Chapter 2. Matrix Algebra

Keywords: Matrix multiplication, transpose of a matrix A^T, inverse matrix A^{-1}

(i) $A\begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & \cdots & Ab_n \end{bmatrix}$	(ii) $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ for 2×2 A with $ad-bc \neq 0$	(iii) $\begin{bmatrix} A \mid I \end{bmatrix} \to \begin{bmatrix} I \mid A^{-1} \end{bmatrix}$ for any square matrix A
(iv) $(AB)^T = B^T A^T$ for any matrices A, B	(v) $(AB)^{-1} = B^{-1}A^{-1}$ for invertible matrices A, B	(vi) $(A^{-1})^T = (A^T)^{-1}$ for an invertible matrix A

Theorem 1. An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n.

Theorem 2. (Invertible Matrix Theorem) For an $n \times n$ matrix A, (a)-(l) are all equivalent
(i) A has n pivot (columns) ⇔ (d) ⇔ (e) ⇔ (f) (ii) A has n pivot (rows) ⇔ (g) ⇔ (h) ⇔ (i)

Chapter 3. Determinants

Keywords: Determinants, Cofactor Expansion across a row or a column, relationship between row operations and determinants, Cramer's Rule, Areas and volumes as determinants.

Defintion 3. Let A be an $n \times n$-matrix.

(a) The submatrix A_{ij} is an $(n-1)\times(n-1)$-matrix obtained from A by deleting ith row and jth column.

(b) **determinant** of A is recursively defined as

$$a_{11}\det A_{11} - a_{12}\det A_{12} + \cdots + (-1)^{1+n} a_{1n}\det A_{1n}$$

Defintion 4. The (i,j)-cofactor of A is

$$C_{ij} = (-1)^{i+j}\det A_{ij}$$

cofactor expansion across row i

$$\det A = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$$

cofactor expansion down column j

$$\det A = a_{1j}C_{1j} + \cdots + a_{nj}C_{nj}$$

Theorem 5. If A is a triangular matrix, then $\det A$ is the product of the entries on the diagonal of A.

Theorem 6. A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 7. $\det A = \det A^T$ and

$$\det(AB) = (\det A)(\det B)$$

(a) $\det B = \det A$ if B is obtained by adding a multiple of another row.

(b) If B is obtained by interchanging two rows, then $\det B = -\det A$.

(c) If B is obtained by multipying k to a row, $\det B = k\det A$.

Cramer's Rule Let A be invertible $n \times n$-matrix and $b \in \mathbb{R}^n$. Then the ith entry x_i of the unique solution is given by

$$x_i = \frac{\det A_i(b)}{\det A}$$

where $A_i(b) = \begin{bmatrix} a_1 & \cdots & b & \cdots & a_n \end{bmatrix}$.

Let A be an invertible $n \times n$ matrix. Then [1]

$$\begin{aligned} A^{-1} &= \frac{1}{\det A}\operatorname{adj} A \\ &= \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \end{aligned}$$

Determinant and Volumes

(a) (Parallelogram) Let $v_1, v_2 \in \mathbb{R}^2$. Then the area of the parallelogram formed by v_1 and v_2 is $\det A$ where $A = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$.

(b) (Parallelopiped) Let $v_1, v_2, v_3 \in \mathbb{R}^3$. Then the volume of the parallelopiped formed by v_1, v_2, v_3 is $\det A$ where $A = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with standard matrix A. If S is a region in $\mathbb{R}^2$ with finite area. Then

$$\text{area of } T(S) = |\det A| \cdot \text{area of } S$$

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation with standard matrix A. If S is a region in $\mathbb{R}^3$ with finite volume. Then

$$\text{volume of } T(S) = |\det A| \cdot \text{volume of } S$$

[1] The (i,j)-entry of A^{-1} is C_{ji} divided by $\det A$, NOT C_{ij}.

Keywords: vector space, subspace, basis, null, column, and row spaces, dimension, coordinate vectors, change-of-coordinate matrix, rank-nullity theorem.

Defintion 8. A **vector space** is a nonempty set V of objects called **vectors** with two operations *addition* and *mutlipliaction by scalars* (real numbers) with ten properties in pg. 202-203.

Defintion 9. A **subspace** of a vector space V is a subset W of V such that (i) $0 \in W$, (ii) is closed under addition, and (iii) is closed under scalar multiplication.

Defintion 10. A **linear transformation** $T : V \to W$ is a function such that $T(u+v) = T(u) + T(v)$ and $T(cu) = cT(u)$.

