

Preliminary Linear Algebra 1

Notation

\forall for all

\exists there exists

\ni such that

\therefore therefore

\because because

\square end of proof (QED)

Notation

$A \implies B$ A implies B

$A \iff B$ A if and only if (iff) B

$a \in B$ a is an element of the set B

$A \subset B$ A is a proper subset of B

$A \subseteq B$ A is a subset of B

\mathbb{R}^n Euclidean n -space

Matrix Notation

• $\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ is a matrix with m rows and n columns.

- The entry in the i^{th} row and j^{th} column of \mathbf{A} is a_{ij} .

Square Matrices and Vectors

- Matrix A is said to be square if $m = n$.
 $m \times n$
- A matrix with one column is called a vector.
- A matrix with one row is called a row vector.

Notation

In these STAT611 notes, bold uppercase letters are typically used to denote matrices, and bold lowercase letters are typically used to denote vectors.

Examples

- $\mathbf{0}$ or $\mathbf{0}_n$ is a vector of zeros.
- $\mathbf{1}$ or $\mathbf{1}_n$ is a vector of ones.
- \mathbf{I} or \mathbf{I}_n or $\mathbf{I}_{n \times n}$ is the identity matrix. For example,

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \text{diag}(1, 1, 1).$$

Special Types of Square Matrices

- A square matrix A is upper triangular if $a_{ij} = 0, \forall i > j$.
- A square matrix A is lower triangular if $a_{ij} = 0, \forall i < j$.
- A square matrix A is diagonal if $a_{ij} = 0, \forall i \neq j$.
- Write one example for each of these types of matrices.

Examples

- Upper triangular
$$\begin{bmatrix} 1 & 0 & 4 & 6 \\ 0 & 2 & -3 & 5 \\ 0 & 0 & 8 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Lower triangular
$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$$

- Diagonal
$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -13 \end{bmatrix}.$$

Matrix Transpose

If $\mathbf{A} = [a_{ij}]$, the transpose of \mathbf{A} , denoted \mathbf{A}' , is the matrix $\mathbf{B} = [b_{ij}]$, where $b_{ij} = a_{ji}$, $\forall i = 1, \dots, m; \quad j = 1, \dots, n$.

That is, $\mathbf{B} = \mathbf{A}'$ is the matrix whose columns are the rows of \mathbf{A} and whose rows are the columns of \mathbf{A} .

A Symmetric Matrix

A square matrix A is symmetric if $A = A'$.

Examples

- Find the transpose of

$$\begin{bmatrix} 4 & -2 \\ 3 & 7 \end{bmatrix}.$$

- Provide an example of a symmetric matrix.

Examples

- $\begin{bmatrix} 4 & -2 \\ 3 & 7 \end{bmatrix}' = \begin{bmatrix} 4 & 3 \\ -2 & 7 \end{bmatrix}$.

- The matrix

$$\begin{bmatrix} 4 & 2 & -1 \\ 2 & 0 & 3 \\ -1 & 3 & 5 \end{bmatrix}$$

is symmetric.

Matrix Multiplication

- Suppose

$$\mathbf{A}_{m \times n} = [a_{ij}] = \begin{bmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_m \end{bmatrix} = [\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(n)}]$$

and

$$\mathbf{B}_{n \times k} = [b_{ij}] = \begin{bmatrix} \mathbf{b}'_1 \\ \vdots \\ \mathbf{b}'_n \end{bmatrix} = [\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(k)}].$$

- Then

$$\begin{aligned} \mathbf{A}_{m \times n} \mathbf{B}_{n \times k} &= \mathbf{C}_{m \times k} = [c_{ij} = \sum_{l=1}^n a_{il}b_{lj}] = [c_{ij} = \mathbf{a}'_i \mathbf{b}^{(j)}] \\ &= [\mathbf{A}\mathbf{b}^{(1)}, \dots, \mathbf{A}\mathbf{b}^{(k)}] = \begin{bmatrix} \mathbf{a}'_1 \mathbf{B} \\ \vdots \\ \mathbf{a}'_m \mathbf{B} \end{bmatrix} = \sum_{l=1}^n \mathbf{a}^{(l)} \mathbf{b}'_l. \end{aligned}$$

Matrix Multiplication

- Suppose

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

- Work out \mathbf{AB} using $\mathbf{AB} = \sum_{l=1}^n \mathbf{a}^{(l)} \mathbf{b}'_l$.

