# MA2001 LINEAR ALGEBRA

## Review

National University of Singapore
Department of Mathematics

## Linear Systems and Their Solutions

- A linear system of $m$ equations and $n$ variables:
  - $$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

  A solution to the linear system is
  - $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$

  which satisfies every equation in the system.
- In practice, we shall always find a **general solution**:
  - all expression that gives **all** the solutions.
- A linear system can have
  - no solution (the system is **inconsistent**);
  - exactly one solution; or
  - infinitely many solutions.



## Elementary Row Operations

- A linear system can be viewed as an **augmented matrix**
  - $$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$
- The three **elementary row operations**:
  - $cR_i$: multiply the $i$th row by a nonzero constant $c$.
  - $R_i \leftrightarrow R_j$: interchange the $i$th row and the $j$th row.
  - $R_i + cR_j$: add $c$ times the $j$th row to the $i$th row.
- If one (augmented) matrix can be obtained from another by a **series of elementary row operations**,
  - they are said to be **row equivalent**, and
  - the corresponding linear systems have the same solution set.

## Row-Echelon Forms

- A matrix is of the **row-echelon form** if
  - Zero rows are place below the nonzero rows; and
  - For any two nonzero rows,
    - the **leading entry** (first nonzero entry) of the higher row is on the left that of the lower row.
- In a matrix of row-echelon form,
  - the leading entry of a nonzero row is a **pivot point**,
  - a column containing a pivot point is a **pivot column**.
- A matrix in row-echelon form is also in **reduced row-echelon form** if in addition,
  - Every leading entry of a nonzero row is $1$;
  - In each pivot column, except the pivot point, all other entries are $0$.



## Row-Echelon Forms

- Given a linear system. Suppose its augmented matrix is in (reduced) row-echelon form.
  1. Set the variables corresponding to non-pivot columns as arbitrary parameters.
  2. Solve the variables corresponding to pivot columns in terms of these arbitrary parameters backwards.
- Criterion to check number of solutions:
  - Last column is pivot
    
    $\Rightarrow$ No solution.
  - Only the last column is non-pivot
    
    $\Rightarrow$ Exactly one solution.
  - Last column and some other columns are non-pivot
    
    $\Rightarrow$ Infinitely many solutions.

## Gaussian Elimination

- **Gaussian Elimination**:
  - Use elementary row operations to get a row-echelon form of a given matrix.
- **Gauss-Jordan Elimination**:
  - Use elementary row operations to get the reduced row-echelon form of a given matrix.
- **Remarks.**
  - To check consistency: Use Gaussian elimination.
  - To solve the system: Use Gauss-Jordan elimination.
  - Discuss case by case if parameters are involved.
  - Using $cR_i$: we must ensure that $c \neq 0$.
  - Using $\frac{1}{c}R_i$ or $R_i + \frac{1}{c}R_j$: we must have $c \neq 0$.



## Homogeneous Linear Systems

- A linear system is called **homogeneous** if it is of the form
  - $$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases}$$
  - Augmented matrix: $\left(\boldsymbol{A} \mid \boldsymbol{0}\right)$.
- **Properties.** Consider a homogeneous linear system.
  - It has the **trivial solution**; so it is consistent.
  - If there are more variables than equations, then the system has infinitely many solutions.
  - The solution set is a vector space (nullspace of $\boldsymbol{A}$).

## Introduction to Matrices

- An $m \times n$ **matrix** is an array of numbers:
  - $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, $(i, j)$-entry is $a_{ij}$.
- **Special Matrices**: Let $\boldsymbol{A} = (a_{ij})_{m\times n}$.
  - Row matrix (vector): $m = 1$.
  - Column matrix (vector): $n = 1$.
  - Square matrix: $m = n$.
  - Upper triangular matrix: $m = n$ and $a_{ij} = 0$ for $i > j$.
  - Diagonal matrix: $m = n$ and $a_{ij} = 0$ for $i \neq j$.
  - Symmetric matrix: $m = n$ and $a_{ij} = a_{ji}$.
  - Scalar matrix: $c\boldsymbol{I}_n$; Identity matrix: $\boldsymbol{I}_n$.
  - Zero matrix: $\mathbf{0}_{m\times n}$.



## Matrix Operations

- Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{m\times n}$.
  - $c\boldsymbol{A} = (ca_{ij})_{m\times n}$, $\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n}$.
  - $\boldsymbol{A}^{\mathrm{T}} = (a_{ji})_{n\times m}$.
- Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{n\times p}$.
  - Then $\boldsymbol{AB}$ is the $m \times p$ matrix whose $(i, j)$-entry is
    - $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

    In other words,

    1. Take the $i$th row of $\boldsymbol{A}$: $a_{i1}, a_{i2}, \ldots, a_{in}$.
    2. Take the $j$th column of $\boldsymbol{B}$: $b_{1j}, b_{2j}, \ldots, b_{nj}$.
    3. Multiply componentwise $a_{i1}b_{1j}, a_{i2}b_{2j}, \ldots, a_{in}b_{nj}$.
    4. Add the products $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.
- If $\boldsymbol{A}$ is a square matrix, $\boldsymbol{A}^m = \overbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}^{m}$, $m \in \mathbb{Z}^+$.

