

21.7 Lidskii's Theorem in Majorization Form

Chapter 17 §07 proved Lidskii's inequality and then translated the whole family of inequalities into one majorization. It could go no further, and said so. This section spends what it banked: the majorization is already in hand, and Sections 2, 3, and 6 turn it into a statement about every convex function and every unitarily invariant norm. The last part of the section carries all of it across to singular values.

21.7.1 What Chapter 17 proved, and what it left

Here is the statement we start from. For Hermitian $\mathbf{A}, \mathbf{B} \in M_n(F)$, Theorem 17.7.2 says that for every index set $1 \leq i_1 < \cdots < i_k \leq n$,

$$\sum_{j=1}^k \lambda_{i_j}(\mathbf{A}) - \sum_{j=1}^k \lambda_{i_j}(\mathbf{B}) \leq \sum_{j=1}^k \lambda_j(\mathbf{A} - \mathbf{B}),$$

and the remark that closes that section observes that, since the largest k entries of a vector are the best choice of k of them, this family of inequalities is exactly clause (M1) of Definition 17.8.1 for the pair of vectors below, while clause (M2) is the identity $\text{tr } \mathbf{A} - \text{tr } \mathbf{B} = \text{tr}(\mathbf{A} - \mathbf{B})$:

$$\lambda(\mathbf{A}) - \lambda(\mathbf{B}) \prec \lambda(\mathbf{A} - \mathbf{B}). \quad (21.7.1)$$

This is proved, in Chapter 17, and we do not prove it again. What Chapter 17 could not do it named exactly: “Not proved here, and left to Chapter 21: the general consequences of a majorization $\mathbf{x} \prec \mathbf{y}$, namely $\sum_i \phi(x_i) \leq \sum_i \phi(y_i)$ for every convex $\phi: \mathbb{R} \rightarrow \mathbb{R}$, together with the matching statement for the unitarily invariant norms of that chapter.”

Both consequences now exist. Section 2 proved the first, as Theorem 21.2.1. Section 3 proved that a symmetric gauge is monotone under majorization, as Theorem 21.3.4, and Section 4 identified the gauges with the unitarily invariant norms. All that is left is to apply them, which takes two short proofs, and then to repeat the exercise for singular values, which takes one idea.

Example. A pair to keep in mind

Let

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

Compute the three eigenvalue lists and check [Eq. 21.7.1](#).

Solution. $\mathbf{A}$ is symmetric with $\text{tr } \mathbf{A} = 4$ and $\det \mathbf{A} = 0$, so $\lambda(\mathbf{A}) = (4, 0)$. $\mathbf{B}$ is symmetric with $\text{tr } \mathbf{B} = 0$ and $\det \mathbf{B} = -1$, so $\lambda(\mathbf{B}) = (1, -1)$. Their difference

$$\mathbf{A} - \mathbf{B} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

has trace 4 and determinant $4 - 9 = -5$, so $\lambda(\mathbf{A} - \mathbf{B}) = (5, -1)$.

The vector of shifts is $\lambda(\mathbf{A}) - \lambda(\mathbf{B}) = (4 - 1, 0 - (-1)) = (3, 1)$, already decreasing. Now

compare with $(5, -1)$: the top sums are $3 \leq 5$, and the totals are $4 = 4$. So $(3, 1) \prec (5, -1)$, as Eq. 21.7.1 promises, and the majorization is strict in its one nontrivial inequality.

21.7.2 Every convex function at once

Theorem 21.7.1. Lidskii's Theorem, Convex Form

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian, $F = \mathbb{R}$ or $\mathbb{C}$. Then for every convex $\phi: \mathbb{R} \rightarrow \mathbb{R}$,

$$\sum_{i=1}^n \phi(\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})) \leq \sum_{i=1}^n \phi(\lambda_i(\mathbf{A} - \mathbf{B})).$$

Proof. By Eq. 21.7.1, $\lambda(\mathbf{A}) - \lambda(\mathbf{B}) \prec \lambda(\mathbf{A} - \mathbf{B})$. Apply Theorem 21.2.1 to this majorization, with $I = \mathbb{R}$ and the given ϕ . This proves the theorem. ■

Three readings are worth having. With $\phi(t) = |t|$, which is convex, the theorem says that the total movement of the eigenvalues is at most the total size of the eigenvalues of the perturbation. With $\phi(t) = t^2$ it is the Hoffman–Wielandt inequality, as the next corollary spells out. With $\phi(t) = \max(t, 0)$ it says that the total **upward** movement of the eigenvalues is at most the total of the positive eigenvalues of the perturbation, and with $\phi(t) = e^{ct}$ it gives bounds no earlier chapter could have stated.