Matrix Multiplication

$$\begin{aligned} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} [5 \ 6] + \begin{bmatrix} 2 \\ 4 \end{bmatrix} [7 \ 8] \\ &= \begin{bmatrix} 5 & 6 \\ 15 & 18 \end{bmatrix} + \begin{bmatrix} 14 & 16 \\ 28 & 32 \end{bmatrix} \\ &= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}. \end{aligned}$$

Transpose of a Matrix Product



$$(AB)' = B'A'$$

- The transpose of a product is the product of the transposes in reverse order.

Scalar Multiplication of a Matrix

If $c \in \mathbb{R}$, then c times the matrix A is the matrix whose i, j^{th} element is c times the i, j^{th} element of A ; i.e.,

$$c\mathbf{A} = c \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.$$

Linear Combination

If $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^m$ and $c_1, \dots, c_n \in \mathbb{R}$, then

$$\sum_{i=1}^n c_i \mathbf{x}_i = c_1 \mathbf{x}_1 + \dots + c_n \mathbf{x}_n$$

is a linear combination (LC) of $\mathbf{x}_1, \dots, \mathbf{x}_n$.

The coefficients of the LC are c_1, \dots, c_n .

Linear Independence and Linear Dependence

- A set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ is linearly independent (LI) if

$$\sum_{i=1}^n c_i \mathbf{x}_i = \mathbf{0} \iff c_1 = \dots = c_n = 0.$$

- A set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ is linearly dependent (LD) if

$$\exists c_1, \dots, c_n \text{ not all } 0 \ni \sum_{i=1}^n c_i \mathbf{x}_i = \mathbf{0}.$$

- Prove or disprove: any set of vectors that contains the vector $\mathbf{0}$ is LD.
- Prove or disprove: The following set of vectors is LI.

$$\begin{bmatrix} 1 \\ -5 \\ 3 \end{bmatrix}, \begin{bmatrix} 7 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ -6 \\ 7 \end{bmatrix} .$$

- Suppose $\mathbf{x}_1 = \mathbf{0}$ and $\mathbf{x}_2, \dots, \mathbf{x}_n$ are any other vectors of the same dimension as \mathbf{x}_1 .
- If we take $c_1 = 1, c_2 = \dots = c_n = 0$, then $\sum_{i=1}^n c_i \mathbf{x}_i = \mathbf{0}$ and c_1, c_2, \dots, c_n are not all zero.
- Thus, any set of vectors containing $\mathbf{0}$ is LD.

- If we take $c_1 = 2, c_2 = 1, c_3 = -1$ then

$$c_1 \begin{bmatrix} 1 \\ -5 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 7 \\ 4 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 9 \\ -6 \\ 7 \end{bmatrix} = \mathbf{0}.$$

Thus the set of vectors is LD.

- One way to arrive at such solution is to search for a solution to the system of the equations:

$$\begin{aligned} c_1 + 7c_2 + 9c_3 &= 0 \\ -5c_1 + 4c_2 - 6c_3 &= 0 \\ 3c_1 + c_2 + 7c_3 &= 0. \end{aligned}$$

Fact V1:

The nonzero vectors x_1, \dots, x_n are LD $\iff x_j$ is a LC of x_1, \dots, x_{j-1} for some $j \in \{2, \dots, n\}$.