## Matrix Operations

- **Remarks**: In general,
  - $\boldsymbol{AB} \neq \boldsymbol{BA}$, $(\boldsymbol{AB})^2 \neq \boldsymbol{A}^2\boldsymbol{B}^2$.
  - $\boldsymbol{AB} = 0 \not\Rightarrow$ "$\boldsymbol{A} = 0$ or $\boldsymbol{B} = 0$".
  - $\boldsymbol{A}^2 = 0 \not\Rightarrow \boldsymbol{A} = 0$, $\boldsymbol{A}^2 = \boldsymbol{I} \not\Rightarrow \boldsymbol{A} = \pm\boldsymbol{I}$.
- Consider $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$
  - $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$.
  - Then the system is $\boldsymbol{Ax} = \boldsymbol{b}$.

  The system is consistent $\Leftrightarrow$ $\boldsymbol{b}$ lies in the column space of $\boldsymbol{A}$.



## Inverses of Square Matrices

- Let $\boldsymbol{A}$ be a square matrix. It is **invertible** if
  - there exists square matrix $\boldsymbol{B}$ s.t. $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$.

  $\boldsymbol{A}$ is **singular** if it is not invertible.
- Suppose that $\boldsymbol{A}$ is invertible.
  - $(\boldsymbol{A} \mid \boldsymbol{I}) \xrightarrow{\text{Gauss-Jordan elimination}} (\boldsymbol{I} \mid \boldsymbol{A}^{-1})$.
- **Properties.** Let $\boldsymbol{A}$ be invertible.
  - $\boldsymbol{Ax} = \boldsymbol{b} \Rightarrow \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.
  - $\boldsymbol{AB}_1 = \boldsymbol{AB}_2 \Rightarrow \boldsymbol{B}_1 = \boldsymbol{B}_2$.
  - $\boldsymbol{C}_1\boldsymbol{A} = \boldsymbol{C}_2\boldsymbol{A} \Rightarrow \boldsymbol{C}_1 = \boldsymbol{C}_2$.
  - $\boldsymbol{AB} = \boldsymbol{I} \Rightarrow \boldsymbol{B} = \boldsymbol{A}^{-1}$.
  - $\boldsymbol{A}^{-m} = (\boldsymbol{A}^{-1})^m = \overbrace{\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}}^{m}$, $m \in \mathbb{Z}^+$.

## Inverses of Square Matrices

- **Theorem for Invertible Matrices**. Let $\boldsymbol{A}$ be a square matrix of order $n$. The following statements are equivalent:
  - $\boldsymbol{A}$ is invertible.
  - $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution $\boldsymbol{x} = \boldsymbol{0}$.
  - $\boldsymbol{Ax} = \boldsymbol{b}$ has exactly one solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.
  - The reduced row-echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}_n$.
  - $\boldsymbol{A}$ is the product of elementary matrices.
  - $\det(\boldsymbol{A}) \neq 0$.
  - $0$ is not an eigenvalue of $\boldsymbol{A}$.
  - $\operatorname{rank}(\boldsymbol{A}) = n$.
  - $\operatorname{nullity}(\boldsymbol{A}) = 0$.
  - The rows of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
  - The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.



## Elementary Matrices

- An **elementary matrix** is obtained from $\boldsymbol{I}$ by applying a single elementary row (or column) operation.
  - $\boldsymbol{I} \xrightarrow{cR_i} \boldsymbol{E}$:
    - Replace the $i$th diagonal entry by $c$.
  - $\boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E}$:
    - Replace the $i$th and $j$th diagonal entries by $0$, and Replace the $(i, j)$- and $(j, i)$-entries by $1$.
  - $\boldsymbol{I} \xrightarrow{R_i + cR_j} \boldsymbol{E}$:
    - Replace the $(i, j)$-entry by $c$.
- **Theorem on the determinant of elementary matrices.**
  - $\det(\boldsymbol{E}) = c$ if $\boldsymbol{I} \xrightarrow{cR_i} \boldsymbol{E}$.
  - $\det(\boldsymbol{E}) = -1$ if $\boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E}$.
  - $\det(\boldsymbol{E}) = 1$ if $\boldsymbol{I} \xrightarrow{R_i + cR_j} \boldsymbol{E}$.

## Elementary Matrices

- **Theorem.** Let $\boldsymbol{A}$ be an $m \times n$ matrix.
  - $\boldsymbol{I}_m \xrightarrow{cR_i} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{EA}$;
    - $\det(\boldsymbol{EA}) = c\det(\boldsymbol{A})$.
  - $\boldsymbol{I}_m \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{EA}$.
    - $\det(\boldsymbol{EA}) = -\det(\boldsymbol{A})$.
  - $\boldsymbol{I}_m \xrightarrow{R_i + cR_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{EA}$.
    - $\det(\boldsymbol{EA}) = \det(\boldsymbol{A})$.
- **Theorem.** $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent $\Leftrightarrow$ there exist elementary matrices $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_k$ such that
  - $\boldsymbol{E}_k\boldsymbol{E}_{k-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} = \boldsymbol{B}$.
- **Theorem.** If $(\boldsymbol{A} \mid \boldsymbol{b})$ is row equivalent to $(\boldsymbol{C} \mid \boldsymbol{d})$, then
  - $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{Cx} = \boldsymbol{d}$ have the same solutions.