Warning. The left side pairs $\lambda_i(\mathbf{A})$ with $\lambda_i(\mathbf{B})$ at the **same index**, both lists being in decreasing order, and the theorem claims nothing about any other pairing. For $\mathbf{A} = \text{diag}(4, 0)$ and $\mathbf{B} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ the spectra are equal, $\lambda(\mathbf{A}) = \lambda(\mathbf{B}) = (4, 0)$, so the sorted pairing gives 0 with $\phi(t) = t^2$. The reversed pairing gives $(4 - 0)^2 + (0 - 4)^2 = 32$, while $\mathbf{A} - \mathbf{B} = \begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix}$ has $\sum_i \lambda_i(\mathbf{A} - \mathbf{B})^2 = \|\mathbf{A} - \mathbf{B}\|_F^2 = 16$. So the same inequality with a different matching is simply false.

21.7.3 Every unitarily invariant norm at once

To state the norm version we need to put a vector of shifts inside a matrix, and Section 4 has already done it. By Lemma 21.4.1 the gauge of a unitarily invariant norm is recovered from the norm by

$$\|\text{diag}_{m,n}(\mathbf{x})\| = \Phi(\mathbf{x}) \quad (\mathbf{x} \in \mathbb{R}^p), \quad (21.7.2)$$

where $\text{diag}_{m,n}(\mathbf{x})$ carries $\mathbf{x}$ down the main diagonal and zeros elsewhere, and is written $\text{diag}(\mathbf{x})$ when $m = n$. One more reading will be used throughout: if $\mathbf{H} \in M_n(F)$ is Hermitian then $\mathbf{H}^* \mathbf{H} = \mathbf{H}^2$ has eigenvalues $\lambda_i(\mathbf{H})^2$, so $\sigma(\mathbf{H}) = |\lambda(\mathbf{H})|^\downarrow$ by Definition 13.9.1, and hence $\|\mathbf{H}\| = \Phi(\sigma(\mathbf{H})) = \Phi(\lambda(\mathbf{H}))$, the last step by the permutation and sign invariance of Definition 21.3.1.

Theorem 21.7.2. Lidskii's Theorem, Norm Form

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian. Then for **every** unitarily invariant norm on $M_n(F)$,

$$\|\text{diag}(\lambda(\mathbf{A}) - \lambda(\mathbf{B}))\| \leq \|\mathbf{A} - \mathbf{B}\|.$$

Idea. Both sides are the same symmetric gauge evaluated at two real vectors: the vector of eigenvalue shifts on the left, the eigenvalue vector of $\mathbf{A} - \mathbf{B}$ on the right. Chapter 17 says the first is majorized by the second, and Section 3 says a symmetric gauge respects that.

Proof. Let Φ be the symmetric gauge of the norm, so that $\|\mathbf{M}\| = \Phi(\sigma(\mathbf{M}))$ (Theorem 21.4.3). Put $\mathbf{x} = \lambda(\mathbf{A}) - \lambda(\mathbf{B})$ and $\mathbf{y} = \lambda(\mathbf{A} - \mathbf{B})$, both in $\mathbb{R}^n$. By Eq. 21.7.2 the left-hand side is $\Phi(\mathbf{x})$; and since $\mathbf{A} - \mathbf{B}$ is Hermitian, the discussion above gives $\|\mathbf{A} - \mathbf{B}\| = \Phi(\mathbf{y})$. By Eq. 21.7.1, $\mathbf{x} \prec \mathbf{y}$, so Theorem 21.3.4 (a) — the part for arbitrary real vectors, which is what majorization rather than weak majorization buys — gives $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$. This proves the theorem. ■

That is the second half of Chapter 17 §07's promise, and it is also what Chapter 20 §11 was waiting for when it listed, among the things that chapter did not do, that its bounds “hold in the Frobenius norm” and that “extending the eigenvalue and singular value bounds to every unitarily invariant norm belongs to Chapter 21”. Two specializations make the point. With the spectral norm, whose gauge is $\|\cdot\|_\infty$, the theorem reads $\max_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2$, which is Weyl's perturbation bound Corollary 17.3.6. With the Frobenius norm, whose gauge is $\|\cdot\|_2$, it is Hoffman–Wielandt. One statement, two classical corollaries, and every norm in between.

