# MATH 325 Q1: LINEAR ALGEBRA III, PART 10

## Spectral Theorem

The following result is called the spectral theorem for unitary and hermitian $\mathbb{C}$-linear maps.

**Theorem.** *Let $(V, \langle -,- \rangle)$ be an inner product space and $\alpha : V \to V$ a unitary or hermitian $\mathbb{C}$-linear map. Then there exists an orthogonal basis of $V$ consisting of eigenvectors of $\alpha$.*

*Proof.* **Step 1.** Let $\lambda$ be an eigenvalue of $\alpha$ and $v \neq 0$ the corresponding eigenvector. Thus $\alpha(v) = \lambda\, v$. Let $W = \mathbb{C}\cdot v$ be the subspace generated by $v$. This is an $\alpha$-invariant subspace of $V$ of dimension 1 since $v$ is an eigenvector. We claim that $W^\perp$ is also $\alpha$-invariant. Indeed, let $w \in W^\perp$. We have to show $\langle \alpha(w), v\rangle = 0$, or equivalently $\langle v, \alpha(w)\rangle = 0$.

**Case 1.** Let $\alpha$ be unitary. Then

$$\bar{\lambda} \cdot \langle \alpha(w), v\rangle \;=\; \langle \alpha(w), \lambda v\rangle \;=\; \langle \alpha(w), \alpha(v)\rangle \;=\; \langle v, w\rangle\,,$$

where the last equation is due to the fact that $\alpha$ is unitary. Since $w \perp v$ we have $\langle v, w\rangle = 0$ and so by above equality $\bar{\lambda}\cdot\langle \alpha(w), v\rangle = 0$. Note that $\lambda \neq 0$ because $\alpha$ being unitary is invertible. Therefore $\langle \alpha(w), v\rangle = 0$ as claimed.

**Case 2.** Assume now that $\alpha$ is hermitian. Then $\alpha = \alpha^*$ and hence we have

$$\langle v, \alpha(w)\rangle \;=\; \langle v, \alpha^*(w)\rangle \;=\; \langle \alpha(v), w\rangle \;=\; \langle \lambda\cdot v, w\rangle \;=\; \lambda\cdot\langle v, w\rangle \;= 0$$

(the last equality holds since $v \perp w$). This implies $\langle v, \alpha(w)\rangle = 0$, as required.

**Step 2.** Remind that we proved before that the restriction of $\alpha$ to $W^\perp$ is also hermitian if $\alpha$ is hermitian and is also unitary if $\alpha$ is unitary. Hence by induction on dimension of vector space, there is an orthogonal basis $v_2, \dots, v_n$ of $W^\perp$ consisting of eigenvectors of $\alpha|_{W^\perp}$. Thus letting $v_1 := v$ we get that the vectors $v_1, v_2, \dots, v_n$ form an orthogonal basis of $V$ consisting of eigenvectors of $\alpha$. $\square$

**Corollary.** *Let $A$ be a hermitian $n \times n$-matrix. Then there exists an unitary $n \times n$-matrix $U$, such that*

$$U^* \cdot A \cdot U = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

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

*Proof.* By the theorem above applied to the linear map

$$\alpha_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad v \mapsto A \cdot v$$

there exists an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of $\alpha_A$ or equivalently of eigenvectors of $A$. Therefore, $A$ is diagonalizable. If $U$ is the matrix whose columns are the vectors of this basis then $U$ is the base change matrix, hence $U^{-1}AU$ is the diagonal matrix, whose diagonal entries are eigenvalues of $A$. On the other hand $U$ is unitary (because its columns is an orthonormal basis of $\mathbb{C}^n$), hence $U^*U = \mathrm{I}_n$ and this implies $U^{-1} = U^*$. Therefore,

$$U^{-1}AU = U^* \cdot A \cdot U$$

is a diagonal matrix with entries the eigenvalues of $A$. □

**Remark.** *Note that all eigenvalues above are real numbers because the matrix $U^*AU$ is also hermitian.*

## Polar decomposition

Recall that any complex number $z \neq 0$ can be uniquely written as

$$z = r \cdot e^{i\,\alpha} = r(\cos(\alpha) + i \cdot \sin(\alpha)),$$

where $r > 0$ is a real number and $\alpha$ is a real number in the interval $[0, 2\pi)$. Our goal is to show that a similar decomposition, called *polar decomposition*, exists also for any invertible complex $n \times n$-matrix.

