Keywords: Eigenvectors, Eigenvalues, algebraic multiplicity, geometric multiplicity, Characteristic Polynomial, Similarity, Diagonalization, Matrix Representation, Complex Eigenvalues.

Defintion 1. Let A be an $n \times n$ matrix. If there exists a (real) scalar λ and a non-zero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$, then λ is called an **eigenvalue** of A and v is called an **eigenvector** of A corresponding to λ.

What is the set E_λ of eigenvectors (and zero vector)? We have $E_\lambda = \text{Null}(A - \lambda I)$ because

$$\begin{aligned} v \in E_\lambda \quad &\Leftrightarrow \quad Av = \lambda v \quad \Leftrightarrow \quad Av = \lambda I v \\ &\Leftrightarrow \quad Av - \lambda I v = 0 \quad \Leftrightarrow \quad (A - \lambda I)v = 0 \\ &\Leftrightarrow \quad v \in \text{Null}(A - \lambda I) \end{aligned}$$

We call E_λ the **eigenspace** of A for λ. The dimension of the eigenspace E_λ is called the **geometric multiplicty** (geo. mul.) of λ.

Theorem 2. The eigenvalues of a triangular matrix are the diagonal entries.

Theorem 3. Let $v_1, \dots, v_r$ be eigenvectors of pair-wise distinct eigenvalues $\lambda_1, \dots, \lambda_r$. Then

$$\{v_1, \dots, v_r\}$$

is a linearly independent set.

Defintion 4. The polynomial

$$\det(A - \lambda I)$$

in variable λ is called the **characteristic polynomial** of A. If λ is a root of the characteristic polynomial of A, then λ is an eigenvalue of A. The multiplicity as a root is called the **algebraic multiplicity** (alg. mul.).

Defintion 5. A is **similar** to B if there is an invertible matrix P such that $A = PBP^{-1}$. If A is similar to B, B is also similar to A.

Theorem 6. If A and B are similar, they have the same characteristic polynomial, hence the same eigenvalues with the same multiplicities.

⚠Two matrices with the same eigenvalues do not have to be similar. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Defintion 7. A $n \times n$ matrix A is **diagonalizable** if A is similar to a diagonal matrix, i.e. $A = PDP^{-1}$ for some invertible matrix P and a diagonal matrix D.

Theorem 8. Let A be $n \times n$ matrix.

A is diagonalizable $\Leftrightarrow$ A has n L.I. eigenvectors

Steps to Diagonalization.

(i) Find the eigenvalue of A.
(ii) Find basis for each eigenspaces.
(iii) Construct P from the vectors in (ii).
(iv) Construct D from the corresponding eigenvalues.

Theorem 10. A $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. This follows from Theorem 3 and Theorem 8.

Theorem 11. Let A be $n \times n$ matrix with distinct eigenvalues $\lambda_1, \dots, \lambda_p$.

(a) geo. mul. of $\lambda_k \le$ alg. mul. of λ_k for $1 \le k \le p$.
(b) A diagonalizable $\Leftrightarrow$ sum of geo. mul. equals n $\Leftrightarrow$ alg. mul. of $\lambda_k =$ geo. mul. of λ_k for all $1 \le k \le p$.
(c) A diagonalizable and $\mathscr{B}_k$ is a basis for E_{λ_k}, then $\mathscr{B}_1 \cup \cdots \cup \mathscr{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$. (1)

The eigenvector and eigenspace of linear transformation is defined the same way from $T(v) = \lambda v$.

Let $T: V \to V$ be a linear transformation. Let $\mathscr{B} = \{b_1, \dots, b_n\}$ be a basis of an n-dim. vector space V. Define the **matrix representation** of T with respect to $\mathscr{B}$ by

$$[T]_{\mathscr{B}} = \begin{bmatrix} [T(b_1)]_{\mathscr{B}} & \cdots & [T(b_n)]_{\mathscr{B}} \end{bmatrix}$$

Then for any $x \in V$, we have

$$[T(x)]_{\mathscr{B}} = [T]_{\mathscr{B}}[x]_{\mathscr{B}}$$

Theorem 9. Let P be the matrix whose columns are given by a basis $\mathscr{B}$. Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation given by $T(x) = Ax$. Then $[T]_{\mathscr{B}} = P^{-1}AP$. In particular, $A = P[T]_{\mathscr{B}}P^{-1}$.

All theory developed so far works well to $\mathbb{C}^n$. Namely, we say that λ and v is a **complex eigenvalue** and a **complex eigenvector** of an $n \times n$ matrix A if there exists $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$ such that $Av = \lambda v$.

⚠(This might not be covered during class) For a $n \times n$ matrix A, if λ is an eigenvalue of A with an eigenvector v of λ. Then $\overline{\lambda}$ is an eigenvector for the eigenvalue $\overline{\lambda}$ where $\overline{\bullet}$ denotes complex conjugation.

Theorem 12. Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi \, (b \neq 0)$ and an asssociated eigenvector $v \in \mathbb{C}^2$. Then

$$A = PCP^{-1} \text{ with } P = \begin{bmatrix} \text{Re } v & \text{Im } v \end{bmatrix} \text{ and } C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

(1) $\cup$ denotes set union. The union $\mathscr{B}_1 \cup \cdots \cup \mathscr{B}_p$ is a new set that contains all elements of $\mathscr{B}_k$ for $1 \le k \le p$.

Keywords: inner product, dot product, length of a vector, distance between two vectors, orthogonality, orthogonal complement, orthogonal set/basis, orthogonal matrix, orthogonal projection, Gram-Schmidt, QR factorization.