Corollary 21.7.3. Hoffman–Wielandt for Hermitian Matrices, Again

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian. Then

$$\sum_{i=1}^n (\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B}))^2 \leq \|\mathbf{A} - \mathbf{B}\|_F^2.$$

Proof. The function $\phi(t) = t^2$ is convex, by Theorem 18.10.6 (a) with $\phi'' = 2 > 0$. By Theorem 21.7.1,

$$\sum_i (\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B}))^2 \leq \sum_i \lambda_i(\mathbf{A} - \mathbf{B})^2.$$

The matrix $\mathbf{E} = \mathbf{A} - \mathbf{B}$ is Hermitian, so the spectral theorem writes $\mathbf{E} = \mathbf{U} \text{diag}(\lambda(\mathbf{E})) \mathbf{U}^*$ with $\mathbf{U}$ unitary, and Lemma 11.7.5 gives $\|\mathbf{E}\|_F^2 = \|\text{diag}(\lambda(\mathbf{E}))\|_F^2 = \sum_i \lambda_i(\mathbf{E})^2$. This proves the corollary. ■

This is Corollary 20.5.4 of Chapter 20 §05, reached by a different road. That section said so in advance: “Chapter 21 shows that a majorization $\mathbf{x} \prec \mathbf{y}$ gives $\sum_i \phi(x_i) \leq \sum_i \phi(y_i)$ for

every convex ϕ ; with $\phi(t) = t^2$ this is Corollary 20.5.4 again, and Chapter 21 turns the same majorization into one bound for every unitarily invariant norm. Nothing here depends on that route; the proof above goes through Birkhoff's theorem . . ." So the two proofs are independent: Chapter 20 argued from Birkhoff's theorem applied to the matrix of squared moduli of the entries of a unitary change of basis, and never used Lidskii's inequality; this one uses Lidskii and never mentions doubly stochastic matrices. Neither is a special case of the other, and Chapter 20 reaches further in one direction: its Corollary 20.5.4 is the Hermitian case of Theorem 20.5.2, which holds for all **normal** $\mathbf{A}, \mathbf{B}$. What that theorem gives there is weaker in kind — some permutation matching the two eigenvalue lists, not the sorted one, since complex eigenvalues have no decreasing order — and the argument here, resting on Eq. 21.7.1, stays with Hermitian matrices throughout.

Quick check. Check Theorem 21.7.2 and Corollary 21.7.3 for the pair of Example (A pair to keep in mind), in the spectral, Frobenius, and trace norms.

Solution. There $\lambda(\mathbf{A}) - \lambda(\mathbf{B}) = (3, 1)$ and $\lambda(\mathbf{A} - \mathbf{B}) = (5, -1)$, so $\text{diag}(3, 1)$ has singular values $(3, 1)$ and $\mathbf{A} - \mathbf{B}$ has singular values $(5, 1)$. The three norms of the left and right sides are: spectral, $3 \leq 5$; Frobenius, $\sqrt{10} \leq \sqrt{26}$; trace, $4 \leq 6$. The Frobenius case squared is $9 + 1 = 10 \leq 26$, which is Corollary 21.7.3.

21.7.4 The same theorem for singular values

Chapter 20 §10 proved Mirsky's theorem in the Frobenius norm and closed with: "Chapter 21 puts both inside one statement covering every unitarily invariant norm", the two being the index-by-index bound Corollary 17.9.5 and the Frobenius bound Theorem 20.10.1. The route is the one that section used and Section 6 used again: **a Hermitian dilation turns singular values into eigenvalues**, and then Lidskii's theorem applies. The only delicate point is the bookkeeping of the zeros, because a rectangular matrix dilates to a matrix with extra zero eigenvalues in the middle of its spectrum.

Theorem 21.7.4. Mirsky's Theorem in Every Unitarily Invariant Norm

Let $\mathbf{A}, \mathbf{B} \in M_{m \times n}(\mathbb{C})$ and $p = \min(m, n)$. Then for **every** unitarily invariant norm on $M_{m \times n}(\mathbb{C})$,

$$\left\| \text{diag}_{m,n}(\sigma(\mathbf{A}) - \sigma(\mathbf{B})) \right\| \leq \|\mathbf{A} - \mathbf{B}\|.$$

Idea. Dilate all three matrices. The eigenvalue list of $\mathcal{H}(\mathbf{M})$ is $\sigma(\mathbf{M})$, then $m + n - 2p$ zeros, then $-\sigma(\mathbf{M})$ reversed, and the **positions** of the three blocks depend only on m and n — so $\mathbf{A}$ and $\mathbf{B}$ have their singular values in the same slots, and subtracting the two lists subtracts them index by index. Lidskii's majorization then compares a symmetric list $\pm(\sigma_i(\mathbf{A}) - \sigma_i(\mathbf{B}))$ with a symmetric list $\pm\sigma_j(\mathbf{A} - \mathbf{B})$ in $\mathbb{R}^{m+n}$. Reading its top- k sums for $k \leq p$ turns it into a weak majorization of the two lists we care about, in $\mathbb{R}^p$, and Section 3 finishes.