16 / 57

## Elementary Matrices

- **Theorem.** Let $\boldsymbol{A}$ be an $m \times n$ matrix.
  - $\boldsymbol{I}_n \xrightarrow{kC_i} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{kC_i} \boldsymbol{AE}$;
    - $\det(\boldsymbol{AE}) = k\det(\boldsymbol{A})$.
  - $\boldsymbol{I}_n \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{AE}$.
    - $\det(\boldsymbol{AE}) = -\det(\boldsymbol{A})$.
  - $\boldsymbol{I}_n \xrightarrow{C_i + kC_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{C_i + kC_j} \boldsymbol{AE}$.
    - $\det(\boldsymbol{AE}) = \det(\boldsymbol{A})$.
- **Find Inverse of Invertible Matrix**:
  - $(\boldsymbol{A} \mid \boldsymbol{I}) \xrightarrow{\text{Gauss-Jordan elimination}} (\boldsymbol{I} \mid \boldsymbol{A}^{-1})$.

  $\boldsymbol{A}$ is singular if its RREF is not $\boldsymbol{I}$.

17 / 57

## Determinants

- Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ be a square matrix.
  - If $n = 1$, $\det(\boldsymbol{A}) = a_{11}$.
  - If $n = 2$, $\det(\boldsymbol{A}) = a_{11}a_{22} - a_{12}a_{21}$.
  - For $n \geq 2$, expansion along the $i$th row:
    - $\det(\boldsymbol{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$.
  - Expansion along the $j$th column:
    - $\det(\boldsymbol{A}) = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$.

    $A_{ij} = (-1)^{i+j}\det(\boldsymbol{M}_{ij})$ is the $(i, j)$-cofactor of $\boldsymbol{A}$.
    - $\boldsymbol{M}_{ij}$ is the submatrix of $\boldsymbol{A}$ by deleting its $i$th row and $j$th column.
- Cofactor expansion is applicable if $n \geq 4$ and if most of the entries along a row (or a column) are zero.



## Determinants

- **Theorem.** Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices.
  - $\boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = c\det(\boldsymbol{A})$.
  - $\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.
  - $\boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = \det(\boldsymbol{A})$.
- **Theorem.** $\det(\boldsymbol{I}_n) = 1$.
  - If $\boldsymbol{A}$ is singular, then $\det(\boldsymbol{A}) = 0$.
  - If $\boldsymbol{A}$ is an upper (or lower) triangular matrix,
    - $\det(\boldsymbol{A})$ is the product of the diagonal entries.
  - $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^{\mathrm{T}})$, $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.
  - $\det(\boldsymbol{A}^{-1}) = 1/\det(\boldsymbol{A})$, if $\boldsymbol{A}$ is invertible.

### Determinants

- Let $\boldsymbol{A}$ be a square matrix of order $n$. The **adjoint matrix**
  - $\mathbf{adj}(\boldsymbol{A}) = (A_{ji})_{n\times n}$.

  Then $\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = \det(\boldsymbol{A})\boldsymbol{I}$.
  - If $\boldsymbol{A}$ is invertible, $\boldsymbol{A}^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\mathbf{adj}(\boldsymbol{A})$.
- **Cramer's Rule**. Let $\boldsymbol{A}$ be an invertible matrix.
  - $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$,
    - $x_j = \dfrac{\det(\boldsymbol{A}_j)}{\det(\boldsymbol{A})}$,

    $\boldsymbol{A}_j$ is obtained by replacing the $j$th column of $\boldsymbol{A}$ by $\boldsymbol{b}$.



## Chapter 3: Vector Spaces



### Euclidean $n$-Spaces

- $\mathbb{R}^n$ is the **Euclidean $n$-space**:
  - Every element in $\mathbb{R}^n$ is called an $n$-vector:
    - $\boldsymbol{v} = (a_1, a_2, \ldots, a_n)$.
  - $c(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$;
  - $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$.
- Implicit and explicit forms of lines and planes.
  - Lines in $\mathbb{R}^3$:
    - $\{(a_0 + at, b_0 + bt, c_0 + ct) \mid t \in \mathbb{R}\}$. implicit
    - Intersection of two planes. explicit.
  - Planes in $\mathbb{R}^3$:
    - $\{(x, y, z) \mid ax + by + cz = d\}$. implicit
    - $\{(s, t, \alpha s + \beta t + \gamma) \mid s, t \in \mathbb{R}\}$. explicit.

## Linear Combinations and Linear Spans

- Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ be vectors in $\mathbb{R}^n$.
  - $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k$ is a **linear combination**.
  - $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is the set of all linear combinations.
    - $\{c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$.
- Criterion to check if $\boldsymbol{v} \in \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$:
  - Consistency of $c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k = \boldsymbol{v}$.
  - Consider the linear system $\left(\begin{array}{ccc|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k & \boldsymbol{v} \end{array}\right)$.
- Criterion to check if $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = \mathbb{R}^n$:
  - A row-echelon form of $\begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$ has no zero row.
  - $\begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$ is of rank $n$ (full rank).
  - If $k < n$, then $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \neq \mathbb{R}^n$.



## Linear Combinations and Linear Spans

- Criterion to check if
  - $\operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\} \subseteq \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$.