Note that the complex number $e^{i\,\alpha} = \cos(\alpha) + i \cdot \sin(\alpha)$ can be viewed as a matrix of size $1 \times 1$ and it has the property

$$e^{i\,\alpha} \cdot \overline{e^{i\,\alpha}} = (\cos(\alpha) + i \cdot \sin(\alpha)) \cdot \overline{(\cos(\alpha) + i \cdot \sin(\alpha))} = 1.$$

As we already know, matrices of an arbitrary size $n \times n$ with the similar property are called unitary. We next want to introduce an analogue of positive real numbers for matrices.

**Definition.** *Let $A$ be a hermitian complex $n \times n$-matrix. We say that $A$ is positive definite if*

$$\langle A \cdot v, v \rangle \ > \ 0$$

*for all vectors $0 \neq v \in \mathbb{C}^n$.*

**Example.** If $A$ is an invertible $n \times n$-matrix then we claim that two matrices

$$A^* \cdot A, \quad A \cdot A^*$$

are hermitian and positive definite. Indeed, we proved before that for arbitrary matrices $A, B$ we have $(AB)^* = B^* A^*$. Therefore, these matrices $A^* \cdot A$, $A \cdot A^*$ are hermitian because for instance

$$(A^*A)^* = (A^*)(A^*)^* = A^*A$$

Furthermore, Since $A$ and $A^*$ are invertible we have $A \cdot v \neq 0$ and $A^* \cdot v \neq 0$ for all $v \in V \setminus \{0\}$. Hence by axiom **(P)** we have $\langle Av, Av \rangle > 0$ and $\langle A^*v, A^*v \rangle > 0$ for all $v \neq 0$ and so

$$0 \ < \ \langle Av, Av \rangle \ = \ \langle v, (A^* \cdot A)v \rangle \ = \ \langle (A^* \cdot A)v, v \rangle \,,$$

and by analogous reasoning $\langle (A \cdot A^*)v, v \rangle > 0$.

**Lemma.** *Let $A$ be a hermitian $n \times n$-matrices. Then*

(i) *$A$ is positive definite if and only if all eigenvalues of $A$ are $> 0$.*

(ii) *If the matrix $A$ is positive definite then $A$ is invertible and the inverse $A^{-1}$ is also positive definite.*

*Proof.* **Part (i):** Since $A$ is hermitian we know that all eigenvalues of $A$ are real numbers. Assume first that $A$ is positive definite and let $\lambda$ be an eigenvalue of $A$ with the corresponding eigenvector $v \in \mathbb{C}^n$. Thus $Av = \lambda\, v$. Since $A$ is positive definite we then get

$$0 < \ \langle Av, v \rangle \ = \ \langle \lambda v, v \rangle \ = \ \lambda \cdot \langle v, v \rangle \,,$$

and so $\lambda > 0$.

Conversely, let $v_1, \ldots, v_n$ be an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of $A$, say $A \cdot v_i = \lambda_i \cdot v_i$ for $i = 1, \ldots, n$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $A$. Assume that all of them are positive.

Let $v \in \mathbb{C}^n \setminus \{0\}$. We have to show that $\langle Av, v\rangle > 0$. We can write $v = \sum_{i=1}^{n} a_i \cdot v_i$ for some complex numbers $a_1, \ldots, a_n$. Then we have

$$\begin{aligned}
\langle A \cdot v, v\rangle &= \langle \sum_{i=1}^{n} a_i \cdot A \cdot v_i\,,\ \sum_{j=1}^{n} a_j \cdot v_j\rangle \\
&= \langle \sum_{i=1}^{n} (a_i\lambda_i) \cdot v_i\,,\ \sum_{i=1}^{n} a_j \cdot v_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} (a_i \cdot \lambda_j \cdot \bar{a}_j) \cdot \langle v_i, v_j\rangle \qquad \text{by } \textbf{(L1)}, \textbf{(L2)} \\
&= \sum_{i=1} \lambda_i \cdot |a_i|^2 \cdot \|v_i\|^2 \qquad \text{since } v_i \perp v_j \text{ for } i \neq j \\
&= \sum_{i=1}^{n} \lambda_i \cdot |a_i|^2 \qquad \text{since } \|v_i\|^2 = 1.
\end{aligned}$$

By assumption all $\lambda_i > 0$ and so we get

$$\langle A \cdot v, v\rangle \;=\; \sum_{i=1}^{n} \lambda_i \cdot |a_i|^2 \;> 0\,.$$

Therefore $\langle A \cdot v, v\rangle > 0$ for all $v \neq 0$ in $\mathbb{C}^n$, and so the matrix $A$ is positive definite.

