

21.2 Convex Functions of a Majorization

Section 1 turned the n inequalities of a majorization into a single identity, $\mathbf{x} = \mathbf{D}\mathbf{y}$ with $\mathbf{D}$ doubly stochastic. This section spends it. Chapter 17 §07 left to this chapter “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 $\phi: \mathbb{R} \rightarrow \mathbb{R}, \dots$ ”; that is the first theorem below, and its converse says that the convex functions see nothing but the majorization. Two more promises are paid from it: Chapter 17 §08’s claim that for nonnegative vectors “ $\mathbf{x} \prec \mathbf{y}$ implies $x_1 \cdots x_n \geq y_1 \cdots y_n$ ”, and Chapters 18 §10 and §11’s claim that “ $\mathbf{A} \mapsto \operatorname{tr} f(\mathbf{A}) = \sum_i f(\lambda_i(\mathbf{A}))$ is convex on the Hermitian matrices with spectrum in that interval” for every convex f .

Conventions. Vectors are real and $\prec, \prec_w$ are Definition 17.8.1. A **convex** function is one in the sense of Definition 18.10.1, and an interval of $\mathbb{R}$ is a convex subset of $\mathbb{R}$, since it contains every number lying between two of its points. In the last subsection $F = \mathbb{R}$ or $\mathbb{C}$, matrices in $M_n(F)$ written $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are **Hermitian**, and $f(\mathbf{A})$ is the spectral functional calculus of Chapter 12 §08, Definition 12.8.2. There $\mathbf{P}_j$, with an integer subscript, is an orthogonal projection of a spectral resolution; $\mathbf{P}_\sigma$ always means a permutation matrix. No analysis is imported here: the only facts used from outside linear algebra reach us through Chapter 18’s already proved results.

21.2.1 Karamata’s inequality

A majorization is a statement about sums of the largest entries. A convex function is one whose graph lies below its chords. The link between them is that an averaging matrix turns one into the other, and §01 says that a majorization *is* an averaging matrix.

Theorem 21.2.1. Karamata’s Inequality

Let $I \subseteq \mathbb{R}$ be an interval, let $\phi: I \rightarrow \mathbb{R}$ be **convex**, and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ satisfy $\mathbf{x} \prec \mathbf{y}$, with every entry of $\mathbf{y}$ lying in I . Then every entry of $\mathbf{x}$ lies in I , and

$$\sum_{i=1}^n \phi(x_i) \leq \sum_{i=1}^n \phi(y_i).$$

Idea. Write $\mathbf{x} = \mathbf{D}\mathbf{y}$ with $\mathbf{D}$ doubly stochastic. Each x_i is then a weighted average of the y_j , with the weights in row i of $\mathbf{D}$; Jensen’s inequality bounds $\phi(x_i)$ by the same average of the $\phi(y_j)$. Add these n inequalities. The coefficient of $\phi(y_j)$ on the right is the sum of column j of $\mathbf{D}$, which is 1. The rows and the columns do two different jobs, and this is where “doubly” earns its keep.

Proof. By Theorem 21.1.3 there is a doubly stochastic $\mathbf{D} = (d_{ij})$ with $\mathbf{x} = \mathbf{D}\mathbf{y}$, that is,

$$x_i = \sum_{j=1}^n d_{ij} y_j \quad (i = 1, \dots, n).$$

Fix i . The numbers $d_{i1}, \dots, d_{in}$ are nonnegative and add up to 1, since row i of $\mathbf{D}$ is a row of a doubly stochastic matrix (Definition 18.8.4). So x_i is a convex combination of $y_1, \dots, y_n$, all of which lie in the convex set I ; by Lemma 18.1.2, $x_i \in I$. Now Theorem 18.10.2, applied

to the convex ϕ on I with the points $y_1, \dots, y_n$ and the weights $d_{i1}, \dots, d_{in}$, gives

$$\phi(x_i) = \phi\left(\sum_j d_{ij} y_j\right) \leq \sum_{j=1}^n d_{ij} \phi(y_j).$$

Summing over i and exchanging the order of the two finite sums,

$$\sum_{i=1}^n \phi(x_i) \leq \sum_{j=1}^n \left(\sum_{i=1}^n d_{ij}\right) \phi(y_j) = \sum_{j=1}^n \phi(y_j),$$

because every column of $\mathbf{D}$ adds up to 1. This proves the theorem. ■

The proof is three lines and uses Theorem 21.1.3 once. That is what §01 was for: without it we would be comparing sorted partial sums with the values of an arbitrary convex function, and there is no route from one to the other.

Example. One majorization, four inequalities

Let $\mathbf{y} = (5, 3, 1, 1)$ and $\mathbf{x} = (3, 3, 2, 2)$. Check that $\mathbf{x} \prec \mathbf{y}$, and write down what [Theorem 21.2.1](#) gives for $\phi(t) = t^2$, for $\phi(t) = |t - 2|$, and for $\phi(t) = \max\{t - 2, 0\}$.

Solution. Both vectors are decreasing with total 10, and the running totals are 3, 6, 8 for $\mathbf{x}$ against 5, 8, 9 for $\mathbf{y}$. So (M1) and (M2) hold, and $\mathbf{x} \prec \mathbf{y}$.

All three functions are convex on $I = \mathbb{R}$: t^2 by Lemma 18.10.5, since its derivative $2t$ is increasing, and the other two as pointwise maxima of affine functions, by Proposition 18.10.3 ($|t - 2| = \max\{t - 2, 2 - t\}$). So:

$$\begin{aligned} \sum_i x_i^2 &= 9 + 9 + 4 + 4 = 26 \leq 25 + 9 + 1 + 1 = 36 = \sum_i y_i^2, \\ \sum_i |x_i - 2| &= 1 + 1 + 0 + 0 = 2 \leq 3 + 1 + 1 + 1 = 6, \\ \sum_i \max\{x_i - 2, 0\} &= 1 + 1 = 2 \leq 3 + 1 = 4. \end{aligned}$$

Each is an instance of the one theorem. The last one measures how much of each vector sits above the level 2, and it is the test function the converse below is built from.

Quick check. Does [Theorem 21.2.1](#) apply to $\phi(t) = t^3$ on $I = \mathbb{R}$? Test it on $\mathbf{x} = (0, 0, 0) \prec \mathbf{y} = (1, 1, -2)$.

Solution. No. On $\mathbb{R}$ the function t^3 is not convex: at $a = -2$, $b = 0$, and $t = \frac{1}{2}$ the midpoint value is $\phi(-1) = -1$, while the chord gives $\frac{1}{2}\phi(-2) + \frac{1}{2}\phi(0) = -4$, and $-1 > -4$. The conclusion of the theorem does fail here. Both vectors have total 0, and the running totals 0, 0 of $\mathbf{x}$ are at most 1, 2, so $\mathbf{x} \prec \mathbf{y}$. But $\sum_i x_i^3 = 0$ while $\sum_i y_i^3 = 1 + 1 - 8 = -6$, and $0 > -6$. On $I = (0, \infty)$, where t^3 is convex, since its derivative $3t^2$ is increasing there (Lemma 18.10.5), the theorem does apply, and no such failure is possible for vectors with strictly positive entries.

21.2.2 The converse, and the weak version

Karamata's inequality is not merely a consequence of majorization; it is equivalent to it. Two families of test functions do the work: the affine ones pin down the totals, and the “hockey sticks” $t \mapsto \max\{t - c, 0\}$ pin down the partial sums.

Proposition 21.2.2. Convex Functions Detect Majorization

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. If

$$\sum_{i=1}^n \phi(x_i) \leq \sum_{i=1}^n \phi(y_i)$$

for **every** convex $\phi: \mathbb{R} \rightarrow \mathbb{R}$, then $\mathbf{x} \prec \mathbf{y}$.