  Check consistency of $\left(\begin{array}{ccc|c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r & \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k \end{array}\right)$.
- **Properties.**
  - Let $S_1$ and $S_2$ be finite subsets of $\mathbb{R}^n$.
    - $S_1 \subseteq S_2 \Rightarrow \operatorname{span}(S_1) \subseteq \operatorname{span}(S_2)$.
  - Let $V = \operatorname{span}(S)$, where $S$ is a finite subset of $\mathbb{R}^n$.
    - $\boldsymbol{0} \in V$.
    - "$c \in \mathbb{R}$ and $\boldsymbol{v} \in V$" $\Rightarrow c\boldsymbol{v} \in V$.
    - "$\boldsymbol{u} \in V$ and $\boldsymbol{v} \in V$" $\Rightarrow \boldsymbol{u} + \boldsymbol{v} \in V$.
  - If $\boldsymbol{v}_k$ is a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}$,
    - $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k\} = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}\}$.

## Subspaces

- A **subspace** of $\mathbb{R}^n$ is $V = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$, $\boldsymbol{v}_i \in \mathbb{R}^n$.
- Criterion to check if a subset $V \subseteq \mathbb{R}^n$ is a subspace.
  - If you guess "yes", try to find $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ such that
    - $V = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.
  - If you guess "no", try to find counterexamples:
    - $\mathbf{0} \notin V$; or
    - $c \in \mathbb{R}$ and $\boldsymbol{v} \in V$, but $c\boldsymbol{v} \notin V$; or
    - $\boldsymbol{u} \in V$ and $\boldsymbol{v} \in V$, but $\boldsymbol{u} + \boldsymbol{v} \in V$.
- **Examples.** $\{\mathbf{0}\}$ and $\mathbb{R}^n$.
  - Lines passing through the origin.
  - Planes containing the origin.
  - The solution set (solution space) of $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$.

25 / 57

## Linear Independency

- $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are said to be **linearly independent** if
  - $c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k = \mathbf{0} \Rightarrow c_1 = \cdots = c_k = 0.$

  Otherwise, they are **linearly dependent**.
- Criterion to check if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are linearly independent.
  - Solve $\left( \begin{array}{ccc|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k & \mathbf{0} \end{array} \right)$.
    - There is only the trivial solution $\Rightarrow$ linearly independent.
    - There is a non-trivial solution $\Rightarrow$ linearly dependent.
  - In particular, if $k > n$, they are linearly dependent.
- **Properties**.
  - If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are linearly dependent,
    - then one of them is a linear combination of others.
  - If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ are linearly independent, and $\boldsymbol{v}_{k+1} \in \mathbb{R}^n$ is not in $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$,
    - then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}$ are linearly independent.

26 / 57

## Bases

- A **vector space** $V$ is a subspace of $\mathbb{R}^n$.
  - A finite subset $S$ of $\mathbb{R}^n$ is a **basis** for $V$ if
    - $\mathrm{span}(S) = V$ and $S$ is linearly independent.
- **Theorem.** Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for $V$.
  - Every $\boldsymbol{v} \in V$ can be uniquely expressed as a linear combination of vectors in $S$:
    - $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$, where $c_i \in \mathbb{R}$.
- Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for $V$.
  - For each $\boldsymbol{v} \in V$, write
    - $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$, where $c_i \in \mathbb{R}$.

    Then $(\boldsymbol{v})_S = (c_1, \ldots, c_k) \in \mathbb{R}^k$ is called the **coordinate vector** of $\boldsymbol{v}$ relative to $S$.



## Bases

- **Properties.** Let $S$ be a basis for a vector space $V$.
  - $\boldsymbol{u} = \boldsymbol{v} \Leftrightarrow (\boldsymbol{u})_S = (\boldsymbol{v})_S$.
  - $(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k)_S = c_1(\boldsymbol{v}_1)_S + \cdots + c_k(\boldsymbol{v}_k)_S$.
- **Theorem.** Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for $V$.
  - $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ are linearly independent in $V$

    $\Leftrightarrow (\boldsymbol{u}_1)_S, \ldots, (\boldsymbol{u}_r)_S$ are linearly independent in $\mathbb{R}^k$.
  - $\mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\} = V$

    $\Leftrightarrow \mathrm{span}\{(\boldsymbol{u}_1)_S, \ldots, (\boldsymbol{u}_r)_S\} = \mathbb{R}^k$.
  - $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ is a basis for $V$

    $\Leftrightarrow \{(\boldsymbol{u}_1)_S, \ldots, (\boldsymbol{u}_r)_S\}$ is a basis for $\mathbb{R}^k$.

## Dimensions

- **Theorem.** Let $V$ be a vector space. Then every basis for $V$ is of the same size.
  - This number is the **dimension** of $V$, $\dim(V)$.
- **Examples.**
  - $\dim(\{\mathbf{0}\}) = 0; \quad \dim(\mathbb{R}^n) = n.$
  - Dimension of the solution space $\boldsymbol{Ax} = \mathbf{0}$:
    - No. of non-pivot colns in a row-echelon form of $\boldsymbol{A}$.
- **Properties.** Let $V$ be a vector space of $\dim(V) = n$.
  - Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k \in V$.
    - If $k > n$, then $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ are linearly dependent.
    - If $k < n$, then $\operatorname{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \neq V$.



## Dimensions

- **Properties.**
  - If $U$ is a subspace of $V$, then $\dim(U) \leq \dim(V)$.
  - If $U$ is a subspace of $V$, then
    - $U = V \Leftrightarrow \dim(U) = \dim(V)$.
    - $U \subsetneqq V \Leftrightarrow \dim(U) < \dim(V)$.
- Criterion to check if a subset $S \subseteq V$ is a basis for $V$.
  - A subset $S \subseteq V$ is a basis for a vector space $V$ if any two of the following three statements are true:
    - $\operatorname{span}(S) = V$;
    - $S$ is linearly independent;
    - $|S| = \dim(V)$.