**To prove (ii)** we note first that by the spectral theorem there exists an unitary matrix $U$, such that

$$U^* \cdot A \cdot U \;=\; \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \;=:\; D\,,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $A$. These eigenvalues are all $> 0$ by (i), and so the diagonal matrix $D$ is invertible. Hence also

$$A = (U^*)^{-1} D U^{-1} = U \cdot D \cdot U^*$$

is invertible (we used here that $U^*U = \mathrm{Id}_n$ which implies $(U^*)^{-1} = U$).

To show that the inverse $A^{-1}$ is positive definite we first observe that the above equality $A = U \cdot D \cdot U^*$ implies

$$A^{-1} = (U^*)^{-1} D^{-1} U^* = U D^{-1} U^*.$$

So $A^{-1}$ is also hermitian. It remains to note that the eigenvalues of $A^{-1}$ are $\lambda_i^{-1}$, $i = 1, \ldots, n$. $\square$

**Corollary.** *Let $A$ be a complex $n \times n$-matrix. If $A$ is unitary and positive definite then $A = \mathrm{I}_n$.*

*Proof.* We proved before that all eigenvalues $\lambda_1, \dots, \lambda_n$ of $A$ have absolute value 1, and by the lemma above they are real numbers $> 0$. Therefore all eigenvalues of $A$ are equal to 1. Also, we proved before that there is a unitary matrix $U$, such that $U^* A U$ is a diagonal matrix whose diagonal entries are the eigenvalues of $A$. Thus

$$U^* \cdot A \cdot U \;=\; \mathrm{Id}_n\,.$$

But this implies that $U^* \cdot A = U^{-1} = U^*$, and so $A = \mathrm{I}_n$ as claimed. $\square$

**Theorem.** *Let $A$ be a positive definite $n \times n$-matrix. Then there exists a unique positive definite $n \times n$-matrix $B$, such that*

$$A \;=\; B^2\,.$$

*Proof.* **Existence.** By the Spectral theorem there exists an unitary matrix $U$ whose columns are the orthonormal basis consisting of eigenvectors $v_1, \dots, v_n$ of $A$, such that

$$U^* \cdot A \cdot U \;=\; \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $A$ with the corresponding eigenvectors $v_1, \dots, v_n$. Since the eigenvalues of $A$ are all $> 0$ (by the lemma above) they all have a positive square root $\sqrt{\lambda_i}$, and we have

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{pmatrix}^2.$$

Setting

$$D \;:=\; \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{pmatrix}$$

and $B := U \cdot D \cdot U^*$ one can easily checks that

$$B^2 \;=\; A\,.$$

The matrix $B$ is hermitian since

$$B^* = (U \cdot D \cdot U^*)^* = (U^*)^* \cdot D^* \cdot U^* = U \cdot D \cdot U^* = B\,.$$

It remains to note that $B$ is also positive definite since all its eigenvalues are $\sqrt{\lambda_i}$ are positive real numbers.

**Uniqueness. Step 1.** It is enough to show that if $E_\lambda(A) \subset \mathbb{C}^n$ is the eigenspace of $A$ corresponding to an eigenvalue $\lambda \in \{\lambda_1, \ldots, \lambda_n\}$, and $C$ is a positive definite matrix, such that $C^2 = A$ then $C \cdot v = \sqrt{\lambda} \cdot v$ for all $v \in E_\lambda(A)$ (because any vector in $\mathbb{C}^n$ can be written uniquely as a sum of vectors from $E_{\lambda_i}$ and hence any matrix can be restored if we know its action on vectors from $E_{\lambda_i}$).

