ for the singular value list of $\mathbf{B}$, and fix a singular value decomposition $\mathbf{B} = \mathbf{U}\Sigma_{\mathbf{B}}\mathbf{V}^*$ (Theorem 13.9.1).

($\leq$) Let $\|\mathbf{A}\| \leq 1$. By Theorem 21.5.2 and then by the defining property of a dual norm (Proposition 18.4.6 (b), applied to the norm Φ of Definition 21.3.3), which gives $\langle \mathbf{x}, \mathbf{y} \rangle \leq \Phi(\mathbf{x})\Phi^*(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$,

$$\begin{aligned} |\mathrm{tr}(\mathbf{B}^*\mathbf{A})| &\leq \sum_{i=1}^p \sigma_i(\mathbf{A})\sigma_i(\mathbf{B}) \leq \Phi(\sigma(\mathbf{A}))\Phi^*(\sigma(\mathbf{B})) \\ &= \|\mathbf{A}\| \Phi^*(\sigma(\mathbf{B})) \leq \Phi^*(\sigma(\mathbf{B})). \end{aligned}$$

($\geq$) By Definition 21.3.3 there is $\mathbf{x} \in \mathbb{R}^p$ with $\Phi(\mathbf{x}) \leq 1$ and $\langle \mathbf{x}, \sigma(\mathbf{B}) \rangle = \Phi^*(\sigma(\mathbf{B}))$. We may assume $x_1 \geq \cdots \geq x_p \geq 0$. Indeed, replacing each x_i by $|x_i|$ leaves $\Phi(\mathbf{x})$ unchanged, by the sign invariance in Definition 21.3.1, and does not decrease $\langle \mathbf{x}, \sigma(\mathbf{B}) \rangle$ because every $\sigma_i(\mathbf{B}) \geq 0$; sorting the result decreasingly leaves Φ unchanged, by the permutation invariance in Definition 21.3.1, and does not decrease $\langle \mathbf{x}, \sigma(\mathbf{B}) \rangle$ by the rearrangement inequality (Lemma 21.5.1 with a permutation matrix). Since $\Phi^*(\sigma(\mathbf{B}))$ is by definition the largest value of $\langle \cdot, \sigma(\mathbf{B}) \rangle$ on $\{\Phi \leq 1\}$, the value is still exactly $\Phi^*(\sigma(\mathbf{B}))$. Now put

$$\mathbf{A} := \mathbf{U}\mathbf{X}\mathbf{V}^*, \quad \mathbf{X} \in M_{m \times n}(F) \text{ with } (\mathbf{X})_{ii} = x_i \ (i \leq p)$$

and all other entries 0. Since $x_1 \geq \cdots \geq x_p \geq 0$, this is a singular value decomposition of $\mathbf{A}$, so $\sigma(\mathbf{A}) = \mathbf{x}$ and $\|\mathbf{A}\| = \Phi(\mathbf{x}) \leq 1$. It is aligned with the chosen decomposition of $\mathbf{B}$, so the equality case of Theorem 21.5.2 gives

$$|\mathrm{tr}(\mathbf{B}^*\mathbf{A})| = \sum_{i=1}^p x_i \sigma_i(\mathbf{B}) = \Phi^*(\sigma(\mathbf{B})).$$

Hence the maximum is at least $\Phi^*(\sigma(\mathbf{B}))$, and the two bounds together prove the formula. (The maximum is attained, by the competitor just built.)

The remaining claims follow. By Theorem 21.3.5, Φ^* is a symmetric gauge, so $\|\cdot\|^*$ is the unitarily invariant norm it corresponds to under Theorem 21.4.3. The trace norm is the one with $\Phi = \|\cdot\|_1$ and the spectral norm the one with $\Phi = \|\cdot\|_\infty$, by Lemma 21.4.1 and Theorem 16.3.3 (c); these gauges are dual to each other by Proposition 18.4.7 (a), (b). The Frobenius norm has $\Phi = \|\cdot\|_2$, which is self-dual by Proposition 18.4.7 (c). This proves the corollary. ■

So the two extreme **normalized** unitarily invariant norms of Section 4 — the smallest, $\|\cdot\|_2$, and the largest, the trace norm, extreme among the normalized ones by Proposition 21.4.4 (c) — are exchanged by duality, and the Frobenius norm sits at the fixed point in the middle. Since $\Phi^{**} = \Phi$ (Theorem 21.3.5), dualizing twice returns the original norm.

