Matrices

Definition 2.2.19 Let $\boldsymbol{A} = \left(a_{ij}\right)$ be $m \times n$. Then the transpose of $\boldsymbol{A}$, $\boldsymbol{A}^{\mathrm{T}} = \left(a_{ji}\right)$ is $n \times m$.

Remark 2.2.21 Let $\boldsymbol{A} = \left(a_{ij}\right)$. It is symmetric if $a_{ij} = a_{ji}$ for all i, j.

Theorem 2.2.22 Let $\boldsymbol{A}$ be $m \times n$. Let c be a scalar.

1. $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.
2. If $\boldsymbol{B}$ is $m \times n$, then $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.
3. $(c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}$.
4. If $\boldsymbol{B}$ is $n \times p$, then $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

Definition 2.3.2 Let $\boldsymbol{A}$ be $n \times n$. It is invertible there exists a $n \times n$ $\boldsymbol{B}$ such that $\boldsymbol{AB} = \boldsymbol{I}$ and $\boldsymbol{BA} = \boldsymbol{I}$. By Theorem 2.3.5, $\boldsymbol{B}$ is uniquely defined by $\boldsymbol{A}$. By Theorem 2.4.12, we only need to verify either one of $\boldsymbol{AB} = \boldsymbol{I}$ or $\boldsymbol{BA} = \boldsymbol{I}$.

Theorem 2.3.9 Let $\boldsymbol{A}, \boldsymbol{B}$ be two invertible matrices of the same size. Let c be a scalar.

1. $c\boldsymbol{A}$ is invertible, $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.
2. $\boldsymbol{A}^{\mathrm{T}}$ is invertible, $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$.
3. $\boldsymbol{A}^{-1}$ is invertible, $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
4. $\boldsymbol{AB}$ is invertible, $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

Remark Let $\boldsymbol{A}, \boldsymbol{B}, \cdots, \boldsymbol{Z}$ be invertible matrices. Then

$$(\boldsymbol{AB}\cdots\boldsymbol{Z})^{-1} = \boldsymbol{Z}^{-1}\cdots\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

Definition 2.4.3 A square matrix is an elementary matrix if it can be obtained from $\boldsymbol{I}$ with a single ero. Elementary matrices are invertible and their inverses are also elementary matrices.

Determinants

Theorem 2.4.14 Let $\boldsymbol{A}, \boldsymbol{B}$ be two square matrices of the same order. If $\boldsymbol{A}$ is singular, then $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are singular.

Definition 2.5.2 The determinant of a $n \times n$ square matrix $\boldsymbol{A}$ is defined as:

$$\det(\boldsymbol{A}) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & \text{if } n > 1 \end{cases}$$

where

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

where $\boldsymbol{M_{ij}}$ is a $(n-1) \times (n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the ith row and jth column. The scalar value A_{ij} is called the (i,j)-cofactor of $\boldsymbol{A}$.

Theorem 2.5.8 If $\boldsymbol{A}$ is triangular, then $\det(\boldsymbol{A})$ is the product of diagonal entries along $\boldsymbol{A}$.

Theorem 2.5.10 If $\boldsymbol{A}$ is a square matrix, then $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^{\mathrm{T}})$.

Theorem 2.5.15 Let $\boldsymbol{A}, \boldsymbol{B}$ be square matrices of the same order.

1. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by multiplying one row of $\boldsymbol{A}$ by a constant k, then $\det(\boldsymbol{B}) = k\det(\boldsymbol{A})$.
2. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by interchanging two rows of $\boldsymbol{A}$, then $\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.
3. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by adding a multiple of one row of $\boldsymbol{A}$ to another row, then $\det(\boldsymbol{B}) = \det(\boldsymbol{A})$.
4. Let $\boldsymbol{E}$ be an elementary matrix of the same size as $\boldsymbol{A}$. Then $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$.

Remark 2.5.18 By Theorem 2.5.10, Theorem 2.5.15 holds for eco.

Theorem 2.5.22 Let $\boldsymbol{A}, \boldsymbol{B}$ be square matrices of order n, and c a scalar. Then

1. $\det(c\boldsymbol{A}) = c^n\det(\boldsymbol{A})$.
2. $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.
3. If $\boldsymbol{A}$ invertible, then $\det(\boldsymbol{A}^{-1}) = \dfrac{1}{\det(\boldsymbol{A})}$.

