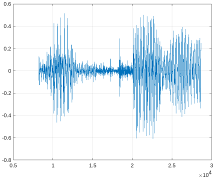
SC301 - Numerical Linear Algebra

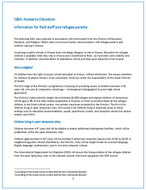
Preliminaries: Linear Algebra

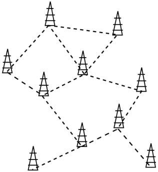
Aditya Tatu

January 7, 2025

Data







	Item 1	Item 2	Item 3	...	Item n
User 1	2	3	?	...	5
User 2	?	4	3	...	?
User 3	3	2	?	...	3
...	...	...	...	...	...
User m	1	?	5	...	4

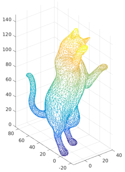


Figure: Examples of Data

- Typical representation of data, raw or feature-based is an m-tuple of real numbers, an element of $\mathbb{R}^m$.

Dataset

- Dataset will be a subset $D \subset \mathbb{R}^m$.
- Operations on $\mathbb{R}^m$:
  1. **Addition** $(+)$: $\forall \mathbf{u} = (u_1, \ldots, u_m), \forall \mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}^m, \mathbf{u}+\mathbf{v} := (u_1 + v_1, \ldots, u_m + v_m)$.
  2. **Scalar Multiplication** $(\cdot)$: $\forall x \in \mathbb{R}, \forall \mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}^m, x \cdot \mathbf{u} := (xu_1, \ldots, xu_m)$.

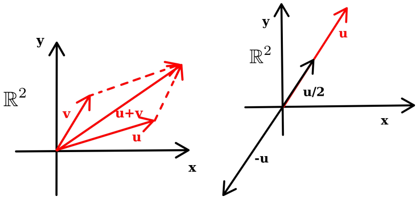


Figure: (left) Addition and (right) Scalar multiplication in $\mathbb{R}^2$

Linear Combination

- Given vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ and $x, y \in \mathbb{R}$, we say that the vector $x \cdot \mathbf{u} + y \cdot \mathbf{v} \in \mathbb{R}^m$ is a **linear combination** of $\mathbf{u}, \mathbf{v}$.
- In general, given vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$, and given scalars $x_1, \ldots, x_n$, we say that the vector $\sum_{k=1}^{n} x_k \mathbf{a}_k$ is a linear combination of $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.
- Explicitly in terms of the m-tuples, $\mathbf{a}_k = (a_{1,k}, a_{2,k}, \ldots, a_{m,k}), k = 1, \ldots, n,$

$$\sum_{k=1}^{n} x_k \mathbf{a}_k = x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vert \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vert \\ a_{m,2} \end{bmatrix} + \ldots + x_m \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vert \\ a_{m,n} \end{bmatrix} \quad (1)$$

$$= \begin{bmatrix} x_1 a_{1,1} \\ x_1 a_{2,1} \\ \vert \\ x_1 a_{m,1} \end{bmatrix} + \begin{bmatrix} x_2 a_{1,2} \\ x_2 a_{2,2} \\ \vert \\ x_2 a_{m,2} \end{bmatrix} + \ldots + \begin{bmatrix} x_n a_{1,n} \\ x_n a_{2,n} \\ \vert \\ x_n a_{m,n} \end{bmatrix} \quad (2)$$

$$\sum_{k=1}^{n} x_k \mathbf{a}_k = \begin{bmatrix} x_1 a_{1,1} \\ x_1 a_{2,1} \\ | \\ x_1 a_{m,1} \end{bmatrix} + \begin{bmatrix} x_2 a_{1,2} \\ x_2 a_{2,2} \\ | \\ x_2 a_{m,2} \end{bmatrix} + \ldots + \begin{bmatrix} x_n a_{1,n} \\ x_n a_{2,n} \\ | \\ x_n a_{m,n} \end{bmatrix} \tag{3}$$

$$= \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & - & a_{1,n} \\ a_{2,1} & a_{2,2} & - & a_{2,n} \\ | & | & | & | \\ a_{m,1} & a_{m,2} & - & a_{m,n} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix}}_{x} \tag{4}$$

- Matrix-Vector Product: $A_{m \times n} x_{n \times 1}$ is a linear combination of columns of A.