**Step 2** (diagonalizability). To show this we first observe that from the existence of Jordan normal forms it follows that the Jordan normal form of $C^2$ has the same number of Jordan blocks as the Jordan normal form of $C$. Since $C^2 = A$ is diagonalizable this implies that $C$ is also diagonalizable. Moreover, all eigenvalues of $C$ are positive real numbers (because $C$ is positive definite.

**Step 3** (counting dimensions of eigenspaces). Let $\rho > 0$ be an eigenvalue of $C$ and $v$ a corresponding eigenvector. Then we have

$$A \cdot v = C \cdot (C \cdot v) = C \cdot (\rho \cdot v) = \rho^2 \cdot v\,,$$

and so $\rho^2 \in \{\lambda_1, \ldots, \lambda_n\}$, *i.e.* since $\rho$ is positive we have $\rho = \sqrt{\lambda_i}$ for some $1 \le i \le n$, and moreover the eigenvector $v$ of $C$ is in the eigenspace of $A$ corresponding to the eigenvector $\lambda_i$.

This implies that if

$$\mu_1, \ldots, \mu_l \in \{\lambda_1, \lambda_2, \ldots . \lambda_n\}$$

are all different eigenvalues of $A$ with the corresponding eigenspaces $E_{\mu_j}(A)$, $1 \le i \le l$, and

$$\rho_1, \rho_2, \ldots, \rho_k$$

are the different eigenvalues of $C$ with corresponding eigenspaces $E_{\rho_j}(C)$, $1 \le j \le k$ then

(i) $\{\rho_1, \ldots, \rho_k\} \subseteq \{\sqrt{\mu_1}, \ldots, \sqrt{\mu_l}\}$, and

(ii) if $\rho_i = \sqrt{\mu_j}$ for some $1 \le i \le k$ and $1 \le j \le l$ then $E_{\rho_i}(C) \subseteq E_{\mu_j}(A)$.

Since both $A$ and $C$ are diagonalizable we have

$$n = \sum_{i=1}^{k} \dim_{\mathbb{C}} E_{\rho_i}(C) = \sum_{j=1}^{l} \dim_{\mathbb{C}} E_{\mu_j}(A)\,,$$

and so (after reordering if necessary) we have $l = k$, $\rho_j = \sqrt{\mu_j}$, and $E_{\rho_j}(C) = E_{\mu_j}(A)$ for all $1 \leq j \leq l$. But this implies that if $\lambda$ is an eigenvalue of $A$ then $C \cdot v = \sqrt{\lambda} \cdot v$ for all $v \in E_\lambda(A)$ as claimed. We are done. □

We are in position to prove the **polar decomposition** for invertible matrices.

**Corollary.** *Let $A$ be an invertible $n \times n$-matrix. Then there exists a positive definite matrix $P$ and a unitary matrix $U$ (both $n \times n$-matrices), such that*

$$A = P \cdot U .$$

*This decomposition is unique, i.e. if $P'$ and $U'$ are other positive definite and unitary matrices, respectively, such that $A = P' \cdot U'$, then $P = P'$ and $U = U'$.*

*Proof.* **Existence.** As observed before the matrix $A \cdot A^*$ is positive definite. By the theorem above there exists a unique positive definite matrix $P$, such that $P^2 = A \cdot A^*$. Set $U = P^{-1} \cdot A$. Then $A = P \cdot U$.

We claim that $U$ is unitary. Indeed, since $P$ is hermitian, so is $P^{-1}$. Therefore we have

$$(P^{-1} \cdot A) \cdot (P^{-1} \cdot A)^* = P^{-1} \cdot (A \cdot A^*) \cdot P^{-1} = P^{-1} \cdot P^2 \cdot P^{-1} = \mathrm{I}_n ,$$

i.e. $U = P^{-1} \cdot A$ is unitary.

**Uniqueness.** Let $P, Q$ be positive definite matrices and $U, V$ unitary matrices, such that

$$P \cdot U = A = Q \cdot V .$$

Then

$$A \cdot A^* = (P \cdot U) \cdot (P \cdot U)^* = P \cdot (U \cdot U^*) \cdot P = P^2$$

and by the analogous computation also $A \cdot A^* = Q^2$. But all matrices $P, Q$, and $A \cdot A^*$ are positive definite and so by uniqueness assertion of the above Theorem we have $P = Q$ which in turn implies $U = V$. We are done. □