Definition 2.5.24 The adjoint of a square matrix $\boldsymbol{A}$ is defined as:

$$\operatorname{adj}(\boldsymbol{A}) = (A_{ij})^{\mathrm{T}}_{n\times n}$$

where A_{ij} is the (i,j)-cofactor of $\boldsymbol{A}$.

Theorem 2.5.25 If $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})}\operatorname{adj}(\boldsymbol{A})$.

Theorem 2.5.27 (Cramer's Rule) Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is a linear system where $\boldsymbol{A}$ is $n \times n$. Let $\boldsymbol{A_i}$ be the matrix obtained from $\boldsymbol{A}$, by replacing the ith column of $\boldsymbol{A}$ by $\boldsymbol{b}$. If $\boldsymbol{A}$ is invertible, then the system has only one solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{pmatrix} \det(\boldsymbol{A_1}) \\ \det(\boldsymbol{A_2}) \\ \vdots \\ \det(\boldsymbol{A_n}) \end{pmatrix}$$

Vector Spaces

Definition 3.2.3 Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ be a set of vectors in $\mathbb{R}^n$. Then the set of all linear combinations of $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}$,

$$\{c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} \mid c_1, c_2, \cdots, c_k \in \mathbb{R}\}$$

is the linear span of S, and is denoted by $\operatorname{span}(S)$.

Theorem 3.2.10 Let $S_1 = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ and $S_2 = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_m}\}$. Then $\operatorname{span}(S_1) \subseteq \operatorname{span}(S_2) \Leftrightarrow$ each $\boldsymbol{u_i}$ is a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_m}$. We verify by ensuring the following augmented matrix is consistent:

$$\left(\begin{array}{cccc|c|c|c|c} \boldsymbol{v_1} & \boldsymbol{v_2} & \cdots & \boldsymbol{v_m} & \boldsymbol{u_1} & \boldsymbol{u_2} & \cdots & \boldsymbol{u_k} \end{array}\right)$$

Definition 3.3.1 Let V be a subset of $\mathbb{R}^n$. Then V is a subspace of $\mathbb{R}^n$ if $V = \operatorname{span}(S)$, where $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ for some vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k} \in \mathbb{R}^n$.

Remark 3.3.8 A subspace is alternatively defined as a non-empty subset of $\mathbb{R}^n$ that is closed under vector addition and scalar multiplication.

Definition 3.4.2 Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ be a set of vectors in $\mathbb{R}^n$. Consider the equation

$$c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0}$$

where $c_1, c_2, \cdots, c_k$ are variables. Then

1. S is a *linearly independent set* and $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}$ are said to be *linearly independent* if the above equation has only the trivial solution.

2. S is a *linearly dependent set* and $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}$ are said to be *linearly dependent* if the above equation has non-trivial solutions.

Definition 3.5.4 Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ be a subset of a vector space V. Then S is a basis for V if S is linearly independent and S spans V.

Theorem 3.6.1 Let V be a vector space which has a basis with k vectors. Then

1. any subset of V with more than k vectors is always linearly dependent;
2. any subset of V with less than k vectors cannot span V.

Definition 3.6.3 The dimension of a vector space V, denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V. In addition, we define the dimension of the zero space to be zero.

Theorem 3.6.7 Let V be a vector space of dimension k and $S \subseteq V$. TFAE:

1. S is a basis for V.
2. S is linearly independent and $|S| = k$.
3. S spans V and $|S| = k$.

Theorem 3.6.9 Let U be a subspace of a vector space V.

1. $\dim(U) \leq \dim(V)$.
2. If $U \neq V$, then $\dim(U) < \dim(V)$.
3. If $\dim(U) = \dim(V)$, then $U = V$.

Notation Let $S = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k}\}$ so that $V = \text{span}(S)$ and $T = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \cdots, \boldsymbol{w_k}\}$ so that $W = \text{span}(T)$. Then

$$V + W = \text{span}(S \cup T)$$

Exercise 3.43 Let V and W be subspaces of $\mathbb{R}^n$. Then

$$\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W)$$

Algorithm Let S be a linearly independent set, consisting of vectors from $\mathbb{R}^n$. Let $|S| < N$. To extend a basis S to $\mathbb{R}^n$,

1. Form a matrix $\boldsymbol{A}$ using the vectors in S as rows
2. Reduce $\boldsymbol{A}$ to a row-echelon form $\boldsymbol{R}$
3. Identify non-pivot columns
4. For each non-pivot column, pick a vector from the standard basis of $\mathbb{R}^n$ such that the '1' is exactly at the position of the non-pivot column
5. $S \cup$ (vectors obtained in Step 4) is a basis for $\mathbb{R}^n$.