Linear Dependence/Independence

- If $\mathbf{b} = x\mathbf{u} + y\mathbf{v}$, we say that $\mathbf{b}$ is **linearly dependent** (or LD) on the set of vectors $\{\mathbf{u}, \mathbf{v}\}$.
  - In general, if $\mathbf{b} = \sum_{k=1}^{n} x_k \mathbf{a}_k$, where $\mathbf{a}_k \in \mathbb{R}^m, x_k \in \mathbb{R}, k = 1, \ldots, n$, we say that $\mathbf{b}$ is linearly dependent on the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.
- Conversely, we say that the vector $\mathbf{b}$ is **linearly independent** (or LI) of the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ if for no scalars $x_1, \ldots, x_n$, $\mathbf{b} = \sum_{k=1}^{n} x_k \mathbf{a}_k$.
- Explicitly, since $\mathbf{a}_k \in \mathbb{R}^m, k = 1, \ldots, n$,

$$x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ | \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ | \\ a_{m,2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ | \\ a_{m,n} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ | \\ b_m \end{bmatrix} \tag{5}$$

$$A\mathbf{x} = \mathbf{b} \tag{6}$$

Solving Linear equations

- Find $\mathbf{x} \in \mathbb{R}^n$ such that $A_{m\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{m\times 1}$
- Column Space of A:

$C(A) = \{\sum_{k=1}^{n} x_k \mathbf{a}_k \mid \forall x_k \in \mathbb{R}, k = 1, \ldots, n\} \subset \mathbb{R}^m$.

- Solution to $A\mathbf{x} = \mathbf{b}$ exists if and only if $\mathbf{b} \in C(A)$.
- Example: Let $A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \\ -3 & 1 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

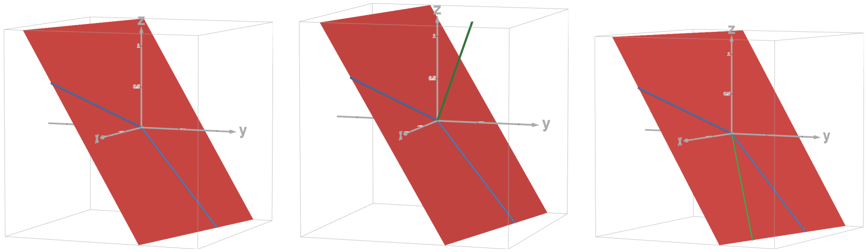


Figure: Column Space and Existence of Solution

- Example: $A_1 = \begin{bmatrix} -1 & 2 \\ 2 & -1 \\ -3 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 1 \\ 2 & -2 \\ -3 & 3 \end{bmatrix}$

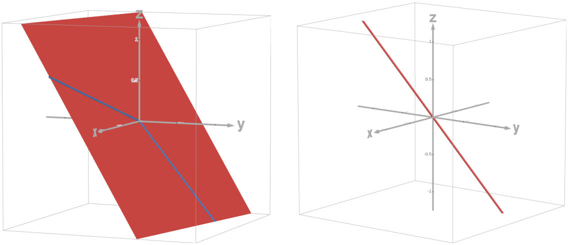


Figure: Column Space of A_1 and A_2.

- Properties of $C(A)$:
  - $\mathbf{0} \in C(A)$
  - If $\mathbf{u}, \mathbf{v} \in C(A), \forall x, y \in \mathbb{R}, x\mathbf{u} + y\mathbf{v} \in C(A)$.

