

21.3 Symmetric Gauge Functions

The last two sections turned a majorization into an inequality between sums of a convex function. This section turns it into an inequality between **norms**. The norms that work are the ones blind to exactly the information a majorization discards — the order of the entries and their signs — and Section 4 will show that these are precisely the norms a matrix inherits from its singular values.

Throughout, the field is $\mathbb{R}$ and the space is $\mathbb{R}^n$ with the dot product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$, $n \geq 1$. Majorization $\prec$ and weak majorization $\prec_w$ are Chapter 17 §08's (Definition 17.8.1), and $\mathbf{x}^\downarrow$ is the decreasing rearrangement. A **permutation matrix $\mathbf{P}_\sigma$ is Definition 3.4.2; $\mathbf{P}_\sigma \mathbf{x}$ rearranges the entries of $\mathbf{x}$. We write $|\mathbf{x}| = (|x_1|, \dots, |x_n|)$, and $\mathbf{x} \leq \mathbf{y}$ means $x_i \leq y_i$ for **every** i — the entrywise order of Chapter 18 §05, never the Loewner order. From §01 we use Corollary 21.1.4, which rests on Theorem 21.1.3; the exercises also use §02's Theorem 21.2.1.**

21.3.1 What a norm must forget

Chapter 17 §06 measured a Hermitian matrix by its Ky Fan partial sum $s_k(\mathbf{A}) = \lambda_1(\mathbf{A}) + \dots + \lambda_k(\mathbf{A})$, and Exercise (C2: Partial sums that behave like a norm) asked whether s_k is a norm. On the positive semidefinite matrices it behaves like one; on all Hermitian matrices it fails, because $s_1(\text{diag}(0, -1)) = 0$ with $\text{diag}(0, -1) \neq \mathbf{O}$. The exercise ended by naming the repair: apply s_k to the **singular** values instead of the eigenvalues. Singular values are nonnegative, so nothing can cancel, and the failure disappears.

That repair suggests the shape of the whole chapter's answer. A measurement of a matrix built from its singular values is a function of a list of n nonnegative numbers. Two things about that list are accidents of bookkeeping. Its **order** is one: the singular values of a matrix form a multiset, and Chapter 13 chose to write them decreasingly. Its **signs** are the other: a formula that is to extend from nonnegative lists to all of $\mathbb{R}^n$, and be a norm there, cannot notice which entries are negative, since $\mathbf{x}$ and the vector obtained by flipping one sign have the same absolute values and must be measured alike.

A symmetric gauge function is a norm that cannot tell the entries of a vector apart by where they sit or which way they point.

Definition 21.3.1. Symmetric Gauge Function

A **symmetric gauge function** on $\mathbb{R}^n$ is a norm $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$ (Definition 16.1.1) such that

- (G1) $\Phi(\mathbf{P}_\sigma \mathbf{x}) = \Phi(\mathbf{x})$ for **every** permutation matrix $\mathbf{P}_\sigma$ and every $\mathbf{x} \in \mathbb{R}^n$ (**permutation invariance**);
- (G2) $\Phi(\varepsilon_1 x_1, \dots, \varepsilon_n x_n) = \Phi(\mathbf{x})$ for **all** signs $\varepsilon_i \in \{+1, -1\}$ and every $\mathbf{x} \in \mathbb{R}^n$ (**sign invariance**).

In words: (G1) says that shuffling the entries does not change the measurement, and (G2) says that flipping any of their signs does not either. The word **every** does the work in both, and the two clauses are genuinely separate: below are a norm with (G1) and not (G2), and one with (G2) and not (G1). Together they say that $\Phi(\mathbf{x})$ depends only on the multiset $\{|x_1|, \dots, |x_n|\}$, that is,

$$\Phi(\mathbf{x}) = \Phi(|\mathbf{x}|^\downarrow) \quad (\mathbf{x} \in \mathbb{R}^n), \quad (21.3.1)$$

since $|\mathbf{x}|$ is obtained from $\mathbf{x}$ by (G2) and $|\mathbf{x}|^\downarrow$ from $|\mathbf{x}|$ by (G1). Note also that (N2) of Definition 16.1.1 already gives $\Phi(-\mathbf{x}) = \Phi(\mathbf{x})$; (G2) is much stronger, because it flips the signs **one entry at a time**.

Warning. The word “gauge” is used here in a different sense from Chapter 18 §04. There, Definition 18.4.1 attached to a convex set K the function $p_K(\mathbf{x}) = \inf\{t > 0 : \mathbf{x} \in tK\}$, which measures a vector by inflating K . That object is not this one. The name “symmetric gauge function” is standard and traditional for Definition 21.3.1, and the two never appear in the same argument in this book; where confusion is possible, this chapter writes Φ and never p_K .

The first examples are the norms already at hand.

Example. The p -norms, the Ky Fan gauges, and size one

Check each against Definition 21.3.1.

- (a) $\|\cdot\|_p$ for $1 \leq p \leq \infty$.
- (b) For $1 \leq k \leq n$, the function $\Phi_k(\mathbf{x}) = \sum_{i \leq k} |x|_i^\downarrow$, the sum of the k largest absolute values.
- (c) On $\mathbb{R}^1$, the function $\Phi(x) = c|x|$ for a constant $c > 0$.

Solution.

- (a) These are norms: Example (The 1-, 2-, and ∞ -norms on F^n) for $p = 1, 2, \infty$, and Corollary 18.11.5 supplies the triangle inequality for the rest. The formulas $(\sum_i |x_i|^p)^{1/p}$ and $\max_i |x_i|$ are built from the absolute values alone and are unchanged by reordering the sum or the maximum, so (G1) and (G2) hold.

- (b) Rewrite the definition as a maximum:

$$\Phi_k(\mathbf{x}) = \max \left\{ \sum_{i \in S} |x_i| : S \subseteq \{1, \dots, n\}, |S| = k \right\},$$

since the largest such sum is the one over the indices of the k largest absolute values. Each function $\mathbf{x} \mapsto \sum_{i \in S} |x_i|$ is nonnegative, absolutely homogeneous, and subadditive, so the maximum of finitely many of them is too, which gives (N2) and (N3). For (N1), every singleton is contained in some S , so $\Phi_k(\mathbf{x}) \geq \max_i |x_i|$, and $\Phi_k(\mathbf{x}) = 0$ forces $\mathbf{x} = \mathbf{0}$. Both (G1) and (G2) are visible in the displayed formula. At the two ends, $\Phi_1 = \|\cdot\|_\infty$ and $\Phi_n = \|\cdot\|_1$.

- (c) A norm on $\mathbb{R}^1$ by inspection. There is only one permutation and $\Phi(-x) = \Phi(x)$, so (G1) and (G2) are automatic: on $\mathbb{R}^1$ **every** norm is a symmetric gauge function. The degenerate case is worth keeping in mind, because it shows that the two clauses carry no information until $n \geq 2$.

