Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

November 29, 2021 (unfinished!)

Contents

Math 504 Lecture 23

1. Positive and nonnegative matrices ([HorJoh, Chapter 8])

1.1. Basics

Recall the triangle inequality:

Proposition 1.1.1 (triangle inequality)**.** Let $z_1, z_2, \ldots, z_n$ be n complex numbers. Then,

$$|z_1| + |z_2| + \cdots + |z_n| \geq |z_1 + z_2 + \cdots + z_n| .$$

Equality holds if and only if $z_1, z_2, \ldots, z_n$ have the same argument (i.e., there exists some $w \in \mathbb{C}$ such that $z_1, z_2, \ldots, z_n$ are nonnegative real multiples of w).

Definition 1.1.2. Let $A \in \mathbb{C}^{n \times m}$ be a matrix.

(a) We say that A is **positive** (and write $A > 0$) if all entries of A are positive reals.

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

(b) We say that A is **nonnegative** (and write $A \geq 0$) if all entries of A are nonnegative reals.

(c) We let $|A| \in \mathbb{R}^{n\times m}$ be the nonnegative matrix obtained by replacing each entry of A by its absolute value. In other words,

$$|A| = \begin{pmatrix} |A_{1,1}| & |A_{1,2}| & \cdots & |A_{1,m}| \\ |A_{2,1}| & |A_{2,2}| & \cdots & |A_{2,m}| \\ \vdots & \vdots & \ddots & \vdots \\ |A_{n,1}| & |A_{n,2}| & \cdots & |A_{n,m}| \end{pmatrix}.$$

Remark 1.1.3. Recall that row vectors and column vectors are matrices. Thus, $v > 0$ and $v \geq 0$ and $|v|$ are defined for them as well. If $v = (v_1, v_2, \ldots, v_k)^T$, then $|v| = (|v_1|, |v_2|, \ldots, |v_k|)^T$.

Warning 1.1.4. Do not mistake $|v|$ (a vector) for $||v||$ (a number). Also, for a matrix A, do not mistake $|A|$ for (an old notation for) the determinant of A.

Exercise 1.1.1. Prove that for any vector $v \in \mathbb{C}^m$, we have $|||v||| = ||v||$, where the left hand side means the length of $|v|$.

Definition 1.1.5. Let $A, B \in \mathbb{R}^{n\times m}$ be two matrices with real entries. Then:

(a) We say that $A \geq B$ if and only if $A - B \geq 0$ (or, equivalently, $A_{i,j} \geq B_{i,j}$ for all $i \in [n]$ and $j \in [m]$).

(b) We say that $A > B$ if and only if $A - B > 0$ (or, equivalently, $A_{i,j} > B_{i,j}$ for all $i \in [n]$ and $j \in [m]$).

(c) We say that $A \leq B$ if and only if $A - B \leq 0$ (or, equivalently, $A_{i,j} \leq B_{i,j}$ for all $i \in [n]$ and $j \in [m]$).

(d) We say that $A < B$ if and only if $A - B < 0$ (or, equivalently, $A_{i,j} < B_{i,j}$ for all $i \in [n]$ and $j \in [m]$).

Example 1.1.6. We have $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \geq \begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix}$.

Remark 1.1.7. Again, recall that row vectors and column vectors are matrices too, so this applies to them as well.

Warning 1.1.8. The relations $\geq$ and $\leq$ and not total orders. For instance, $(2, 1)$ is neither $\geq$ nor $\leq$ to $(3, 0)$.

Warning 1.1.9. Do not mistake $\geq$ for $\succcurlyeq$ (majorization).

Warning 1.1.10. For $n = 0$, the trivial vector $v = () \in \mathbb{R}^0$ satisfies $v > v$ and $v < v$ and $v \geq v$ and $v \leq v$, because the "for all" statements are vacuously true.

Warning 1.1.11. Given two matrices A and B, the relation $A \geq B$ is **not** equivalent to "$A > B$ or $A = B$". For example, $(2,1) \geq (3,1)$, but neither $>$ nor $=$.

