Math 504: Advanced Linear Algebra

Hugo Woerdeman, with edits by Darij Grinberg*

October 25, 2021 (unfinished!)

Contents

1. Hermitian matrices **1**
1.1. Basics 1
1.2. Definiteness 4

Math 504 Lecture 15

1. Hermitian matrices

Recall: A **Hermitian matrix** is an $n \times n$-matrix $A \in \mathbb{C}^{n \times n}$ such that $A^* = A$.

Note that this is the complex analogue of real symmetric matrices ($A \in \mathbb{R}^{n \times n}$ such that $A^T = A$).

If A is a Hermitian matrix, then $A_{i,i} \in \mathbb{R}$ and $A_{i,j} = \overline{A_{j,i}}$.

For instance, the matrix $\begin{pmatrix} -1 & i & 2 \\ -i & 5 & 1-i \\ 2 & 1+i & 0 \end{pmatrix}$ is Hermitian.

1.1. Basics

Theorem 1.1.1. Let $A \in \mathbb{C}^{n \times n}$ be an $n \times n$-matrix. Then, the following are equivalent:

- $\mathcal{A}$: The matrix A is Hermitian (i.e., we have $A^* = A$).

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

- $\mathcal{B}$: We have $A = UDU^*$ for some unitary U and some real diagonal D (that is, D is a diagonal matrix with real entries).
- $\mathcal{C}$: The matrix A is normal and its eigenvalues are real.
- $\mathcal{D}$: We have $\langle Ax, x\rangle \in \mathbb{R}$ for each $x \in \mathbb{C}^n$.
- $\mathcal{E}$: The matrix S^*AS is Hermitian for all $S \in \mathbb{C}^{n\times k}$ (for all $k \in \mathbb{N}$).

To prove this, we will need a lemma:

Lemma 1.1.2. Let $M \in \mathbb{C}^{n\times n}$ be an $n \times n$-matrix. Assume that $\langle Mx, x\rangle = 0$ for each $x \in \mathbb{C}^n$. Then, $M = 0$.

Proof. If $x \in \mathbb{C}^n$ has entries $x_1, x_2, \ldots, x_n$, then

$$\begin{aligned}
\langle Mx, x\rangle &= \left\langle \begin{pmatrix} M_{1,1}x_1 + M_{1,2}x_2 + \cdots + M_{1,n}x_n \\ M_{2,1}x_1 + M_{2,2}x_2 + \cdots + M_{2,n}x_n \\ \vdots \\ M_{n,1}x_1 + M_{n,2}x_2 + \cdots + M_{n,n}x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\rangle \\
&= \sum_{i=1}^{n} \left(M_{i,1}x_1 + M_{i,2}x_2 + \cdots + M_{i,n}x_n\right) \overline{x_i} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} M_{i,j}x_j\overline{x_i} = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{i,j}\overline{x_i}x_j.
\end{aligned}$$

So this is always $= 0$ by assumption, no matter what x is. So we have shown that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} M_{i,j}\overline{x_i}x_j = 0 \qquad \text{for every } x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n.$$

In particular:

- We can apply this to $x = e_1 = (1, 0, 0, \ldots, 0)^T$, and we obtain $M_{1,1} \cdot \overline{1} \cdot 1 = 0$, which means $M_{1,1} = 0$. Similarly, we can find $M_{i,i} = 0$ for all $i \in [n]$.
- We can apply this to $x = e_1 + e_2 = (1, 1, 0, 0, \ldots, 0)^T$, and we obtain
  $$M_{1,1} \cdot \overline{1} \cdot 1 + M_{1,2} \cdot \overline{1} \cdot 1 + M_{2,1} \cdot \overline{1} \cdot 1 \cdot M_{2,2} \cdot \overline{1} \cdot 1 = 0.$$
  This simplifies to
  $$M_{1,1} + M_{1,2} + M_{2,1} + M_{2,2} = 0.$$
  However, the previous bullet point yields $M_{1,1} = 0$ and $M_{2,2} = 0$, so this simplifies further to
  $$M_{1,2} + M_{2,1} = 0.$$