The functions in (b) will be used constantly, so they get a name of their own.

Definition 21.3.2. Ky Fan Gauge

For $1 \leq k \leq n$, the k -th **Ky Fan gauge** on $\mathbb{R}^n$ is

$$\Phi_k(\mathbf{x}) = \sum_{i=1}^k |x_i|^\downarrow,$$

the sum of the k largest absolute values of the entries of $\mathbf{x}$.

Nonexamples by minimal change. Each drops exactly one clause.

- A norm with (G1) and not (G2). On $\mathbb{R}^2$ let

$$\Phi(\mathbf{x}) = \max(|x_1 + x_2|, |x_1|, |x_2|).$$

Each of the three expressions is nonnegative, absolutely homogeneous, and subadditive, so their maximum is too; and $\Phi(\mathbf{x}) = 0$ forces $x_1 = x_2 = 0$, so Φ is a norm. Swapping x_1 and x_2 permutes the three expressions, so (G1) holds. But $\Phi(1, 1) = 2$ while $\Phi(1, -1) = 1$: (G2) fails.

- A norm with (G2) and not (G1). On $\mathbb{R}^2$ let $\Phi(\mathbf{x}) = |x_1| + 2|x_2|$. This is a norm, and it is built from absolute values, so (G2) holds. But $\Phi(1, 0) = 1$ and $\Phi(0, 1) = 2$: (G1) fails.

Why this definition. The two clauses are exactly what a norm needs in order to be computable from an unordered list of nonnegative numbers, and nothing more is demanded — in particular, no relation to any other norm, and no smoothness. Dropping (G1) allows a norm to prefer one coordinate to another, which no measurement of singular values can do, since relabeling the singular values of a matrix is not an operation on the matrix at all. Dropping (G2) allows a norm to see cancellation between coordinates, which is fatal: Section 7 needs Φ evaluated at differences of eigenvalue lists and of singular value lists, whose entries have both signs (Theorem 21.7.2 and Theorem 21.7.4).

Quick check. Is $\Phi(\mathbf{x}) = \max_i x_i$ (no absolute values) a symmetric gauge function on $\mathbb{R}^n$ for $n \geq 2$? Which clause fails first?

Solution. It is not, and it fails before (G1) or (G2) are reached: it is not a norm. Absolute homogeneity (N2) fails at $c = -1$, since $\max_i (-x_i) = -\min_i x_i$, which for $\mathbf{x} = (1, 0)$ is 0 while $\max_i x_i = 1$. Positive definiteness (N1) fails too: $\Phi(-1, -1) = -1 < 0$. Restricted to vectors with nonnegative entries the formula is perfectly reasonable, and it equals $\|\mathbf{x}\|_\infty$ there; the definition insists on a norm on all of $\mathbb{R}^n$ precisely so that this kind of one-sided formula is excluded.

21.3.2 A symmetric gauge is monotone

The first payoff is that a symmetric gauge respects the crudest comparison there is: if every entry of $\mathbf{x}$ is smaller in absolute value than the matching entry of $\mathbf{y}$, then $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$. This is not automatic for a norm. The function $N(\mathbf{x}) = |x_1 - x_2| + |x_1|$ is a norm on $\mathbb{R}^2$, and $N(1, 0) = 2$ is **larger** than $N(1, 1) = 1$ although $(1, 0)$ is dominated entrywise by $(1, 1)$. Sign invariance is what rules this out.

Proposition 21.3.1. Symmetric Gauges Are Monotone

Let Φ be a symmetric gauge function on $\mathbb{R}^n$, and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ satisfy $|x_i| \leq |y_i|$ for every i .

Then $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$.

Idea. Change one coordinate at a time. Shrinking a single entry towards 0 is taking an average of two vectors: the vector itself and the one with that entry's sign flipped. Sign invariance says those two have the same Φ , and the triangle inequality says the average has no larger one.

Proof. Step 1: one coordinate. Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ agree in every coordinate except the j -th, where $|u_j| \leq |v_j|$. If $v_j = 0$ then $u_j = 0$ and $\mathbf{u} = \mathbf{v}$. Otherwise put

$$\theta = \frac{1}{2} \left(1 + \frac{u_j}{v_j} \right),$$

which lies in $[0, 1]$ because $|u_j/v_j| \leq 1$. Let $\mathbf{v}'$ be $\mathbf{v}$ with its j -th coordinate negated. Then $\theta v_j + (1 - \theta)(-v_j) = u_j$, while in every other coordinate $\mathbf{v}$ and $\mathbf{v}'$ agree with $\mathbf{u}$, so

$$\mathbf{u} = \theta \mathbf{v} + (1 - \theta) \mathbf{v}'.$$

By (N3) and (N2) of Definition 16.1.1, $\Phi(\mathbf{u}) \leq \theta \Phi(\mathbf{v}) + (1 - \theta) \Phi(\mathbf{v}')$, and $\Phi(\mathbf{v}') = \Phi(\mathbf{v})$ by (G2). Hence $\Phi(\mathbf{u}) \leq \Phi(\mathbf{v})$.

Step 2: all coordinates. Put $\mathbf{z}^{(0)} = \mathbf{y}$ and, for $1 \leq j \leq n$,

$$\mathbf{z}^{(j)} = (x_1, \dots, x_j, y_{j+1}, \dots, y_n),$$

so that $\mathbf{z}^{(n)} = \mathbf{x}$. Consecutive vectors $\mathbf{z}^{(j-1)}$ and $\mathbf{z}^{(j)}$ agree except in coordinate j , where $|x_j| \leq |y_j|$. By Step 1, $\Phi(\mathbf{z}^{(j)}) \leq \Phi(\mathbf{z}^{(j-1)})$ for each j . Chaining the n inequalities gives $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$, as claimed. ■

Only (G2) was used. Permutation invariance is needed for the next step, and there it does all the work.

Quick check. Deduce from Proposition 21.3.1 that $\Phi(\mathbf{x}) \geq |x_j| \Phi(\mathbf{e}_j)$ for every j . What does the same proposition give as an upper bound, using $\mathbf{x} = \sum_j x_j \mathbf{e}_j$?

Solution. The vector $x_j \mathbf{e}_j$ has j -th entry x_j and all others 0, so its entries are dominated in absolute value by those of $\mathbf{x}$; Proposition 21.3.1 gives $|x_j| \Phi(\mathbf{e}_j) = \Phi(x_j \mathbf{e}_j) \leq \Phi(\mathbf{x})$. For the upper bound the triangle inequality suffices: $\Phi(\mathbf{x}) \leq \sum_j |x_j| \Phi(\mathbf{e}_j)$. By (G1) all the numbers $\Phi(\mathbf{e}_j)$ are equal, to $\Phi(\mathbf{e}_1)$ say, so

$$\Phi(\mathbf{e}_1) \|\mathbf{x}\|_\infty \leq \Phi(\mathbf{x}) \leq \Phi(\mathbf{e}_1) \|\mathbf{x}\|_1.$$

Every symmetric gauge is trapped between multiples of $\|\cdot\|_\infty$ and $\|\cdot\|_1$, with the same constant at both ends. Section 4 turns this into the statement that every **normalized** unitarily invariant norm lies between the spectral norm and the trace norm.