Proof. Write $N = m + n$, let $\delta_i = \sigma_i(\mathbf{A}) - \sigma_i(\mathbf{B})$ for $i \leq p$ and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_p)$, and let Φ be the symmetric gauge of the norm.

Step 1: what has to be proved. By Eq. 21.7.2 and the sign invariance of Definition 21.3.1, $\|\text{diag}_{m,n}(\delta)\| = \Phi(\delta) = \Phi(|\delta|)$; and $\|\mathbf{A} - \mathbf{B}\| = \Phi(\sigma(\mathbf{A} - \mathbf{B}))$ by Theorem 21.4.3. Both $|\delta|$ and $\sigma(\mathbf{A} - \mathbf{B})$ have nonnegative entries, so by Theorem 21.3.4 (b) it is enough to prove the weak majorization

$$|\delta| \prec_w \sigma(\mathbf{A} - \mathbf{B}) \quad \text{in } \mathbb{R}^p. \quad (21.7.3)$$

Step 2: the dilations and their spectra. For $\mathbf{M} \in M_{m \times n}(\mathbb{C})$ let $\mathcal{H}(\mathbf{M}) = \begin{pmatrix} \mathbf{O} & \mathbf{M} \\ \mathbf{M}^* & \mathbf{O} \end{pmatrix} \in M_N(\mathbb{C})$ be the Hermitian dilation of Proposition 17.9.3. Conjugate transposition is additive, so

$$\mathcal{H}(\mathbf{A}) - \mathcal{H}(\mathbf{B}) = \mathcal{H}(\mathbf{A} - \mathbf{B}),$$

and all three matrices are Hermitian. By Proposition 17.9.3 the eigenvalues of $\mathcal{H}(\mathbf{M})$ are $\pm\sigma_1(\mathbf{M}), \dots, \pm\sigma_p(\mathbf{M})$ together with $N - 2p$ zeros; since the $\sigma_i(\mathbf{M})$ are nonnegative and decreasing, the decreasing list is

$$\lambda_i(\mathcal{H}(\mathbf{M})) = \begin{cases} \sigma_i(\mathbf{M}) & 1 \leq i \leq p, \\ 0 & p < i \leq N - p, \\ -\sigma_{N+1-i}(\mathbf{M}) & N - p < i \leq N. \end{cases} \quad (21.7.4)$$

The three ranges depend only on m and n , so the formula applies to $\mathbf{A}$, to $\mathbf{B}$, and to $\mathbf{A} - \mathbf{B}$ **with the same ranges**. This is the bookkeeping the rectangular case needs: the zeros of $\mathcal{H}(\mathbf{A})$ sit in exactly the positions of the zeros of $\mathcal{H}(\mathbf{B})$, and so cancel when the two lists are subtracted.

Step 3: Lidskii. By Eq. 21.7.1 applied to the Hermitian pair $\mathcal{H}(\mathbf{A}), \mathcal{H}(\mathbf{B})$,

$$\lambda(\mathcal{H}(\mathbf{A})) - \lambda(\mathcal{H}(\mathbf{B})) \prec \lambda(\mathcal{H}(\mathbf{A} - \mathbf{B})).$$

By Eq. 21.7.4 the vector on the left has entries δ_i in positions $i \leq p$, zeros in the middle range, and $-\delta_{N+1-i}$ in the last p positions; the vector on the right has $\sigma_j(\mathbf{A} - \mathbf{B})$, zeros, and the negatives of the $\sigma_j(\mathbf{A} - \mathbf{B})$ in the same three ranges.

Step 4: read the running totals. Fix $k \leq p$. The left-hand vector is the symmetric list of Lemma 21.6.3 with $q = p$, $r = N - 2p$, and $z_i = \delta_i$, so by that lemma its k -th running total — the largest sum of k of its entries — is $\sum_{j \leq k} |\delta|_j^\downarrow$. The right-hand vector is already decreasing, so its k -th running total is $\sum_{j \leq k} \sigma_j(\mathbf{A} - \mathbf{B})$. Since $k \leq p < N$, clause (M1) of Definition 17.8.1 applies at this k and gives

$$\sum_{j=1}^k |\delta|_j^\downarrow \leq \sum_{j=1}^k \sigma_j(\mathbf{A} - \mathbf{B}).$$