Linearly Independent Set of vectors

● We say that $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is LI if none of the a_i's can be expressed as a linear combination of the remaining $\mathbf{a}_k$'s.

▶ Re-writing, if the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is LD then there must be at least one vector, say $\mathbf{a}_i$ from the set such that $\mathbf{a}_i = \sum_{k=1,k\neq i}^{n} x_k \mathbf{a}_k$. Then,

$$\sum_{k=1,k\neq i}^{n} x_k \mathbf{a}_k - \mathbf{a}_i = \mathbf{0} \tag{7}$$

$$\sum_{k=1}^{n} x_k \mathbf{a}_k = \mathbf{0}, \text{ with } x_i = -1 \tag{8}$$

● Thus, there is at least one non-zero scalar in the set $\{x_1, \ldots, x_n\}$ such that $\sum_{k=1}^{n} x_k \mathbf{a}_k = \mathbf{0}$.

▶ So, if cols of A are LD, there exists $\mathbf{x} \neq \mathbf{0}$ such that $Ax = \mathbf{0}$. And $A\mathbf{0} = \mathbf{0}$.

LI set of Vectors

- How to verify LI of $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$?
  - Find $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$.

Nullspace

- For $A \in \mathbb{R}^{m \times n}$, its Nullspace, denoted by $N(A)$ is defined as:

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

- Properties of $N(A)$:
  - $\mathbf{0} \in N(A)$
  - If $\mathbf{u}, \mathbf{v} \in N(A), \forall x, y \in \mathbb{R}, x\mathbf{u} + y\mathbf{v} \in N(A)$.

- If $\mathbf{w} \neq \mathbf{0}, \mathbf{w} \in N(A)$, cols of A are not LI.
- If $\mathbf{y} \in \mathbb{R}^n$ satisfies/solves $A\mathbf{x} = \mathbf{b}$, and $\mathbf{w} \in N(A)$, then $\mathbf{y} + \mathbf{w}$ also solves $A\mathbf{x} = \mathbf{b} \Rightarrow$ Multiple Solutions!
- If there are multiple solutions to $A\mathbf{x} = \mathbf{b}$, then what can we conclude about $N(A)$?

Matrix Multiplication

- Let $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times k}$, then $AB \in \mathbb{R}^{m\times k}$:

$$AB = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ | & | & | \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} b_{1,1} & \dots & b_{n,1} \\ | & | & | \\ b_{n,1} & \dots & b_{n,k} \end{bmatrix} \tag{9}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} a_{1,i} b_{i,1} & \dots & \sum_{i=1}^{n} a_{1,i} b_{i,n} \\ | & | & | \\ \sum_{i=1}^{n} a_{m,i} b_{i,1} & \dots & \sum_{i=1}^{n} a_{m,i} b_{i,n} \end{bmatrix} \tag{10}$$

  - Cols. of AB are LC of cols of A
  - Rows of AB are LC of rows of B
  - Also,

$$AB = \sum_{i=1}^{n} \begin{bmatrix} a_{1,i} b_{i,1} & \dots & a_{1,i} b_{i,n} \\ | & | & | \\ a_{m,i} b_{i,1} & \dots & a_{m,i} b_{i,n} \end{bmatrix}$$

Matrix Multiplication

- Let $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times k}$, then $AB \in \mathbb{R}^{m\times k}$:

$$AB = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ | & | & | \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} b_{1,1} & \dots & b_{n,1} \\ | & | & | \\ b_{n,1} & \dots & b_{n,k} \end{bmatrix} \tag{9}$$

$$= \begin{bmatrix} \sum_{i=1}^n a_{1,i} b_{i,1} & \dots & \sum_{i=1}^n a_{1,i} b_{i,n} \\ | & | & | \\ \sum_{i=1}^n a_{m,i} b_{i,1} & \dots & \sum_{i=1}^n a_{m,i} b_{i,n} \end{bmatrix} \tag{10}$$