Proposition 1.1.12. Let $A \in \mathbb{C}^{n\times m}$ and $B \in \mathbb{C}^{m\times p}$ be two matrices. Then,

$$|A| \cdot |B| \geq |AB| .$$

Proof. We must prove that $(|A| \cdot |B|)_{i,k} \geq |AB|_{i,k}$ for all i and k.
So let $i \in [n]$ and $k \in [p]$. Then,

$$\begin{aligned}(|A| \cdot |B|)_{i,k} &= \sum_{j=1}^{m} \underbrace{|A|_{i,j}}_{=|A_{i,j}|} \cdot \underbrace{|B|_{j,k}}_{=|B_{j,k}|} = \sum_{j=1}^{m} \underbrace{|A_{i,j}| \cdot |B_{j,k}|}_{=|A_{i,j}B_{j,k}|} \\ &= \sum_{j=1}^{m} |A_{i,j}B_{j,k}| \geq \left| \sum_{j=1}^{m} A_{i,j}B_{j,k} \right| .\end{aligned}$$

In view of

$$|AB|_{i,k} = \left|(AB)_{i,k}\right| = \left| \sum_{j=1}^{m} A_{i,j}B_{j,k} \right| ,$$

we can rewrite this as $(|A| \cdot |B|)_{i,k} \geq |AB|_{i,k}$, qed. □

Corollary 1.1.13. Let $A \in \mathbb{C}^{n\times n}$ and $k \in \mathbb{N}$. Then, $|A|^k \geq \left|A^k\right|$.

Proof. Induction on k, using the previous proposition (and the fact that $|I_n| = I_n$). □

Proposition 1.1.14. Let $A \in \mathbb{C}^{n\times m}$ and $x \in \mathbb{C}^m$. Then:
(a) We have $|A| \cdot |x| \geq |Ax|$.
(b) If at least one row of A is positive and we have $A \geq 0$ and $|Ax| = A \cdot |x|$, then $|x| = \omega x$ for some $\omega \in \mathbb{C}$ satisfying $|\omega| = 1$.
(c) If $x > 0$ and $Ax = |A|\, x$, then $A = |A|$ (so that $A \geq 0$).

Proof. **(a)** follows from the previous proposition.
(b) Assume that at least one row of A is positive and we have $A \geq 0$ and $|Ax| = A \cdot |x|$.

We have assumed that at least one row of A is positive. Let the i-th row of A be positive. Write $x = (x_1, x_2, \ldots, x_n)^T$. From $|Ax| = A \cdot |x|$, we have

$$(\text{the } i\text{-th entry of } |Ax|) = (\text{the } i\text{-th entry of } A \cdot |x|) = \sum_{j=1}^{n} A_{i,j} \cdot \left|x_j\right|,$$

so that

$$\sum_{j=1}^{n} A_{i,j} \cdot \left|x_j\right| = (\text{the } i\text{-th entry of } |Ax|) = |\text{the } i\text{-th entry of } Ax|$$
$$= \left|\sum_{j=1}^{n} A_{i,j} x_j\right|$$

(since the i-th entry of Ax is $\sum_{j=1}^{n} A_{i,j} x_j$). This is an equality case of the triangle inequality (since $A_{i,j} = \left|A_{i,j}\right|$ for all j). Thus, the complex numbers $A_{i,j} x_j$ for all $j \in [n]$ have the same argument. In other words, the x_j for all $j \in [n]$ have the same argument. But this means that we can multiply x by a complex number on the unit circle and obtain a vector of positive reals. That latter vector, of course, will be $|x|$.

(c) Suppose $x > 0$ and $Ax = |A|\, x$. We must show that $A = |A|$ (so that $A \geq 0$).

Write $x = (x_1, x_2, \ldots, x_n)^T$. Fix $i \in [n]$. Then,

$$(\text{the } i\text{-th entry of } Ax) = (\text{the } i\text{-th entry of } |A|\, x),$$

i.e.

$$\sum_{j=1}^{n} A_{i,j} x_j = \sum_{j=1}^{n} \underbrace{\left|A_{i,j}\right| x_j}_{=\left|A_{i,j} x_j\right|} = \sum_{j=1}^{n} \left|A_{i,j} x_j\right| \geq \left|\sum_{j=1}^{n} A_{i,j} x_j\right| \geq \sum_{j=1}^{n} A_{i,j} x_j.$$