As $k \leq p$ was arbitrary, this is Eq. 21.7.3, and Step 1 finishes the proof. ■

Remark. Where the factor of two went. Chapter 20 §10 proved the Frobenius case by taking a **norm** of the dilation, and paid for it: $\|\mathcal{H}(\mathbf{M})\|_F^2 = 2\|\mathbf{M}\|_F^2$, so a factor 2 appeared on each side and canceled. That cancellation is special to a norm which adds squares, and

for a general unitarily invariant norm there is no such identity — indeed $\|\mathcal{H}(\mathbf{M})\|$ is a norm on a different matrix space, computed by a gauge on $\mathbb{R}^N$ rather than on $\mathbb{R}^p$. The proof above never takes a norm of a dilation. It uses the dilation only to obtain a **majorization of eigenvalue lists**, and then reads off the p running totals it needs. No factor of 2 arises, and none has to cancel.

Specializing the gauge recovers both statements that Chapter 20 §10 wanted inside one: $\Phi = \|\cdot\|_\infty$ gives $\max_i |\sigma_i(\mathbf{A}) - \sigma_i(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2$, which is Corollary 17.9.5, and $\Phi = \|\cdot\|_2$ gives Theorem 20.10.1. Chapter 20 §11's ledger entry — “extending the eigenvalue and singular value bounds to every unitarily invariant norm belongs to Chapter 21” — is now paid for the eigenvalue bounds by Theorem 21.7.2 and for the singular value bounds by Theorem 21.7.4. Chapter 20 §10's square-root and polar bounds (Theorem 20.10.2, Theorem 20.10.3, Theorem 20.10.4) are **not** extended here: they still hold only in the two norms Chapter 20 proved them in, the spectral and the Frobenius.

The pair of Example (A pair to keep in mind) illustrates the singular value form as well. There $\mathbf{A} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ is positive semidefinite with eigenvalues 4, 0, so $\sigma(\mathbf{A}) = (4, 0)$; $\mathbf{B}$ has eigenvalues ± 1 , so $\sigma(\mathbf{B}) = (1, 1)$; and $\mathbf{A} - \mathbf{B}$ has eigenvalues 5, -1 , so $\sigma(\mathbf{A} - \mathbf{B}) = (5, 1)$. The vector of differences is $\sigma(\mathbf{A}) - \sigma(\mathbf{B}) = (3, -1)$, whose absolute values (3, 1) have running totals $3 \leq 5$ and $4 \leq 6$: the weak majorization of Eq. 21.7.3, and with it the norm inequality in every unitarily invariant norm.

Warning. Both Theorem 21.7.2 and Theorem 21.7.4 compare **sorted** lists, and the matching they use is the sorted one. No claim is made for any other pairing, and Theorem 21.7.4 in particular is not a statement about eigenvalues: for non-Hermitian matrices the eigenvalues can be wildly unstable while the singular values are not, which is the whole content of Chapter 20 §10's remark that Mirsky's theorem needs no normality hypothesis at all.

21.7.5 Exercises

21.7.5.1 A. Check your understanding

Exercise. A1

- State the majorization Eq. 21.7.1, and say which chapter proves it.
- State Theorem 21.7.2, and say what $\text{diag}(\lambda(\mathbf{A}) - \lambda(\mathbf{B}))$ means.
- Determine whether the following is true for Hermitian $\mathbf{A}, \mathbf{B}$, and justify your answer:
 $\sum_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq \sum_i |\lambda_i(\mathbf{A} - \mathbf{B})|$.
- Determine whether the following is true, and justify your answer: Theorem 21.7.4 holds with the singular values of $\mathbf{A}$ and $\mathbf{B}$ listed in **increasing** order on the left.
- Where in the proof of Theorem 21.7.4 is it used that $\mathbf{A}$ and $\mathbf{B}$ have the same size?

Solution.

- For Hermitian $\mathbf{A}, \mathbf{B} \in M_n(F)$, $\lambda(\mathbf{A}) - \lambda(\mathbf{B}) \prec \lambda(\mathbf{A} - \mathbf{B})$. It is proved in Chapter 17 §07, from Theorem 17.7.2 over all index sets together with the equality of traces.
- For every unitarily invariant norm, $\|\text{diag}(\lambda(\mathbf{A}) - \lambda(\mathbf{B}))\| \leq \|\mathbf{A} - \mathbf{B}\|$. The matrix

$\text{diag}(\mathbf{x})$ is the diagonal matrix in $M_n(F)$ whose i -th diagonal entry is x_i ; here $x_i = \lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})$, the shift of the i -th eigenvalue, the lists being in decreasing order.

