Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

November 17, 2021 (unfinished!)

Contents

Math 504 Lecture 20

1. Hermitian matrices (cont'd)

1.1. Introduction to majorization theory (cont'd)

Recall:

> **Convention 1.1.1.** Let $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ be a column vector with real entries. Then, for each $i \in [n]$, we let $x_i^{\downarrow}$ denote the i-th largest entry of x. So $\left(x_1^{\downarrow}, x_2^{\downarrow}, \ldots, x_n^{\downarrow}\right)$ is the unique permutation of the tuple $(x_1, x_2, \ldots, x_n)$ that satisfies
> $$x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_n^{\downarrow}.$$

For example, if $x = (3, 5, 2)^T$, then $x_1^{\downarrow} = 5$ and $x_2^{\downarrow} = 3$ and $x_3^{\downarrow} = 2$.
Similarly, we define $x_i^{\uparrow}$ to be the i-th smallest entry of x.

*Drexel University, Korman Center, 15 S 33rd Street, Philadelphia PA, 19104, USA

Definition 1.1.2. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ be two column vectors with real entries. Then, we say that x **majorizes** y (and we write $x \succcurlyeq y$) if and only if we have

$$\sum_{i=1}^{m} x_i^{\downarrow} \geq \sum_{i=1}^{m} y_i^{\downarrow} \qquad \text{for each } m \in [n],$$

with equality for $m = n$ (and possibly for other m's). In other words, x majorizes y if and only if

$$\begin{aligned}
x_1^{\downarrow} &\geq y_1^{\downarrow}; \\
x_1^{\downarrow} + x_2^{\downarrow} &\geq y_1^{\downarrow} + y_2^{\downarrow}; \\
x_1^{\downarrow} + x_2^{\downarrow} + x_3^{\downarrow} &\geq y_1^{\downarrow} + y_2^{\downarrow} + y_3^{\downarrow}; \\
&\ldots; \\
x_1^{\downarrow} + x_2^{\downarrow} + \cdots + x_{n-1}^{\downarrow} &\geq y_1^{\downarrow} + y_2^{\downarrow} + \cdots + y_{n-1}^{\downarrow}; \\
x_1^{\downarrow} + x_2^{\downarrow} + \cdots + x_n^{\downarrow} &= y_1^{\downarrow} + y_2^{\downarrow} + \cdots + y_n^{\downarrow}.
\end{aligned}$$

Last time, we gave an intuition for majorization: We said that x majorizes y if and only if you can obtain y from x by "having the entries come closer together (while keeping the average equal)". Let us now turn this into an actual theorem. First, some definitions:

Definition 1.1.3. **(a)** If $x \in \mathbb{R}^n$ is any column vector, then x_i will mean the i-th coordinate of x (for any $i \in [n]$).

(b) If $x \in \mathbb{R}^n$ is any column vector, then $x^{\downarrow}$ will mean the column vector obtained from x by sorting the coordinates in weakly decreasing order. Thus,

$$x^{\downarrow} = \left(x_1^{\downarrow}, x_2^{\downarrow}, \ldots, x_n^{\downarrow}\right)^T.$$

(c) A vector $x \in \mathbb{R}^n$ is said to be **weakly decreasing** if $x_1 \geq x_2 \geq \cdots \geq x_n$.

Lemma 1.1.4. Let $x, y \in \mathbb{R}^n$. Then, $x \succcurlyeq y$ if and only if $x^{\downarrow} \succcurlyeq y^{\downarrow}$.

Proof. The definition of $\succcurlyeq$ only involves $x^{\downarrow}$ and $y^{\downarrow}$. In other words, whether or not we have $x \succcurlyeq y$ does not depend on the order of the coordinates of x or of those of y. Thus, replacing x and y by $x^{\downarrow}$ and $y^{\downarrow}$ doesn't make any difference. $\square$

Definition 1.1.5. Let $x \in \mathbb{R}^n$. Let $i, j \in [n]$ be such that $x_i \leq x_j$. Let $t \in [x_i, x_j]$ (that is, $t \in \mathbb{R}$ and $x_i \leq t \leq x_j$). Let $y \in \mathbb{R}^n$ be the column vector obtained from x by

replacing the coordinates x_i and x_j by u and v

for some $u, v \in [x_i, x_j]$ satisfying $u + v = x_i + x_j$.

Then, we say that y is obtained from x by a **Robin Hood move** (short: **RH move**), and we write

$$x \xrightarrow{\text{RH}} y.$$

Moreover, if x and y are weakly decreasing, then this RH move is said to be an **order-preserving RH move** (short **OPRH move**), and we write

$$x \xrightarrow{\text{OPRH}} y.$$

Example 1.1.6. **(a)** Replacing two coordinates of a vector x by their average is an RH move.

(b) Swapping two coordinates of a vector x is an RH move.

(c) If $x \in \mathbb{R}^n$ is weakly decreasing, then replacing two adjacent entries of x by their average is an OPRH move.