## Transition Matrices

- Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for a vector space $V$.
  - Let $\boldsymbol{v} \in V$ and write $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$.

  The **coordinate vector** of $\boldsymbol{v}$ relative to $S$:
  - $(\boldsymbol{v})_S = (c_1, \ldots, c_k), \quad [\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}.$
- Find coordinate vector:
  - Solve $c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k = \boldsymbol{v}$.
  - Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$. Then $\boldsymbol{A}[\boldsymbol{v}]_S = \boldsymbol{v}$.
    - Solve $\left( \begin{array}{ccc|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k & \boldsymbol{v} \end{array} \right)$.



## Transition Matrices

- Let $V$ be a vector space having bases
  - $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.

  The **transition matrix** $\boldsymbol{P}$ from $S$ to $T$ is an $k \times k$ matrix
  - such that $\boldsymbol{P}[\boldsymbol{v}]_S = [\boldsymbol{v}]_T$ for all $\boldsymbol{v} \in V$.

  $\boldsymbol{P} = \begin{pmatrix} [\boldsymbol{u}_1]_T & \cdots & [\boldsymbol{u}_k]_T \end{pmatrix}$. It is obtained by solving
  - $\left( \begin{array}{ccc|c|c|c} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k & \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k \end{array} \right)$.
- **Properties.** Let $S, T, R$ be bases for a vector space $V$.
  - If $\boldsymbol{P}$ is the transition matrix from $S$ to $T$,
    - then $\boldsymbol{P}^{-1}$ is the transition matrix from $T$ to $S$.
  - $\boldsymbol{Q}$ is also the transition matrix from $T$ to $R$,
    - then $\boldsymbol{QP}$ is the transition matrix from $S$ to $R$.

## Row Spaces and Column Spaces

- Let $\boldsymbol{A}$ be an $m \times n$ matrix.
  - View $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{r}_1 \\ \vdots \\ \boldsymbol{r}_m \end{pmatrix}$, where $\boldsymbol{r}_i \in \mathbb{R}^n$.

  $\operatorname{span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m\} \subseteq \mathbb{R}^n$ is the **row space** of $\boldsymbol{A}$.
- **Properties**.
  - If $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent, then they have the same row space. (In other words, the row space is preserved by elementary row operations.)
  - Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$.
    - Then the nonzero rows of $\boldsymbol{R}$ form a basis for the row space of $\boldsymbol{A}$ (and of $\boldsymbol{R}$).



## Row Spaces and Column Spaces

- Let $\boldsymbol{A}$ be an $m \times n$ matrix.
  - View $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_n \end{pmatrix}$, where $\boldsymbol{c}_j \in \mathbb{R}^m$.

  $\operatorname{span}\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\} \subseteq \mathbb{R}^m$ is the **column space** of $\boldsymbol{A}$.
- **Properties**.
  - Suppose $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_n \end{pmatrix}$ are row equivalent.
    - If $\boldsymbol{a}_k = c_1\boldsymbol{a}_{j_1} + \cdots + c_r\boldsymbol{a}_{j_r}$,
      - then $\boldsymbol{b}_k = c_1\boldsymbol{b}_{j_1} + \cdots + c_r\boldsymbol{b}_{j_r}$.
    - If $\boldsymbol{a}_{j_1}, \ldots, \boldsymbol{a}_{j_r}$ are linearly independent,
      - then $\boldsymbol{b}_{j_1}, \ldots, \boldsymbol{b}_{j_r}$ are linearly independent.
    - If $\{\boldsymbol{a}_{j_1}, \ldots, \boldsymbol{a}_{j_r}\}$ is a basis for the coln space of $\boldsymbol{A}$,
      - then $\{\boldsymbol{b}_{j_1}, \ldots, \boldsymbol{b}_{j_r}\}$ is a basis for that of $\boldsymbol{B}$.

## Row Spaces and Column Spaces

- Find a basis for $V = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.

Method 1. View each $\boldsymbol{v}_i$ as a row vector and let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_k \end{pmatrix}$.

- ○ Find a row-echelon form $\boldsymbol{R}$ of $\boldsymbol{A}$.
- ○ The nonzero rows of $\boldsymbol{R}$ is a basis for $V$.

Method 2. View $\boldsymbol{v}_i$ as column vectors, set $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$.

- ○ Find a row-echelon form $\boldsymbol{R}'$ of $\boldsymbol{B}$.
- ○ The columns of $\boldsymbol{B}$ which are corresponding to the pivot columns of $\boldsymbol{R}'$ then form a basis for $V$.

Using the 2nd method, we can find a basis for $V$ by selecting vectors from $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$.



## Row Spaces and Column Spaces

- Extend a linearly independent set to a basis for $\mathbb{R}^n$.
  - ○ Suppose $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are linearly independent in $\mathbb{R}^n$.
    - i) View each $\boldsymbol{v}_i$ as a row vector.
    - ii) Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_k \end{pmatrix}$.
    - iii) Find a row-echelon form $\boldsymbol{R}$ of $\boldsymbol{A}$.
    - iv) Find the non-pivot columns of $\boldsymbol{R}$: $j_1, \ldots, j_{n-k}$.
    - v) $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{e}_{j_1}, \ldots, \boldsymbol{e}_{j_{n-k}}\}$ is a basis for $\mathbb{R}^n$.
- **Theorem.** $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent

  $\Leftrightarrow$ $\boldsymbol{b}$ lies in the column space of $\boldsymbol{A}$.