- We can apply this to $x = e_1 + ie_2 = (1, i, 0, 0, \ldots, 0)^T$, and we obtain

$$M_{1,1} \cdot \overline{1} \cdot 1 + M_{1,2} \cdot \overline{1} \cdot i + M_{2,1} \cdot \overline{i} \cdot 1 \cdot M_{2,2} \cdot \overline{i} \cdot i = 0.$$

  This simplifies to

$$M_{1,1} + iM_{1,2} - iM_{2,1} + M_{2,2} = 0.$$

  However, we know that $M_{1,1} = 0$ and $M_{2,2} = 0$, so this simplifies further to

$$iM_{1,2} - iM_{2,1} = 0.$$

  Thus,

$$M_{1,2} - M_{2,1} = 0.$$

  Adding this to

$$M_{1,2} + M_{2,1} = 0,$$

  we obtain $2M_{1,2} = 0$. In other words, $M_{1,2} = 0$. Similarly, we can show that $M_{i,j} = 0$ for all $i \neq j$.

So we have now shown that all entries of M are 0. In other words, $M = 0$. This proves the lemma. $\square$

Now we can prove the theorem:

Proof of Theorem. The implication $\mathcal{A} \Longrightarrow \mathcal{B}$ follows from the spectral theorem. So does the implication $\mathcal{A} \Longrightarrow \mathcal{C}$. The implication $\mathcal{B} \Longrightarrow \mathcal{A}$ follows from a corollary of the spectral theorem. Finally, $\mathcal{C} \Longrightarrow \mathcal{B}$ also follows from the spectral theorem. So we only need to prove the equivalence $\mathcal{A} \Longleftrightarrow \mathcal{D} \Longleftrightarrow \mathcal{E}$.

- *Proof of $\mathcal{A} \Longrightarrow \mathcal{D}$:* Assume that $\mathcal{A}$ holds. Thus, $A = A^*$. Now, let $x \in \mathbb{C}^n$. Then, $\langle Ax, x \rangle = \overline{\langle x, Ax \rangle}$ (by the rules for inner products). However, by the formula $\langle u, v \rangle = v^* u$, we have

$$\langle Ax, x \rangle = x^* Ax \qquad \text{and} \qquad \langle x, Ax \rangle = \underbrace{(Ax)^*}_{=x^*A^*} x = x^* \underbrace{A^*}_{=A} x = x^* Ax.$$

  Comparing these two equalities, we see that $\langle Ax, x \rangle = \langle x, Ax \rangle$. Comparing this with $\langle Ax, x \rangle = \overline{\langle x, Ax \rangle}$, we obtain $\langle x, Ax \rangle = \overline{\langle x, Ax \rangle}$. This entails $\langle x, Ax \rangle \in \mathbb{R}$ (since the only complex numbers $z \in \mathbb{C}$ that satisfy $z = \overline{z}$ are the real numbers). Therefore, $\langle Ax, x \rangle = \langle x, Ax \rangle \in \mathbb{R}$. Thus, the statement $\mathcal{D}$ is proved.

- *Proof of $\mathcal{D} \Longrightarrow \mathcal{A}$:* Assume that $\mathcal{D}$ holds. Thus, $\langle Ax, x \rangle \in \mathbb{R}$ for each $x \in \mathbb{C}^n$. Again, we can see that each $x \in \mathbb{C}^n$ satisfies

$$\langle Ax, x \rangle = x^* Ax \qquad \text{and} \qquad \langle x, Ax \rangle = \underbrace{(Ax)^*}_{=x^*A^*} x = x^* A^* x.$$

Thus, each $x \in \mathbb{C}^n$ satisfies

$$\begin{aligned} x^* A x &= \langle Ax, x \rangle = \overline{\langle Ax, x \rangle} \qquad \text{(since } \langle Ax, x \rangle \in \mathbb{R}\text{)} \\ &= \langle x, Ax \rangle = x^* A^* x, \end{aligned}$$

so that

$$x^* A x - x^* A^* x = 0,$$

so that

$$x^* \left(A^* - A \right) x = 0.$$

Applying our Lemma to $M = A^* - A$, we thus conclude that $A^* - A = 0$. In other words, $A^* = A$. This proves $\mathcal{A}$.