Transition Matrices

Definition 3.5.8 Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ be a basis for a vector space V, and let $\boldsymbol{v} \in V$. By Theorem 3.5.7, $\boldsymbol{v}$ is expressed uniquely as a linear combination

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}$$

and $c_1, c_2, \cdots, c_k$ are the coordinates of $\boldsymbol{v}$ relative to the basis S.

$$(\boldsymbol{v})_S = (c_1, c_2, \cdots, c_k) \in \mathbb{R}^k$$

is the coordinate vector of $\boldsymbol{v}$ relative to the basis S.

Remark 3.5.10 Let S be a basis for a vector space V.

1. For any $\boldsymbol{u}, \boldsymbol{v} \in V$, $\boldsymbol{u} = \boldsymbol{v} \Leftrightarrow (\boldsymbol{u})_S = (\boldsymbol{v})_S$.
2. For any $\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_r} \in V$ and $c_1, c_2, \cdots, c_r \in \mathbb{R}$,

$$\begin{aligned}&(c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_r\boldsymbol{v_r})_S \\ &= c_1(\boldsymbol{v_1})_S + c_2(\boldsymbol{v_2})_S + \cdots + c_r(\boldsymbol{v_r})_S\end{aligned}$$

Notation 3.7.1 Sometimes, it is more conveient to write the coordinate vector in the form of a column vector. Thus we define

$$[\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$

and it is also the coordinate vector of $\boldsymbol{v}$ relative to S. Note the difference in notation from Definition 3.5.8.

Discussion 3.7.2 (Excerpt) Let $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ and $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k}\}$ be two bases for a vector space V. Since they are both bases, then we can write each $\boldsymbol{u_i}$ as a linear combination of $\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k}$, i.e.

$$\begin{aligned}\boldsymbol{u_1} &= a_{11}\boldsymbol{v_1} + a_{21}\boldsymbol{v_2} + \cdots + a_{k1}\boldsymbol{v_k} \\ \boldsymbol{u_2} &= a_{12}\boldsymbol{v_1} + a_{22}\boldsymbol{v_2} + \cdots + a_{k2}\boldsymbol{v_k} \\ &\vdots \\ \boldsymbol{u_k} &= a_{1k}\boldsymbol{v_1} + a_{2k}\boldsymbol{v_2} + \cdots + a_{kk}\boldsymbol{v_k}\end{aligned}$$

Then

$$\begin{aligned}\boldsymbol{P} &= \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \vdots & \vdots & & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{kk} \end{pmatrix} \\ &= \begin{pmatrix} [\boldsymbol{u_1}]_T & [\boldsymbol{u_2}]_T & \cdots & [\boldsymbol{u_k}]_T \end{pmatrix}\end{aligned}$$

is the transition matrix from S to T, and for every $\boldsymbol{w} \in V$,

$$[\boldsymbol{w}]_T = \boldsymbol{P}[\boldsymbol{w}]_S$$

Remark Alternatively, we can do the following to find $\boldsymbol{P}$.

$$\begin{aligned}&\left(\begin{array}{cccc|c|c|c|c} \boldsymbol{v_1} & \boldsymbol{v_2} & \cdots & \boldsymbol{v_k} & \boldsymbol{u_1} & \boldsymbol{u_2} & \cdots & \boldsymbol{u_k} \end{array}\right) \\ \xrightarrow{\text{GJE}} &\left(\begin{array}{ccc|ccc} & \boldsymbol{I} & & & \boldsymbol{P} & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right)\end{aligned}$$

There may or may not be zero rows at the bottom of the augmented matrix after GJE.

1. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}, \boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k} \in \mathbb{R}^m$ where $m > k$, then there are zero rows. Just take the square matrix bounded to the right by the augmented line and the number of columns.
2. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}, \boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k} \in \mathbb{R}^k$ then there are no zero rows.