## Ranks

- **Theorem.** Let $\boldsymbol{A}$ be a matrix. Then its row space and its column space have the same dimension, $\operatorname{rank}(\boldsymbol{A})$.
- Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$.
  - $\operatorname{rank}(\boldsymbol{A}) =$ number of nonzero rows of $\boldsymbol{R}$.
  - $\operatorname{rank}(\boldsymbol{A}) =$ number of pivot columns of $\boldsymbol{R}$.
- **Properties.**
  - $\operatorname{rank}(\boldsymbol{A}) = 0 \Leftrightarrow \boldsymbol{A} = \boldsymbol{0}$.
  - $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^{\mathrm{T}})$.
  - If $\boldsymbol{A}$ is $m \times n$, then $\operatorname{rank}(\boldsymbol{A}) \leq m$, $\operatorname{rank}(\boldsymbol{A}) \leq n$.
  - $\operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{A})$, $\operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{B})$.
  - $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent

    $\Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b})$.



## Nullspaces and Nullities

- Let $\boldsymbol{A}$ be an $m \times n$ matrix.
  - The solution space to $\boldsymbol{Ax} = \boldsymbol{0}$ is the **nullspace** of $\boldsymbol{A}$.
  - The dimension of the nullspace is $\operatorname{nullity}(\boldsymbol{A})$.
- **Dimension Theorem for Matrices**. For $m \times n$ matrix $\boldsymbol{A}$,
  - $\operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) = n$.
- **Theorem.** Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{x}_0$.
  - Every solution to $\boldsymbol{Ax} = \boldsymbol{b}$ is of the form
    - $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{v}$,

    where $\boldsymbol{v} \in$ nullspace of $\boldsymbol{A}$, i.e., $\boldsymbol{Av} = \boldsymbol{0}$.

## The Dot Product

- Let $\boldsymbol{u} = (u_1, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, \ldots, v_n)$ be in $\mathbb{R}^n$.
  - **Dot product**: $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \cdots + u_n v_n$.
  - **Norm**: $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$.
  - **Distance**: $d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$.
  - **Angle**: $\cos\theta = \dfrac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}$, $\boldsymbol{u} \neq \boldsymbol{0}, \boldsymbol{v} \neq \boldsymbol{0}$.
- **Properties**. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.
  - $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.
  - $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$.
  - $(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v})$.
  - $\|c\boldsymbol{v}\| = |c| \, \|\boldsymbol{v}\|$.
  - $\|\boldsymbol{v}\| \geq 0$ and $\|\boldsymbol{v}\| = 0 \Leftrightarrow \boldsymbol{v} = \boldsymbol{0}$.
  - $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\|$.



## Orthogonal and Orthonormal Bases

- **Definitions.**
  - $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$ are **orthogonal** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
  - $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$ is **orthogonal** if
    - $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for all $i \neq j$.
  - $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$ is **orthonormal** if
    - $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for $i \neq j$ and $\|\boldsymbol{v}_i\| = 1$ for all $i$.
- **Theorem.**
  - If $S \subseteq \mathbb{R}^n$ is an orthogonal set of nonzero vectors,
    - then $S$ is linearly independent.
  - If $S \subseteq \mathbb{R}^n$ is an orthonormal set of vectors,
    - then $S$ is linearly independent.

## Orthogonal and Orthonormal Bases

- **Definitions.** Let $S$ be a basis for a vector space $V$.
  - $S$ is an **orthogonal basis** for $V$ if $S$ is orthogonal.
  - $S$ is an **orthonormal basis** for $V$ if $S$ is orthonormal.
- **Theorem.** Let $V$ be a vector space.
  - Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be an orthogonal basis for $V$.
    - $\boldsymbol{v} \in V \Rightarrow (\boldsymbol{v})_S = \dfrac{\boldsymbol{v} \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \cdots + \dfrac{\boldsymbol{v} \cdot \boldsymbol{v}_k}{\|\boldsymbol{v}_k\|^2}\boldsymbol{v}_k.$
  - Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be an orthonormal basis for $V$.
    - $\boldsymbol{v} \in V \Rightarrow (\boldsymbol{v})_S = (\boldsymbol{v} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + \cdots + (\boldsymbol{v} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k.$
- **Theorem.** Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for $V \subseteq \mathbb{R}^n$.
  - $\boldsymbol{v}$ is orthogonal to $V$ (i.e., $\boldsymbol{v} \cdot \boldsymbol{w} = 0$ for all $\boldsymbol{w} \in V$)

    $\Leftrightarrow$ $\boldsymbol{v}$ is orthogonal to $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$.



## Orthogonal and Orthonormal Bases

- **Definition.** Let $V$ be a subspace of $\mathbb{R}^n$.
  - For every $\boldsymbol{v} \in \mathbb{R}^n$, there exists a unique $\boldsymbol{p} \in V$
    - such that $\boldsymbol{v} - \boldsymbol{p}$ is orthogonal to $V$.