- (c) True. Take $\phi(t) = |t|$ in [Theorem 21.7.1](#); it is convex, being a norm on $\mathbb{R}$. Alternatively it is [Theorem 21.7.2](#) for the trace norm, using [Eq. 21.7.2](#).
- (d) True, and for a cheap reason: reversing both lists changes the vector δ only by reversing it, and the gauge of a unitarily invariant norm is permutation-invariant ([Definition 21.3.1](#)), so the left-hand side is unchanged. What would **not** be true is to sort one list increasingly and the other decreasingly; see the warning after [Theorem 21.7.1](#).
- (e) Twice. The two matrices must lie in the same space for $\mathbf{A} - \mathbf{B}$ to be defined and for one norm to apply to all three matrices; and in Step 2, m, n , and p being the same for $\mathbf{A}$ and $\mathbf{B}$ is what puts the zeros of the two dilated spectra in the same positions, so that the difference of the two lists has the shape used in Step 4.

21.7.5.2 B. Practice

Exercise. B1: Running the three consequences

Let $\mathbf{A} = \text{diag}(5, 1, 0)$ and let $\mathbf{B} = \mathbf{A} - \mathbf{E}$ with

$$\mathbf{E} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

- (a) Compute $\lambda(\mathbf{B})$ and $\lambda(\mathbf{E})$.
- (b) Verify [Eq. 21.7.1](#) for this pair.
- (c) Verify [Theorem 21.7.2](#) in the spectral, Frobenius, and trace norms, and [Corollary 21.7.3](#).

Solution.

- (a) $\mathbf{E}$ acts as $\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ on $\text{span}(\mathbf{e}_1, \mathbf{e}_3)$ and as 0 on $\mathbf{e}_2$, so $\lambda(\mathbf{E}) = (2, 0, -2)$. Next,

$$\mathbf{B} = \mathbf{A} - \mathbf{E} = \begin{pmatrix} 5 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 0 \end{pmatrix}.$$

On $\text{span}(\mathbf{e}_1, \mathbf{e}_3)$ it is $\begin{pmatrix} 5 & -2 \\ -2 & 0 \end{pmatrix}$, with trace 5 and determinant -4 , so its eigenvalues there are $(5 \pm \sqrt{41})/2$; together with the eigenvalue 1 from $\mathbf{e}_2$,

$$\lambda(\mathbf{B}) = \left(\frac{5+\sqrt{41}}{2}, 1, \frac{5-\sqrt{41}}{2} \right) \approx (5.702, 1, -0.702).$$

- (b) The shifts are

$$\begin{aligned} \lambda(\mathbf{A}) - \lambda(\mathbf{B}) &= \left(5 - \frac{5+\sqrt{41}}{2}, 0, \frac{\sqrt{41}-5}{2} \right) \\ &= \left(\frac{5-\sqrt{41}}{2}, 0, \frac{\sqrt{41}-5}{2} \right) \approx (-0.702, 0, 0.702). \end{aligned}$$

Sorted decreasingly this is $(0.702, 0, -0.702)$, with running totals 0.702, 0.702, 0. The list $\lambda(\mathbf{E}) = (2, 0, -2)$ has running totals 2, 2, 0. Each of the first two is larger, and the totals agree at 0. So the majorization holds.

- (c) Write $c = \frac{\sqrt{41}-5}{2} \approx 0.702$. The left-hand matrix is $\text{diag}(-c, 0, c)$, whose singular values are $(c, c, 0)$; the right-hand matrix is $\mathbf{E}$, whose singular values are $(2, 2, 0)$. Spectral: $c \leq 2$. Frobenius: $c\sqrt{2} \approx 0.993 \leq 2\sqrt{2} \approx 2.828$. Trace: $2c \approx 1.403 \leq 4$. Squaring the Frobenius comparison gives

$$2c^2 = \frac{(\sqrt{41}-5)^2}{2} = \frac{66-10\sqrt{41}}{2} = 33 - 5\sqrt{41} \approx 0.984$$

on the left and $\|\mathbf{E}\|_F^2 = 8$ on the right, which is [Corollary 21.7.3](#).

Exercise. B2: Singular values of a one-entry change

Let $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$. Compute both sides of [Theorem 21.7.4](#) in the trace norm, and compare with what [Corollary 17.9.5](#) gives on its own.