Both bounds are cited later, so they are recorded here.

Corollary 21.3.2. A Symmetric Gauge Is Trapped Between Two Norms

Let Φ be a symmetric gauge function on $\mathbb{R}^n$. Then $\Phi(\mathbf{e}_j) = \Phi(\mathbf{e}_1)$ for every j , and

$$\Phi(\mathbf{e}_1)\|\mathbf{x}\|_\infty \leq \Phi(\mathbf{x}) \leq \Phi(\mathbf{e}_1)\|\mathbf{x}\|_1 \quad (\mathbf{x} \in \mathbb{R}^n).$$

Proof. Fix j . Some permutation matrix carries $\mathbf{e}_j$ to $\mathbf{e}_1$, so $\Phi(\mathbf{e}_j) = \Phi(\mathbf{e}_1)$ by (G1). The entries of $x_j\mathbf{e}_j$ are dominated in absolute value by those of $\mathbf{x}$, so Proposition 21.3.1 gives

$$|x_j| \Phi(\mathbf{e}_1) = \Phi(x_j\mathbf{e}_j) \leq \Phi(\mathbf{x}),$$

where the equality is (N2) of Definition 16.1.1; taking the largest $|x_j|$ gives the left-hand inequality. For the right-hand one, $\mathbf{x} = \sum_j x_j\mathbf{e}_j$, so (N3) and (N2) give $\Phi(\mathbf{x}) \leq \sum_j |x_j| \Phi(\mathbf{e}_j) = \Phi(\mathbf{e}_1)\|\mathbf{x}\|_1$. This proves the corollary. ■

21.3.3 From weak majorization to a dominating vector

Now for the bridge. A majorization $\mathbf{x} \prec \mathbf{y}$ will be handled by Corollary 21.1.4, which writes $\mathbf{x}$ as an average of rearrangements of $\mathbf{y}$; a norm cannot grow under an average. Weak majorization loses the equality of totals, so $\mathbf{x}$ need not be such an average — $(1, 0) \prec_w (2, 0)$, but $(1, 0)$ is no average of $(2, 0)$ and $(0, 2)$. What survives is that $\mathbf{x}$ is dominated **entrywise** by something that is such an average, and that is exactly the hypothesis Proposition 21.3.1 wants.

Lemma 21.3.3. Weak Majorization Is Domination plus Majorization

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then $\mathbf{x} \prec_w \mathbf{y}$ if and only if there is a vector $\mathbf{z} \in \mathbb{R}^n$ with

$$\mathbf{x} \leq \mathbf{z} \text{ entrywise,} \quad \mathbf{z} \prec \mathbf{y}.$$

If $\mathbf{x}$ has **nonnegative** entries, then $\mathbf{z}$ may be taken with nonnegative entries as well.

Idea. $(\Leftarrow)$ is two easy monotonicity steps. For $(\Rightarrow)$, sort both vectors and think of the deficit $\sum_i y_i - \sum_i x_i$ as water to be poured into $\mathbf{x}$. Pouring it into the **top** entries would break the partial-sum inequalities immediately, so pour it into the bottom: raise every entry of $\mathbf{x}$ that lies below a level c up to c , and choose c so that the totals match. Choosing c needs no analysis: on each interval between consecutive entries of $\mathbf{x}$ the total is an affine function of c , so the right c can be written down.

Proof. $(\Leftarrow)$ Suppose $\mathbf{x} \leq \mathbf{z}$ entrywise and $\mathbf{z} \prec \mathbf{y}$. Fix k and let I be a set of k indices carrying

the k largest entries of $\mathbf{x}$, so that $\sum_{i \leq k} x_i^\downarrow = \sum_{i \in I} x_i$. Then

$$\sum_{i \leq k} x_i^\downarrow = \sum_{i \in I} x_i \leq \sum_{i \in I} z_i \leq \sum_{i \leq k} z_i^\downarrow,$$

the last step because the sum of **any** k entries of $\mathbf{z}$ is at most the sum of its k largest. Hence $\mathbf{x} \prec_w \mathbf{z}$, and $\mathbf{z} \prec \mathbf{y}$ gives $\sum_{i \leq k} z_i^\downarrow \leq \sum_{i \leq k} y_i^\downarrow$ for $k < n$ and equality of totals at $k = n$; in both cases $\sum_{i \leq k} z_i^\downarrow \leq \sum_{i \leq k} y_i^\downarrow$. Combining, $\mathbf{x} \prec_w \mathbf{y}$.

($\Rightarrow$) Suppose $\mathbf{x} \prec_w \mathbf{y}$. Majorization and weak majorization depend only on the decreasing rearrangements (Definition 17.8.1 and the remark after it), and the entrywise relation $\mathbf{x} \leq \mathbf{z}$ is preserved by applying one permutation to both sides. So it is enough to construct $\mathbf{z}$ when $\mathbf{x} = \mathbf{x}^\downarrow$ and $\mathbf{y} = \mathbf{y}^\downarrow$: for general $\mathbf{x}$, choose $\mathbf{P}_\sigma$ with $\mathbf{P}_\sigma \mathbf{x} = \mathbf{x}^\downarrow$, build $\mathbf{z}'$ for the pair $(\mathbf{x}^\downarrow, \mathbf{y}^\downarrow)$, and take $\mathbf{z} = \mathbf{P}_\sigma^{-1} \mathbf{z}'$.

So let $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$ with $\sum_{i \leq k} x_i \leq \sum_{i \leq k} y_i$ for every k . Write $S = \sum_{i=1}^n y_i$ and, for $0 \leq k \leq n-1$,

$$c_k = \frac{S - (x_1 + \cdots + x_k)}{n - k},$$

the level at which the last $n - k$ entries would have to sit for the total to be S . At $k = n - 1$ we have $c_{n-1} = S - \sum_{i \leq n-1} x_i \geq \sum_{i \leq n} x_i - \sum_{i \leq n-1} x_i = x_n$, using $\mathbf{x} \prec_w \mathbf{y}$ at $k = n$. So the set of indices $k \in \{0, \dots, n-1\}$ with $c_k \geq x_{k+1}$ is nonempty; let j be its **smallest** element and put $c = c_j$. Define

$$\mathbf{z} = (x_1, \dots, x_j, \underbrace{c, \dots, c}_{n-j}).$$

The vector $\mathbf{z}$ dominates $\mathbf{x}$. For $i \leq j$ the entries agree. For $i > j$, $z_i = c \geq x_{j+1} \geq x_i$, since $\mathbf{x}$ is decreasing and $c \geq x_{j+1}$ by the choice of j . If in addition $\mathbf{x} \geq \mathbf{0}$, then $c \geq x_{j+1} \geq 0$ and every entry of $\mathbf{z}$ is nonnegative, which is the last claim of the lemma.