Vector Spaces of Matrices

Definition 4.1.2 Let $\boldsymbol{A} = \begin{pmatrix} a_{ij} \end{pmatrix}$ be $m \times n$. The row space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\boldsymbol{A}$. The column space of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^m$ spanned by the columns of $\boldsymbol{A}$.

Theorem 4.1.17 Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices. Then the row space of $\boldsymbol{A}$ and the row space of $\boldsymbol{B}$ are identical, i.e. elementary row operations preserve the row space of a matrix.

Theorem 4.1.11 Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices. Then

1. A given set of columns of $\boldsymbol{A}$ is linearly independent if and only if the set of corresponding columns of $\boldsymbol{B}$ is linearly independent.
2. A given set of columns of $\boldsymbol{A}$ forms a basis for the column space of $\boldsymbol{A}$ if and only if the set of corresponding columns of $\boldsymbol{B}$ forms a basis for the column space of $\boldsymbol{B}$.

Theorem 4.1.16 Let $\boldsymbol{A}$ be $m \times n$. Then

$$\text{the column space of } \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{u} \mid \boldsymbol{u} \in \mathbb{R}^n\}$$

Hence a system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}$ lies in the column space of $\boldsymbol{A}$.

Theorem 4.2.1 The row space and column space of a matrix have the same dimension.

Definition 4.2.3 The rank of a matrix is the dimension of its row space (or column space). We denote the rank of a matrix $\boldsymbol{A}$ by $\operatorname{rank}(\boldsymbol{A})$. Note that $\operatorname{rank}(\boldsymbol{A})$ is equal to the number of nonzero rows as well as the number of pivot columns in a row-echelon form of $\boldsymbol{A}$.

Remark 4.2.5

1. For a $m \times n$ matrix $\boldsymbol{A}$, $\operatorname{rank}(\boldsymbol{A}) \leq \min\{m, n\}$. If equal, then $\boldsymbol{A}$ has full rank.
2. A square matrix $\boldsymbol{A}$ has full rank if and only if $\det \boldsymbol{A} \neq 0$.
3. $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^{\mathrm{T}})$ because the row space of $\boldsymbol{A}$ is the column space of $\boldsymbol{A}^{\mathrm{T}}$.

Theorem 4.2.8 Let $\boldsymbol{A}$ be $m \times n$, and $\boldsymbol{B}$ be $n \times p$. Then

$$\operatorname{rank}(\boldsymbol{A}\boldsymbol{B}) \leq \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}$$

Definition 4.3.1 Let $\boldsymbol{A}$ be $m \times n$. The solution space of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is the null space of $\boldsymbol{A}$.

Theorem 4.3.4 Let $\boldsymbol{A}$ be a matrix with n columns. Then

$$\operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) = n$$

Theorem 4.3.6 Suppose $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution $\boldsymbol{v}$. Then the solution set of the system is given by

$$M = \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \text{ is an element of the null space of } \boldsymbol{A}\}$$

Exercise 4.22 Let $\boldsymbol{A}$ be $m \times n$ and $\boldsymbol{P}$ be an invertible matrix of order m. Then $\operatorname{rank}(\boldsymbol{P}\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$.

Exercise 4.23 Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two matrices of the same size. Then

$$\operatorname{rank}(\boldsymbol{A} + \boldsymbol{B}) \leq \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B})$$

Exercise 4.25 Let $\boldsymbol{A}$ be $m \times n$.

1. null space of $\boldsymbol{A}$ = null space of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$
2. $\operatorname{nullity}(\boldsymbol{A}) = \operatorname{nullity}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$
3. $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$

Orthogonality

Definition 5.1.2 Let $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \cdots, v_n)$ be two vectors in $\mathbb{R}^n$.

1. The dot product (or inner product) of $\boldsymbol{u}$ and $\boldsymbol{v}$ is
$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$
2. The norm (or length) of $\boldsymbol{u}$ is
$$||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$
Vectors of norm 1 are called unit vectors.
3. The distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is
$$\begin{aligned} d(\boldsymbol{u}, \boldsymbol{v}) &= ||\boldsymbol{u} - \boldsymbol{v}|| \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2} \end{aligned}$$
4. The angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is
$$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||\,||\boldsymbol{v}||}\right)$$
The angle is well defined because $-1 \leq \dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||\,||\boldsymbol{v}||} \leq 1$.