Proof of Fact V1:

(\implies) Suppose there exist c_1, \dots, c_n such that $\sum_{i=1}^n c_i \mathbf{x}_i = \mathbf{0}$. Let

$$j = \max\{i : c_i \neq 0\}.$$

Since $\mathbf{x}_1, \dots, \mathbf{x}_n$ are nonzero, $j > 1$. Then

$$\begin{aligned} \sum_{i=1}^j c_i \mathbf{x}_i = \mathbf{0} &\implies \sum_{i=1}^{j-1} c_i \mathbf{x}_i = -c_j \mathbf{x}_j. \\ &\implies \sum_{i=1}^{j-1} \frac{-c_i}{c_j} \mathbf{x}_i = \mathbf{x}_j. \end{aligned}$$

(\impliedby) Suppose $\mathbf{x}_j = \sum_{i=1}^{j-1} d_i \mathbf{x}_i$, then $\sum_{i=1}^n d_i \mathbf{x}_i = \mathbf{0}$, where

$$d_i = \begin{cases} c_i & \text{if } i < j \\ -1 & \text{if } i = j \\ 0 & \text{if } i > j. \end{cases}$$



Orthogonality

- The two vectors x, y are orthogonal to each other if their inner product is zero, i.e.,

$$\mathbf{x}'\mathbf{y} = \mathbf{y}'\mathbf{x} = \sum_{i=1}^n x_i y_i = 0.$$

- The length of a vector, also known as its Euclidean norm, is

$$\|\mathbf{x}\| := \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{\sum_{i=1}^n x_i^2}.$$

- The vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are mutually orthogonal if

$$\mathbf{x}_i' \mathbf{x}_j = 0, \quad \forall i \neq j.$$

- The vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are mutually orthonormal if

$$\mathbf{x}_i' \mathbf{x}_j = 0 \quad \forall i \neq j, \text{ and } \|\mathbf{x}_i\| = 1 \quad \forall i = 1, \dots, n.$$

- Write down a set of mutually orthogonal but not mutually orthonormal vectors.
- Write down a set of mutually orthonormal vectors.

- $\begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are mutually orthogonal but not mutually orthonormal.
- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are mutually orthonormal.
- $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ are mutually orthonormal.

Orthogonal Matrix

- A square matrix with mutually orthonormal columns is called an orthogonal matrix.

- Show that if Q is orthogonal, then $Q'Q = I$.
- Show that if Q is orthogonal and x is any vector of appropriate dimension, then $\|Qx\| = \|x\|$.

- $Q'Q = [q_i'q_j]$, where $Q = [q_1, \dots, q_n]$.
- By orthogonality of Q , q_1, \dots, q_n are mutually orthonormal.
- Thus,

$$q_i'q_j = 0 \quad \forall i \neq j$$

and

$$\|q_i\| = 1 \quad \forall i = 1, \dots, n.$$

$$\therefore Q'Q = I.$$

$$\begin{aligned}\|Q\mathbf{x}\| &= \sqrt{(Q\mathbf{x})'Q\mathbf{x}} \\ &= \sqrt{\mathbf{x}'Q'Q\mathbf{x}} \\ &= \sqrt{\mathbf{x}'I\mathbf{x}} \\ &= \sqrt{\mathbf{x}'\mathbf{x}} \\ &= \|\mathbf{x}\|.\end{aligned}$$

An orthogonal matrix Q is sometimes called a rotation matrix because if a vector x is premultiplied by Q , the result (Qx) is the vector x rotated to a new position in \mathbb{R}^n .

Vector Space

A vector space \mathcal{S} is a set of vectors that is closed under addition (i.e., if $\mathbf{x}_1 \in \mathcal{S}, \mathbf{x}_2 \in \mathcal{S}$, then $\mathbf{x}_1 + \mathbf{x}_2 \in \mathcal{S}$) and closed under scalar multiplication (i.e., if $c \in \mathbb{R}, \mathbf{x} \in \mathcal{S}$, then $c\mathbf{x} \in \mathcal{S}$).

In other words,

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 \in \mathcal{S} \quad \forall c_1, c_2 \in \mathbb{R}; \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}.$$

- Is $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$ a vector space?
- Is $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{1}'\mathbf{x} = 0\}$ a vector space?
- Is $\{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^m\}$ a vector space?
 $n \times m$

- Suppose $\mathbf{y} \in \mathbb{R}^n$, $c \in \mathbb{R}$ and $\|\mathbf{y}\| = 1$, then

$$\begin{aligned}\|\mathbf{c}\mathbf{y}\| &= \sqrt{(\mathbf{c}\mathbf{y})'\mathbf{c}\mathbf{y}} \\ &= \sqrt{c^2\mathbf{y}'\mathbf{y}} \\ &= |c|\|\mathbf{y}\| = |c|.\end{aligned}$$

- Thus $\mathbf{y} \in \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$ does not imply that $c\mathbf{y} \in \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$. Therefore, this set is not a vector space.