- *Proof of* $\mathcal{A} \Longrightarrow \mathcal{E}$*:* If A is Hermitian, then $A^* = A$, so that

$$\left(S^* A S \right)^* = S^* \underbrace{A^*}_{=A} \underbrace{\left(S^* \right)^*}_{=S} = S^* A S,$$

  and therefore $S^* A S$ is again Hermitian. This proves $\mathcal{A} \Longrightarrow \mathcal{E}$.

- *Proof of* $\mathcal{E} \Longrightarrow \mathcal{A}$*:* If statement $\mathcal{E}$ holds, then we can apply it to $S = I_n$ (and $k = n$), and conclude that $I_n^* A I_n$ is Hermitian; but this is simply saying that A is Hermitian. So $\mathcal{E} \Longrightarrow \mathcal{A}$ follows.

The theorem is proved. □

As a reminder: Sums of Hermitian matrices are Hermitian, but products are not (in general).

1.2. Definiteness

Definition 1.2.1. Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix.

(a) We say that A is **positive semidefinite** if it satisfies

$$\langle Ax, x \rangle \geq 0 \qquad \text{for all } x \in \mathbb{C}^n.$$

(b) We say that A is **positive definite** if it satisfies

$$\langle Ax, x \rangle > 0 \qquad \text{for all nonzero } x \in \mathbb{C}^n.$$

(c) We say that A is **negative semidefinite** if it satisfies

$$\langle Ax, x \rangle \leq 0 \qquad \text{for all } x \in \mathbb{C}^n.$$

(d) We say that A is **negative definite** if it satisfies

$$\langle Ax, x \rangle < 0 \qquad \text{for all nonzero } x \in \mathbb{C}^n.$$

(e) We say that A is **indefinite** if it is neither positive semidefinite nor negative semidefinite, i.e., if there exist vectors $x, y \in \mathbb{C}^n$ such that

$$\langle Ax, x \rangle < 0 < \langle Ay, y \rangle .$$

Here are some examples of definiteness:

Example 1.2.2. Let $n \in \mathbb{N}$. Let $J = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$. This matrix J is real symmetric, thus Hermitian. Is it positive definite? Positive semidefinite?

Let $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$. Then,

$$\begin{aligned} \langle Jx, x \rangle &= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{x_i} x_j = \left(\sum_{i=1}^{n} \overline{x_i} \right) \left(\sum_{j=1}^{n} x_j \right) = \left(\overline{\sum_{i=1}^{n} x_i} \right) \left(\sum_{j=1}^{n} x_j \right) \\ &= \left(\overline{\sum_{i=1}^{n} x_i} \right) \left(\sum_{i=1}^{n} x_i \right) = \left| \sum_{i=1}^{n} x_i \right|^2 \geq 0. \end{aligned}$$

So J is positive semidefinite.

Is J positive definite? To have $\langle Jx, x \rangle = 0$ is equivalent to having $\sum\limits_{i=1}^{n} x_i = 0$. When $n = 1$ (or $n = 0$), this is equivalent to having $x = 0$, so J is positive definite in this case. However, if $n > 1$, then this is not equivalent to having $x = 0$, and in fact the vector $e_1 - e_2$ is an example of a nonzero vector $x \in \mathbb{C}^n$ such that $\langle Jx, x \rangle = 0$. So J is not positive definite unless $n \leq 1$.

Example 1.2.3. Consider a diagonal matrix

$$D := \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

with $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$. When is D positive semidefinite?

We want $\langle Dx, x \rangle \geq 0$ for all $x \in \mathbb{C}^n$. Let $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$. Then,

$$\langle Dx, x \rangle = \sum_{i=1}^{n} \lambda_i \overline{x_i} x_i = \sum_{i=1}^{n} \lambda_i |x_i|^2 .$$

If $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$, then we therefore conclude that $\langle Dx, x \rangle \geq 0$, so that D is positive semidefinite. Otherwise, D is not positive semidefinite, since we can pick an $x = e_j$ where j satisfies $\lambda_i < 0$. So D is positive semidefinite if and only if $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$. A similar argument shows that D is positive definite if and only if $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$.