  - Cols. of AB are LC of cols of A
  - Rows of AB are LC of rows of B
  - Also,

$$AB = \sum_{i=1}^{n} \begin{bmatrix} a_{1,i} b_{i,1} & \dots & a_{1,i} b_{i,n} \\ | & | & | \\ a_{m,i} b_{i,1} & \dots & a_{m,i} b_{i,n} \end{bmatrix} = \sum_{i=1}^{n} C_i \tag{11}$$

Rank of a Matrix

- The number of LI cols of a matrix A is called the Column rank of A.
- The number of LI rows of a matrix A is called the Row rank of A.
- **Assignment 1:** Try to show that they are the same.
  - Notation: $r(A)$.
- **Assignment 2:** Let $A \in \mathbb{R}^{m\times 1}, B \in \mathbb{R}^{1\times n}$. Find the rank of AB.

Under & Over determined system of equations

- Let $A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$. The system $A\mathbf{x} = \mathbf{b}$ is said to be
  - Under-determined if there are fewer equations than variables, i.e., $m < n$.
  - Over-determined if there are more equations than variables, i.e., $m > n$.

Span

- Let $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathbb{R}^m$ be a set of vectors. Then

$$span(S) = \{\sum_{k=1}^{n} x_k \mathbf{a}_k \mid \forall x_k \in \mathbb{R}, k = 1, \ldots, n\}$$

- Let $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \ldots \quad \mathbf{a}_n] \in \mathbb{R}^{m \times n}$. Then, $span(S) = C(A)$.

$\mathbb{R}^n$ as a Vector Space

● $\mathbb{R}^n$ with the set $\mathbb{R}$, and two binary operations, addition $+$ and scalar multiplication $\cdot$ satisfy:

- **Closure w.r.t Addition:** $\forall u, v \in \mathbb{R}^n, u + v \in \mathbb{R}^n$
- **Additive identity:** $\mathbf{0} \in \mathbb{R}^n, \forall u \in \mathbb{R}^n, u + \mathbf{0} = \mathbf{0} + u = u$
- **Additive inverse:** $\forall u \in \mathbb{R}^n, \exists v \in \mathbb{R}^n, u + v = u + v = \theta$. We will denote v by $-u$.
- **Associativity:** $\forall u, v, w \in \mathbb{R}^n, (u + v) + w = u + (v + w)$.
- **Commutative:** $\forall u, v \in \mathbb{R}^n, u + v = v + u$.
- **Closure w.r.t Scalar multiplication:** $\forall a \in \mathbb{R}, \forall u \in \mathbb{R}^n, a \cdot u \in \mathbb{R}^n$.
- **Scalar Multiplication identity:** $\exists 1 \in \mathbb{R}$ such that $1 \cdot u = u, \forall u \in \mathbb{R}^n$.
- **Distributivity:** $\forall a \in \mathbb{R}, \forall u, v \in \mathbb{R}^n, a \cdot (u + v) = a \cdot u + a \cdot v$, and $\forall a, b \in \mathbb{R}, \forall u \in \mathbb{R}^n, (a + b) \cdot u = a \cdot u + b \cdot u$.
- **Compatibility of multiplication of real nos. and scalar multiplication:** $\forall a, b \in \mathbb{R}, \forall u \in \mathbb{R}^n, (a \times b) \cdot u = a \cdot (b \cdot u)$.

- **Subspace:** A set $W \subseteq \mathbb{R}^n$ is said to be a subspace of $\mathbb{R}^n$ if W is also a vector space. A non-empty subset W of $\mathbb{R}^n$ is a subspace iff it is closed with respect to $+$ and $\cdot$.

● **Subspace:** A set $W \subseteq \mathbb{R}^n$ is said to be a subspace of $\mathbb{R}^n$ if W is also a vector space. A non-empty subset W of $\mathbb{R}^n$ is a subspace iff it is closed with respect to $+$ and $\cdot$.