Quick check. What does Corollary 21.5.3 say for $\mathbf{B} = \mathbf{e}_1\mathbf{e}_1^*$, a single matrix unit, and $\|\cdot\|$ the spectral norm?

Solution. Here $\sigma(\mathbf{B}) = (1, 0, \dots, 0)$, and the gauge of the spectral norm is $\|\cdot\|_\infty$, whose dual is $\|\cdot\|_1$. So $\|\mathbf{B}\|^* = \|(1, 0, \dots, 0)\|_1 = 1$: the largest value of $|\operatorname{tr}(\mathbf{B}^*\mathbf{A})| = |a_{11}|$ over matrices with $\|\mathbf{A}\|_2 \leq 1$ is 1, attained at $\mathbf{A} = \mathbf{e}_1\mathbf{e}_1^*$. An entry of a matrix never exceeds its spectral norm, and for a matrix unit that is sharp.

21.5.5 Exercises

21.5.5.1 A. Check your understanding

Exercise. A1

- Define a doubly substochastic matrix, and say which doubly substochastic matrices are doubly stochastic.
- State Theorem 21.5.2, with its hypotheses, and say what “aligned” means.
- Determine whether the following is true for all $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$, and justify your answer: $|\operatorname{tr}(\mathbf{A}\mathbf{B})| \leq \sum_i \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$.
- Where in the proof of Theorem 21.5.2 is the unitarity of $\mathbf{M}$ and $\mathbf{N}$ used, and what exactly is deduced from it?
- Determine whether the following is true, and justify your answer: if $\operatorname{tr}(\mathbf{A}^*\mathbf{B}) = \sum_i \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$ then $\mathbf{A} = \mathbf{B}$.

Solution.

- A square real matrix with nonnegative entries whose every row sum and every column sum is at most 1 (Definition 21.5.1). It is doubly stochastic exactly when all those sums equal 1.
- For $\mathbf{A}, \mathbf{B} \in M_{m \times n}(F)$ and $p = \min(m, n)$, $|\operatorname{tr}(\mathbf{A}^*\mathbf{B})| \leq \sum_{i \leq p} \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$. The pair is aligned when there are unitary $\mathbf{U}, \mathbf{V}$ with $\mathbf{A} = \mathbf{U}\Sigma_{\mathbf{A}}\mathbf{V}^*$ and $\mathbf{B} = \mathbf{U}\Sigma_{\mathbf{B}}\mathbf{V}^*$, the **same** $\mathbf{U}$ and $\mathbf{V}$ for both; then equality holds.
- True. Apply Theorem 21.5.2 to the pair $\mathbf{A}^*$ and $\mathbf{B}$, using $\operatorname{tr}(\mathbf{A}\mathbf{B}) = \operatorname{tr}((\mathbf{A}^*)^*\mathbf{B})$ and $\sigma(\mathbf{A}^*) = \sigma(\mathbf{A})$, which is Proposition 21.4.4 (a).
- Only in the claim, and only to say that each row of $\mathbf{M}$ and each column of $\mathbf{N}$ is a unit vector. Cauchy–Schwarz then makes the weight array doubly substochastic. Nothing else about $\mathbf{M}$ and $\mathbf{N}$ is used.
- False. Take $\mathbf{A} = \mathbf{I}_2$ and $\mathbf{B} = 2\mathbf{I}_2$: both sides are 4, and $\mathbf{A} \neq \mathbf{B}$. The theorem says that alignment forces equality, and $\mathbf{A}$ and $\mathbf{B}$ are aligned, with $\mathbf{U} = \mathbf{V} = \mathbf{I}$; it says nothing that would force the two matrices to be equal.

21.5.5.2 B. Practice

Exercise. B1: Both sides of the inequality

Let

$$\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix}.$$

Compute $\text{tr}(\mathbf{A}^\top \mathbf{B})$, the singular values of both matrices, and both sides of [Theorem 21.5.2](#). Then find a matrix $\mathbf{B}''$ with the same singular values as $\mathbf{B}$ for which equality holds.

Solution. $\mathbf{A}^\top \mathbf{B} = \begin{pmatrix} 0 & 12 \\ 2 & 0 \end{pmatrix}$, so $\text{tr}(\mathbf{A}^\top \mathbf{B}) = 0$. Since $\mathbf{A}$ is diagonal with positive decreasing entries, $\sigma(\mathbf{A}) = (3, 1)$. For $\mathbf{B}$, $\mathbf{B}^\top \mathbf{B} = \text{diag}(4, 16)$, so $\sigma(\mathbf{B}) = (4, 2)$. The right-hand side is $3 \cdot 4 + 1 \cdot 2 = 14$, and $0 \leq 14$ with a lot to spare: $\mathbf{B}$ sends the direction $\mathbf{A}$ stretches most to the one it stretches least.