The vector $\mathbf{z}$ is decreasing. Only the step from x_j to c is in question, so assume $j \geq 1$. By the minimality of j we have $c_{j-1} < x_j$. Multiplying out the two definitions,

$$(n - j + 1)c_{j-1} = S - \sum_{i \leq j-1} x_i = (n - j)c_j + x_j,$$

so $(n - j)c = (n - j + 1)c_{j-1} - x_j < (n - j + 1)x_j - x_j = (n - j)x_j$, and $n - j \geq 1$ gives $c < x_j$. Hence $\mathbf{z} = \mathbf{z}^\downarrow$.

The vector $\mathbf{z}$ is majorized by $\mathbf{y}$. Its total is $\sum_{i \leq j} x_i + (n - j)c = S$ by the definition of c_j , which is (M2). For (M1) with $k \leq j$, the first k entries of $\mathbf{z}$ are those of $\mathbf{x}$, so $\sum_{i \leq k} z_i = \sum_{i \leq k} x_i \leq \sum_{i \leq k} y_i$. For $k > j$, subtract from the totals: since $\sum_i z_i = \sum_i y_i = S$,

$$\sum_{i \leq k} z_i - \sum_{i \leq k} y_i = \sum_{i > k} y_i - \sum_{i > k} z_i = \sum_{i > k} y_i - (n - k)c,$$

so it suffices to show $\sum_{i > k} y_i \leq (n - k)c$. First, at $k = j$ this holds: from $(n - j)c = S - \sum_{i \leq j} x_i$ and $\sum_{i \leq j} x_i \leq \sum_{i \leq j} y_i$,

$$(n - j)c \geq S - \sum_{i \leq j} y_i = \sum_{i > j} y_i.$$

So c is at least the average of the $n - j$ smallest entries of $\mathbf{y}$. Now let $k > j$. Each of the entries $y_{j+1}, \dots, y_k$ is at least each of $y_{k+1}, \dots, y_n$, because $\mathbf{y}$ is decreasing; hence each of

them is at least the average A of $y_{k+1}, \dots, y_n$, and therefore

$$\sum_{i>j} y_i = \sum_{i=j+1}^k y_i + (n-k)A \geq (k-j)A + (n-k)A = (n-j)A.$$

Combining, $(n-j)c \geq \sum_{i>j} y_i \geq (n-j)A$, so $c \geq A$ and $(n-k)c \geq (n-k)A = \sum_{i>k} y_i$, as required. This proves $\mathbf{z} \prec \mathbf{y}$ and the lemma. ■

The construction is short enough to run by hand.

Example. Filling from the bottom

Let $\mathbf{x} = (4, 1, 1)$ and $\mathbf{y} = (5, 3, 2)$. Verify $\mathbf{x} \prec_w \mathbf{y}$ and produce a $\mathbf{z}$ with $\mathbf{x} \leq \mathbf{z}$ and $\mathbf{z} \prec \mathbf{y}$.

Solution. Both vectors are already decreasing. The running totals are 4, 5, 6 for $\mathbf{x}$ and 5, 8, 10 for $\mathbf{y}$, and $4 \leq 5$, $5 \leq 8$, $6 \leq 10$, so $\mathbf{x} \prec_w \mathbf{y}$. The totals differ, so $\mathbf{x} \not\prec \mathbf{y}$.

Here $S = 10$ and $n = 3$. The levels are $c_0 = 10/3$ and $c_1 = (10 - 4)/2 = 3$. Since $c_0 = 10/3 < 4 = x_1$, the index $k = 0$ fails; since $c_1 = 3 \geq 1 = x_2$, the smallest admissible index is $j = 1$, and $c = 3$. Hence

$$\mathbf{z} = (4, 3, 3).$$

Check: $(4, 1, 1) \leq (4, 3, 3)$ entrywise; $\mathbf{z}$ is decreasing; its running totals 4, 7, 10 satisfy $4 \leq 5$, $7 \leq 8$, $10 = 10$, so $\mathbf{z} \prec \mathbf{y}$. The deficit of 4 between the totals was poured entirely into the two bottom entries, which is what keeps the top partial sums under control.

21.3.4 The inequality

Everything is now in place.

Theorem 21.3.4. Symmetric Gauges Respect Majorization

Let Φ be a symmetric gauge function on $\mathbb{R}^n$.

- (a) If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x} \prec \mathbf{y}$, then $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$.
- (b) If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ have **nonnegative** entries and $\mathbf{x} \prec_w \mathbf{y}$, then $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$.

Idea. For (a), §01 turns $\mathbf{x} \prec \mathbf{y}$ into “ $\mathbf{x}$ is an average of rearrangements of $\mathbf{y}$ ”, and a norm never exceeds an average of the values it takes on the points averaged — those values all being $\Phi(\mathbf{y})$, by permutation invariance. For (b), [Lemma 21.3.3](#) inserts a middle vector $\mathbf{z}$: from $\mathbf{x}$ to $\mathbf{z}$ use monotonicity, from $\mathbf{z}$ to $\mathbf{y}$ use (a).

Proof.

- (a) By Corollary 21.1.4, $\mathbf{x} \in \text{conv}\{\mathbf{P}_\sigma \mathbf{y} : \sigma \in S_n\}$, so there are weights $t_\sigma \geq 0$ with $\sum_\sigma t_\sigma =$

1 and $\mathbf{x} = \sum_{\sigma} t_{\sigma} \mathbf{P}_{\sigma} \mathbf{y}$. By the triangle inequality (N3) and absolute homogeneity (N2) of Definition 16.1.1, applied to this finite sum,

$$\Phi(\mathbf{x}) \leq \sum_{\sigma} t_{\sigma} \Phi(\mathbf{P}_{\sigma} \mathbf{y}) = \sum_{\sigma} t_{\sigma} \Phi(\mathbf{y}) = \Phi(\mathbf{y}),$$

where the middle equality is (G1).

- (b) By Lemma 21.3.3 there is a $\mathbf{z}$ with nonnegative entries satisfying $\mathbf{x} \leq \mathbf{z}$ entrywise and $\mathbf{z} \prec \mathbf{y}$. Since $\mathbf{x}$ and $\mathbf{z}$ have nonnegative entries, $|x_i| = x_i \leq z_i = |z_i|$ for every i , so Proposition 21.3.1 gives $\Phi(\mathbf{x}) \leq \Phi(\mathbf{z})$. By part (a), $\Phi(\mathbf{z}) \leq \Phi(\mathbf{y})$. This proves the theorem. ■

Part (a) needs no sign condition, and part (b) cannot do without one.