Remark 5.1.3 Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$. Then $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$.

Definition 5.2.1

1. Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$ are orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
2. A set S of vectors in $\mathbb{R}^n$ is called orthogonal if every pair of distinct vectors in S are orthogonal.
3. A set S of vectors in $\mathbb{R}^n$ is called orthonormal if S is orthogonal and every vector in S is a unit vector.

Theorem 5.2.4 Let S be an orthogonal set of nonzero vectors in a vector space. Then S is linearly independent.

Definition 5.2.5

1. A basis S for a vector space is called an orthogonal basis if S is orthogonal.
2. A basis S for a vector space is called an orthonormal basis if S is orthonormal.

Theorem 5.2.8

1. If $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ is an orthogonal basis for a vector space V, then for any $\boldsymbol{w}$ in V,
$$\boldsymbol{w} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\boldsymbol{u_k}$$
i.e. $(\boldsymbol{w})_S = \left(\dfrac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}, \dfrac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}, \cdots, \dfrac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\right)$.
2. If $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k}\}$ is an orthonormal basis for a vector space V, then for any vector $\boldsymbol{w}$ in V,
$$\boldsymbol{w} = (\boldsymbol{w} \cdot \boldsymbol{v_1})\boldsymbol{v_1} + (\boldsymbol{w} \cdot \boldsymbol{v_2})\boldsymbol{v_2} + \cdots + (\boldsymbol{w} \cdot \boldsymbol{v_k})\boldsymbol{v_k}$$
i.e. $(\boldsymbol{w})_T = (\boldsymbol{w} \cdot \boldsymbol{v_1}, \boldsymbol{w} \cdot \boldsymbol{v_2}, \cdots, \boldsymbol{w} \cdot \boldsymbol{v_k})$.

Remark Declare two orthonormal bases S and T, with $S = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ and $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \cdots, \boldsymbol{v_k}\}$. By Discussion 3.7.2,

$$\boldsymbol{P} = \begin{pmatrix} [\boldsymbol{u_1}]_T & [\boldsymbol{u_2}]_T & \cdots & [\boldsymbol{u_k}]_T \end{pmatrix}$$

is the transition matrix from S to T. By Theorem 5.2.8.2, we can write $\boldsymbol{P}$ in the following manner:

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{u_1} \cdot \boldsymbol{v_1} & \boldsymbol{u_2} \cdot \boldsymbol{v_1} & \cdots & \boldsymbol{u_k} \cdot \boldsymbol{v_1} \\ \boldsymbol{u_1} \cdot \boldsymbol{v_2} & \boldsymbol{u_2} \cdot \boldsymbol{v_2} & \cdots & \boldsymbol{u_k} \cdot \boldsymbol{v_2} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{u_1} \cdot \boldsymbol{v_k} & \boldsymbol{u_2} \cdot \boldsymbol{v_k} & \cdots & \boldsymbol{u_k} \cdot \boldsymbol{v_k} \end{pmatrix} = C_S^{\mathrm{T}} C_T$$

where C_S, C_T are matrices whose columns are vectors from S, T respectively. Also note that $\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}^{-1}$ so the transition matrix from T to S can easily be found.

Definition 5.2.10 Let V be a subspace of $\mathbb{R}^n$. A vector $\boldsymbol{u} \in \mathbb{R}^n$ is orthogonal to V if $\boldsymbol{u}$ is orthogonal to all vectors in V.

Definition 5.2.13 Let V be a subspace of $\mathbb{R}^n$. Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written uniquely as $\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$ such that $\boldsymbol{n}$ is orthogonal to V, and $\boldsymbol{p} \in V$. The vector $\boldsymbol{p}$ is the (orthogonal) projection of $\boldsymbol{u}$ onto V.

Theorem 5.2.15 Let V be a subspace of $\mathbb{R}^n$, and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$. If $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ is an orthogonal basis for V, then

$$\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \frac{\boldsymbol{w} \cdot \boldsymbol{u_2}}{\boldsymbol{u_2} \cdot \boldsymbol{u_2}}\boldsymbol{u_2} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\boldsymbol{u_k}$$

is the projection of $\boldsymbol{w}$ onto V.