For equality, align: take $\mathbf{B}'' = \text{diag}(4, 2)$. Then $\text{tr}(\mathbf{A}^\top \mathbf{B}'') = 12 + 2 = 14$, and $\sigma(\mathbf{B}'') = (4, 2)$.

Exercise. B2: The trace against the trace norm

Let $\mathbf{A} \in M_n(\mathbb{C})$.

- (a) Prove that $|\text{tr } \mathbf{A}| \leq \sum_{i=1}^n \sigma_i(\mathbf{A})$.
- (b) Determine all $\mathbf{A}$ for which equality holds with $\mathbf{A}$ positive semidefinite.

Solution.

- (a) Apply [Theorem 21.5.2](#) with the pair $\mathbf{I}_n$ and $\mathbf{A}$, in that order: $\text{tr}(\mathbf{I}^* \mathbf{A}) = \text{tr } \mathbf{A}$, and $\sigma_i(\mathbf{I}) = 1$ for every i , so $|\text{tr } \mathbf{A}| \leq \sum_i 1 \cdot \sigma_i(\mathbf{A})$.
- (b) For $\mathbf{A} \succeq \mathbf{O}$ the eigenvalues are nonnegative ([Theorem 13.1.4](#)) and, by the spectral theorem, $\mathbf{A} = \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^*$ with $\lambda_i \geq 0$ decreasing; this is a singular value decomposition, so $\sigma_i(\mathbf{A}) = \lambda_i(\mathbf{A})$. Hence $\text{tr } \mathbf{A} = \sum_i \lambda_i = \sum_i \sigma_i(\mathbf{A})$ by [Corollary 12.1.4](#): equality holds for **every** positive semidefinite $\mathbf{A}$. (Consistently with the theorem, $\mathbf{I}$ and such an $\mathbf{A}$ are aligned, with $\mathbf{V} = \mathbf{U}$.)

Exercise. B3: A weight matrix

Let $\mathbf{A} = \text{diag}(5, 1)$ and let $\mathbf{B} = \mathbf{U} \text{diag}(2, 1) \mathbf{U}^*$ with $\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Write down the matrices $\mathbf{M}$ and $\mathbf{N}$ of the proof of [Theorem 21.5.2](#) for this pair, the weight matrix $\mathbf{D}$, and verify [Eq. 21.5.1](#) and the inequality numerically.

Solution. Take $\mathbf{U}_1 = \mathbf{V}_1 = \mathbf{I}$ for $\mathbf{A}$, and $\mathbf{U}_2 = \mathbf{U}$, $\mathbf{V}_2 = \mathbf{U}$ for $\mathbf{B}$ (the decomposition given is a singular value decomposition, since $\text{diag}(2, 1)$ is diagonal with positive decreasing entries). Then $\mathbf{M} = \mathbf{U}_1^* \mathbf{U}_2 = \mathbf{U}$ and $\mathbf{N} = \mathbf{V}_2^* \mathbf{V}_1 = \mathbf{U}^*$, so

$$m_{kl} n_{lk} = u_{kl} \overline{u_{kl}} = |u_{kl}|^2 = \frac{1}{2}$$

for all k, l , and $\mathbf{D} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, which is doubly stochastic, hence doubly substochastic. Now Eq. 21.5.1 gives

$$\operatorname{tr}(\mathbf{A}^* \mathbf{B}) = \frac{1}{2}(5 \cdot 2 + 5 \cdot 1 + 1 \cdot 2 + 1 \cdot 1) = \frac{18}{2} = 9,$$

which checks against $\mathbf{B} = \frac{1}{2} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and $\operatorname{tr}(\mathbf{A} \mathbf{B}) = \frac{1}{2}(5 \cdot 3 + 1 \cdot 3) = 9$. The bound is $5 \cdot 2 + 1 \cdot 1 = 11$, and $9 \leq 11$.

21.5.5.3 C. Going deeper

Exercise. C1: The triangle inequality for the trace norm

Write $\|\mathbf{A}\|_{(p)} = \sum_i \sigma_i(\mathbf{A})$ for the trace norm of Definition 21.4.2, $\mathbf{A} \in M_{m \times n}(\mathbb{C})$.