Idea. Use $\phi(t) = t$ and $\phi(t) = -t$ to get the equal totals. For the k -th partial sum, use the function that is 0 below a level c and rises with slope 1 above it. Summed over a vector it always exceeds “top k sum minus kc ”, and it *equals* that when c is the k -th largest entry. Setting $c = y_k^\downarrow$ makes the right-hand side exact and the left-hand side an over-estimate, which is the direction we need.

Proof. Both $\phi(t) = t$ and $\phi(t) = -t$ are affine, hence convex by Example (First convex functions) (a). They give $\sum_i x_i \leq \sum_i y_i$ and $-\sum_i x_i \leq -\sum_i y_i$, so the totals are equal. This is (M2).

Fix k with $1 \leq k \leq n - 1$, put $c = y_k^\downarrow$, and let

$$\phi_c(t) = \max\{t - c, 0\},$$

the pointwise maximum of the two affine functions $t \mapsto t - c$ and $t \mapsto 0$, hence convex by Proposition 18.10.3. For any $\mathbf{z} \in \mathbb{R}^n$ the sum $\sum_i \phi_c(z_i)$ does not depend on the order of the entries, so it equals $\sum_i \phi_c(z_i^\downarrow)$. Every term is ≥ 0 , and each of the first k terms satisfies $\phi_c(z_i^\downarrow) \geq z_i^\downarrow - c$. Hence

$$\sum_{i=1}^n \phi_c(z_i) \geq \sum_{i=1}^k z_i^\downarrow - kc. \quad (*)$$

For $\mathbf{z} = \mathbf{y}$ and this particular c , $(*)$ is an equality: for $i \leq k$ we have $y_i^\downarrow \geq y_k^\downarrow = c$, so $\phi_c(y_i^\downarrow) = y_i^\downarrow - c$; and for $i > k$ we have $y_i^\downarrow \leq c$, so $\phi_c(y_i^\downarrow) = 0$. Therefore

$$\sum_{i=1}^k x_i^\downarrow - kc \stackrel{(*)}{\leq} \sum_{i=1}^n \phi_c(x_i) \leq \sum_{i=1}^n \phi_c(y_i) = \sum_{i=1}^k y_i^\downarrow - kc,$$

the middle step being the hypothesis applied to ϕ_c . Canceling kc gives (M1) at k . As k was arbitrary, $\mathbf{x} \prec \mathbf{y}$. ■

Together with Theorem 21.2.1 this is an equivalence: $\mathbf{x} \prec \mathbf{y}$ if and only if $\sum_i \phi(x_i) \leq \sum_i \phi(y_i)$ for every convex $\phi: \mathbb{R} \rightarrow \mathbb{R}$. So majorization is exactly the order that all convex functions agree on. The proof also shows how few test functions are needed: the two affine ones and the $n - 1$

hockey sticks at the levels $y_1^\downarrow, \dots, y_{n-1}^\downarrow$.

Weak majorization has a matching statement, with the class of test functions cut down.

Proposition 21.2.3. Karamata for Weak Majorization

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\mathbf{x} \prec_w \mathbf{y}$, and let $\phi: \mathbb{R} \rightarrow \mathbb{R}$ be convex and **increasing**. Then

$$\sum_{i=1}^n \phi(x_i) \leq \sum_{i=1}^n \phi(y_i).$$

Idea. The only thing wrong with $\mathbf{y}$ is that its total may be too big. Take the excess away from the **smallest** entry of $\mathbf{y}$. That keeps the vector sorted, leaves all the partial sums before the last one untouched, and produces a genuine majorization, to which [Theorem 21.2.1](#) applies. Lowering an entry can only lower ϕ , which is where “increasing” is used.

Proof. Let $s = \sum_i y_i - \sum_i x_i$. Taking $k = n$ in (M1) for $\mathbf{x} \prec_w \mathbf{y}$ gives $s \geq 0$. Define $\mathbf{w} \in \mathbb{R}^n$ by

$$w_i = y_i^\downarrow \quad (i \leq n-1), \quad w_n = y_n^\downarrow - s.$$

Only the last entry has changed, and it has been lowered, so $\mathbf{w}$ is still decreasing and $\mathbf{w}^\downarrow = \mathbf{w}$. Its total is $\sum_i y_i - s = \sum_i x_i$, which is (M2) for $\mathbf{x} \prec \mathbf{w}$. For $k \leq n-1$, $\sum_{i \leq k} w_i = \sum_{i \leq k} y_i^\downarrow \geq \sum_{i \leq k} x_i^\downarrow$ by (M1) for $\mathbf{x} \prec_w \mathbf{y}$. So $\mathbf{x} \prec \mathbf{w}$.

By [Theorem 21.2.1](#) with $I = \mathbb{R}$, $\sum_i \phi(x_i) \leq \sum_i \phi(w_i)$. Since $w_i \leq y_i^\downarrow$ for every i and ϕ is increasing, $\phi(w_i) \leq \phi(y_i^\downarrow)$, so

$$\sum_i \phi(w_i) \leq \sum_i \phi(y_i^\downarrow) = \sum_i \phi(y_i).$$

Chaining the two inequalities proves the proposition. ■

Warning. “**Increasing**” cannot be dropped from [Proposition 21.2.3](#). Take $\mathbf{x} = (0, 0)$ and $\mathbf{y} = (1, 0)$. The running totals 0, 0 of $\mathbf{x}$ are at most 1, 1, so $\mathbf{x} \prec_w \mathbf{y}$. The function $\phi(t) = -t$ is convex but decreasing, and $\sum_i \phi(x_i) = 0$ while $\sum_i \phi(y_i) = -1$. A weak majorization allows $\mathbf{y}$ to carry more total than $\mathbf{x}$, and a decreasing ϕ punishes exactly that.

21.2.3 Schur-convex functions

[Theorem 21.2.1](#) is a statement about one particular shape of function, $\mathbf{x} \mapsto \sum_i \phi(x_i)$. The property it establishes deserves a name of its own, because many functions have it that are not of that shape.

A function is Schur-convex when spreading a vector out can only increase it.

The domain needs a moment’s thought. If G is to be compared at $\mathbf{x}$ and $\mathbf{y}$ whenever $\mathbf{x} \prec \mathbf{y}$,

then the domain must contain $\mathbf{x}$ as soon as it contains $\mathbf{y}$. Call a set $A \subseteq \mathbb{R}^n$ **symmetric** if $\mathbf{P}_\sigma \mathbf{x} \in A$ for every $\mathbf{x} \in A$ and every $\sigma \in S_n$. A symmetric **convex** set has the property we need: if $\mathbf{y} \in A$ and $\mathbf{x} \prec \mathbf{y}$, then $\mathbf{x} \in \text{conv}\{\mathbf{P}_\sigma \mathbf{y}\}$ by Corollary 21.1.4, every $\mathbf{P}_\sigma \mathbf{y}$ lies in A by symmetry, and a convex set contains the convex combinations of its points (Lemma 18.1.2), so $\mathbf{x} \in A$.

Definition 21.2.1. Schur-Convex Function

Let $A \subseteq \mathbb{R}^n$ be **symmetric and convex**, and let $G: A \rightarrow \mathbb{R}$. Then G is **Schur-convex** if

$$\mathbf{x} \prec \mathbf{y} \implies G(\mathbf{x}) \leq G(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$. It is **strictly Schur-convex** if the inequality is **strict** whenever $\mathbf{x}^\downarrow \neq \mathbf{y}^\downarrow$.