Solution. $\mathbf{A}$ is symmetric with eigenvalues 4 and 2, both positive, so $\sigma(\mathbf{A}) = (4, 2)$. For $\mathbf{B}$, $\mathbf{B}^\top \mathbf{B} = \begin{pmatrix} 9 & 3 \\ 3 & 10 \end{pmatrix}$, with trace 19 and determinant 81, so its eigenvalues are $\frac{19 \pm \sqrt{37}}{2}$; these are the squares of $\frac{\sqrt{37}+1}{2}$, since $(\frac{\sqrt{37}+1}{2})^2 = \frac{38 \pm 2\sqrt{37}}{4}$. Hence

$$\sigma(\mathbf{B}) = \left(\frac{\sqrt{37}+1}{2}, \frac{\sqrt{37}-1}{2} \right) \approx (3.541, 2.541).$$

The differences are $\sigma(\mathbf{A}) - \sigma(\mathbf{B}) = \left(\frac{7-\sqrt{37}}{2}, \frac{5-\sqrt{37}}{2} \right) \approx (0.459, -0.541)$, so the left side in the trace norm is

$$\frac{7-\sqrt{37}}{2} + \frac{\sqrt{37}-5}{2} = 1.$$

On the right, $\mathbf{A} - \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ has singular values $(1, 0)$, so $\|\mathbf{A} - \mathbf{B}\| = 1$ in the trace norm. The inequality holds with **equality** here.

[Corollary 17.9.5](#) gives only $|\sigma_i(\mathbf{A}) - \sigma_i(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2 = 1$ for each i separately, hence at best 2 for the sum. [Mirsky's theorem](#) halves that, and here it is exact.

Exercise. B3: A bound with no norm in sight

Let $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$ be Hermitian with $\|\mathbf{A} - \mathbf{B}\|_2 \leq \varepsilon$. Prove that $\sum_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq n\varepsilon$, and give a pair for which this is an equality.

Solution. By [Theorem 21.7.1](#) with $\phi(t) = |t|$,

$$\sum_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq \sum_i |\lambda_i(\mathbf{A} - \mathbf{B})| \leq n \max_i |\lambda_i(\mathbf{A} - \mathbf{B})| = n\|\mathbf{A} - \mathbf{B}\|_2 \leq n\varepsilon,$$

the last equality by [Lemma 17.3.5](#). For equality take $\mathbf{B} = \mathbf{A} - \varepsilon \mathbf{I}$: then $\lambda_i(\mathbf{B}) = \lambda_i(\mathbf{A}) - \varepsilon$ for every i , the left side is $n\varepsilon$, and $\|\mathbf{A} - \mathbf{B}\|_2 = \varepsilon$.

21.7.5.3 C. Going deeper

Exercise. C1: The Schatten norms

Recall the Schatten r -norm $\|\mathbf{M}\|_{S_r} = (\sum_i \sigma_i(\mathbf{M})^r)^{1/r}$ of Example (The Schatten norms), the unitarily invariant norm with gauge $\|\cdot\|_r$, and let $1 \leq r < \infty$.

- (a) Prove that for Hermitian $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$, $\sum_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})|^r \leq \sum_i |\lambda_i(\mathbf{A} - \mathbf{B})|^r$, in two ways: from [Theorem 21.7.1](#) and from [Theorem 21.7.2](#).
- (b) Deduce the same statement for singular values, with σ in place of λ and $\mathbf{A}, \mathbf{B}$ arbitrary of the same size.

Solution.

- (a) From [Theorem 21.7.1](#). First, $\phi(t) = |t|^r$ is convex on $\mathbb{R}$. For $r = 1$ this is $|t| = \max\{t, -t\}$, convex by Proposition 18.10.3. Let $r > 1$, let $a, b \in \mathbb{R}$, and let $0 < \theta < 1$ (the cases $\theta = 0, 1$ are equalities); write $\theta' = 1 - \theta$ and let $q = r/(r-1)$ be the exponent conjugate to r . Apply Theorem 18.11.4 in $\mathbb{R}^2$ to

$$\mathbf{x} = (\theta^{1/r}|a|, (\theta')^{1/r}|b|), \quad \mathbf{y} = (\theta^{1/q}, (\theta')^{1/q}).$$

Since $\frac{1}{r} + \frac{1}{q} = 1$, the rules for powers of Chapter 18 §11 give $\langle \mathbf{x}, \mathbf{y} \rangle = \theta|a| + \theta'|b|$ and $\|\mathbf{y}\|_q = (\theta + \theta')^{1/q} = 1$, so

$$\theta|a| + \theta'|b| \leq \|\mathbf{x}\|_r = (\theta|a|^r + \theta'|b|^r)^{1/r}.$$

The function $u \mapsto u^r$ is increasing on $[0, \infty)$ (Chapter 18 §11), and $|\theta a + \theta' b| \leq \theta|a| + \theta'|b|$ by Theorem 0.6.4, so raising the displayed inequality to the power r gives

$$|\theta a + \theta' b|^r \leq (\theta|a| + \theta'|b|)^r \leq \theta|a|^r + \theta'|b|^r,$$

which is Definition 18.10.1 for ϕ on $\mathbb{R}$. Now apply [Theorem 21.7.1](#).