Example 11. The set $\mathbb{R}^n$ of column vectors with n entries and the set $\mathbb{P}_n$ of polynomials of degree at most n are vector spaces.(2) Also, a subspace of a vector space is a vector space.

Example 12. Let $v_1, \ldots, v_p \in V$. Then the **span**

$$\text{Span}\{v_1, \ldots, v_p\} = \{c_1 v_1 + \cdots + c_p v_p \mid c_i \in \mathbb{R}\}$$

of $\{v_1, \ldots, v_n\}$ is a subspace of V. This is called the **subspace** of V generated by $\{v_1, \ldots, v_p\}$.

(2) The addition and scalar multiplications of $\mathbb{R}^n$ and $\mathbb{P}_n$ are defined differently.

Defintion 13. A set of vectors $\{v_1, \ldots, v_p\}$ in V is **linearly independent** if the **linear dependence relation** $c_1 v_1 + \cdots + c_p v_p = 0$ has only trivial solution.

Theorem 14. $\{v_1, \ldots, v_p\}$ with $v_1 \neq 0$ is linearly dependent if and only if some v_j $(j > 1)$ is a linear combination of $v_1, \ldots, v_{j-1}$.

Defintion 15. A indexed set $\mathscr{B}$ of vectors in a vector space V is called a **basis** if (i) $\mathscr{B}$ is linearly independent and (ii) $\text{Span}\,\mathscr{B} = V$.

Defintion 16. If a vector space V is spanned by a finite set, then V is said to be **finite-dimensional** and the number of vectors in a basis is called a **dimension** of V. (3)

(3) A basis of a vector space is not unique, but they all have the same number of vectors by Theorem 18.

Theorem 17. (Spanning Set Theorem) Let $S = \{v_1, \ldots, v_p\}$ be a subset of V and let $W = \text{Span}\,S$. Then (i) if $v_k \in S$ is a linear combination of the remaining vectors in S, then the set $\{v_1, \ldots, \cancel{v_k}, \ldots, v_p\}$ formed by removing v_k from S still spans W and (ii) if $S \neq \{0\}$, then a subset of S is a basis of W.

Theorem 18. (a) If a vector space V has a basis $\mathscr{B}$ with n vectors, then any set in V containing more than n vectors must be linearly dependent. Also, every basis of V must consist of exactly n vectors. (b) Every vector can be written uniquely as a linear combinations of vectors in $\mathscr{B}$.

Theorem 19. Let W be a subspace of a finite-dimensional vector space. Then $\dim W \leq \dim V$.

Theorem 20. For $\dim V = n$, (i) any linearly independent set of V with n-elements or (ii) any spanning set of V with n-elements is a basis.

Let A be a $m \times n$-matrix. Write $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ where a_is are the column vectors of A. Also, let $r_1, \ldots, r_m$ be its row vectors. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation defined by A.

Subspace	A basis(4)	Dimension
$\text{Nul}\,A = \{x \in \mathbb{R}^n \mid Ax = 0\}$ $= \text{Ker}(T)$ subspace of $\mathbb{R}^n$	$\mathscr{B}$ is the set of vectors appearing in the general solution in parametric vector form	$\text{nullity}\,A = \dim \text{Nul}\,A$
$\text{Col}\,A = \text{Span}\,\{a_1, \ldots, a_n\}$ $= \text{Range}(T)$ subspace of $\mathbb{R}^m$	$\mathscr{B} = \{$ pivot columns of A $\}$	$\text{rank}\,A = \dim \text{Col}\,A$
$\text{Row}\,A = \text{Span}\,\{r_1, \ldots, r_n\}$ is a subspace of $\mathbb{R}^n$	$\mathscr{B}$ is the set of nonzero row vectors of an echelon form B of A	$\text{Row}\,A = \text{Row}\,B$ $A \to B$ row. eq.

Theorem 21 (Rank-Nullity).

$$\text{rank}\,A + \text{nullity}\,A = \#\text{ of cols. of } A$$

Defintion 22. The **standard basis** is the set $\{e_1, \ldots, e_n\}$ in $\mathbb{R}^n$ where e_i is the vector whose entries are all zero except 1 at the ith entry.

Defintion 23. $\mathscr{B} = \{b_1, \ldots, b_n\}$ be a basis.

$$[x]_{\mathscr{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

given $x = c_1 b_1 + \cdots + c_n b_n$, is called **coordinate vector** of x relative to $\mathscr{B}$.