Definition 21.2.2. Schur-Concave Function

With A as in Definition 21.2.1, a function $G: A \rightarrow \mathbb{R}$ is **Schur-concave** if $-G$ is Schur-convex, that is, if

$$\mathbf{x} \prec \mathbf{y} \implies G(\mathbf{x}) \geq G(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in A$; and **strictly Schur-concave** if $-G$ is strictly Schur-convex.

In words: a Schur-convex function is one that respects the majorization order, and a Schur-concave one reverses it. The two standing examples of A below are $\mathbb{R}^n$ itself and the open orthant $\{\mathbf{x} : x_i > 0 \text{ for every } i\}$, both symmetric and convex.

A Schur-convex function is automatically **symmetric**, meaning $G(\mathbf{P}_\sigma \mathbf{x}) = G(\mathbf{x})$ for every permutation. Indeed $\mathbf{x}$ and $\mathbf{P}_\sigma \mathbf{x}$ are rearrangements of each other, so each majorizes the other (Proposition 17.8.1 (a) and the remark after Definition 17.8.1), and two applications of the definition give $G(\mathbf{x}) \leq G(\mathbf{P}_\sigma \mathbf{x}) \leq G(\mathbf{x})$.

Example. Four functions on $\mathbb{R}^n$

Decide which of the following are Schur-convex, and which are Schur-concave, on $A = \mathbb{R}^n$.

- (a) $G(\mathbf{x}) = \sum_i x_i^2$.
- (b) $G(\mathbf{x}) = \max_i x_i$.
- (c) $G(\mathbf{x}) = \sum_i x_i$.
- (d) $G(\mathbf{x}) = \sum_i x_i^3$, for $n = 3$.

Solution.

- (a) Schur-convex. $\phi(t) = t^2$ is convex on $\mathbb{R}$, as in Example (One majorization, four inequalities), so Theorem 21.2.1 gives $\sum_i x_i^2 \leq \sum_i y_i^2$ whenever $\mathbf{x} \prec \mathbf{y}$.
- (b) Schur-convex. $\max_i x_i = x_1^\downarrow$, and (M1) at $k = 1$ says exactly $x_1^\downarrow \leq y_1^\downarrow$. No convexity is needed here; the definition of majorization already contains the statement.
- (c) Both Schur-convex and Schur-concave, since (M2) makes it constant on each majorization class. This is the degenerate case, and it shows that the two notions are not exclusive.
- (d) Neither, and one pair of comparisons settles both. The flat vector is majorized by every vector with the same total (Example (A chain, the two extremes, and size one) (b)),

so

$$(0, 0, 0) \prec (2, -1, -1), \quad (0, 0, 0) \prec (1, 1, -2),$$

both totals being 0. Now $G(0, 0, 0) = 0$, while $G(2, -1, -1) = 8 - 1 - 1 = 6$ and $G(1, 1, -2) = 1 + 1 - 8 = -6$. Schur-convexity applied to the second majorization would demand $0 \leq -6$, and Schur-concavity applied to the first would demand $0 \geq 6$. Both fail.

Part (d) is the nonexample by minimal change. The function $\sum_i \phi(x_i)$ with $\phi(t) = t^3$ differs from the Schur-convex $\sum_i \phi(x_i)$ with $\phi(t) = t^2$ only in the exponent, and t^3 fails the one clause that matters, convexity on the interval where the entries live.

The next theorem explains why so many natural functions turn out Schur-convex: convexity plus symmetry is enough, and there is no need to be of the form $\sum_i \phi(x_i)$.

Theorem 21.2.4. A Symmetric Convex Function Is Schur-Convex

Let $A \subseteq \mathbb{R}^n$ be symmetric and convex, and let $G: A \rightarrow \mathbb{R}$ be **convex** and **symmetric**, that is, $G(\mathbf{P}_\sigma \mathbf{x}) = G(\mathbf{x})$ for every $\mathbf{x} \in A$ and every $\sigma \in S_n$. Then G is Schur-convex.

Idea. $\mathbf{x} \prec \mathbf{y}$ puts $\mathbf{x}$ inside the permutation polytope of $\mathbf{y}$, whose corners all carry the **same** value of G , by symmetry. A convex function on a convex combination is at most the corresponding combination of its values, and an average of copies of one number is that number.

Proof. Let $\mathbf{x}, \mathbf{y} \in A$ with $\mathbf{x} \prec \mathbf{y}$. By Corollary 21.1.4 and Theorem 18.1.3,

$$\mathbf{x} = \sum_{i=1}^m t_i \mathbf{P}_{\sigma_i} \mathbf{y}$$

for some $\sigma_1, \dots, \sigma_m \in S_n$ and weights $t_i \geq 0$ with $\sum_i t_i = 1$. Each $\mathbf{P}_{\sigma_i} \mathbf{y}$ lies in A , since A is symmetric. By Theorem 18.10.2 applied to the convex G on A ,

$$G(\mathbf{x}) \leq \sum_{i=1}^m t_i G(\mathbf{P}_{\sigma_i} \mathbf{y}) = \sum_{i=1}^m t_i G(\mathbf{y}) = G(\mathbf{y}),$$

the middle equality being the symmetry of G and the last one $\sum_i t_i = 1$. This shows that G is Schur-convex. ■

Now the promise from Chapter 17 §08. There the second proof of Hadamard's inequality used only the trace, and a remark recorded that a proof "by majorization in earnest" would need the fact that "for vectors with nonnegative entries, $\mathbf{x} \prec \mathbf{y}$ implies $x_1 \cdots x_n \geq y_1 \cdots y_n$ (the product is **Schur-concave**)". Here it is.

Corollary 21.2.5. The Product Is Schur-Concave

Let $\mathbf{y} \in \mathbb{R}^n$ have **nonnegative** entries and let $\mathbf{x} \prec \mathbf{y}$. Then every entry of $\mathbf{x}$ is nonnegative,

and

$$x_1 x_2 \cdots x_n \geq y_1 y_2 \cdots y_n.$$

Idea. Take $-\log$, which is convex, and apply [Theorem 21.2.1](#): the product becomes a sum and the inequality flips. The logarithm needs positive entries, so the case of a zero entry is dealt with first — and it is the easy case, because a zero entry makes the right-hand side 0 while the left-hand side is a product of nonnegative numbers.

Proof. First, every entry of $\mathbf{x}$ is nonnegative. For $n = 1$, (M2) gives $\mathbf{x} = \mathbf{y}$. For $n \geq 2$, subtracting (M1) at $k = n - 1$ from (M2) gives

$$x_n^\downarrow \geq y_n^\downarrow \geq 0,$$

and $x_n^\downarrow$ is the smallest entry of $\mathbf{x}$.

Case 1: some $y_i = 0$. Then $y_1 \cdots y_n = 0$, while $x_1 \cdots x_n \geq 0$ as a product of nonnegative numbers. The inequality holds.

Case 2: every $y_i > 0$. Then $y_n^\downarrow > 0$, so $x_n^\downarrow > 0$ by the display above, and every entry of both vectors lies in $I = (0, \infty)$. By Lemma 18.11.1 (c) the function $\log$ is concave on I , so $\phi = -\log$ is convex there. By [Theorem 21.2.1](#),

$$\sum_{i=1}^n (-\log x_i) \leq \sum_{i=1}^n (-\log y_i), \quad \text{that is,} \quad \sum_{i=1}^n \log y_i \leq \sum_{i=1}^n \log x_i.$$

By $\log(ab) = \log a + \log b$ (Lemma 18.11.1 (c)) and induction on n , $\sum_i \log x_i = \log(x_1 \cdots x_n)$ and likewise for $\mathbf{y}$. So $\log(y_1 \cdots y_n) \leq \log(x_1 \cdots x_n)$. Since $\log$ is strictly increasing on I (Lemma 18.11.1 (c)), two positive numbers with $\log u \leq \log v$ satisfy $u \leq v$; otherwise $u > v$ would give $\log u > \log v$. Hence $y_1 \cdots y_n \leq x_1 \cdots x_n$. This proves the corollary. ■

So the flat vector, the bottom of the majorization order, has the largest product among non-negative vectors with a fixed total — which is Lemma 17.8.6 again, now as a special case rather than a separate induction.