- Let

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{1}'\mathbf{x} = 0\}.$$

- Suppose $c_1, c_2 \in \mathbb{R}$ and $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$, then

$$\mathbf{1}'(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1\mathbf{1}'\mathbf{x}_1 + c_2\mathbf{1}'\mathbf{x}_2 = 0.$$

- Thus $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 \in \mathcal{S}$ and it follows that \mathcal{S} is a vector space.

- Let

$$\mathcal{S} = \left\{ \mathbf{Ax} : \mathbf{x} \in \mathbb{R}^m \right\}.$$

- Suppose $c_1, c_2 \in \mathbb{R}$ and $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{S}$.

- $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{S} \implies \exists \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^m \ni$

$$\mathbf{y}_1 = \mathbf{Ax}_1 \text{ and } \mathbf{y}_2 = \mathbf{Ax}_2.$$

- Thus,

$$c_1\mathbf{y}_1 + c_2\mathbf{y}_2 = c_1\mathbf{Ax}_1 + c_2\mathbf{Ax}_2 = \mathbf{A}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2).$$

$\therefore c_1\mathbf{x}_1 + c_2\mathbf{x}_2 \in \mathbb{R}^m, c_1\mathbf{y}_1 + c_2\mathbf{y}_2 \in \mathcal{S}$. It follows that \mathcal{S} is a vector space.

Generators of a Vector Space

A vector space \mathcal{S} is said to be generated by a set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ if

$$\mathbf{x} \in \mathcal{S} \implies \mathbf{x} = \sum_{i=1}^n c_i \mathbf{x}_i \text{ for some } c_1, \dots, c_n \in \mathbb{R}.$$

Span of a Set of Vectors

- The span of vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ is the set of all LC of $\mathbf{x}_1, \dots, \mathbf{x}_n$, i.e.,

$$\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_n\} = \left\{ \sum_{i=1}^n c_i \mathbf{x}_i : c_1, \dots, c_n \in \mathbb{R} \right\}.$$

- $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is the vector space generated by $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Find a set of vectors that generates the space

$$\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{1}'\mathbf{x} = 0\};$$

i.e., find a set of vectors whose span is

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{1}'\mathbf{x} = 0\}.$$

- Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. Note that $\mathbf{1}'\mathbf{x}_1 = 0$ and $\mathbf{1}'\mathbf{x}_2 = 0$.

Thus, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}$ so that $\text{span}\{\mathbf{x}_1, \mathbf{x}_2\} \subseteq \mathcal{S}$.

- Now suppose $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathcal{S}$.

Then $0 = \mathbf{1}'\mathbf{y} = y_1 + y_2 + y_3 \implies y_3 = -y_1 - y_2$ so that

$$y_1\mathbf{x}_1 + y_2\mathbf{x}_2 = \begin{bmatrix} y_1 \\ 0 \\ -y_1 \end{bmatrix} + \begin{bmatrix} 0 \\ y_2 \\ -y_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ -y_1 - y_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathbf{y}.$$

$\therefore \mathcal{S} \subseteq \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$, and $\mathcal{S} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$.

Basis of a Vector Space

If a vector space \mathcal{S} is generated by LI vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, then $\mathbf{x}_1, \dots, \mathbf{x}_n$ form a basis for \mathcal{S} .

Fact V2:

Suppose $\mathbf{a}_1, \dots, \mathbf{a}_n$ form a basis for a vector space \mathcal{S} . If $\mathbf{b}_1, \dots, \mathbf{b}_k$ are LI vectors in \mathcal{S} , then $k \leq n$.