● Examples: For $A \in \mathbb{R}^{m \times n}$, $C(A)$ is a subspace of $\mathbb{R}^m$, while $N(A)$ is a subspace of $\mathbb{R}^n$.

● **Subspace:** A set $W \subseteq \mathbb{R}^n$ is said to be a subspace of $\mathbb{R}^n$ if W is also a vector space. A non-empty subset W of $\mathbb{R}^n$ is a subspace iff it is closed with respect to $+$ and $\cdot$.

● Examples: For $A \in \mathbb{R}^{m \times n}$, $C(A)$ is a subspace of $\mathbb{R}^m$, while $N(A)$ is a subspace of $\mathbb{R}^n$.

● **Assignment 3:** Given that U, W are subspaces of $\mathbb{R}^n$,

- Is $U \cap W$ a subspace of $\mathbb{R}^n$?
- Is $U \cup W$ a subspace of $\mathbb{R}^n$?
- Let $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathbb{R}^m$. Show that $span(S)$ is a subspace of $\mathbb{R}^m$.

● **Subspace:** A set $W \subseteq \mathbb{R}^n$ is said to be a subspace of $\mathbb{R}^n$ if W is also a vector space. A non-empty subset W of $\mathbb{R}^n$ is a subspace iff it is closed with respect to $+$ and $\cdot$.

● Examples: For $A \in \mathbb{R}^{m \times n}$, $C(A)$ is a subspace of $\mathbb{R}^m$, while $N(A)$ is a subspace of $\mathbb{R}^n$.

● **Assignment 3:** Given that U, W are subspaces of $\mathbb{R}^n$,

- Is $U \cap W$ a subspace of $\mathbb{R}^n$?
- Is $U \cup W$ a subspace of $\mathbb{R}^n$?
- Let $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subset \mathbb{R}^m$. Show that $span(S)$ is a subspace of $\mathbb{R}^m$.

● **Assignment 4:** Let $w \in \mathbb{R}^3, w \neq 0$, and $S = \{a \in \mathbb{R}^3 \mid \langle a, w \rangle = w^T a = 0\}$. Is S a subspace of $\mathbb{R}^3$?

Basis

- **Basis:** Let V be a subspace of $\mathbb{R}^n$. A set $\beta \subset V$ is said to be a **basis** of V if (1) $span(\beta) = V$, and (2) β is a set of linearly independent vectors.
- **Assignment 5:** A subset $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of V is a basis of V if and only if every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \ldots + x_k\mathbf{b}_k, x_i \in \mathbb{R}, i = 1, \ldots, k.$

Basis

● **Basis:** Let V be a subspace of $\mathbb{R}^n$. A set $\beta \subset V$ is said to be a **basis** of V if (1) $span(\beta) = V$, and (2) β is a set of linearly independent vectors.

● **Assignment 5:** A subset $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of V is a basis of V if and only if every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \ldots + x_k\mathbf{b}_k, x_i \in \mathbb{R}, i = 1, \ldots, k$.

- The vector $\begin{bmatrix} x_1 \\ x_2 \\ | \\ x_k \end{bmatrix}$ of coefficients, denoted by $[\mathbf{v}]_\beta$ is called the **representation of v in basis** β.

Basis

● **Basis:** Let V be a subspace of $\mathbb{R}^n$. A set $\beta \subset V$ is said to be a **basis** of V if (1) $span(\beta) = V$, and (2) β is a set of linearly independent vectors.

● **Assignment 5:** A subset $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of V is a basis of V if and only if every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \ldots + x_k\mathbf{b}_k, x_i \in \mathbb{R}, i = 1, \ldots, k$.

- The vector $\begin{bmatrix} x_1 \\ x_2 \\ | \\ x_k \end{bmatrix}$ of coefficients, denoted by $[\mathbf{v}]_\beta$ is called the **representation of v in basis** β.