Warning. Schur-convex does not mean convex. By [Corollary 21.2.5](#) the function $G(\mathbf{x}) = -x_1 x_2$ is Schur-convex on the symmetric convex set $A = \{\mathbf{x} \in \mathbb{R}^2 : x_1, x_2 \geq 0\}$. It is not convex: at $\mathbf{x} = (1, 1)$, $\mathbf{y} = (2, 2)$, and $t = \frac{1}{2}$,

$$G\left(\frac{3}{2}, \frac{3}{2}\right) = -\frac{9}{4} = -2.25, \quad \frac{1}{2}G(1, 1) + \frac{1}{2}G(2, 2) = \frac{1}{2}(-1) + \frac{1}{2}(-4) = -2.5,$$

and $-2.25 > -2.5$. [Theorem 21.2.4](#) is a one-way street.

Quick check. Is the function $G(\mathbf{x}) = x_1^\downarrow + x_2^\downarrow$ on $\mathbb{R}^3$ Schur-convex? Is it convex?

Solution. Schur-convex: the inequality $x_1^\downarrow + x_2^\downarrow \leq y_1^\downarrow + y_2^\downarrow$ is (M1) at $k = 2$. Convex: yes, and this gives a second proof. The two largest entries have the largest sum among all pairs, so $G(\mathbf{x}) = \max_{i < j} (x_i + x_j)$, a maximum of three linear functions, convex by Proposition 18.10.3. It is also symmetric, so Theorem 21.2.4 applies. (This G is the Ky Fan sum s_2 of Chapter 17 §06, read on vectors instead of eigenvalues.)

21.2.4 Convex functions of a Hermitian matrix

The last payoff is about matrices. For a Hermitian $\mathbf{A}$ and a convex f on an interval containing $\text{spec}(\mathbf{A})$, the number $\text{tr } f(\mathbf{A})$ turns out to be $\sum_i f(\lambda_i(\mathbf{A}))$. Chapter 18 §10 stated, and deferred, that this is a convex function of $\mathbf{A}$. The proof needs two lemmas: a bookkeeping one about traces, and one saying that the quadratic form of $f(\mathbf{A})$ dominates f of the quadratic form of $\mathbf{A}$.

Lemma 21.2.6. Reading a Trace on an Orthonormal Basis

Let $F = \mathbb{R}$ or $\mathbb{C}$.

- (a) Let $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ be an orthonormal basis of F^n and $\mathbf{M} \in M_n(F)$. Then $\text{tr } \mathbf{M} = \sum_{i=1}^n \langle \mathbf{M}\mathbf{u}_i, \mathbf{u}_i \rangle$.
- (b) Let $I \subseteq \mathbb{R}$, let $\mathbf{A} \in M_n(F)$ be Hermitian with $\text{spec}(\mathbf{A}) \subseteq I$, and let $g: I \rightarrow \mathbb{R}$. Then $g(\mathbf{A})\mathbf{u} = g(\lambda)\mathbf{u}$ for every eigenvector $\mathbf{u}$ of $\mathbf{A}$ with eigenvalue λ , and $\text{tr } g(\mathbf{A}) = \sum_{i=1}^n g(\lambda_i(\mathbf{A}))$.

Proof.

- (a) Let $\mathbf{U} \in M_n(F)$ be the matrix whose i -th column is $\mathbf{u}_i$. Then $(\mathbf{U}^*\mathbf{U})_{ij} = \mathbf{u}_i^*\mathbf{u}_j = \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \delta_{ij}$, so $\mathbf{U}$ is unitary, and $(\mathbf{U}^*\mathbf{M}\mathbf{U})_{ii} = \mathbf{u}_i^*\mathbf{M}\mathbf{u}_i = \langle \mathbf{M}\mathbf{u}_i, \mathbf{u}_i \rangle$. By Theorem 0.9.6 (3), used with the factors $\mathbf{U}^*$ and $\mathbf{M}\mathbf{U}$, $\text{tr}(\mathbf{U}^*\mathbf{M}\mathbf{U}) = \text{tr}(\mathbf{M}\mathbf{U}\mathbf{U}^*) = \text{tr } \mathbf{M}$. Summing the diagonal of $\mathbf{U}^*\mathbf{M}\mathbf{U}$ gives (a).
- (b) Let $\mathbf{A} = \sum_j \mu_j \mathbf{P}_j$ be the spectral resolution of Definition 12.8.1, with $\mu_1, \dots, \mu_k$ the distinct eigenvalues; it exists by Theorem 12.8.1, since a Hermitian matrix has an orthonormal basis of eigenvectors (Corollary 12.4.3 over $\mathbb{C}$, Corollary 12.5.2 over $\mathbb{R}$). Let $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ with $\mathbf{u} \neq \mathbf{0}$. Then $\lambda = \mu_{j_0}$ for exactly one j_0 , and $\mathbf{u}$ lies in the eigenspace of μ_{j_0} , which is the image of $\mathbf{P}_{j_0}$; so $\mathbf{P}_{j_0}\mathbf{u} = \mathbf{u}$, and $\mathbf{P}_j\mathbf{u} = \mathbf{P}_j\mathbf{P}_{j_0}\mathbf{u} = \mathbf{0}$ for $j \neq j_0$ by Theorem 12.8.1 (b). Hence $g(\mathbf{A})\mathbf{u} = \sum_j g(\mu_j)\mathbf{P}_j\mathbf{u} = g(\lambda)\mathbf{u}$, by Definition 12.8.2.

Now take an orthonormal basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ of F^n with $\mathbf{A}\mathbf{u}_i = \lambda_i(\mathbf{A})\mathbf{u}_i$, which exists by the spectral theorem just cited. By the previous paragraph, $\langle g(\mathbf{A})\mathbf{u}_i, \mathbf{u}_i \rangle = g(\lambda_i(\mathbf{A}))\langle \mathbf{u}_i, \mathbf{u}_i \rangle = g(\lambda_i(\mathbf{A}))$, and (a) adds these up to $\text{tr } g(\mathbf{A})$. ■

Lemma 21.2.7. Peierls' Inequality

Let $I \subseteq \mathbb{R}$ be an interval, let $\mathbf{A} \in M_n(F)$ be Hermitian with $\text{spec}(\mathbf{A}) \subseteq I$, let $f: I \rightarrow \mathbb{R}$ be convex, and let $\mathbf{u} \in F^n$ with $\|\mathbf{u}\| = 1$. Then $\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle \in I$ and

$$f(\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle) \leq \langle f(\mathbf{A})\mathbf{u}, \mathbf{u} \rangle,$$

with **equality** when $\mathbf{u}$ is an eigenvector of $\mathbf{A}$.

Idea. Write $\mathbf{A} = \sum_j \mu_j \mathbf{P}_j$, its spectral resolution. Then $\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle = \sum_j \mu_j \|\mathbf{P}_j \mathbf{u}\|^2$ is a weighted average of the eigenvalues, with weights $\|\mathbf{P}_j \mathbf{u}\|^2$ that add up to $\|\mathbf{u}\|^2 = 1$. The same weights appear in $\langle f(\mathbf{A})\mathbf{u}, \mathbf{u} \rangle = \sum_j f(\mu_j) \|\mathbf{P}_j \mathbf{u}\|^2$. So the two sides of the inequality are f of an average against the average of f , which is Jensen.