Defintion 13. For $u, v \in \mathbb{R}^n$, the **dot product** (or the **inner product**) of u and v is $u^T v$ and is written $u \cdot v$. If $u = [u_1, \dots, u_n]^T$ and $v = [v_1, \dots, v_n]^T$, then $u \cdot v = u_1 v_1 + \cdots + u_n v_n$.

Defintion 14. The **length** $||v||$ of a vector is defined by $\sqrt{v \cdot v} = \sqrt{v_1^2 + \cdots + v_n^2}$. In particular, $||v||^2 = v \cdot v$. For a scalar $c \in \mathbb{R}$, we have $||cv|| = |c|||v||$. If $||v|| = 1$, v is called a **unit vector**.

Defintion 15. The **distance** between u and v is defined by $\text{dist}(u, v) = ||u - v|| = ||v - u||$.

Defintion 16. Two vectors u and v are **orthogonal** if $u \cdot v = 0$. We sometimes denote it by $u \perp v$.

Theorem 17 (Pythagorean)**.** If $u \perp v$, then

$$||u + v||^2 = ||u||^2 + ||v||^2$$

Defintion 18. For a subspace $W \subset \mathbb{R}^n$, a vector v is **orthogonal** to W if for all $w \in W$, $v \perp w$. The set of all vectos v that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^{\perp}$. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

$$(\text{Row } A)^{\perp} = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

Defintion 19. The set $\{u_1, \dots, u_p\}$ of vectors in $\mathbb{R}^n$ is **orthogonal set** if every pair of distinct vectors are orthogonal. An **orthogonal basis** is a orthogonal set that is also a basis.

Theorem 20. Let $\{u_1, \dots, u_p\}$ be a orthogonal set with all u_i nonzero vectors, then it is linearly independent.

Defintion 21. Let $W \subset \mathbb{R}^n$ be a subspace with orthogonal basis $\{w_1, \dots, w_p\}$ and $y \in \mathbb{R}^n$. The **projection** $\text{Proj}_W y$ of y onto W is defined by

$$\text{Proj}_W y = \frac{y \cdot w_1}{w_1 \cdots w_1} w_1 + \cdots + \frac{y \cdot w_p}{w_p \cdot w_p} w_p.$$

Theorem 22. Let $W \subset \mathbb{R}^n$ be a subspace of $\mathbb{R}^n$. Let $y \in \mathbb{R}^n$. Then y can be uniquely written as

$$y = \hat{y} + z$$

where $\hat{y} \in W$ and $z \in W^{\perp}$. In fact, $\hat{y} = \text{Proj}_W y = UU^T y$ where U is the matrix whose columns are a orthonormal basis of W. Furthermore,

(a) $y \in W$ if and only if $y = \text{Proj}_W y$.
(b) $\hat{y}$ is the closest point to y in W in the sense that $||y - \hat{y}|| < ||y - w||$ for all $w \in W$.

Theorem 23. A matrix U is **orthogonal** (i.e. $U^T U = I$) if and only if the columns of U form an orthonormal basis of $\mathbb{R}^n$. If U is square, U orthogonal if and only if $U^T = U^{-1}$.

Theorem 24 (Gram-Schmidt)**.** Let $\{x_1, \dots, x_p\}$ be a basis for a nonzero subspace W of $\mathbb{R}^n$. Then we can construct an orthogonal basis $\{u_1, \dots, u_p\}$ via

$$\begin{aligned} u_1 &= x_1 \\ u_2 &= x_2 - \tfrac{x_2 \cdot u_1}{u_1 \cdot u_1} u_1 \\ &\vdots \\ u_p &= x_p - \tfrac{x_p \cdot u_1}{u_1 \cdot u_1} u_1 - \cdots - \tfrac{x_p \cdot u_{p-1}}{u_{p-1} \cdot u_{p-1}} u_{p-1} \end{aligned}$$

and $\text{Span}\{x_1, \dots, x_k\} = \text{Span}\{u_1, \dots, u_k\}$ for $1 \le k \le p$. In addition, one can obtain orthonormal basis via normalization, i.e. $\left\{\frac{u_1}{||u_1||}, \dots, \frac{u_p}{||u_p||}\right\}$.

Theorem 25 (QR Factorization)**.** Let A be an $m \times n$ matrix with linearly independent columns. Then $A = QR$ where $Q = [u_1 \ \cdots \ u_n]$ is an $m \times n$ orthogonal matrix for some orthonormal basis $\{u_1, \dots, u_n\}$ for Col A, and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal with $R = Q^T A$.

⚠ If one chooses an arbitrary orthonormal basis of Col A, $Q^T A$ may not be have positive diagonal entry. If the kth diagonal entry r_{kk} of R is negative, we can replace both r_{kk} and u_k by $-r_{kk}$ and $-u_k$ respectively.

Defintion 26. For $m \times n$ A and $b \in \mathbb{R}^m$, a **least-squares solution** of $Ax = b$ is $\hat{x} \in \mathbb{R}^n$ such that $||b - A\hat{x}|| \le ||b - Ax||$ for all $x \in \mathbb{R}^n$.

To find $\hat{x}$, we solve the *normal equation* for $Ax = b$, $A^T Ax = A^T b$ which is always consistent. When $A^T A$ is invertible (this is not always the case), we have

$$\hat{x} = (A^T A)^{-1} A^T b.$$

Least-squares solution of $Ax = b$ may not be unique. However, it is unique in the following situation.

Theorem 27. Let A be an $m \times n$ matrix with linearly independent columns. Then we have a QR factorization $A = QR$. Then for each $b \in \mathbb{R}^m$, the equation $Ax = b$ has a unique least-square solution,

$$\hat{x} = R^{-1} Q^T b.$$