Proof of Fact V2:

- Because $\mathbf{a}_1, \dots, \mathbf{a}_n$ form a basis for \mathcal{S} and $\mathbf{b}_1 \in \mathcal{S}$, $\mathbf{b}_1 = \sum_{i=1}^n c_i \mathbf{a}_i$ for some $c_1, \dots, c_n \in \mathbb{R}$. Thus, $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}_1$ are LD by Fact V1.
- Again, using V1, we have \mathbf{a}_j a LC of $\mathbf{b}_1, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}$ for some $j \in \{1, 2, \dots, n\}$.

- Thus, $\mathbf{b}_1, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n$ generate \mathcal{S} . It follows that $\mathbf{b}_1, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n, \mathbf{b}_2$ is a LD set of vectors by V1.
- Again by V1, one of the vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n$ is a LC of the preceding vectors. It is not $\mathbf{b}_2 \because \mathbf{b}_1, \dots, \mathbf{b}_k$ are LI.

- Thus $\mathbf{b}_1, \mathbf{b}_2$ and $n - 2$ of $\mathbf{a}_1, \dots, \mathbf{a}_n$ generate \mathcal{S} .
- If $k > n$, we can continue adding \mathbf{b} vectors and deleting \mathbf{a} vectors to get $\mathbf{b}_1, \dots, \mathbf{b}_n$ generates \mathcal{S} . However, then V1 would imply $\mathbf{b}_1, \dots, \mathbf{b}_{n+1}$ are LD. This contradicts LI of $\mathbf{b}_1, \dots, \mathbf{b}_k \therefore k \leq n$. □

Fact V3:

If $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ and $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ each provide a basis for a vector space \mathcal{S} , then $n = k$.

Proof: By V2, we have $k \leq n$ and $n \leq k$. $\therefore k = n$. □

Dimension of a Vector Space

A basis for a vector space is not unique, but the number of vectors in the basis, known as dimension of the vector space, is unique.

Find $\dim(\mathcal{S})$, the dimension of vector space \mathcal{S} , for

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{1}'\mathbf{x} = 0\}.$$

- As demonstrated previously,

$$\text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\} = \mathcal{S}.$$

- Because $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ is a LI set of vectors, $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ forms a basis for \mathcal{S} .

- Thus,

$$\dim(\mathcal{S}) = 2 \quad (\text{even though dimension of vectors in } \mathcal{S} \text{ is } 3).$$

Consider the set $\{\mathbf{0}\}_{n \times 1}$. Is this a vector space? If so, what is its dimension?

- $\{\mathbf{0}\}$ is a vector space because

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0} \in \{\mathbf{0}\} \quad \forall c_1, c_2 \in \mathbb{R} \text{ and } \forall \mathbf{x}_1, \mathbf{x}_2 \in \{\mathbf{0}\}.$$

- The set $\{\mathbf{0}\}$ generates the vector space $\{\mathbf{0}\}$. However, $\{\mathbf{0}\}$ is not a LI set of vectors. $\dim(\{\mathbf{0}\}) = 0$.

Fact V4:

Suppose $\mathbf{a}_1, \dots, \mathbf{a}_n$ are LI vectors in a vector space \mathcal{S} with dimension n .
Then $\mathbf{a}_1, \dots, \mathbf{a}_n$ form a basis for \mathcal{S} .

Proof of Fact V4:

- It suffices to show that $\mathbf{a}_1, \dots, \mathbf{a}_n$ generate \mathcal{S} .
- Let \mathbf{a} denote an arbitrary vector in \mathcal{S} .
- By V2, $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{a}$ are LD. By V1, $\mathbf{a} = \sum_{i=1}^n c_i \mathbf{a}_i$ for some $c_1, \dots, c_n \in \mathbb{R}$.
- Thus

$$\mathcal{S} = \left\{ \sum_{i=1}^n c_i \mathbf{x}_i : c_1, \dots, c_n \in \mathbb{R} \right\},$$

and the result follows. □

Fact V5:

If $\mathbf{a}_1, \dots, \mathbf{a}_k$ are LI vectors in an n -dimensional vector space \mathcal{S} , then there exists a basis for \mathcal{S} that contains $\mathbf{a}_1, \dots, \mathbf{a}_k$.