Proof. Let $\mathbf{A} = \mu_1 \mathbf{P}_1 + \cdots + \mu_k \mathbf{P}_k$ be the spectral resolution of Definition 12.8.1, with $\mu_1, \dots, \mu_k$ the **distinct** eigenvalues of $\mathbf{A}$, which exists by Theorem 12.8.1 since a Hermitian matrix has an orthonormal basis of eigenvectors (Corollary 12.4.3 over $\mathbb{C}$, Corollary 12.5.2 over $\mathbb{R}$). Put

$$w_j = \langle \mathbf{P}_j \mathbf{u}, \mathbf{u} \rangle.$$

By Theorem 12.8.1 (b), $\mathbf{P}_j = \mathbf{P}_j^2$ and $\mathbf{P}_j^* = \mathbf{P}_j$, so

$$w_j = \langle \mathbf{P}_j \mathbf{P}_j \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{P}_j \mathbf{u}, \mathbf{P}_j^* \mathbf{u} \rangle = \langle \mathbf{P}_j \mathbf{u}, \mathbf{P}_j \mathbf{u} \rangle = \|\mathbf{P}_j \mathbf{u}\|^2 \geq 0,$$

and by Theorem 12.8.1 (c), $\sum_j w_j = \langle \mathbf{u}, \mathbf{u} \rangle = 1$. Hence, by Theorem 12.8.1 (d) and Definition 12.8.2,

$$\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle = \sum_{j=1}^k \mu_j w_j, \quad \langle f(\mathbf{A})\mathbf{u}, \mathbf{u} \rangle = \sum_{j=1}^k f(\mu_j) w_j.$$

The first is a convex combination of $\mu_1, \dots, \mu_k$, all of which lie in the convex set I , so $\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle \in I$ by Lemma 18.1.2. Applying Theorem 18.10.2 to the convex f on I , with points μ_j and weights w_j ,

$$f(\langle \mathbf{A}\mathbf{u}, \mathbf{u} \rangle) = f\left(\sum_j w_j \mu_j\right) \leq \sum_j w_j f(\mu_j) = \langle f(\mathbf{A})\mathbf{u}, \mathbf{u} \rangle.$$

For the equality clause, let $\mathbf{A}\mathbf{u} = \lambda \mathbf{u}$ with $\|\mathbf{u}\| = 1$. By Lemma 21.2.6 (b), $f(\mathbf{A})\mathbf{u} = f(\lambda)\mathbf{u}$, so the right side is $f(\lambda)\langle \mathbf{u}, \mathbf{u} \rangle = f(\lambda)$; and the left side is $f(\langle \lambda \mathbf{u}, \mathbf{u} \rangle) = f(\lambda)$ as well. This proves the lemma. ■

Quick check. Take $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $f(t) = t^2$, and $\mathbf{u} = \mathbf{e}_1$. Compute both sides of Lemma 21.2.7, and then repeat with $\mathbf{u} = \frac{1}{\sqrt{2}}(1, 1)$.

Solution. $\langle \mathbf{A}\mathbf{e}_1, \mathbf{e}_1 \rangle = a_{11} = 2$, so the left side is $f(2) = 4$. Here $f(\mathbf{A}) = \mathbf{A}^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$, so the right side is $(\mathbf{A}^2)_{11} = 5$. Indeed $4 \leq 5$, and the inequality is strict because $\mathbf{e}_1$ is not an eigenvector of $\mathbf{A}$.

For $\mathbf{u} = \frac{1}{\sqrt{2}}(1, 1)$, an eigenvector with $\mathbf{A}\mathbf{u} = 3\mathbf{u}$, the left side is $f(3) = 9$ and the right side is $\langle \mathbf{A}^2 \mathbf{u}, \mathbf{u} \rangle = 9\langle \mathbf{u}, \mathbf{u} \rangle = 9$. Equality, as the lemma says.

Theorem 21.2.8. The Trace of a Convex Function Is Convex

Let $I \subseteq \mathbb{R}$ be an interval and let $f: I \rightarrow \mathbb{R}$ be **convex**. Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian with $\text{spec}(\mathbf{A}) \subseteq I$ and $\text{spec}(\mathbf{B}) \subseteq I$, and let $0 \leq \theta \leq 1$. Then $\mathbf{C} = \theta\mathbf{A} + (1 - \theta)\mathbf{B}$ is Hermitian with $\text{spec}(\mathbf{C}) \subseteq I$, and

$$\text{tr } f(\mathbf{C}) \leq \theta \text{tr } f(\mathbf{A}) + (1 - \theta) \text{tr } f(\mathbf{B}).$$

In words: the Hermitian matrices with spectrum in I form a **convex** set, and $\mathbf{A} \mapsto \text{tr } f(\mathbf{A})$ is a **convex** function on it.

Idea. Fix an orthonormal eigenbasis of the midpoint matrix $\mathbf{C} = \theta\mathbf{A} + (1 - \theta)\mathbf{B}$. Along each of those n directions the value of f at the eigenvalue of $\mathbf{C}$ is exactly the left side of Peierls' inequality, because an eigenvector gives equality; and the eigenvalue itself splits as θ times a quadratic form of $\mathbf{A}$ plus $(1 - \theta)$ times one of $\mathbf{B}$. Ordinary convexity of f handles that split, and Peierls' inequality, now as an inequality, converts the two quadratic forms into quadratic forms of $f(\mathbf{A})$ and $f(\mathbf{B})$. Summing over the basis turns quadratic forms into traces.

Proof. Step 1: the spectrum of $\mathbf{C}$. The matrix $\mathbf{C} = \theta\mathbf{A} + (1 - \theta)\mathbf{B}$ is Hermitian, being a real combination of Hermitian matrices. Let μ be an eigenvalue of $\mathbf{C}$, with unit eigenvector $\mathbf{x}$. Then

$$\mu = \langle \mathbf{C}\mathbf{x}, \mathbf{x} \rangle = \theta \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle + (1 - \theta) \langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle.$$

By Lemma 13.5.2, $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$ lies in $[\lambda_n(\mathbf{A}), \lambda_1(\mathbf{A})]$, an interval whose endpoints are eigenvalues of $\mathbf{A}$, hence contained in I ; likewise $\langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle \in I$. So μ is a convex combination of two points of I and lies in I . Hence $\text{spec}(\mathbf{C}) \subseteq I$, and $f(\mathbf{C})$ is defined.