Defintion 24. $\mathscr{B} = \{b_1, \ldots, b_n\}$, $\mathscr{C} = \{c_1, \ldots, c_n\}$ be bases of V.

$$\underset{\mathscr{C} \leftarrow \mathscr{B}}{P} = \begin{bmatrix} [b_1]_{\mathscr{C}} & \cdots & [b_n]_{\mathscr{C}} \end{bmatrix}$$

is called the **change-of-coordinates matrix** from $\mathscr{B}$ to $\mathscr{C}$. Also,

$$\left[c_1 \cdots c_n \middle| b_1 \cdots b_n \right] \to \left[I_n \middle| \underset{\mathscr{C} \leftarrow \mathscr{B}}{P} \right]$$

(4) There are infinitely many basis to a vector space. This is just one of them.

CHAPTER 5. EIGENVALUES AND EIGENVECTORS

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 $v_1, \ldots, v_r$ be eigenvectors of pair-wise distinct eigenvalues $\lambda_1, \ldots, \lambda_r$. Then

$$\{v_1, \ldots, v_r\}$$

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 $\lambda_1, \ldots, \lambda_p$.

(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 λ_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 $\mathscr{B} = \{b_1, \ldots, b_n\}$ 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 $u = \left[u_1, \ldots, u_n\right]^T$ and $v = \left[v_1, \ldots, v_n\right]^T$, 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 $\{u_1, \ldots, u_p\}$ 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 $\{u_1, \ldots, u_p\}$ 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 $\{w_1, \ldots, w_p\}$ 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 $\{x_1, \ldots, x_p\}$ be a basis for a nonzero subspace W of $\mathbb{R}^n$. Then we can construct an orthogonal basis $\{u_1, \ldots, u_p\}$ 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 $\text{Span}\{x_1, \ldots, x_k\} = \text{Span}\{u_1, \ldots, u_k\}$ for $1 \le k \le p$. In addition, one can obtain orthonormal basis via normalization, i.e. $\left\{\frac{u_1}{||u_1||}, \ldots, \frac{u_p}{||u_p||}\right\}$.

Theorem 25 (QR Factorization)**.** Let A be an $m \times n$ matrix with linearly independent columns. Then $A = QR$ where $Q = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$ is an $m \times n$ orthogonal matrix for some orthonormal basis $\{u_1, \ldots, u_n\}$ 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,

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

Chapter 6.6 "Least-squares Line"

Given an experimental data

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

we want to find a line $y = \beta_0 + \beta_1 x$ that best fits the data. In particular, we want β_0 and β_1 such that

$$\begin{array}{rcl} \beta_0 + \beta_1 x_1 & = & y_1 \\ \vdots & & \vdots \\ \beta_0 + \beta_1 x_n & = & y_n \end{array}$$

This is same as trying to solve the linear system

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{=A} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}}_{=\beta} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{=b}$$

In many real life applications, $A\beta = b$ is inconsistent. The least-squares solution $\begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix}^T$ defines a line $y = \beta_0 + \beta_1 x$ which we call **least-squares line** that best fits the data point $(x_1, y_1), \dots, (x_n, y_n)$. To recall, $\begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix}^T$ is the solution to the equation $A^T A\beta = A^T b$.

Chapter 7.1 Diagonalization of Symmetric Matrices

Defintion 1. A matrix A is **symmetric** if $A = A^T$. Equivalently, the matrix has arbitrary entries along the main diagonal, and its entries are symmetric with respect to the main diagonal.

Defintion 2. A is **orthogonally diagonalizable** if there exists an orthogonal matrix P ($P^{-1} = P^D$) and a diagonal matrix D such that $A = PDP^T$.

Theorem 3 (Spectral Theorem)**.** Let A be an symmetric $n \times n$-matrix. Then

(a) A has n real eigenvalues, counting multiplicities.

(b) The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation. (i.e. diagonalizable)

(c) The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

Furthermore, an $n \times n$-matrix A is symmetric if and only if A is orthogonally diagonalizable.

How do we orthogonally diagonalize an $n \times n$-matrix A? You can do this when you can find an orthonormal basis consisting $\{u_1, \dots, u_n\}$ of eigenvectors of A (not always possible). Let

$$Q = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$$

which is an orthogonal matrix. Then

$$A = QDQ^T$$

where D is the diagonal matrix with eigenvalues corresponding to $\{u_1, \dots, u_n\}$. What is **amazing** about the spectral theorem is that it says that for a symmetric matrix A, you can always find an orthonormal basis of eigenvectors of A.