Warning. Part (b) is false for vectors with negative entries. Take $\mathbf{x} = (0, -10)$ and $\mathbf{y} = (0, 0)$ in $\mathbb{R}^2$. The decreasing rearrangements are $(0, -10)$ and $(0, 0)$, and the running totals $0, -10$ are at most $0, 0$, so $\mathbf{x} \prec_w \mathbf{y}$. But $\Phi(\mathbf{x}) > 0 = \Phi(\mathbf{y})$ for every norm Φ . The lemma is not at fault — $\mathbf{z} = \mathbf{0}$ does satisfy $\mathbf{x} \leq \mathbf{z}$ and $\mathbf{z} \prec \mathbf{y}$. What fails is the step $\Phi(\mathbf{x}) \leq \Phi(\mathbf{z})$, because Proposition 21.3.1 compares **absolute values**, and $|-10| > |0|$. Weak majorization lets a vector run far in the negative direction, and a norm measures that; only equality of totals, as in $\prec$, or nonnegativity of the entries rules it out. This is why Section 4 applies the theorem to singular values and never to eigenvalues.

The converse of part (b) says that the family of symmetric gauges is strong enough to *detect* weak majorization, and it costs nothing: the Ky Fan gauges Φ_k of Definition 21.3.2 are themselves symmetric gauges (Example (The p-norms, the Ky Fan gauges, and size one) (b)), so the inequalities $\Phi_k(\mathbf{x}) \leq \Phi_k(\mathbf{y})$ for every k are among those hypothesized, and for nonnegative vectors they are the definition of $\mathbf{x} \prec_w \mathbf{y}$. Exercise C1 below states the two directions together, and Section 6 turns them into a theorem about matrices.

21.3.5 The dual gauge

Chapter 18 §04 gave every norm on $\mathbb{R}^n$ a dual, $\|\mathbf{y}\|_* = \max\{\langle \mathbf{x}, \mathbf{y} \rangle : \|\mathbf{x}\| \leq 1\}$ (Definition 18.4.4), and proved that taking the dual twice returns the norm (Theorem 18.4.8). Applied to a symmetric gauge, the construction stays inside the class.

Definition 21.3.3. Dual Gauge

Let Φ be a symmetric gauge function on $\mathbb{R}^n$. Its **dual gauge** is the dual norm of Definition 18.4.4,

$$\Phi^*(\mathbf{y}) = \max\{\langle \mathbf{x}, \mathbf{y} \rangle : \mathbf{x} \in \mathbb{R}^n, \Phi(\mathbf{x}) \leq 1\}.$$

Theorem 21.3.5. The Dual of a Symmetric Gauge

Let Φ be a symmetric gauge function on $\mathbb{R}^n$. Then Φ^* is a symmetric gauge function, and $\Phi^{**} = \Phi$.

Idea. Being a norm and being its own double dual are Chapter 18's, quoted. Only (G1) and (G2) need an argument, and each is the same one: the transformation can be moved off $\mathbf{y}$ and onto $\mathbf{x}$, where it permutes the constraint set $\{\Phi \leq 1\}$ onto itself.

Proof. That Φ^* is a norm is Proposition 18.4.6 (a), and $\Phi^{**} = \Phi$ is Theorem 18.4.8; both apply to any norm on $\mathbb{R}^n$.

(G1). Let $\mathbf{P} = \mathbf{P}_\sigma$ be a permutation matrix. By Lemma 3.4.7 (b), $\mathbf{P}^\top = \mathbf{P}^{-1}$, so $\langle \mathbf{x}, \mathbf{P}\mathbf{y} \rangle = \mathbf{x}^\top \mathbf{P}\mathbf{y} = (\mathbf{P}^\top \mathbf{x})^\top \mathbf{y} = \langle \mathbf{P}^\top \mathbf{x}, \mathbf{y} \rangle$. As $\mathbf{x}$ runs over $\{\Phi \leq 1\}$, so does $\mathbf{x}' = \mathbf{P}^\top \mathbf{x}$: indeed $\Phi(\mathbf{P}^\top \mathbf{x}) = \Phi(\mathbf{x})$ by (G1) for Φ , and $\mathbf{x} \mapsto \mathbf{P}^\top \mathbf{x}$ is a bijection of $\mathbb{R}^n$. Hence the two maxima are over the same set of numbers, and $\Phi^*(\mathbf{P}\mathbf{y}) = \Phi^*(\mathbf{y})$.

(G2). Fix signs $\varepsilon_i = \pm 1$ and write $\mathbf{y}_\varepsilon = (\varepsilon_1 y_1, \dots, \varepsilon_n y_n)$. Then $\langle \mathbf{x}, \mathbf{y}_\varepsilon \rangle = \sum_i x_i \varepsilon_i y_i = \langle \mathbf{x}_\varepsilon, \mathbf{y} \rangle$, and $\mathbf{x} \mapsto \mathbf{x}_\varepsilon$ is a bijection of $\mathbb{R}^n$ with $\Phi(\mathbf{x}_\varepsilon) = \Phi(\mathbf{x})$ by (G2) for Φ . So again the two maxima range over the same numbers, and $\Phi^*(\mathbf{y}_\varepsilon) = \Phi^*(\mathbf{y})$. This proves the theorem. ■

So duality is an involution of the set of symmetric gauge functions on $\mathbb{R}^n$. The examples are Chapter 18's: by Proposition 18.4.7 and Theorem 18.11.6,

$$(\|\cdot\|_p)^* = \|\cdot\|_q \quad \left(\frac{1}{p} + \frac{1}{q} = 1, \quad 1 \leq p \leq \infty \right),$$

so $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are dual to each other and $\|\cdot\|_2$ is self-dual. In the notation of Definition 21.3.2 that says $\Phi_1^* = \Phi_n$ and $\Phi_n^* = \Phi_1$; exercise C2 below computes Φ_k^* for the k in between. Section 5 will find the matrix meaning of all of this: the dual of a unitarily invariant norm, taken with respect to the trace pairing, is the norm attached to the dual gauge.

21.3.6 Exercises

21.3.6.1 A. Check your understanding

Exercise. A1

- State Definition 21.3.1, naming both invariance clauses and the space the function is defined on.
- Explain why $\Phi(\mathbf{x}) = \Phi(|\mathbf{x}|^\downarrow)$ for every symmetric gauge and every $\mathbf{x}$.
- True or false, with a reason: if $\mathbf{x} \prec_w \mathbf{y}$ in $\mathbb{R}^n$, then $\|\mathbf{x}\|_2 \leq \|\mathbf{y}\|_2$.
- Name the hypothesis in Theorem 21.3.4 (b) that part (a) does without, and say where it is used in the proof.