(d) More generally: If $x \in \mathbb{R}^n$ is weakly decreasing, then replacing its coordinates x_i and x_{i+1} by u and $x_i + x_{i+1} - u$ is an OPRH move if and only if $u \in \left[\dfrac{x_i + x_{i+1}}{2}, x_i\right]$.

Proposition 1.1.7. If $x \xrightarrow{\text{RH}} y$, then the sum of the entries of x equals the sum of the entries of y.

Proof. Clear. □

Lemma 1.1.8. Let $x, y \in \mathbb{R}^n$ be weakly decreasing column vectors such that y is obtained from x by a (finite) sequence of OPRH moves. Then, $x \succcurlyeq y$.

Proof. Recall that the relation $\succcurlyeq$ is reflexive and transitive. Thus, if $x_{[0]} \succcurlyeq x_{[1]} \succcurlyeq \cdots \succcurlyeq x_{[m]}$, then $x_{[0]} \succcurlyeq x_{[m]}$. Therefore, it suffices to prove the proposition in the case when y is obtained from x by a **single** OPRH move.

So let us assume that y is obtained from x by a **single** RH move. Let this move be replacing x_i and x_j by u and v, where $x_i \leq x_j$ and $u, v \in [x_i, x_j]$ with $u + v = x_i + x_j$. WLOG we have $x_i < x_j$ (since otherwise, the OPRH move changes nothing). Therefore, $i > j$ (since x is weakly decreasing). Thus,

$$y = (x_1, x_2, \ldots, x_{j-1}, v, x_{j+1}, x_{j+2}, \ldots, x_{i-1}, u, x_{i+1}, x_{i+2}, \ldots, x_n)$$

(since y is obtained from x by replacing x_i and x_j by u and v).

Now, we must prove that $x \succcurlyeq y$. In other words, we must prove that

$$x_1 + x_2 + \cdots + x_m \geq y_1 + y_2 + \cdots + y_m$$

for each $m \in [n]$ (since x and y are weakly decreasing), and we must prove that

$$x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_n.$$

The latter equality follows from $u+v=x_i+x_j$. So we only need to prove the former inequality. So let us fix an $m \in [n]$. We must show that

$$x_1+x_2+\cdots+x_m \geq y_1+y_2+\cdots+y_m$$

We are in one of the following cases:

1. We have $m<j$.
2. We have $j \leq m < i$.
3. We have $i \leq m$.

In Case 1, we have $x_1+x_2+\cdots+x_m = y_1+y_2+\cdots+y_m$, because $x_p = y_p$ for all $p \leq m$ in this case.

In Case 2, we have

$$\begin{aligned} y_1+y_2+\cdots+y_m &= x_1+x_2+\cdots+x_{j-1}+v+x_{j+1}+x_{j+2}+\cdots+x_m \\ &= (x_1+x_2+\cdots+x_m) + \underbrace{v-x_j}_{\substack{\leq 0 \\ (\text{since } v\in[x_i,x_j])}} \\ &\leq x_1+x_2+\cdots+x_m. \end{aligned}$$

In Case 3, we have

$$\begin{aligned} &y_1+y_2+\cdots+y_m \\ &= x_1+x_2+\cdots+x_{j-1}+v+x_{j+1}+x_{j+2}+\cdots+x_{i-1}+u+x_{i+1}+x_{i+2}+\cdots+x_m \\ &= (x_1+x_2+\cdots+x_m) + \underbrace{(u-x_i)+(v-x_j)}_{\substack{=0 \\ (\text{since } u+v=x_i+x_j)}} \\ &= x_1+x_2+\cdots+x_m. \end{aligned}$$

So we have proved $x_1+x_2+\cdots+x_m \geq y_1+y_2+\cdots+y_m$ in all cases, and we are done. $\square$

Theorem 1.1.9 (RH criterion for majorization)**.** Let $x,y \in \mathbb{R}^n$ be two weakly decreasing column vectors. Then, $x \succcurlyeq y$ if and only if y can be obtained from x by a (finite) sequence of OPRH moves.

Example 1.1.10. (a) We have $(4,1,1) \succcurlyeq (2,2,2)$, and indeed $(2,2,2)$ can be obtained from $(4,1,1)$ by OPRH moves as follows:

$$(4,1,1) \overset{\text{OPRH}}{\longrightarrow} (3,2,1) \overset{\text{OPRH}}{\longrightarrow} (2,2,2)\,.$$

(b) We have $(7,5,2,0) \succcurlyeq (4,4,3,3)$, and indeed $(4,4,3,3)$ can be obtained from $(7,5,2,0)$ by OPRH moves as follows:

$$(7,5,2,0) \overset{\text{OPRH}}{\longrightarrow} (6,6,2,0) \overset{\text{OPRH}}{\longrightarrow} (6,5,3,0) \overset{\text{OPRH}}{\longrightarrow} (6,4,3,1) \overset{\text{OPRH}}{\longrightarrow} (4,4,3,3).$$

Here is another way to do this:

$$(7,5,2,0) \overset{\text{OPRH}}{\longrightarrow} (7,4,3,0) \overset{\text{OPRH}}{\longrightarrow} (4,4,3,3).$$

Proof of Theorem. $\Longleftarrow$: This follows from the lemma above.