Step 2: the inequality. Choose an orthonormal basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ of F^n consisting of eigenvectors of $\mathbf{C}$, which exists by Corollary 12.4.3 over $\mathbb{C}$ and Corollary 12.5.2 over $\mathbb{R}$. Fix i . By Lemma 21.2.7 applied to $\mathbf{C}$ and the eigenvector $\mathbf{u}_i$, with equality,

$$\langle f(\mathbf{C})\mathbf{u}_i, \mathbf{u}_i \rangle = f(\langle \mathbf{C}\mathbf{u}_i, \mathbf{u}_i \rangle).$$

By Step 1, $\langle \mathbf{C}\mathbf{u}_i, \mathbf{u}_i \rangle = \theta\alpha_i + (1 - \theta)\beta_i$ with $\alpha_i = \langle \mathbf{A}\mathbf{u}_i, \mathbf{u}_i \rangle \in I$ and $\beta_i = \langle \mathbf{B}\mathbf{u}_i, \mathbf{u}_i \rangle \in I$. So Definition 18.10.1 and then Lemma 21.2.7, now applied to $\mathbf{A}$ and to $\mathbf{B}$ with the unit vector $\mathbf{u}_i$, give

$$\begin{aligned} \langle f(\mathbf{C})\mathbf{u}_i, \mathbf{u}_i \rangle &= f(\theta\alpha_i + (1 - \theta)\beta_i) \leq \theta f(\alpha_i) + (1 - \theta)f(\beta_i) \\ &\leq \theta \langle f(\mathbf{A})\mathbf{u}_i, \mathbf{u}_i \rangle + (1 - \theta) \langle f(\mathbf{B})\mathbf{u}_i, \mathbf{u}_i \rangle, \end{aligned}$$

where the last step uses $\theta \geq 0$ and $1 - \theta \geq 0$. Summing over $i = 1, \dots, n$ and using Lemma 21.2.6 (a) three times, once for each of $f(\mathbf{C})$, $f(\mathbf{A})$, and $f(\mathbf{B})$,

$$\text{tr } f(\mathbf{C}) \leq \theta \text{tr } f(\mathbf{A}) + (1 - \theta) \text{tr } f(\mathbf{B}).$$

Since $\mathbf{A}$, $\mathbf{B}$, and θ were arbitrary, this is Definition 18.10.1 for $\mathbf{A} \mapsto \text{tr } f(\mathbf{A})$ on the set of Hermitian matrices with spectrum in I , which Step 1 shows is convex. This proves the theorem. ■

This pays Chapter 18 §10's "it is proved in Chapter 21, and nothing in this chapter uses it", and Chapter 18 §11's closing note that "the convexity of $\mathbf{A} \mapsto \operatorname{tr} f(\mathbf{A})$ for convex f , and the consequences of a majorization for every convex function, belong to Chapter 21".

Example. The logarithm of the determinant

Let $f(t) = -\log t$ on $I = (0, \infty)$, so that the matrices in play are the positive definite Hermitian ones. Identify $\operatorname{tr} f(\mathbf{A})$, and check the resulting inequality for

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

Solution. A Hermitian $\mathbf{A}$ has $\operatorname{spec}(\mathbf{A}) \subseteq (0, \infty)$ exactly when $\mathbf{A} \succ \mathbf{O}$, by Theorem 13.1.5 (b). For such an $\mathbf{A}$, Lemma 21.2.6 (b), then $\log(ab) = \log a + \log b$ (Lemma 18.11.1 (c)) with induction on n , and then Theorem 9.3.3 give

$$\operatorname{tr} f(\mathbf{A}) = -\sum_i \log \lambda_i(\mathbf{A}) = -\log \prod_i \lambda_i(\mathbf{A}) = -\log \det \mathbf{A}.$$

By Theorem 21.2.8, $\mathbf{A} \mapsto -\log \det \mathbf{A}$ is **convex on the positive definite matrices**; equivalently, $\det$ is log-concave there.

For the data given, $\det \mathbf{A} = 4$ and $\det \mathbf{B} = 4 - 1 = 3$, and

$$\mathbf{C} = \frac{1}{2}(\mathbf{A} + \mathbf{B}) = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & 3 \end{pmatrix}, \quad \det \mathbf{C} = \frac{9}{2} - \frac{1}{4} = \frac{17}{4}.$$

Both $\mathbf{A}$ and $\mathbf{B}$ are positive definite ($\mathbf{A}$ is diagonal with positive entries; $\mathbf{B}$ has eigenvalues 3 and 1), so the theorem says $-\log \frac{17}{4} \leq -\frac{1}{2} \log 4 - \frac{1}{2} \log 3$, that is, after multiplying by -1 and exponentiating,

$$\det \mathbf{C} = \frac{17}{4} = 4.25 \geq \sqrt{4 \cdot 3} = \sqrt{12} \approx 3.464,$$

which holds, since $4.25^2 = 18.0625 \geq 12$.

Remark. Theorem 21.2.8 holds for **every** convex f , which is a very large class. It is nonetheless much weaker than the operator convexity of Definition 17.11.1 (b), which asks for $f(\theta \mathbf{A} + (1 - \theta) \mathbf{B}) \preceq \theta f(\mathbf{A}) + (1 - \theta) f(\mathbf{B})$ in the Loewner order, a statement about matrices rather than about one number extracted from them. Taking traces turns the Loewner inequality into the one proved here, since the trace of a positive semidefinite matrix is the sum of its nonnegative eigenvalues (Theorem 13.1.4 (b), Theorem 9.3.3); so operator convexity implies trace convexity, but the converse fails badly, and a later section of this chapter exhibits convex functions that are not operator convex.

21.2.5 Exercises

21.2.5.1 A. Check your understanding

Exercise. A1

- State Karamata's inequality, with all its hypotheses.
- Which test functions prove [Proposition 21.2.2](#), and which clause of Definition 17.8.1 does each one give?
- Define Schur-convex and Schur-concave.
- True or false: every Schur-convex function on $\mathbb{R}^n$ is convex. Justify your answer.
- True or false: every convex function on $\mathbb{R}^n$ is Schur-convex. Justify your answer.
- State what [Theorem 21.2.8](#) says for $f(t) = t^2$, and identify the resulting function of $\mathbf{A}$.

Solution.

- If I is an interval, $\phi: I \rightarrow \mathbb{R}$ is convex, $\mathbf{x} \prec \mathbf{y}$ in $\mathbb{R}^n$, and every entry of $\mathbf{y}$ lies in I , then every entry of $\mathbf{x}$ lies in I and $\sum_i \phi(x_i) \leq \sum_i \phi(y_i)$ ([Theorem 21.2.1](#)).
- The affine functions $\phi(t) = t$ and $\phi(t) = -t$ give the equal totals, (M2). The hockey stick $\phi_c(t) = \max\{t - c, 0\}$ with $c = y_k^\downarrow$ gives (M1) at k .
- For a symmetric convex $A \subseteq \mathbb{R}^n$, $G: A \rightarrow \mathbb{R}$ is Schur-convex if $\mathbf{x} \prec \mathbf{y}$ implies $G(\mathbf{x}) \leq G(\mathbf{y})$ ([Definition 21.2.1](#)), and Schur-concave if $\mathbf{x} \prec \mathbf{y}$ implies $G(\mathbf{x}) \geq G(\mathbf{y})$ ([Definition 21.2.2](#)).
- False. $G(\mathbf{x}) = -x_1x_2$ on the nonnegative quadrant is Schur-convex by [Corollary 21.2.5](#), but the warning after that corollary shows that $G(\frac{3}{2}, \frac{3}{2}) = -2.25$ exceeds the average -2.5 of $G(1, 1)$ and $G(2, 2)$.
- False. A Schur-convex function must be symmetric, as shown after [Definition 21.2.2](#), and $G(\mathbf{x}) = x_1$ is linear, hence convex, but not symmetric: $G(1, 0) = 1 \neq 0 = G(0, 1)$, while $(1, 0)$ and $(0, 1)$ majorize each other. Symmetry is exactly the missing hypothesis of [Theorem 21.2.4](#).
- For $I = \mathbb{R}$ and $f(t) = t^2$, it says that $\mathbf{A} \mapsto \text{tr}(\mathbf{A}^2)$ is convex on the Hermitian matrices. Since $\mathbf{A}$ is Hermitian, $\text{tr}(\mathbf{A}^2) = \text{tr}(\mathbf{A}^* \mathbf{A}) = \|\mathbf{A}\|_F^2$, the squared Frobenius norm.

21.2.5.2 B. Practice
Exercise. B1: Four consequences of one majorization

Let $\mathbf{y} = (6, 2, 1, 1)$ and $\mathbf{x} = (4, 3, 2, 1)$.