Proof of Fact V5:

- $k \leq n$ by V2.
- If $k = n$, the result follows from V4.
- Suppose $k < n$. Then, there exist $\mathbf{a}_{k+1} \in \mathcal{S}$ such that $\mathbf{a}_1, \dots, \mathbf{a}_{k+1}$ are LI. Because if not, $\mathbf{a}_1, \dots, \mathbf{a}_k$ would generate \mathcal{S} (by V1), and thus be a basis of dimension $k < n$, which is impossible by V3. Similarly, we can continue to add vectors to $\{\mathbf{a}_1, \dots, \mathbf{a}_{k+1}\}$ until we have $\mathbf{a}_1, \dots, \mathbf{a}_n$ LI vectors. The result follows from V4. \square

Fact V6:

If $\mathbf{a}_1, \dots, \mathbf{a}_k$ are LI and orthonormal vectors in \mathbb{R}^n , then there exist $\mathbf{a}_{k+1}, \dots, \mathbf{a}_n$ such that $\mathbf{a}_1, \dots, \mathbf{a}_n$ are LI and orthonormal.

Proof: HW problem.



Rank of a Matrix

It can be shown that

- the (maximum) number of LI rows of a matrix \mathbf{A} is the same as the (maximum) number of LI columns of \mathbf{A} .
 $m \times n$
- This number of LI rows or columns is known as the rank of \mathbf{A} and is denoted $rank(\mathbf{A})$ or $r(\mathbf{A})$.
 $m \times n$

- If $r(\mathbf{A}) = m$, \mathbf{A} is said to have full row rank.
- If $r(\mathbf{A}) = n$, \mathbf{A} is said to have full column rank.

Inverse of a Matrix

- If $r(\mathbf{A}) = n$, there exists a matrix \mathbf{B} such that $\mathbf{A} \mathbf{B} = \mathbf{I}$.
 $n \times n$ $n \times n$ $n \times n$ $n \times n$
- Such a matrix \mathbf{B} is called the inverse of \mathbf{A} and is denoted \mathbf{A}^{-1} .

• Prove that $r(\mathbf{A}) = n \iff \exists \mathbf{B} \ni \mathbf{A} \mathbf{B} = \mathbf{I}$.

• Prove that $\mathbf{A} \mathbf{B} = \mathbf{I} \implies \mathbf{B} \mathbf{A} = \mathbf{I}$

• Thus $\mathbf{A} \mathbf{A}^{-1} = \mathbf{A}^{-1} \mathbf{A} = \mathbf{I}$.

Proof:

(\implies):

- The columns of A form a basis for \mathbb{R}^n by V4. Thus, there exists a LC of columns of A that equals e_i for all $i = 1, \dots, n$, where e_i is the i^{th} column of the identity matrix I .
- Let b_i denote the coefficients of the LC of the columns of A that yields e_i . Then, with $B = [b_1, \dots, b_n]$, we have $AB = [Ab_1, \dots, Ab_n] = [e_1, \dots, e_n] = I$.

(\Leftarrow):

- If $\exists \underset{n \times n}{\mathbf{B}} \ni \underset{n \times n}{\mathbf{A}} \underset{n \times n}{\mathbf{B}} = \underset{n \times n}{\mathbf{I}}$, then the columns of \mathbf{A} generate \mathbb{R}^n
 $\therefore \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{B}\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$.
- If the columns of \mathbf{A} were LD, then a subset of the columns of \mathbf{A} would be LI and also generate \mathbb{R}^n .

- However, such a subset would be a basis for \mathbb{R}^n and thus must have n elements.

- Thus, the columns of A must be LI. Hence, $r_{n \times n}(A) = n$. □



$$\begin{aligned}
 \underset{n \times n}{A} \underset{n \times n}{B} = \underset{n \times n}{I} &\implies \text{Columns of } \underset{n \times n}{A} \text{ are LI} \\
 &\implies \text{Rows of } \underset{n \times n}{A} \text{ are LI} \\
 &\implies \text{Rows of } \underset{n \times n}{A} \text{ are a basis for } \mathbb{R}^n \\
 &\implies \exists \underset{n \times n}{C} \ni \underset{n \times n}{C} \underset{n \times n}{A} = \underset{n \times n}{I}.
 \end{aligned}$$

- Thus,

$$AB = I \implies CAB = CI$$

$$\implies IB = C$$

$$\implies B = C.$$

- $\therefore AA^{-1} = A^{-1}A = I$



Singular / Nonsingular Matrix

- If $r(\mathbf{A}) = n$, \mathbf{A} is said to be nonsingular.
- If $r(\mathbf{A}) < n$, \mathbf{A} is said to be singular.