From [Theorem 21.7.2](#). Take the Schatten r -norm, whose gauge is $\|\cdot\|_r$. The left-hand side of [Theorem 21.7.2](#) is $(\sum_i |\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})|^r)^{1/r}$ by [Eq. 21.7.2](#), and the right-hand side is $(\sum_i |\lambda_i(\mathbf{A} - \mathbf{B})|^r)^{1/r}$, since $\mathbf{A} - \mathbf{B}$ is Hermitian. Raise to the power r .

- (b) Apply [Theorem 21.7.4](#) with the Schatten r -norm and raise to the power r , reading both sides through Theorem 21.4.3 as above.

Exercise. C2: What the dilation costs

Let $\mathbf{A}, \mathbf{B} \in M_{m \times n}(\mathbb{C})$ and let $\mathcal{H}(\cdot)$ be the Hermitian dilation of Proposition 17.9.3.

- (a) Prove that $\|\mathcal{H}(\mathbf{A})\|$ and $\|\mathbf{A}\|$ cannot be compared in general, by computing both for the Frobenius norm and for the spectral norm with $\mathbf{A} = \mathbf{I}_2$.
- (b) Explain, in one or two sentences, why the proof of [Theorem 21.7.4](#) is nevertheless free of any such factor.
- (c) Deduce Corollary 21.6.2 from [Theorem 21.7.4](#), and say why Section 6 nevertheless proved it another way.

Hint: for (c), a matrix of rank at most k has $\sigma_i = 0$ for $i > k$.

Solution.

- (a) For $\mathbf{A} = \mathbf{I}_2$, $\mathcal{H}(\mathbf{A}) \in M_4(\mathbb{C})$ has singular values $(1, 1, 1, 1)$, so $\|\mathcal{H}(\mathbf{A})\|_F = 2$ while $\|\mathbf{A}\|_F = \sqrt{2}$: the Frobenius norm grows by $\sqrt{2}$. The spectral norms, however, are both 1. So the ratio $\|\mathcal{H}(\mathbf{A})\| / \|\mathbf{A}\|$ depends on the norm, and there is no single factor to divide out. (The two norms also live on different spaces, M_4 and M_2 .)
- (b) Because the dilation is used only to produce a majorization between **eigenvalue lists**, from which the required inequality between running totals of the p numbers $|\delta_i|$ and the p numbers $\sigma_j(\mathbf{A} - \mathbf{B})$ is read off directly. No norm is ever applied to a dilated matrix.
- (c) Let $\text{rank } \mathbf{B} \leq k$, so $\sigma_i(\mathbf{B}) = 0$ for $i > k$ and hence $|\delta_i| = \sigma_i(\mathbf{A})$ for $i > k$, in the notation of the proof of [Theorem 21.7.4](#). Let $\mathbf{z} \in \mathbb{R}^p$ have $z_i = 0$ for $i \leq k$ and $z_i = \sigma_i(\mathbf{A})$ for $i > k$. Then $|z_i| \leq |\delta_i|$ for every i , so Proposition 21.3.1 gives $\Phi(\mathbf{z}) \leq \Phi(\boldsymbol{\delta})$; and $\Phi(\mathbf{z}) = \Phi(\sigma_{k+1}, \dots, \sigma_p, 0, \dots, 0)$ by the permutation invariance of Definition 21.3.1. Combining with [Theorem 21.7.4](#),

$$\Phi(\sigma_{k+1}, \dots, \sigma_p, 0, \dots, 0) \leq \Phi(\boldsymbol{\delta}) = \|\text{diag}_{m,n}(\boldsymbol{\delta})\| \leq \|\mathbf{A} - \mathbf{B}\|,$$

which is the bound of Corollary 21.6.2, attained at $\mathbf{A}_k$. Section 6 proved it earlier in the chapter, so it could not use this section; and its route gives more, namely $\sigma_i(\mathbf{A} - \mathbf{A}_k) \leq \sigma_i(\mathbf{A} - \mathbf{B})$ index by index, which is the route Chapter 13 §11 predicted.