Solution.

- A symmetric gauge function on $\mathbb{R}^n$ is a norm $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying (G1) $\Phi(\mathbf{P}_\sigma \mathbf{x}) = \Phi(\mathbf{x})$ for every permutation matrix, and (G2) $\Phi(\varepsilon_1 x_1, \dots, \varepsilon_n x_n) = \Phi(\mathbf{x})$ for all choices of signs $\varepsilon_i = \pm 1$.
- This is Eq. 21.3.1. Replacing each x_i by $|x_i|$ is a choice of signs, so (G2) gives $\Phi(\mathbf{x}) = \Phi(|\mathbf{x}|)$; sorting $|\mathbf{x}|$ decreasingly is applying a permutation matrix, so (G1) gives $\Phi(|\mathbf{x}|) =$

$$\Phi(|\mathbf{x}|^\downarrow).$$

- (c) False in general. Take $\mathbf{x} = (0, -10)$ and $\mathbf{y} = (0, 0)$: then $\mathbf{x} \prec_w \mathbf{y}$ but $\|\mathbf{x}\|_2 = 10 > 0$. It is true when $\mathbf{x}$ and $\mathbf{y}$ have nonnegative entries, by [Theorem 21.3.4](#) (b) with $\Phi = \|\cdot\|_2$.
- (d) Nonnegativity of the entries. It is used to pass from $\mathbf{x} \leq \mathbf{z}$ entrywise to $|x_i| \leq |z_i|$, which is what [Proposition 21.3.1](#) requires; without it, $x_i \leq z_i$ says nothing about absolute values.

21.3.6.2 B. Practice

Exercise. B1: Which of these are symmetric gauges

Determine which of the following are symmetric gauge functions on $\mathbb{R}^3$. Justify your answer, naming the exact clause that fails when one does.

- (a) $\Phi(\mathbf{x}) = |x_1| + |x_2| + |x_3|$.
- (b) $\Phi(\mathbf{x}) = |x_1|^\downarrow + |x_2|^\downarrow$.
- (c) $\Phi(\mathbf{x}) = (|x_1|^{1/2} + |x_2|^{1/2} + |x_3|^{1/2})^2$.
- (d) $\Phi(\mathbf{x}) = \max(|x_1|, |x_2|) + |x_3|$.
- (e) $\Phi(\mathbf{x}) = |x_1 + x_2 + x_3| + |x_1 - x_2| + |x_2 - x_3|$.

Solution.

- (a) Yes: this is $\|\cdot\|_1 = \Phi_3$, [Example \(The p-norms, the Ky Fan gauges, and size one\)](#) (a).
- (b) Yes: this is the Ky Fan gauge Φ_2 of [Definition 21.3.2](#), checked in [Example \(The p-norms, the Ky Fan gauges, and size one\)](#) (b).
- (c) No. It is nonnegative, absolutely homogeneous, permutation invariant, and sign invariant, but (N3) fails: with $\mathbf{x} = (1, 0, 0)$ and $\mathbf{y} = (0, 1, 0)$ we get $\Phi(\mathbf{x}) = \Phi(\mathbf{y}) = 1$ while $\Phi(\mathbf{x} + \mathbf{y}) = (1 + 1)^2 = 4 > 2$. It is not a norm, so it is not a symmetric gauge.
- (d) No. It is a norm — a sum of two seminorms, positive definite because each coordinate appears — and it is sign invariant, so (G2) holds. But (G1) fails: $\Phi(1, 1, 0) = 1$ while $\Phi(1, 0, 1) = 2$.
- (e) No. It is a norm: it is nonnegative and absolutely homogeneous, subadditive as a sum of three seminorms, and $\Phi(\mathbf{x}) = 0$ forces $x_1 = x_2 = x_3$ and $|3x_1| = 0$, hence $\mathbf{x} = \mathbf{0}$. Both invariance clauses fail. For (G1), $\Phi(1, 1, 0) = 2 + 0 + 1 = 3$ while $\Phi(1, 0, 1) = 2 + 1 + 1 = 4$. For (G2), $\Phi(1, 1, 1) = 3 + 0 + 0 = 3$ while $\Phi(1, -1, 1) = 1 + 2 + 2 = 5$.

Exercise. B2: Filling from the bottom

Let $\mathbf{x} = (6, 2, 2, 0)$ and $\mathbf{y} = (7, 4, 3, 2)$ in $\mathbb{R}^4$.

- (a) Verify that $\mathbf{x} \prec_w \mathbf{y}$ and that $\mathbf{x} \not\prec \mathbf{y}$.
- (b) Carry out the construction in the proof of [Lemma 21.3.3](#) to produce $\mathbf{z}$ with $\mathbf{x} \leq \mathbf{z}$ and $\mathbf{z} \prec \mathbf{y}$.
- (c) Hence bound $\|\mathbf{x}\|_2$ by $\|\mathbf{y}\|_2$ without computing either, and then check the bound numerically.

Solution.

- (a) Both are decreasing. Running totals: 6, 8, 10, 10 for $\mathbf{x}$ and 7, 11, 14, 16 for $\mathbf{y}$. Each is at most the corresponding one, so $\mathbf{x} \prec_w \mathbf{y}$; the totals 10 and 16 differ, so $\mathbf{x} \not\prec \mathbf{y}$.
- (b) Here $S = 16$ and $n = 4$, so $c_0 = 16/4 = 4$, $c_1 = (16 - 6)/3 = 10/3$, $c_2 = (16 - 8)/2 = 4$, $c_3 = 16 - 10 = 6$. Test $c_k \geq x_{k+1}$: $c_0 = 4 < 6 = x_1$, so $k = 0$ fails; $c_1 = 10/3 \geq 2 = x_2$, so $j = 1$ and $c = 10/3$. Hence

$$\mathbf{z} = (6, \frac{10}{3}, \frac{10}{3}, \frac{10}{3}).$$

Check: $\mathbf{x} \leq \mathbf{z}$ entrywise; $\mathbf{z}$ is decreasing; running totals 6, $\frac{28}{3}$, $\frac{38}{3}$, 16, and $6 \leq 7$, $\frac{28}{3} \approx 9.33 \leq 11$, $\frac{38}{3} \approx 12.67 \leq 14$, $16 = 16$. So $\mathbf{z} \prec \mathbf{y}$.

- (c) Both vectors have nonnegative entries, so [Theorem 21.3.4](#) (b) with $\Phi = \|\cdot\|_2$ gives $\|\mathbf{x}\|_2 \leq \|\mathbf{y}\|_2$. Numerically $\|\mathbf{x}\|_2 = \sqrt{36 + 4 + 4} = \sqrt{44}$ and $\|\mathbf{y}\|_2 = \sqrt{49 + 16 + 9 + 4} = \sqrt{78}$, and $44 \leq 78$. The intermediate vector satisfies $\|\mathbf{z}\|_2^2 = 36 + 3 \cdot \frac{100}{9} = \frac{208}{3} \approx 69.3$, between the two, as the proof predicts.