● **Assignment 6:** Let β_1, β_2 be two basis of V. Then, show that $|\beta_1| = |\beta_2|$.

Basis

- **Basis:** Let V be a subspace of $\mathbb{R}^n$. A set $\beta \subset V$ is said to be a **basis** of V if (1) $span(\beta) = V$, and (2) β is a set of linearly independent vectors.
- **Assignment 5:** A subset $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of V is a basis of V if and only if every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \ldots + x_k\mathbf{b}_k, x_i \in \mathbb{R}, i = 1, \ldots, k$.
  - The vector $\begin{bmatrix} x_1 \\ x_2 \\ | \\ x_k \end{bmatrix}$ of coefficients, denoted by $[\mathbf{v}]_\beta$ is called the **representation of v in basis** β.
- **Assignment 6:** Let β_1, β_2 be two basis of V. Then, show that $|\beta_1| = |\beta_2|$.
- **Dimension:** Let $V \subset \mathbb{R}^n$ be a subspace with basis β. Then dimension of V, denoted by $dim(V)$, is $dim(V) = |\beta|$.

Basis

- **Basis:** Let V be a subspace of $\mathbb{R}^n$. A set $\beta \subset V$ is said to be a **basis** of V if (1) $span(\beta) = V$, and (2) β is a set of linearly independent vectors.
- **Assignment 5:** A subset $\beta = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of V is a basis of V if and only if every $\mathbf{v} \in V$ can be uniquely written as $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \ldots + x_k\mathbf{b}_k, x_i \in \mathbb{R}, i = 1, \ldots, k$.
  - The vector $\begin{bmatrix} x_1 \\ x_2 \\ | \\ x_k \end{bmatrix}$ of coefficients, denoted by $[\mathbf{v}]_\beta$ is called the **representation of v in basis** β.
- **Assignment 6:** Let β_1, β_2 be two basis of V. Then, show that $|\beta_1| = |\beta_2|$.
- **Dimension:** Let $V \subset \mathbb{R}^n$ be a subspace with basis β. Then dimension of V, denoted by $dim(V)$, is $dim(V) = |\beta|$.
- **Rank-Nullity Theorem:** For any matrix $A \in \mathbb{R}^{m \times n}$, $r(A) + dim(N(A)) = n$. $dim(N(A))$ is called the **nullity** of A.

Norm on $\mathbb{R}^n$

- Geometric interpretation - measure of length of a vector.

Norm on $\mathbb{R}^n$

- Geometric interpretation - measure of length of a vector.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sqrt{\sum_{k=1}^{n} u_k^2}$ - Euclidean norm.

Norm on $\mathbb{R}^n$

- Geometric interpretation - measure of length of a vector.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sqrt{\sum_{k=1}^{n} u_k^2}$ - Euclidean norm.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sum_{k=1}^{n} |u_k|$ - l_1 norm.

Norm on $\mathbb{R}^n$

- Geometric interpretation - measure of length of a vector.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sqrt{\sum_{k=1}^{n} u_k^2}$ - Euclidean norm.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sum_{k=1}^{n} |u_k|$ - l_1 norm.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \max_{1 \leq k \leq n} |u_k|$ - max norm.

Norm on $\mathbb{R}^n$

- Geometric interpretation - measure of length of a vector.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sqrt{\sum_{k=1}^{n} u_k^2}$ - Euclidean norm.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \sum_{k=1}^{n} |u_k|$ - l_1 norm.
  - Example: $\forall \mathbf{u} \in \mathbb{R}^n, ||\mathbf{u}|| = \max_{1 \leq k \leq n} |u_k|$ - max norm.

Inner Product on $\mathbb{R}^n$

- Geometric interpretation - measures angle between vectors.

Inner Product on $\mathbb{R}^n$

- Geometric interpretation - measures angle between vectors.
  - Example: $\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^T \mathbf{u} = \sum_{k=1}^{n} v_i u_i$.