Inverse of a Nonsingular 2×2 Matrix

- $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ if $ad - bc \neq 0$.

- $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is singular if $ad - bc = 0$.

Column Space of a Matrix

The column space of a matrix \mathbf{A} , denoted by $\mathcal{C}(\mathbf{A})$, is the vector space generated by the columns of \mathbf{A} ; i.e.,

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{Ax} : \mathbf{x} \in \mathbb{R}^n\}.$$

$\dim(\mathcal{C}(\mathbf{A})) = r(\mathbf{A})$ because the largest possible subset of LI columns of \mathbf{A} is a basis for $\mathcal{C}(\mathbf{A})$.

- Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 4 & 2 & 8 \end{bmatrix} .$$

- Find $r(\mathbf{A})$.
- Give a basis for $\mathcal{C}(\mathbf{A})$.
- Characterize $\mathcal{C}(\mathbf{A})$.

- $$3 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 8 \end{bmatrix} .$$

- Thus, the columns of A are LD and $r(A) < 3$.

- $$c_1 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} c_1 \\ 2c_1 + c_2 \\ 4c_1 + 2c_2 \end{bmatrix} = \mathbf{0} \implies c_1 = c_2 = 0. \therefore r(A) = 2.$$

- A basis for $\mathcal{C}(\mathbf{A})$ is given by $\left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$.

- $\mathbf{x} \in \mathcal{C}(\mathbf{A}) \implies \mathbf{x} = c_1 \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ for some $c_1, c_2 \in \mathbb{R}^n$.

- Note

$$\left\{ \mathbf{x} = \begin{bmatrix} c_1 \\ 2c_1 + c_2 \\ 4c_1 + 2c_2 \end{bmatrix} : c_1, c_2 \in \mathbb{R} \right\} = \mathcal{C}(\mathbf{A})$$

is the set of vectors in \mathbb{R}^3 where the first component is arbitrary and the third component is twice the second component, i.e.,

$$\{\mathbf{x} \in \mathbb{R}^3 : 2x_2 = x_3\}.$$

Result A.1:

$$\mathit{rank}(\mathbf{AB}) \leq \min\{\mathit{rank}(\mathbf{A}), \mathit{rank}(\mathbf{B})\}$$

Proof of Result A.1:

- Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ denote the columns of \mathbf{B} so that $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_n]$.
- Then $\mathbf{AB} = [\mathbf{Ab}_1, \dots, \mathbf{Ab}_n]$. This implies that the columns of \mathbf{AB} are in $\mathcal{C}(\mathbf{A})$.

- $\dim(\mathcal{C}(\mathbf{A}))$ is $\text{rank}(\mathbf{A})$.
- There does not exist a set of LI vectors in $\mathcal{C}(\mathbf{A})$ with more than $\text{rank}(\mathbf{A})$ vectors by Fact V2.
- It follows that $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$.

- It remains to show that

$$\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{B}).$$

- Note that $\text{rank}(\mathbf{AB})$ is the same as $\text{rank}((\mathbf{AB})') = \text{rank}(\mathbf{B}'\mathbf{A}')$.
- Our previous argument shows that

$$\text{rank}(\mathbf{B}'\mathbf{A}') \leq \text{rank}(\mathbf{B}') = \text{rank}(\mathbf{B}).$$



Provide an example where

$$\text{rank}(\mathbf{AB}) < \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}.$$



$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

- Then

$$\mathbf{AB} = \mathbf{0}.$$

- Therefore,

$$\text{rank}(\mathbf{AB}) = 0, \quad \text{rank}(\mathbf{A}) = 1, \quad \text{rank}(\mathbf{B}) = 1.$$

Result A.2:

- (a) If $A = BC$, then $\mathcal{C}(A) \subseteq \mathcal{C}(B)$,
- (b) If $\mathcal{C}(A) \subseteq \mathcal{C}(B)$, then there exist C such that $A = BC$.