This is a chain of inequalities in which the first and the last side are equal. Thus, all inequalities in it must be equalities. In particular, we thus have equality in the triangle inequality $\sum_{j=1}^{n} \left|A_{i,j} x_j\right| \geq \left|\sum_{j=1}^{n} A_{i,j} x_j\right|$. Hence, the complex numbers $A_{i,j} x_j$ for all $j \in [n]$ have the same argument. Moreover, $\left|\sum_{j=1}^{n} A_{i,j} x_j\right| \geq \sum_{j=1}^{n} A_{i,j} x_j$ must also become an equality, so this argument has to be 0. This shows that the complex numbers $A_{i,j} x_j$ for all $j \in [n]$ are nonnegative reals. Since $x > 0$, this means that the $A_{i,j}$ are nonnegative reals. Since we have proved this for all i, we thus conclude that all entries of A are nonnegative reals. Hence, $A = |A|$. □

Proposition 1.1.15. Let A, B, C, D be four complex matrices.

(a) If $A \leq B$ and $C \leq D$, then $A + C \leq B + D$ if $A + C$ and $B + D$ are well-defined.

(b) If $A \leq B$ and $0 \leq C$, then $AC \leq BC$ if AC and BC are well-defined.
(c) If $A \leq B$ and $0 \leq C$, then $CA \leq CB$ if CA and CB are well-defined.
(d) If $0 \leq A \leq B$ and $0 \leq C \leq D$, then $0 \leq AC \leq BD$ if AC and BD are well-defined.
(e) If $0 \leq A \leq B$ and $k \in \mathbb{N}$, then $0 \leq A^k \leq B^k$.

Proof. **(a)** Straightforward.
(b) Assume $A \leq B$ and $0 \leq C$. Then, compare

$$(AC)_{i,k} = \sum_j A_{i,j} C_{j,k} \qquad \text{with} \qquad (BC)_{i,k} = \sum_j B_{i,j} C_{j,k}.$$

The right hand side of the first equation is $\leq$ to the right hand side of the second, because $A_{i,j} \leq B_{i,j}$ and $C_{j,k} \geq 0$ for all j. So $(AC)_{i,k} \leq (BC)_{i,k}$. Thus, $AC \leq BC$.
(c) Similar.
(d) Part **(b)** yields $AC \leq BC$. Part **(c)** yields $BC \leq BD$. Since $\leq$ is transitive, this entails $AC \leq BD$.
(e) Follows from **(d)** by induction on k. □

1.2. The spectral radius

Definition 1.2.1. The **spectral radius** $\rho(A)$ of a matrix $A \in \mathbb{C}^{n \times n}$ (with $n > 0$) is defined to be the largest absolute value of an eigenvalue of A. That is,

$$\rho(A) := \max\{|\lambda| \mid \lambda \in \sigma(A)\}.$$

Note that $\rho(A)$ is always a nonnegative real.

By Exercise 3.4.2 (equivalence $\mathcal{A} \iff \mathcal{C}$), a square matrix A satisfies $\rho(A) = 0$ if and only if A is nilpotent.
If $A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, then $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$.

Theorem 1.2.2. Let $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ be such that $B \geq |A|$. Then, $\rho(A) \leq \rho(B)$.

Proof. If $\rho(A) = 0$, then this is obvious. So, WLOG assume that $\rho(A) > 0$.
We can thus scale both matrices A and B by $\dfrac{1}{\rho(A)}$. This does not break the inequality $B \geq |A|$, and also does not break the claim $\rho(A) \leq \rho(B)$ (since $\rho(\lambda A) = |\lambda| \rho(A)$ for any $\lambda \in \mathbb{C}$).
So we WLOG assume that $\rho(A) = 1$. (This is achieved by the scaling we just mentioned.)
This yields that A has an eigenvalue λ with $|\lambda| = 1$. Let v be a nonzero eigenvector at this eigenvalue λ. Thus,

$$A^m v = \lambda^m v \qquad \text{for any } m \in \mathbb{N}.$$

Now, we must prove $\rho(A) \leq \rho(B)$. In other words, we must prove that $1 \leq \rho(B)$. Assume the contrary. Thus, $\rho(B) < 1$. Hence, all eigenvalues of B have absolute value < 1. Hence, Corollary 3.5.2 shows that $\lim\limits_{m\to\infty} B^m = 0$. Therefore, $\lim\limits_{m\to\infty} B^m |v| = 0$.