$\Longrightarrow$: Let $x \succcurlyeq y$. We must show that y can be obtained from x by a finite sequence of OPRH moves.

If $x = y$, then this is clear (just take the empty sequence). So we WLOG assume that $x \neq y$.. We claim now that there is a further weakly decreasing vector $z \in \mathbb{R}^n$ such that

1. we have $x \overset{\text{OPRH}}{\longrightarrow} z$;
2. we have $z \succcurlyeq y$;
3. the vector z has more entries in common with y than x does; in other words, we have
$$|\{i \in [n] \mid z_i = y_i\}| > |\{i \in [n] \mid x_i = y_i\}|.$$

In other words, we claim that by making a strategic OPRH move starting at x, we can reach a vector z that still majorizes y but has at least one more entry in common with y than x does. If we can prove this claim, then we will automatically obtain a recursive procedure to transform x into y by a sequence of OPRH moves. (And in fact, this procedure will use at most n moves, because each move makes the vector agree with y in at least one more position.)

So let us prove our claim.

Since x is weakly decreasing, we have $x = x^{\downarrow}$. Similarly, $y = y^{\downarrow}$. Thus, from $x \succcurlyeq y$, we obtain

$$x_1 + x_2 + \cdots + x_m \geq y_1 + y_2 + \cdots + y_m \qquad \text{for all } m \in [n],$$

as well as $x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_n$.

Combining $x \neq y$ with $x_1 + x_2 + \cdots + x_n = y_1 + y_2 + \cdots + y_n$, we see that there exists some $a \in [n]$ such that $x_a > y_a$ (why?). Moreover, there exists some pair $(a,b) \in [n] \times [n]$ such that

$$x_a > y_a \qquad \text{and} \qquad x_b < y_b \qquad \text{and} \qquad a < b.$$

(*Proof:* Pick the smallest a such that $x_a \neq y_a$, then the inequality $x_1 + x_2 + \cdots + x_a \geq y_1 + y_2 + \cdots + y_a$ shows that $x_a > y_a$. Now, pick the smallest $b > a$ such that

$x_1 + x_2 + \cdots + x_b = y_1 + y_2 + \cdots + y_b$, then comparing this with $x_1 + x_2 + \cdots + x_{b-1} \geq y_1 + y_2 + \cdots + y_{b-1}$ yields $x_b < y_b$, and thus we have found our pair (a, b).)

So let us pick such a pair (a, b) with smallest possible $b - a$. Then,

$$\begin{aligned} x_a &> y_a, \\ x_j &= y_j \qquad \text{for all } a < j < b, \\ x_b &< y_b \end{aligned}$$

(here, the equalities $x_j = y_j$ come from the "smallest possible $b - a$" condition). Since y is weakly decreasing, we thus have

$$x_a > y_a \geq \left(\text{all of the } x_j \text{ and } y_j \text{ with } a < j < b\right) \geq y_b > x_b.$$

(If there are no $a < j < b$, then this is supposed to read $x_a > y_a \geq y_b > x_b$.) This shows, in particular, that y_a, y_b lie in the open interval (x_a, x_b).

Now, we make an RH move that "squeezes x_a and x_b together" until either x_a reaches y_a or x_b reaches y_b (whatever happens first). In formal terms, this means that we

replace x_a and x_b by y_a and $x_a + x_b - y_a$ if $x_a - y_a \leq y_b - x_b$, and
replace x_a and x_b by $x_a + x_b - y_b$ and y_b if $x_a - y_a \geq y_b - x_b$.

Let $z \in \mathbb{R}^n$ be the resulting n-tuple. We claim that z is weakly decreasing and satisfies the three requirements 1, 2, 3 above:

1. we have $x \overset{\text{OPRH}}{\longrightarrow} z$;
2. we have $z \succcurlyeq y$;
3. the vector z has more entries in common with y than x does; in other words, we have
$$\left|\{i \in [n] \mid z_i = y_i\}\right| > \left|\{i \in [n] \mid x_i = y_i\}\right|.$$

Indeed, the chain of inequalities

$$x_a > y_a \geq \left(\text{all of the } x_j \text{ and } y_j \text{ with } a < j < b\right) \geq y_b > x_b$$

reveals that z is weakly decreasing; thus, our RH move is an OPRH move. So requirement 1 holds.

Requirement 2 is not hard to check (distinguish between the cases $m < a$, $a \leq m < b$ and $m \geq b$). Requirement 3 is easy: $x_a > y_a$ and $x_b < y_b$ but one of z_a and z_b equals the corresponding y_a or y_b.

This completes the proof, as explained above. □