- (a) Prove that $\|\mathbf{A}\|_{(p)} = \max\{|\operatorname{tr}(\mathbf{A}^* \mathbf{Q})| : \mathbf{Q} \in M_{m \times n}(\mathbb{C}), \|\mathbf{Q}\|_2 \leq 1\}$.
- (b) Hence deduce that $\|\cdot\|_{(p)}$ satisfies the triangle inequality, without using Lemma 21.4.2.

Hint: for (a), Corollary 21.5.3 with the spectral norm, read the other way round.

Solution.

- (a) The spectral norm is unitarily invariant, with gauge $\Phi = \|\cdot\|_\infty$ (Lemma 21.4.1, Theorem 16.3.3 (c)). By Corollary 21.5.3 applied to that norm, the maximum in the statement equals $\Phi^*(\sigma(\mathbf{A})) = \|\sigma(\mathbf{A})\|_1 = \sum_i \sigma_i(\mathbf{A})$, using Proposition 18.4.7 (b).
- (b) Conjugate transposition is additive, so $\operatorname{tr}((\mathbf{A} + \mathbf{B})^* \mathbf{Q}) = \operatorname{tr}(\mathbf{A}^* \mathbf{Q}) + \operatorname{tr}(\mathbf{B}^* \mathbf{Q})$ for every $\mathbf{Q}$. Given $\mathbf{A}, \mathbf{B}$, choose by (a) a $\mathbf{Q}$ with $\|\mathbf{Q}\|_2 \leq 1$ attaining the maximum for $\mathbf{A} + \mathbf{B}$; then, by the triangle inequality for the modulus (Theorem 0.6.4) and (a) again,

$$\|\mathbf{A} + \mathbf{B}\|_{(p)} = |\operatorname{tr}((\mathbf{A} + \mathbf{B})^* \mathbf{Q})| \leq |\operatorname{tr}(\mathbf{A}^* \mathbf{Q})| + |\operatorname{tr}(\mathbf{B}^* \mathbf{Q})| \leq \|\mathbf{A}\|_{(p)} + \|\mathbf{B}\|_{(p)}.$$

No step used Lemma 21.4.2: the route ran through the trace inequality and duality alone.

Exercise. C2: The nearest unitary, again

Let $\mathbf{A} \in M_n(\mathbb{C})$ with polar decomposition $\mathbf{A} = \mathbf{W}|\mathbf{A}|$ (Theorem 13.10.1).

- (a) Prove that $\max\{\operatorname{Re} \operatorname{tr}(\mathbf{Q}^* \mathbf{A}) : \mathbf{Q} \in M_n(\mathbb{C}) \text{ unitary}\} = \sum_i \sigma_i(\mathbf{A})$, attained at $\mathbf{Q} = \mathbf{W}$.
- (b) Deduce Corollary 13.10.2, that $\mathbf{W}$ minimizes $\|\mathbf{A} - \mathbf{Q}\|_F$ over unitary $\mathbf{Q}$.

Solution.

- (a) For unitary $\mathbf{Q}$, $\sigma_i(\mathbf{Q}) = 1$ for every i , so Theorem 21.5.2 gives $\operatorname{Re} \operatorname{tr}(\mathbf{Q}^* \mathbf{A}) \leq |\operatorname{tr}(\mathbf{Q}^* \mathbf{A})| \leq \sum_i \sigma_i(\mathbf{A})$. For attainment, fix a singular value decomposition $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*$ and recall from the proof of Theorem 13.10.1 that $\mathbf{W} = \mathbf{U} \mathbf{V}^*$. Then $\mathbf{W}^* \mathbf{A} = \mathbf{V} \mathbf{U}^* \mathbf{U} \mathbf{\Sigma} \mathbf{V}^* = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^*$, whose trace is $\operatorname{tr} \mathbf{\Sigma} = \sum_i \sigma_i(\mathbf{A})$ by Theorem 0.9.6 (3); it is real.
- (b) For unitary $\mathbf{Q}$, expanding in the Frobenius inner product,

$$\|\mathbf{A} - \mathbf{Q}\|_F^2 = \|\mathbf{A}\|_F^2 + n - 2 \operatorname{Re} \operatorname{tr}(\mathbf{Q}^* \mathbf{A}),$$

since $\|\mathbf{Q}\|_F^2 = \operatorname{tr}(\mathbf{Q}^* \mathbf{Q}) = n$. Only the last term depends on $\mathbf{Q}$, and it enters with a

minus sign, so minimizing the left side is maximizing $\operatorname{Re} \operatorname{tr}(\mathbf{Q}^* \mathbf{A})$. By (a) the maximum is attained at $\mathbf{Q} = \mathbf{W}$, which is Corollary 13.10.2.