Inner Product on $\mathbb{R}^n$

- Geometric interpretation - measures angle between vectors.
  - Example: $\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^T \mathbf{u} = \sum_{k=1}^{n} v_i u_i$.
  - Norm induced by inner-product: $||\mathbf{u}||^2 = \langle \mathbf{u}, \mathbf{u} \rangle$.
  - Angle between vectors: $\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}||||\mathbf{v}||}$.
  - **Cauchy-Schwartz Inequality:** $|\cos\theta| \leq 1$

Inner Product on $\mathbb{R}^n$

- Geometric interpretation - measures angle between vectors.
  - Example: $\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^T \mathbf{u} = \sum_{k=1}^{n} v_i u_i$.
  - Norm induced by inner-product: $||\mathbf{u}||^2 = \langle \mathbf{u}, \mathbf{u} \rangle$.
  - Angle between vectors: $\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{u}||||\mathbf{v}||}$.
  - **Cauchy-Schwartz Inequality:** $|\cos\theta| \leq 1 \Rightarrow |\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}||||\mathbf{v}||$.
- **Orthogonal vectors:** For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, the vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be orthogonal.
  - **Orthogonal vectors:** The set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} \subset \mathbb{R}^n$ is said to be orthogonal if $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \forall 1 \leq i \neq j \leq k$.
  - **Orthonormal vectors:** The set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} \subset \mathbb{R}^n$ is said to be orthonormal if apart from being orthogonal, they are all unit vectors, i.e., $\langle \mathbf{a}_i, \mathbf{a}_i \rangle = 1, \forall 1 \leq i \leq k$.
- **Orthogonal Matrix:** A matrix $A \in \mathbb{R}^{n \times n}$ is called Orthogonal if its columns are orthonormal.
- **Orthonormal basis:** For a subspace $V \subseteq \mathbb{R}^n$, an orthonormal set of vectors β that also forms a basis of V is called an Orthonormal basis (ONB).
- **Assignment 7:** What is the advantage of having an ONB? [Hint: Finding representation of a vector in that basis is easy].

Eigenvalues and Eigenvectors

● **Eigenvalues & Eigenvectors:** Let $A \in \mathbb{R}^{n\times n}$. If $A\mathbf{v} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}, \mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$, then $\lambda, \mathbf{v}$ is said to be eigenvalue-eigenvector pair.

- Computing eigenvalues and eigenvectors (in theory) $A\mathbf{v} = \lambda\mathbf{v}$.

● For $A \in \mathbb{R}^{n\times n}$, if there are n linearly independent eigenvectors, say $\mathbf{e}_1, \ldots, \mathbf{e}_n$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, then:

- $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ forms a basis of $\mathbb{R}^n$, and
- **Matrix diagonalization:**

$$A[\mathbf{e}_1 \quad \mathbf{e}_2 \quad \ldots \quad \mathbf{e}_n] = [A\mathbf{e}_1 \quad A\mathbf{e}_2 \quad \ldots \quad A\mathbf{e}_n] \tag{12}$$

$$= [\lambda_1\mathbf{e}_1 \quad \lambda_2\mathbf{e}_2 \quad \ldots \quad \lambda_n\mathbf{e}_n] \tag{13}$$

$$= [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \ldots \quad \mathbf{e}_n]\mathit{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{14}$$

$$\Rightarrow AE = E\Lambda$$

Eigenvalues and Eigenvectors

● **Eigenvalues & Eigenvectors:** Let $A \in \mathbb{R}^{n \times n}$. If $A\mathbf{v} = \lambda \mathbf{v}$ for some $\lambda \in \mathbb{R}, \mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$, then $\lambda, \mathbf{v}$ is said to be eigenvalue-eigenvector pair.