  Then $\boldsymbol{p}$ is the **projection** of $\boldsymbol{v}$ onto $V$.
- **Theorem.** Let $V$ be a subspace of $\mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$.
  - Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be an orthogonal basis for $V$.
    - Projection of $\boldsymbol{v}$ on $V$:
      - $\dfrac{\boldsymbol{v} \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \cdots + \dfrac{\boldsymbol{v} \cdot \boldsymbol{v}_k}{\|\boldsymbol{v}_k\|^2}\boldsymbol{v}_k.$

    Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be an orthonormal basis for $V$.
    - Projection of $\boldsymbol{v}$ on $V$:
      - $(\boldsymbol{v} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + \cdots + (\boldsymbol{v} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k.$

## Orthogonal and Orthonormal Bases

- **Gram-Schmidt Process.** Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be a basis for a vector space $V$.

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{u}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 \\
\cdots &= \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\boldsymbol{v}_k &= \boldsymbol{u}_k - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \cdots - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_{k-1}}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}.
\end{aligned}$$

  - Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for $V$.

  Let $\boldsymbol{w}_1 = \boldsymbol{v}_1/\|\boldsymbol{v}_1\|, \ldots, \boldsymbol{w}_k = \boldsymbol{v}_k/\|\boldsymbol{v}_k\|$.

  - Then $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ is an orthonormal basis for $V$.



## Best Approximations

- **Theorem.** Let $V$ be a subspace of $\mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$.
  - Let $\boldsymbol{p}$ be the projection of $\boldsymbol{v}$ onto $V$.
    - $d(\boldsymbol{v}, \boldsymbol{p}) \leq d(\boldsymbol{v}, \boldsymbol{w})$ for all $\boldsymbol{w} \in V$.

    Hence, $\boldsymbol{p}$ is the **best approximation** of $\boldsymbol{v}$ in $V$.
- **Properties.** Let $V$ be a subspace of $\mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$.
  - Let $\boldsymbol{p}$ be the projection of $\boldsymbol{v}$ onto $V$.
    - $\boldsymbol{v} - \boldsymbol{p}$ is orthogonal to $V$.
    - $\|\boldsymbol{v} - \boldsymbol{p}\|$ is the shortest distance from $\boldsymbol{v}$ to $V$.
- A linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is not necessarily consistent.
  - The **least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b}$ is $\boldsymbol{v}$
    - such that $\|\boldsymbol{Av} - \boldsymbol{b}\|$ is minimized.

## Best Approximations

- **Theorem.** The least squares solutions to $\boldsymbol{Ax} = \boldsymbol{b}$ are precisely all the solutions to
  - $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$, which is always consistent.
- **Note**. Let $\boldsymbol{v}$ be a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$.
  - $\|\boldsymbol{Av} - \boldsymbol{b}\|$ is minimized.
  - $\boldsymbol{Av}$ is the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$.
  - $\|\boldsymbol{Av} - \boldsymbol{b}\|$ is the shortest distance from $\boldsymbol{b}$ to the column space of $\boldsymbol{A}$.
- **Find projection onto a vector space**.
  - Suppose $V = \mathrm{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$.
    - i) Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$.
    - ii) Solve $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.
    - iii) Suppose $\boldsymbol{v}$ is any solution. Then $\boldsymbol{Av}$ is the projection of $\boldsymbol{b}$ onto $V$.



## Orthogonal Matrices

- Let $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$ and $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$.
  - $S$ is orthogonal $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is diagonal.
  - $S$ is orthonormal $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}_k$.
- **Definition.** A square matrix $\boldsymbol{A}$ is **orthogonal** if
  - $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}_n$, or equivalently, $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$.
- **Theorem.** Let $\boldsymbol{A}$ be a square matrix of order $n$.
  - $\boldsymbol{A}$ is an orthogonal matrix

    $\Leftrightarrow$ the rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$

    $\Leftrightarrow$ the columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$.
- **Theorem.** Let $S$ be an orthonormal basis for $V$.
  - Let $T$ be another basis for $V$ and $\boldsymbol{P}$ be the transition matrix from $S$ to $T$. Then
    - $T$ is orthonormal $\Leftrightarrow$ $\boldsymbol{P}$ is orthogonal.

### Eigenvalues and Eigenvectors

- **Definition.** Let $\boldsymbol{A}$ be a square matrix of order $n$.
  - Suppose $\boldsymbol{Av} = \lambda\boldsymbol{v}$ for some $\lambda \in \mathbb{R}$ and $\boldsymbol{0} \neq \boldsymbol{v} \in \mathbb{R}^n$.
    - Then $\lambda$ is an **eigenvalue** of $\boldsymbol{A}$, and $\lambda$ is an **eigenvector** of $\boldsymbol{A}$ associated to $\lambda$.
  - **Characteristic polynomial**: $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$.
  - **Characteristic equation**: $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$.
- **Properties.** Let $\boldsymbol{A}$ be a square matrix of order $n$.
  - $\lambda$ is an eigenvalue of $\boldsymbol{A} \Leftrightarrow \det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$.
  - If $\boldsymbol{A}$ is upper (or lower) triangular,
    - the diagonal entries are the eigenvalues of $\boldsymbol{A}$.
  - The set of all eigenvectors associated to eigenvalue $\lambda$:
    - Nonzero vectors in the nullspace of $\lambda\boldsymbol{I} - \boldsymbol{A}$.