Theorem 5.2.19 Let $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ be a basis for a vector space V. Let

$$\begin{aligned}
\boldsymbol{v_1} &= \boldsymbol{u_1} \\
\boldsymbol{v_2} &= \boldsymbol{u_2} - \frac{\boldsymbol{u_2} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1} \\
&\vdots \\
\boldsymbol{v_k} &= \boldsymbol{u_k} - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1} - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_2}}{\boldsymbol{v_2} \cdot \boldsymbol{v_2}}\boldsymbol{v_2} - \cdots - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_{k-1}}}{\boldsymbol{v_{k-1}} \cdot \boldsymbol{v_{k-1}}}\boldsymbol{v_{k-1}}
\end{aligned}$$

Then $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k}\}$ is an orthogonal basis for V. We can normalize the vectors if we want an orthonormal basis.

Theorem 5.3.2 Let V be a subspace in $\mathbb{R}^n$. If $\boldsymbol{u}$ is a vector in $\mathbb{R}^n$ and $\boldsymbol{p}$ is the projection of $\boldsymbol{u}$ onto V, then

$$d(\boldsymbol{u}, \boldsymbol{p}) \leq d(\boldsymbol{u}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in V$$

i.e. $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.

Theorem 5.3.8 Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system where $\boldsymbol{A}$ is $m \times n$. Let $\boldsymbol{p}$ be the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$. Then

$$||\boldsymbol{b} - \boldsymbol{p}|| \leq ||\boldsymbol{b} - \boldsymbol{Av}|| \quad \text{for all } \boldsymbol{v} \in V$$

i.e. $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{Au} = \boldsymbol{p}$.

Theorem 5.3.10 Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system. Then $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{u}$ is a solution to $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.

Definition 5.4.3 A square matrix $\boldsymbol{A}$ is orthogonal if $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$.

Theorem 5.4.6 Let $\boldsymbol{A}$ be a square matrix of order n. TFAE:

1. $\boldsymbol{A}$ is orthogonal.
2. The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$.
3. The columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$.

Exercise 5.7 Let W be a subspace of $\mathbb{R}^n$. Define $W^{\perp} = \{\boldsymbol{u} \in \mathbb{R}^n \mid \boldsymbol{u} \text{ is orthogonal to } W\}$. Then $W^{\perp}$ is a subspace of $\mathbb{R}^n$. From HW3, W and $W^{\perp}$ are disjoint and $W + W^{\perp} = \mathbb{R}^n$.

Eigenvalues and Eigenvectors

Definition 6.1.3 Let $\boldsymbol{A}$ be a square matrix of order n. Let $\boldsymbol{u} \in \mathbb{R}^n$ be a non-zero column vector. If $\boldsymbol{Au} = \lambda\boldsymbol{u}$ for some scalar λ, then $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue λ.

Definition 6.1.6 Let $\boldsymbol{A}$ be a square matrix of order n. The equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ is the characteristic equation of $\boldsymbol{A}$. The polynomial $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ is the characteristic polynomial of $\boldsymbol{A}$.

Theorem 6.1.8 Let $\boldsymbol{A}$ be $n \times n$. TFAE:

1. $\boldsymbol{A}$ is invertible.
2. $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution.
3. RREF of $\boldsymbol{A}$ is the identity matrix.
4. $\boldsymbol{A}$ can be expressed as a product of elementary matrices.
5. $\det(\boldsymbol{A}) \neq 0$.
6. The rows of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
7. The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
8. $\operatorname{rank}(\boldsymbol{A}) = n$ (i.e. $\operatorname{nullity}(\boldsymbol{A}) = 0$).
9. 0 is not an eigenvalue of $\boldsymbol{A}$.

Theorem 6.1.9 If $\boldsymbol{A}$ is triangular, the eigenvalues of $\boldsymbol{A}$ are the diagonal entries of $\boldsymbol{A}$.

Definition 6.1.11 Let $\boldsymbol{A}$ be a square matrix of order n and λ an eigenvalue of $\boldsymbol{A}$. Then the solution space of $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is the eigenspace of $\boldsymbol{A}$ associated with the eigenvalue λ and it is denoted by E_{λ}.

Definition 6.2.1 A square matrix $\boldsymbol{A}$ is diagonalizable if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix.

Theorem 6.2.3 Let $\boldsymbol{A}$ be a square matrix of order n. Then $\boldsymbol{A}$ is diagonalizable if and only if $\boldsymbol{A}$ has n linearly independent eigenvectors.