- Computing eigenvalues and eigenvectors (in theory) $A\mathbf{v} = \lambda \mathbf{v}$.

● For $A \in \mathbb{R}^{n \times n}$, if there are n linearly independent eigenvectors, say $\mathbf{e}_1, \dots, \mathbf{e}_n$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$, then:

- $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ forms a basis of $\mathbb{R}^n$, and
- **Matrix diagonalization:**

$$
\begin{aligned}
A[\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n] &= [A\mathbf{e}_1 \quad A\mathbf{e}_2 \quad \dots \quad A\mathbf{e}_n] && (12)\\
&= [\lambda_1 \mathbf{e}_1 \quad \lambda_2 \mathbf{e}_2 \quad \dots \quad \lambda_n \mathbf{e}_n] && (13)\\
&= [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n] diag(\lambda_1, \lambda_2, \dots, \lambda_n) && (14)\\
&\Rightarrow AE = E\Lambda \Rightarrow A = E\Lambda E^{-1} && (15)
\end{aligned}
$$

- Let $A = \begin{bmatrix} 2.8 & -2.4 \\ -2.4 & -0.8 \end{bmatrix}$.
- Using basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$,

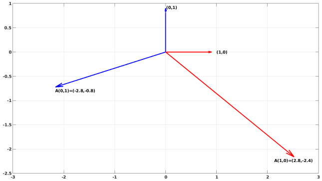


Figure: Usual basis

- Using eigenbasis $\left\{ \begin{bmatrix} 0.44 \\ 0.89 \end{bmatrix}, \begin{bmatrix} 0.89 \\ -0.44 \end{bmatrix} \right\}$,

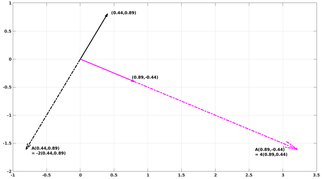


Figure: Eigenbasis

Eigenbasis

- Do all matrices yield an eigenbasis?
  - No, examples: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$,

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \theta \neq k\pi, k \in \mathbb{Z}$.

- **Spectral Theorem for Real-Symmetric Matrices:** Let $A \in \mathbb{R}^{n\times n}$ be any real-symmetric matrix, i.e., $A^T = A$. Then,
  - its eigenvalues are real,and
  - A has n-orthogonal eigenvectors.
- **Symmetric Positive-Definite Matrix:** A matrix $A \in \mathbb{R}^{n\times n}$ is said to be Symmetric Positive-Definite (SPD), if (1) $A^T = A$, and (2) $x^T Ax > 0, \forall x \neq 0, x \in \mathbb{R}^n$.
  - For a SPD matrix, all eigenvalues are strictly positive.
  - A matrix $A \in \mathbb{R}^{n\times n}$ is **Symmetric Positive-Semidefinite** if (1) $A^T = A$, and (2) $x^T Ax \geq 0, \forall x \neq 0, x \in \mathbb{R}^n$.

Singular Value Decomposition

- For $A \in \mathbb{R}^{m \times n}$ with $r(A) = r$.

▶ $A^T A \in \mathbb{R}^{n \times n}$ and $AA^T \in \mathbb{R}^{m \times m}$ are real-symmetric positive semi-definite matrices.

▶ **Assignment 8:** Show that $A^T A$ and AA^T have the same non-zero eigenvalues.

▶ $A^T A = V^T \Lambda_n V$ and $AA^T = U^T \Lambda_m U$, where $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ are orthogonal matrices, and $\Lambda_n = diag(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0) \in \mathbb{R}^{n \times n}, \Lambda_m = diag(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0) \in \mathbb{R}^{m \times m}$.

- Then $A = U \Sigma V^T$, with $\Sigma = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_r}, 0, \ldots, 0) \in \mathbb{R}^{m \times n}$. Notation: $\sigma_k = \sqrt{\lambda_k}, k = 1, \ldots, r$ are called the **Singular values**.