- Verify that $\mathbf{x} \prec \mathbf{y}$.
- Compute $\sum_i x_i^2$ and $\sum_i y_i^2$, and $\sum_i \max\{x_i - 3, 0\}$ and $\sum_i \max\{y_i - 3, 0\}$. Name the theorem each comparison illustrates.
- Compute $\prod_i x_i$ and $\prod_i y_i$. Hence check [Corollary 21.2.5](#).
- Give a convex $\phi: \mathbb{R} \rightarrow \mathbb{R}$ for which the two sides are **equal**.

Solution.

- Both vectors are decreasing with total 10; the running totals are 4, 7, 9 for $\mathbf{x}$ against 6, 8, 9 for $\mathbf{y}$. So (M1) and (M2) hold.
- $\sum_i x_i^2 = 16 + 9 + 4 + 1 = 30$ and $\sum_i y_i^2 = 36 + 4 + 1 + 1 = 42$, so $30 \leq 42$. Also

$\sum_i \max\{x_i - 3, 0\} = 1 + 0 + 0 + 0 = 1$ and $\sum_i \max\{y_i - 3, 0\} = 3$, so $1 \leq 3$. Both illustrate [Theorem 21.2.1](#), the first with $\phi(t) = t^2$ and the second with the hockey stick at level 3.

- (c) $\prod_i x_i = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ and $\prod_i y_i = 6 \cdot 2 \cdot 1 \cdot 1 = 12$. Indeed $24 \geq 12$, as [Corollary 21.2.5](#) requires for nonnegative vectors.
- (d) Any affine ϕ , for instance $\phi(t) = t$: both sides are the common total 10, by (M2). Affine functions are the ones on which majorization gives no information.

Exercise. B2: Which are Schur-convex?

Determine which of the following functions are Schur-convex on the stated symmetric convex set. Justify your answer, naming the theorem used or giving a counterexample.

- (a) $G(\mathbf{x}) = \max_i x_i - \min_i x_i$ on $\mathbb{R}^n$.
- (b) $G(\mathbf{x}) = \sum_i |x_i|$ on $\mathbb{R}^n$.
- (c) $G(\mathbf{x}) = x_1 + 2x_2$ on $\mathbb{R}^2$.
- (d) $G(\mathbf{x}) = \sum_i \sqrt{x_i}$ on $A = \{\mathbf{x} \in \mathbb{R}^n : x_i > 0 \text{ for every } i\}$.

Solution.

- (a) Schur-convex. $G(\mathbf{x}) = x_1^\downarrow - x_n^\downarrow$, and $\mathbf{x} \prec \mathbf{y}$ gives $x_1^\downarrow \leq y_1^\downarrow$ by (M1) at $k = 1$ and $x_n^\downarrow \geq y_n^\downarrow$ by subtracting (M1) at $k = n - 1$ from (M2). Subtracting the second from the first gives $G(\mathbf{x}) \leq G(\mathbf{y})$.
- (b) Schur-convex, by [Theorem 21.2.1](#) with $\phi(t) = |t|$, convex as a maximum of two affine functions (Proposition 18.10.3).
- (c) Not Schur-convex. It is not symmetric, and a Schur-convex function must be. Concretely $(1, 0)$ and $(0, 1)$ are rearrangements of each other, so $(1, 0) \prec (0, 1)$ and $(0, 1) \prec (1, 0)$. Schur-convexity would give $G(1, 0) \leq G(0, 1)$ from the first and $G(0, 1) \leq G(1, 0)$ from the second, hence $G(1, 0) = G(0, 1)$. But $G(1, 0) = 1$ and $G(0, 1) = 2$.
- (d) Not Schur-convex; it is Schur-**concave**. First, $t \mapsto \sqrt{t}$ is concave on $(0, \infty)$: for $a, b > 0$ and $0 \leq \theta \leq 1$, Lemma 17.8.6 gives $2\sqrt{ab} \leq a + b$, so

$$\begin{aligned} (\theta\sqrt{a} + (1-\theta)\sqrt{b})^2 &= \theta^2 a + (1-\theta)^2 b + 2\theta(1-\theta)\sqrt{ab} \\ &\leq \theta^2 a + (1-\theta)^2 b + \theta(1-\theta)(a+b) \\ &= \theta a + (1-\theta)b, \end{aligned}$$

using $\theta^2 + \theta(1-\theta) = \theta$ and $(1-\theta)^2 + \theta(1-\theta) = 1-\theta$; taking square roots of two nonnegative numbers preserves the inequality. So $\phi(t) = -\sqrt{t}$ is convex on $(0, \infty)$, and [Theorem 21.2.1](#) turns $\mathbf{x} \prec \mathbf{y}$ into $\sum_i (-\sqrt{x_i}) \leq \sum_i (-\sqrt{y_i})$, that is, $G(\mathbf{x}) \geq G(\mathbf{y})$. For a concrete instance, $(2, 2) \prec (3, 1)$ and $G(2, 2) = 2\sqrt{2} \approx 2.83$, while $G(3, 1) = \sqrt{3} + 1 \approx 2.73$, so G went **down** as the vector spread out.

Exercise. B3: Trace convexity on two-by-two matrices

- (a) Let $f(t) = t^2$, $\mathbf{A} = \text{diag}(2, 0)$, $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, and $\theta = \frac{1}{2}$. Compute $\text{tr } f(\mathbf{A})$, $\text{tr } f(\mathbf{B})$, and $\text{tr } f(\mathbf{C})$ for $\mathbf{C} = \frac{1}{2}(\mathbf{A} + \mathbf{B})$, and verify [Theorem 21.2.8](#).

- (b) Let $f(t) = -\log t$, $\mathbf{A} = \mathbf{I}_2$, $\mathbf{B} = \text{diag}(4, \frac{1}{4})$, and $\theta = \frac{1}{2}$. Do the same, and state the inequality in terms of determinants.

Solution.

- (a) Both matrices are real symmetric, hence Hermitian. $\mathbf{A}^2 = \text{diag}(4, 0)$, so $\text{tr } f(\mathbf{A}) = 4$; and $\mathbf{B}^2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, so $\text{tr } f(\mathbf{B}) = 4$. Next

$$\mathbf{C} = \frac{1}{2} \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \mathbf{C}^2 = \begin{pmatrix} \frac{5}{2} & 1 \\ 1 & \frac{1}{2} \end{pmatrix},$$

so $\text{tr } f(\mathbf{C}) = \frac{5}{2} + \frac{1}{2} = 3$. The theorem asserts $3 \leq \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 4 = 4$, which holds.

- (b) Both are positive definite, so their spectra lie in $(0, \infty)$. By [Example \(The logarithm of the determinant\)](#), $\text{tr } f(\mathbf{M}) = -\log \det \mathbf{M}$. Here $\det \mathbf{A} = 1$ and $\det \mathbf{B} = 4 \cdot \frac{1}{4} = 1$, so both traces are $-\log 1 = 0$. The midpoint is $\mathbf{C} = \text{diag}(\frac{5}{2}, \frac{5}{8})$, with $\det \mathbf{C} = \frac{25}{16}$, so $\text{tr } f(\mathbf{C}) = -\log \frac{25}{16}$. The theorem asserts

$$-\log \frac{25}{16} \leq \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 0 = 0,$$

which holds because $\frac{25}{16} > 1$ and $\log$ is increasing with $\log 1 = 0$ (Lemma 18.11.1 (c)). In determinant form: $\det \mathbf{C} = \frac{25}{16} \geq 1 = \sqrt{\det \mathbf{A} \det \mathbf{B}}$.