    The nullspace of $\lambda\boldsymbol{I} - \boldsymbol{A}$ is the **eigenspace** $E_\lambda$.



### Diagonalization

- **Definition.** Let $\boldsymbol{A}$ be a square matrix. Suppose there exists an invertible matrix $\boldsymbol{P}$ and a diagonal matrix $\boldsymbol{D}$ s.t.
  - $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$.

  Then $\boldsymbol{A}$ is called **diagonalizable**.
  - $\boldsymbol{A}^m = \boldsymbol{PD}^m\boldsymbol{P}^{-1}$ for any positive integer $m$.
- **Properties.** Let $\boldsymbol{A}$ be a square matrix of order $n$.
  - Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ and $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.

  Then $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$ if and only if
  - $\boldsymbol{Av}_1 = \lambda\boldsymbol{v}_1, \ldots, \boldsymbol{Av}_n = \lambda_n\boldsymbol{v}_n$, and
  - $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form a basis for $\mathbb{R}^n$.

## Diagonalization

- **Theorem.** Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of $\boldsymbol{A}$.
  - Let $\boldsymbol{v}_i$ be an eigenvector of $\boldsymbol{A}$ associated to $\lambda_i$.
    - Then $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ are linearly independent.
- **Algorithm.** Let $\boldsymbol{A}$ be a square matrix.
  1. Solve $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ to find distinct eigenvalues:
     - $\lambda_1, \dots, \lambda_k$.
  2. For each $\lambda_i$, find a basis $S_{\lambda_i}$ for the eigenspace $E_{\lambda_i}$.
  3. Let $S = S_{\lambda_1} \cup \cdots \cup S_{\lambda_k}$.
     - If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.
     - If $|S| = n$, then $\boldsymbol{A}$ is diagonalizable.
- **Theorem.** If $\boldsymbol{A}$ has order $n$, and $\boldsymbol{A}$ has $n$ distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable.

52 / 57

## Orthogonal Diagonalization

- **Definition.** Let $\boldsymbol{A}$ be a square matrix of order $n$.
  - $\boldsymbol{A}$ is **orthogonally diagonalizable** if there exists an orthogonal matrix $\boldsymbol{P}$ and a diagonal matrix $\boldsymbol{D}$ s.t.
    - $\boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P} = \boldsymbol{D}$.
- **Theorem.** Let $\boldsymbol{A}$ is a square matrix.
  - $\boldsymbol{A}$ is orthogonally diagonalizable $\Leftrightarrow$ $\boldsymbol{A}$ is symmetric.
- **Algorithm.** Let $\boldsymbol{A}$ be a symmetric matrix.
  1. Set $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ to find distinct eigenvalues $\lambda_i$.
  2. For each $\lambda_i$, find a basis $S_{\lambda_i}$ for the eigenspace $E_{\lambda_i}$.
  3. Use Gram-Schmidt process to find an orthonormal basis $T_{\lambda_i}$ for $E_{\lambda_i}$.
  4. Let $T = T_{\lambda_1} \cup \cdots \cup T_{\lambda_k} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$.
  5. $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ is orthogonal, & diagonalizes $\boldsymbol{A}$.

53 / 57

### Linear Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

- A **linear transformation** is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$:
  - $T\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$.
- **Properties.** Let $T$ be a linear transformation.
  - $T(\mathbf{0}) = \mathbf{0}$.
  - $T(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) = c_1T(\boldsymbol{v}_1) + \cdots + c_kT(\boldsymbol{v}_k)$.
- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
  - Let $\boldsymbol{A} = \begin{pmatrix} T(\boldsymbol{e}_1) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix}$.
    - Then $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

    $\boldsymbol{A}$ is called the **standard matrix** for $T$.



### Linear Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

- **Algorithm.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ a basis for $T$.
  1. Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$.
  2. Let $\boldsymbol{B} = \begin{pmatrix} T(\boldsymbol{v}_1) & \cdots & T(\boldsymbol{v}_n) \end{pmatrix}$.
  3. The standard matrix for $T$ is $\boldsymbol{B}\boldsymbol{P}^{-1}$.
- **Theorem.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ be linear transformations.
  - Let $\boldsymbol{A}$ be the standard matrix for $T$, and $\boldsymbol{B}$ the standard matrix for $S$.
  - Then $\boldsymbol{B}\boldsymbol{A}$ is the standard matrix for $S \circ T$.

## Ranges and Kernels

- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\boldsymbol{A}$ be the standard matrix for $T$.
  - The **range** of $T$ is
    - $\mathrm{R}(T) = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$.
    - $\mathrm{R}(T) =$ column space of $\boldsymbol{A}$.
    - $\mathrm{rank}(T) = \dim \mathrm{R}(T) = \mathrm{rank}(\boldsymbol{A})$.
  - The **kernel** of $T$ is
    - $\mathrm{Ker}(T) = \{\boldsymbol{v} \in \mathbb{R}^n \mid T(\boldsymbol{v}) = \boldsymbol{0}\} \subseteq \mathbb{R}^n$.
    - $\mathrm{Ker}(T) =$ nullspace of $\boldsymbol{A}$.
    - $\mathrm{nullity}(T) = \dim \mathrm{Ker}(T) = \mathrm{nullity}(\boldsymbol{A})$.
- **Dimension Theorem for Linear Transformation.**
  - Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
    - $\mathrm{rank}(T) + \mathrm{nullity}(T) = n$.