Exercise. B3: Monotonicity used twice

Let Φ be a symmetric gauge function on $\mathbb{R}^3$ with $\Phi(\mathbf{e}_1) = 1$.

- (a) Prove that $1 \leq \Phi(1, 1, 0) \leq 2$.
- (b) Prove that $\Phi(2, 1, 1) \leq \Phi(3, 1, 0)$.
- (c) Give a symmetric gauge for which the inequality in (b) is strict, and one for which it is an equality.

Solution.

- (a) The Quick check following [Proposition 21.3.1](#) gives $\Phi(\mathbf{e}_1)\|\mathbf{x}\|_\infty \leq \Phi(\mathbf{x}) \leq \Phi(\mathbf{e}_1)\|\mathbf{x}\|_1$. With $\mathbf{x} = (1, 1, 0)$ and $\Phi(\mathbf{e}_1) = 1$ this reads $1 \leq \Phi(1, 1, 0) \leq 2$.
- (b) Both vectors are decreasing with total 4, and the running totals 2, 3 of $(2, 1, 1)$ are at most the running totals 3, 4 of $(3, 1, 0)$, so $(2, 1, 1) \prec (3, 1, 0)$. Now [Theorem 21.3.4](#) (a) applies.
- (c) Strict for $\Phi = \|\cdot\|_2$: $\sqrt{6} < \sqrt{10}$. It is also strict for $\Phi = \Phi_2$, where the values are 3 and 4. Equality holds for $\Phi = \|\cdot\|_1 = \Phi_3$, where both values are 4: any two nonnegative vectors with the same total have the same 1-norm, so a majorization can never be detected by Φ_3 alone. This is the general shape of the equality case — the top gauge sees only the total, which a majorization preserves.

21.3.6.3 C. Going deeper

Exercise. C1: Symmetric gauges detect weak majorization

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ have nonnegative entries.

- (a) Prove that $\mathbf{x} \prec_w \mathbf{y}$ if and only if $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$ for **every** symmetric gauge function Φ on $\mathbb{R}^n$.
- (b) Prove that $\mathbf{x} \prec \mathbf{y}$ if and only if $\mathbf{x} \prec_w \mathbf{y}$ and $\sum_i x_i = \sum_i y_i$.

- (c) Deduce that $\mathbf{x} \prec \mathbf{y}$ if and only if $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$ for every symmetric gauge and $\sum_i x_i = \sum_i y_i$.

Hint for (a): only finitely many gauges are needed.

Solution.

- (a) $(\Rightarrow)$ is [Theorem 21.3.4](#) (b). $(\Leftarrow)$ Apply the hypothesis to the Ky Fan gauges Φ_k of [Definition 21.3.2](#), $k = 1, \dots, n$. Since $\mathbf{x}$ and $\mathbf{y}$ have nonnegative entries, $|\mathbf{x}|^\downarrow = \mathbf{x}^\downarrow$ and likewise for $\mathbf{y}$, so $\Phi_k(\mathbf{x}) = \sum_{i \leq k} x_i^\downarrow$ and $\Phi_k(\mathbf{y}) = \sum_{i \leq k} y_i^\downarrow$. The inequalities $\Phi_k(\mathbf{x}) \leq \Phi_k(\mathbf{y})$ for $k = 1, \dots, n$ are exactly (M1) extended to $k = n$, which is Definition 17.8.1's $\mathbf{x} \prec_w \mathbf{y}$.
- (b) This is a restatement of Definition 17.8.1: $\mathbf{x} \prec \mathbf{y}$ means (M1) for $k \leq n-1$ together with equality of totals, and $\mathbf{x} \prec_w \mathbf{y}$ means (M1) for $k \leq n$. Given equality of totals, the case $k = n$ of (M1) is automatic, so the two lists of conditions agree.
- (c) Combine (a) and (b): $\mathbf{x} \prec \mathbf{y}$ holds exactly when $\mathbf{x} \prec_w \mathbf{y}$ and the totals are equal, and by (a) the first of these is equivalent to $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$ for every symmetric gauge.

Exercise. C2: The dual of a Ky Fan gauge

Fix $1 \leq k \leq n$ and let Φ_k be the Ky Fan gauge of [Definition 21.3.2](#).

- (a) Prove that $\Phi_k^*(\mathbf{y}) \geq \max(\|\mathbf{y}\|_\infty, \frac{1}{k}\|\mathbf{y}\|_1)$ by exhibiting two competitors.
- (b) Let $\Phi_k(\mathbf{x}) \leq 1$ and let $\mathbf{y}$ have nonnegative, decreasing entries. Put $t = |x|_k^\downarrow$ and split each $|x_i|$ as $(|x_i| - t)_+ + \min(|x_i|, t)$. Prove that $\sum_i (|x_i| - t)_+ \leq 1 - kt$ and that $0 \leq t \leq 1/k$.
- (c) Deduce $\Phi_k^*(\mathbf{y}) = \max(\|\mathbf{y}\|_\infty, \frac{1}{k}\|\mathbf{y}\|_1)$, and check the cases $k = 1$ and $k = n$ against Proposition 18.4.7.

Solution.

- (a) The vector $\mathbf{e}_1$ has $\Phi_k(\mathbf{e}_1) = 1$; replacing it by $\pm \mathbf{e}_j$ with the sign of y_j , for a j where $|y_j|$ is largest, gives $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{y}\|_\infty$. The vector $\mathbf{x}$ with $x_i = \pm 1/k$, the sign of y_i , has $\Phi_k(\mathbf{x}) = k \cdot \frac{1}{k} = 1$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{k}\|\mathbf{y}\|_1$. So $\Phi_k^*(\mathbf{y})$ is at least both numbers.
- (b) Since $|x|^\downarrow$ is decreasing, $(|x|_i^\downarrow - t)_+ = 0$ for $i > k$, so

$$\sum_i (|x_i| - t)_+ = \sum_{i \leq k} (|x|_i^\downarrow - t) = \Phi_k(\mathbf{x}) - kt \leq 1 - kt.$$

Also $kt \leq \sum_{i \leq k} |x|_i^\downarrow = \Phi_k(\mathbf{x}) \leq 1$, so $0 \leq t \leq 1/k$.