21.2.5.3 C. Going deeper

Exercise. C1: Entropy

For $\mathbf{x}$ in $A = \{\mathbf{x} \in \mathbb{R}^n : x_i > 0 \text{ for every } i\}$, let $H(\mathbf{x}) = -\sum_i x_i \log x_i$.

- (a) Prove that $g(t) = t \log t$ is convex on $(0, \infty)$. *Hint: apply Theorem 18.10.2 to the concave $\log$, with the points $1/a$, $1/b$ and cleverly chosen weights.*
 (b) Deduce that H is Schur-concave on A .
 (c) Deduce that $H(\mathbf{x}) \leq \log n$ for every $\mathbf{x} \in A$ with $\sum_i x_i = 1$, with equality when $\mathbf{x} = (\frac{1}{n}, \dots, \frac{1}{n})$.

Solution.

- (a) Let $a, b > 0$ and $0 \leq \theta \leq 1$, and put $c = \theta a + (1 - \theta)b > 0$. Set

$$w_1 = \frac{\theta a}{c}, \quad w_2 = \frac{(1 - \theta)b}{c},$$

which are nonnegative and satisfy $w_1 + w_2 = (\theta a + (1 - \theta)b)/c = 1$. The function $-\log$ is convex on $(0, \infty)$ by Lemma 18.11.1 (c), so Theorem 18.10.2 with the points $1/a$ and $1/b$ gives

$$-\log\left(w_1 \frac{1}{a} + w_2 \frac{1}{b}\right) \leq -w_1 \log \frac{1}{a} - w_2 \log \frac{1}{b}.$$

Now $w_1/a + w_2/b = \theta/c + (1 - \theta)/c = 1/c$, and $\log(1/a) = -\log a$ since $\log a + \log(1/a) =$

$\log 1 = 0$ (Lemma 18.11.1 (c)). So the display reads

$$\log c \leq w_1 \log a + w_2 \log b = \frac{\theta a \log a + (1 - \theta)b \log b}{c}.$$

Multiplying by $c > 0$ gives $c \log c \leq \theta a \log a + (1 - \theta)b \log b$, which is Definition 18.10.1 for g .

- (b) By (a) and [Theorem 21.2.1](#) with $I = (0, \infty)$, $\mathbf{x} \prec \mathbf{y}$ in A implies $\sum_i x_i \log x_i \leq \sum_i y_i \log y_i$, that is, $H(\mathbf{x}) \geq H(\mathbf{y})$. By [Definition 21.2.2](#), H is Schur-concave. (The set A is symmetric and convex, so [Definition 21.2.2](#) applies.)
- (c) Let $\mathbf{x} \in A$ with $\sum_i x_i = 1$. Its mean is $\frac{1}{n}$, so the flat vector $\mathbf{f} = (\frac{1}{n}, \dots, \frac{1}{n})$ satisfies $\mathbf{f} \prec \mathbf{x}$ by Example (A chain, the two extremes, and size one) (b), and $\mathbf{f} \in A$. By (b), $H(\mathbf{f}) \geq H(\mathbf{x})$. Finally

$$H(\mathbf{f}) = - \sum_{i=1}^n \frac{1}{n} \log \frac{1}{n} = - \log \frac{1}{n} = \log n,$$

using $\log \frac{1}{n} = -\log n$. So $H(\mathbf{x}) \leq \log n$, with equality at $\mathbf{x} = \mathbf{f}$.

Exercise. C2: Three norms under majorization

Let $\mathbf{x} \prec \mathbf{y}$ in $\mathbb{R}^n$. Prove that

$$\|\mathbf{x}\|_1 \leq \|\mathbf{y}\|_1, \quad \|\mathbf{x}\|_2 \leq \|\mathbf{y}\|_2, \quad \|\mathbf{x}\|_\infty \leq \|\mathbf{y}\|_\infty.$$

Which of the three needs [Theorem 21.2.1](#), and which follows from Definition 17.8.1 alone?

Solution. For $\|\cdot\|_1$, apply [Theorem 21.2.1](#) with $\phi(t) = |t|$, convex on $\mathbb{R}$ by Proposition 18.10.3: $\sum_i |x_i| \leq \sum_i |y_i|$.

For $\|\cdot\|_2$, apply [Theorem 21.2.1](#) with $\phi(t) = t^2$, convex on $\mathbb{R}$ by Lemma 18.10.5, since its derivative $2t$ is increasing: $\sum_i x_i^2 \leq \sum_i y_i^2$. Both sides are nonnegative and $t \mapsto \sqrt{t}$ is increasing on $[0, \infty)$, so $\|\mathbf{x}\|_2 \leq \|\mathbf{y}\|_2$.

For $\|\cdot\|_\infty$, no convexity is needed. Since $\|\mathbf{z}\|_\infty = \max_i |z_i| = \max\{z_1^\downarrow, -z_n^\downarrow\}$, and (M1) at $k = 1$ gives $x_1^\downarrow \leq y_1^\downarrow$ while subtracting (M1) at $k = n - 1$ from (M2) gives $x_n^\downarrow \geq y_n^\downarrow$, hence $-x_n^\downarrow \leq -y_n^\downarrow$. The maximum of the two smaller numbers is at most the maximum of the two larger ones.

So the ∞ -norm follows from Definition 17.8.1 directly, and the other two from [Theorem 21.2.1](#). A later section of this chapter proves the same for **every** norm that is invariant under permuting and changing the signs of the coordinates.

Exercise. C3: One convex function is not enough

Let $\mathbf{x} = (2, -1, -1)$ and $\mathbf{y} = (1, 1, -2)$.

- (a) Show that $\sum_i x_i = \sum_i y_i$ and $\sum_i x_i^2 = \sum_i y_i^2$, yet neither $\mathbf{x} \prec \mathbf{y}$ nor $\mathbf{y} \prec \mathbf{x}$.
- (b) Exhibit a convex ϕ with $\sum_i \phi(x_i) > \sum_i \phi(y_i)$, and another with the inequality reversed.

Explain how this is consistent with [Proposition 21.2.2](#).

Solution.

- (a) Both totals are 0, and $\sum_i x_i^2 = 4 + 1 + 1 = 6 = 1 + 1 + 4 = \sum_i y_i^2$. Both vectors are already decreasing, with running totals 2, 1 for $\mathbf{x}$ and 1, 2 for $\mathbf{y}$. (M1) at $k = 1$ fails for $\mathbf{x} \prec \mathbf{y}$, since $2 > 1$, and (M1) at $k = 2$ fails for $\mathbf{y} \prec \mathbf{x}$, since $2 > 1$.
- (b) Take $\phi(t) = \max\{t - 1, 0\}$, convex by Proposition 18.10.3. Then $\sum_i \phi(x_i) = 1 + 0 + 0 = 1$ and $\sum_i \phi(y_i) = 0 + 0 + 0 = 0$, so $1 > 0$. This is the test function of [Proposition 21.2.2](#) at $c = y_1^\downarrow = 1$, and it certifies $\mathbf{x} \not\prec \mathbf{y}$.

Take $\psi(t) = \max\{t + 1, 0\}$. Then $\sum_i \psi(x_i) = 3 + 0 + 0 = 3$ and $\sum_i \psi(y_i) = 2 + 2 + 0 = 4$, so $3 < 4$, certifying $\mathbf{y} \not\prec \mathbf{x}$ by the same proposition applied with the roles exchanged, at $c = x_2^\downarrow = -1$.

There is no contradiction: [Proposition 21.2.2](#) asks the inequality to hold for **every** convex ϕ . Agreeing on the affine functions and on t^2 is far from agreeing on all of them, and here two hockey sticks pull in opposite directions.