However, $B \geq |A| \geq 0$ entails $B^m \geq |A|^m$, so that

$$\begin{aligned} B^m |v| &\geq |A|^m |v| \geq |A^m| \cdot |v| \qquad \left(\text{since } |A|^m \geq |A^m|\right) \\ &\geq |A^m v| = |\lambda^m v| = \underbrace{|\lambda^m|}_{\substack{=1 \\ (\text{since } |\lambda|=1)}} \cdot |v| = |v|. \end{aligned}$$

In view of $\lim\limits_{m\to\infty} B^m |v| = 0$, this is only possible if $|v| = 0$. However, $|v| \neq 0$ (since v is nonzero). Contradiction, qed. □

Corollary 1.2.3. Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ be such that $B \geq |A|$. Then, $\rho(A) \leq \rho(|A|) \leq \rho(B)$.

Proof. Applying the above theorem to $|A|$ instead of B, we get $\rho(A) \leq \rho(|A|)$.

Applying the above theorem to $|A|$ instead of A, we get $\rho(|A|) \leq \rho(B)$ (since $||A|| = |A|$). □

Corollary 1.2.4. Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ satisfy $B \geq A \geq 0$, then $\rho(A) \leq \rho(B)$.

Proof. Apply the theorem, noticing that $|A| = A$. □

Corollary 1.2.5. Let $A \in \mathbb{R}^{n\times n}$ satisfy $A \geq 0$.

(a) If $\widetilde{A}$ is a principal submatrix of A (that is, a matrix obtained from A by removing a bunch of rows along with the corresponding columns), then $\rho\left(\widetilde{A}\right) \leq \rho(A)$.

(b) We have $\max\{A_{i,i} \mid i \in [n]\} \leq \rho(A)$.

(c) If $A_{i,i} > 0$ for some $i \in [n]$, then $\rho(A) > 0$.

Proof. **(a)** Let $\widetilde{A}$ be a principal submatrix of A. For simplicity, I assume that $\widetilde{A}$ is A without the n-th row and the n-th column. Thus,

$$A = \begin{pmatrix} \widetilde{A} & y \\ x & \lambda \end{pmatrix} \qquad \text{(in block-matrix notation)}$$

for some nonnegative x, y and λ. Now, it is easy to see that

$$\rho\left(\widetilde{A}\right) = \rho\left(\begin{pmatrix} \widetilde{A} & 0 \\ 0 & 0 \end{pmatrix}\right)$$

(since $\sigma\left(\left(\begin{array}{cc}\widetilde{A} & 0\\ 0 & 0\end{array}\right)\right)=\sigma\left(\widetilde{A}\right)\cup\{0\}$). However, $0\leq\left(\begin{array}{cc}\widetilde{A} & 0\\ 0 & 0\end{array}\right)\leq\left(\begin{array}{cc}\widetilde{A} & y\\ x & \lambda\end{array}\right)=A$, so the previous corollary yields

$$\rho\left(\left(\begin{array}{cc}\widetilde{A} & 0\\ 0 & 0\end{array}\right)\right)\leq\rho\left(A\right).$$

So $\rho\left(\widetilde{A}\right)\leq\rho\left(A\right)$. This proves part **(a)**.

(b) We must show that $A_{i,i}\leq\rho\left(A\right)$ for all $i\in[n]$.

So let $i\in[n]$. Then, the 1×1-matrix $\left(\begin{array}{c}A_{i,i}\end{array}\right)$ is a principal submatrix of A (obtained by removing all rows of A other than the i-th one, and all columns of A other than the i-th one). Hence, part **(a)** yields $\rho\left(\left(\begin{array}{c}A_{i,i}\end{array}\right)\right)\leq\rho\left(A\right)$. However, $\rho\left(\left(\begin{array}{c}A_{i,i}\end{array}\right)\right)=\left|A_{i,i}\right|=A_{i,i}$ (since $A\geq 0$). Therefore, $A_{i,i}\leq\rho\left(A\right)$, qed.

(c) Follows from **(b)**. □