- (c) By [Theorem 21.3.5](#), Φ_k^* is a symmetric gauge, so its value at $\mathbf{y}$ depends only on $|\mathbf{y}|^\downarrow$; and replacing $\mathbf{x}$ by $|\mathbf{x}|$ leaves $\Phi_k(\mathbf{x})$ unchanged and does not decrease $\langle \mathbf{x}, \mathbf{y} \rangle$ when $\mathbf{y} \geq \mathbf{0}$. So we may assume $\mathbf{y}$ has nonnegative decreasing entries and $\mathbf{x} \geq \mathbf{0}$. With t as in (b), and using $y_i \leq y_1 = \|\mathbf{y}\|_\infty$ in the first sum and $\min(x_i, t) \leq t$ in the second,

$$\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{y}\|_\infty(1 - kt) + t\|\mathbf{y}\|_1 = \|\mathbf{y}\|_\infty + t(\|\mathbf{y}\|_1 - k\|\mathbf{y}\|_\infty).$$

If $\|\mathbf{y}\|_1 \leq k\|\mathbf{y}\|_\infty$, the bracket is at most 0 and $t \geq 0$, so the right side is at most

$\|\mathbf{y}\|_\infty$. If $\|\mathbf{y}\|_1 > k\|\mathbf{y}\|_\infty$, the right side increases with t , and $t \leq 1/k$ gives the bound $\|\mathbf{y}\|_\infty + \frac{1}{k}\|\mathbf{y}\|_1 - \|\mathbf{y}\|_\infty = \frac{1}{k}\|\mathbf{y}\|_1$. In both cases $\langle \mathbf{x}, \mathbf{y} \rangle \leq \max(\|\mathbf{y}\|_\infty, \frac{1}{k}\|\mathbf{y}\|_1)$, which with (a) gives equality.

For $k = 1$, $\Phi_1 = \|\cdot\|_\infty$ and the formula gives $\max(\|\mathbf{y}\|_\infty, \|\mathbf{y}\|_1) = \|\mathbf{y}\|_1$, matching Proposition 18.4.7 (b). For $k = n$, $\Phi_n = \|\cdot\|_1$ and the formula gives $\max(\|\mathbf{y}\|_\infty, \frac{1}{n}\|\mathbf{y}\|_1) = \|\mathbf{y}\|_\infty$, matching part (a) of the same proposition.

Exercise. C3: Where each hypothesis is used

- (a) Determine whether there is a norm on $\mathbb{R}^2$ that is permutation invariant and satisfies $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$ whenever $|x_i| \leq |y_i|$ for all i , but is **not** a symmetric gauge function. Justify your answer.
- (b) Theorem 21.3.4 (b) needs nonnegative entries, while Lemma 21.3.3 does not. Locate the exact step of the proof of the theorem that uses the hypothesis, and explain why $\mathbf{x} = (0, -10) \prec_w (0, 0) = \mathbf{y}$ breaks it even though the lemma supplies a perfectly good $\mathbf{z}$.
- (c) Let $\mathbf{x} \prec_w \mathbf{y}$ with $\mathbf{x} = \mathbf{x}^\downarrow$, and let $\mathbf{z}$ be the vector built in the proof of Lemma 21.3.3. Prove that $\mathbf{z} \prec \mathbf{z}'$ for **every** $\mathbf{z}'$ with $\mathbf{x} \leq \mathbf{z}'$ and $\sum_i z'_i = \sum_i y_i$, and deduce that $\sum_i \phi(z_i) \leq \sum_i \phi(z'_i)$ for every convex ϕ . *Hint: compare top partial sums; Theorem 21.2.1 does the last step.*

Solution.

- (a) There is none: monotonicity in the stated sense already forces (G2). If $|x_i| \leq |y_i|$ implies $\Phi(\mathbf{x}) \leq \Phi(\mathbf{y})$, then for any signs the vectors $\mathbf{x}$ and $(\varepsilon_1 x_1, \dots, \varepsilon_n x_n)$ dominate each other entrywise in absolute value, so their Φ -values are equal, which is (G2). With permutation invariance this is Definition 21.3.1. So the conclusion of Proposition 21.3.1 is not merely implied by (G2): for any norm the two are equivalent.
- (b) The step is $\Phi(\mathbf{x}) \leq \Phi(\mathbf{z})$, which invokes Proposition 21.3.1 and therefore needs $|x_i| \leq |z_i|$, not $x_i \leq z_i$. For $\mathbf{x} = (0, -10)$ and $\mathbf{y} = (0, 0)$ the lemma does supply a $\mathbf{z}$, namely $\mathbf{z} = \mathbf{0}$: indeed $\mathbf{z} \prec \mathbf{y}$ and $\mathbf{x} \leq \mathbf{z}$. But $|x_2| = 10 > 0 = |z_2|$, so Proposition 21.3.1 does not apply, and it must not: $\Phi(\mathbf{x}) > 0 = \Phi(\mathbf{z})$ for every norm. When the entries are nonnegative, $x_i \leq z_i$ and $|x_i| \leq |z_i|$ say the same thing, and the gap closes.
- (c) Write $S = \sum_i y_i$ and recall $\mathbf{z} = (x_1, \dots, x_j, c, \dots, c)$, decreasing, with $\sum_i z_i = S$. Let $\mathbf{z}'$ satisfy $\mathbf{x} \leq \mathbf{z}'$ and $\sum_i z'_i = S$, and write $M_1 \geq \dots \geq M_n$ for its entries in decreasing order. Since the totals agree, (M2) holds, and it remains to check (M1) for $1 \leq k \leq n-1$.

Case $k \leq j$. Then $\sum_{i \leq k} z_i = \sum_{i \leq k} x_i \leq \sum_{i \leq k} z'_i \leq \sum_{i \leq k} M_i$, the first inequality from $\mathbf{x} \leq \mathbf{z}'$ and the second because no k entries of $\mathbf{z}'$ add up to more than its k largest.

Case $k > j$. Here $\sum_{i \leq k} z_i = S - (n-k)c$, so the claim is $\sum_{i > k} M_i \leq (n-k)c$. Suppose not. Then the average of $M_{k+1}, \dots, M_n$ exceeds c , and M_{k+1} , being the largest of them, satisfies $M_{k+1} > c$; hence $M_i > c$ for every $i \leq k+1$. Also $\sum_{i \leq j} M_i \geq \sum_{i \leq j} x_i$, by the argument of

the previous case. Therefore

$$\begin{aligned} S &= \sum_i M_i = \sum_{i \leq j} M_i + \sum_{i=j+1}^k M_i + \sum_{i > k} M_i \\ &> \sum_{i \leq j} x_i + (k-j)c + (n-k)c = S, \end{aligned}$$

the last equality being $\sum_{i \leq j} x_i + (n-j)c = S$, which defined c . This contradiction proves the case.

So $\mathbf{z} \prec \mathbf{z}'$, and Theorem 21.2.1 gives $\sum_i \phi(z_i) \leq \sum_i \phi(z'_i)$ for every convex ϕ . In words: among all admissible ways of raising $\mathbf{x}$ to the required total, the flat filling is the most even one.