Remark 6.2.5.2 Suppose the characteristic polynomial of the matrix $\boldsymbol{A}$ can be factorized as

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of $\boldsymbol{A}$. For each eigenvalue λ_i,

$$\dim(E_{\lambda_i}) \leq r_i$$

and $\boldsymbol{A}$ is diagonalizable if and only if for each i, $\dim(E_{\lambda_i}) = r_i$. Note that r_i is called the multiplicity of eigenvalue λ_i.

Theorem 6.2.7 Let $\boldsymbol{A}$ be a square matrix of order n. If $\boldsymbol{A}$ has n distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable.

Example 6.2.11.2 Let $a_n = a_{n-1} + a_{n-2}$, where $a_0 = 0, a_1 = 1$. Then

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$$

Let $\boldsymbol{x_n} = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Then

$$\boldsymbol{x_n} = \boldsymbol{Ax_{n-1}} = \boldsymbol{A}^2\boldsymbol{x_{n-2}} = \cdots = \boldsymbol{A}^n\boldsymbol{x_0} = \boldsymbol{A}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Then diagonalize $\boldsymbol{A}$ to obtained closed form for $\boldsymbol{x_n}$ and thus a_n.

Definition 6.3.2 A square matrix $\boldsymbol{A}$ is orthogonally diagonalizable if there exists an orthogonal matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ is a diagonal matrix.

Theorem 6.3.4 A square matrix is orthogonally diagonalizable if and only if it is symmetric.

Linear Transformations

Definition 7.1.1 A linear transformation is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix}$$

where a_{ij} are scalars. In particular, if $n = m$, T is also called a linear operator on $\mathbb{R}^n$. We can rewrite the formula of T as

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The matrix $(a_{ij})_{m\times n}$ above is called the standard matrix for T.

Remark If we can express $T(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$ for all $\boldsymbol{v} \in \mathbb{R}^n$, then it is a linear transformation.

Theorem 7.1.4 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

1. $T(\boldsymbol{0}) = \boldsymbol{0}$.
2. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \cdots, \boldsymbol{u_k} \in \mathbb{R}^n$ and $c_1, c_2, \cdots, c_k \in \mathbb{R}$, then

$$\begin{aligned} &T(c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}) \\ &= c_1T(\boldsymbol{u_1}) + c_2T(\boldsymbol{u_2}) + \cdots + c_kT(\boldsymbol{u_k}) \end{aligned}$$

Definition 7.1.10 Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be linear transformations. The composition of T with S, denoted by $T \circ S$, is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^k$ such that

$$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) \quad \text{for } \boldsymbol{u} \in \mathbb{R}^n$$

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be the standard matrices for S and T. Then the standard matrix for $T \circ S$ is $\boldsymbol{BA}$.

Definition 7.2.1 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The range of T is the set of images of T, i.e.

$$\mathrm{R}(T) = \{T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Let $\boldsymbol{A}$ be the standard matrix for T. Then

$$\mathrm{R}(T) = \text{the column space of } \boldsymbol{A}$$

Definition 7.2.5 Let T be a linear transformation with standard matrix $\boldsymbol{A}$. The dimension of $\mathrm{R}(T)$ is the rank of T, and $\mathrm{rank}(T) = \mathrm{rank}(\boldsymbol{A})$.

Definition 7.2.7 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The kernel of T is the set of vectors whose image is the zero vector in $\mathbb{R}^m$, i.e.

$$\mathrm{Ker}(T) = \{\boldsymbol{u} \mid T(\boldsymbol{u}) = \boldsymbol{0}\} \subseteq \mathbb{R}^n$$

Let $\boldsymbol{A}$ be the standard matrix for T. Then

$$\mathrm{Ker}(T) = \text{the null space of } \boldsymbol{A}$$

Definition 7.2.10 Let T be a linear transformation with standard matrix $\boldsymbol{A}$. The dimension of $\mathrm{Ker}(T)$ is the nullity of T, and $\mathrm{nullity}(T) = \mathrm{nullity}(\boldsymbol{A})$.