Example 1.2.4. The Hilbert matrix

$$\begin{pmatrix} \frac{1}{1} & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n} \end{pmatrix}$$

(i.e., the $n \times n$-matrix whose (i,j)-th entry is $\dfrac{1}{i+j-1}$) is positive definite. In other words, for any $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\overline{x_i} x_j}{i+j-1} \geq 0.$$

This is not obvious at all, and will be a HW exercise (with hints). More generally, if $a_1, a_2, \ldots, a_n$ are positive reals, then the $n \times n$-matrix whose (i,j)-th entry is $\dfrac{1}{a_i + a_j}$ is positive definite.

As an application of positive semidefiniteness, the Schoenberg theorem generalizes the triangle inequality. Recall that the triangle inequality says that three nonnegative real numbers x, y, z are the mutual distances of 3 points in the plane if and only if $x \leq y + z$ and $y \leq z + x$ and $z \leq x + y$. In higher dimensions, the analogous criterion is the following:

Theorem 1.2.5 (Schoenberg's theorem)**.** Let $n \in \mathbb{N}$ and $r \in \mathbb{N}$. Let $d_{i,j}$ be a nonnegative real for each $i, j \in [n]$. Assume that $d_{i,i} = 0$ for all $i \in [n]$, and furthermore $d_{i,j} = d_{j,i}$ for all $i, j \in [n]$. Then, there exist n points $P_1, P_2, \ldots, P_n \in \mathbb{R}^r$ satisfying

$$\left|P_i - P_j\right| = d_{i,j} \qquad \text{for all } i, j \in [n]$$

if and only if the $(n-1) \times (n-1)$-matrix whose (i,j)-th entry is

$$d_{i,n}^2 + d_{j,n}^2 - d_{i,j}^2 \qquad \text{for all } i, j \in [n-1]$$

is positive semidefinite and has rank $\leq r$.

We will not prove this here.

Remark 1.2.6. If $A \in \mathbb{R}^{n \times n}$ and $\langle Ax, x \rangle \geq 0$ for all $x \in \mathbb{R}^n$, then we **cannot** conclude that A is positive semidefinite. The reason is that it does not follow that A is symmetric. For example, $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ satisfies

$$\langle Ax, x \rangle = 2x_1^2 + x_1 x_2 + 2x_2^2 = \frac{1}{2}(x_1 + x_2)^2 + \frac{3}{2}\left(x_1^2 + x_2^2\right) \geq 0$$

for each $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$, but it is not symmetric.

Theorem 1.2.7. Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix. Then:

(a) The matrix A is positive semidefinite if and only if all eigenvalues of A are nonnegative.

(b) The matrix A is positive definite if and only if all eigenvalues of A are positive.

(Recall that the eigenvalues of A are reals by the spectral theorem.)

Proof. **(a)** Assume that A is positive semidefinite. Let λ be an eigenvalue of A. Let $x \neq 0$ be a corresponding eigenvector. Then, $Ax = \lambda x$. However, $\langle Ax, x \rangle \geq 0$ since A is positive semidefinite. So $\langle \lambda x, x \rangle \geq 0$. However, $\langle \lambda x, x \rangle = \lambda \langle x, x \rangle$, so that $\lambda \langle x, x \rangle \geq 0$. We can cancel $\langle x, x \rangle$ (since $\langle x, x \rangle > 0$). Thus, we get $\lambda \geq 0$. Therefore, all eigenvalues of A are ≥ 0.

Conversely, assume that all eigenvalues of A are ≥ 0. By the spectral theorem, we can write A as

$$A = UDU^*, \qquad \text{where } D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. However,

$$D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \left(\operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\right)\right)^2$$

(the square roots here are well-defined, since the λ_i are nonnegative by assumption). .... $\square$