Theorem 7.2.12 If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = n$$

Remark Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation. Let $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ be vectors in $\mathbb{R}^3$. Given

$$T(\boldsymbol{u_1}) = \boldsymbol{v_1} \quad T(\boldsymbol{u_2}) = \boldsymbol{v_2} \quad T(\boldsymbol{u_3}) = \boldsymbol{v_3}$$

If $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ form a basis for $\mathbb{R}^3$, then we can find $\boldsymbol{A}$:

$$\begin{aligned} \boldsymbol{A}\begin{pmatrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} \end{pmatrix} \\ \boldsymbol{A} &= \begin{pmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} & \boldsymbol{v_3} \end{pmatrix}\begin{pmatrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{pmatrix}^{-1} \end{aligned}$$

Otherwise, there is insufficient information to determine $\boldsymbol{A}$. A similar approach works for $\mathbb{R}^n$.

Miscellaneous

Remark $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ is consistent. Let $\boldsymbol{A}$ be $m \times n$.

1. By Theorem 4.1.16, it is consistent if $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \in$ column space of $\boldsymbol{A}$.
2. By rank-nullity, $\mathrm{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = n - \mathrm{nullity}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.
3. By Exercise 4.25, $\mathrm{nullity}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \mathrm{nullity}(\boldsymbol{A})$.
4. By rank-nullity, $\mathrm{nullity}(\boldsymbol{A}) = n - \mathrm{rank}(\boldsymbol{A})$.
5. By Remark 4.2.5, $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^{\mathrm{T}})$.
6. Combining 2-5, $\mathrm{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^{\mathrm{T}})$.
7. Let $\boldsymbol{v} \in$ column space of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. Then $\boldsymbol{v} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{w} \in$ column space of $\boldsymbol{A}^{\mathrm{T}}$. Thus
$$\text{column space of } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} \subseteq \text{column space of } \boldsymbol{A}^{\mathrm{T}}$$
8. Combining 6 and 7,
$$\text{column space of } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \text{column space of } \boldsymbol{A}^{\mathrm{T}}$$
9. Since $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \in$ column space of $\boldsymbol{A}^{\mathrm{T}}$, then also $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \in$ column space of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

Theorem Let $\boldsymbol{u}, \boldsymbol{v}$ be eigenvectors belonging to different eigenspaces of a square matrix $\boldsymbol{A}$.

1. $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly independent.
2. If $\boldsymbol{A}$ is symmetric, then $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v}$.

Proof of 1. (Exercise 6.22)

1. Let $\boldsymbol{u}, \boldsymbol{v}$ be associated with distinct eigenvalues λ, μ respectively.
2. Assume otherwise that $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly dependent.
3. Then $\boldsymbol{v} = k\boldsymbol{u}$ for some scalar k.
4. Then $\boldsymbol{A}\boldsymbol{v} = \mu\boldsymbol{v} = k\mu\boldsymbol{u} = k\frac{\mu}{\lambda}\boldsymbol{A}\boldsymbol{u} = \frac{\mu}{\lambda}\boldsymbol{A}\boldsymbol{v}$.
5. Then $\frac{\mu}{\lambda} - 1 = 0$, i.e. $\lambda = \mu$.
6. Hence contradiction with line 2.

Proof of 2. (Exercise 6.26)

1. Let $\boldsymbol{u}, \boldsymbol{v}$ be associated with distinct eigenvalues λ, μ respectively.
2. Then we have $\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$ and $\boldsymbol{A}\boldsymbol{v} = \mu\boldsymbol{v}$.
3. Then, $\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}\boldsymbol{u}\cdot\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{u})^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u}\cdot\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{v}$.
4. $\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{u} = \boldsymbol{v}\cdot(\lambda\boldsymbol{u}) = \lambda(\boldsymbol{v}\cdot\boldsymbol{u}) = \lambda(\boldsymbol{u}\cdot\boldsymbol{v})$.
5. $\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{v} = \boldsymbol{u}\cdot(\mu\boldsymbol{v}) = \mu(\boldsymbol{u}\cdot\boldsymbol{v})$.
6. Combining 3-5, $\lambda(\boldsymbol{u}\cdot\boldsymbol{v}) = \mu(\boldsymbol{u}\cdot\boldsymbol{v})$.
7. Then, $(\lambda - \mu)(\boldsymbol{u}\cdot\boldsymbol{v}) = 0$
8. Since $\lambda \neq \mu$, then $\boldsymbol{u}\cdot\boldsymbol{v} = 0$.

Remark By Remark 5.1.3, we can regard dot product as matrix multiplication.