Proof of A.2(a):

- Suppose $x \in \mathcal{C}(A)$ then $\exists y \ni x = Ay$.
- Now $A = BC \implies x = BCy$.
- Thus, $\exists z \ni x = Bz$ (namely, $z = Cy$). $\therefore x \in \mathcal{C}(B)$.
- We have shown $x \in \mathcal{C}(A) \implies x \in \mathcal{C}(B)$. $\therefore \mathcal{C}(A) \subseteq \mathcal{C}(B)$. □

Proof of A.2(b):

- Let $\mathbf{a}_1, \dots, \mathbf{a}_n$ denote the columns of \mathbf{A} .
 $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B}) \implies \mathbf{a}_1, \dots, \mathbf{a}_n \in \mathcal{C}(\mathbf{B})$.
- Let \mathbf{c}_i be such that $\mathbf{B}\mathbf{c}_i = \mathbf{a}_i \forall i = 1, \dots, n$. Then denote $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_n]$.
- It follows that

$$\begin{aligned}\mathbf{BC} &= \mathbf{B}[\mathbf{c}_1, \dots, \mathbf{c}_n] \\ &= [\mathbf{B}\mathbf{c}_1, \dots, \mathbf{B}\mathbf{c}_n] \\ &= [\mathbf{a}_1, \dots, \mathbf{a}_n] = \mathbf{A}.\end{aligned}$$



Null Space of a Matrix

- The null space of a matrix \mathbf{A} , denoted $\mathcal{N}(\mathbf{A})$ is defined as

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{y} : \mathbf{A}\mathbf{y} = \mathbf{0}\}.$$

- Note that $\mathcal{N}(\mathbf{A})$ is the set of vectors orthogonal to every row of \mathbf{A} .

A vector in $\mathcal{N}(\mathbf{A})$ can also be seen as a vector of coefficients corresponding to a LC of the columns of \mathbf{A} that is $\mathbf{0}$.

Note that if \mathbf{A} has dimension $m \times n$, then the vectors in $\mathcal{C}(\mathbf{A})$ have dimension m and the vectors in $\mathcal{N}(\mathbf{A})$ have dimension n .

Is the null space of a matrix A a vector space?
 $m \times n$

- Yes.
- Suppose $\mathbf{x} \in \mathcal{N}(\mathbf{A})$. Then $\forall c \in \mathbb{R}, \mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = c\mathbf{0} = \mathbf{0}$. Thus $\mathbf{x} \in \mathcal{N}(\mathbf{A}) \implies c\mathbf{x} \in \mathcal{N}(\mathbf{A}) \quad \forall c \in \mathbb{R}$.
- Suppose $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$. Then $\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}\mathbf{x}_1 + \mathbf{A}\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Thus, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(\mathbf{A}) \implies \mathbf{x}_1 + \mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$.

Find the null space of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ -1 & -2 \\ 3 & 6 \end{bmatrix}.$$

$$\mathbf{A}\mathbf{y} = \mathbf{0} \implies \begin{cases} y_1 + 2y_2 = 0 \\ 2y_1 + 4y_2 = 0 \\ -y_1 - 2y_2 = 0 \\ 3y_1 + 6y_2 = 0 \end{cases}$$

$$\implies y_1 = -2y_2 \implies \mathcal{N}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^2 : y_1 = -2y_2\}.$$

Result A.3:

$$\text{rank}_{m \times n}(\mathbf{A}) = n \iff \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}.$$

Proof of Result A.3:

- Let $[\mathbf{a}_1, \dots, \mathbf{a}_n] = \mathbf{A}$. Then

$$\begin{aligned}\mathbf{A}\mathbf{y} &= [\mathbf{a}_1, \dots, \mathbf{a}_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\ &= y_1\mathbf{a}_1 + \dots + y_n\mathbf{a}_n.\end{aligned}$$



$$\begin{aligned}r(\mathbf{A})_{m \times n} = n &\iff \mathbf{a}_1, \dots, \mathbf{a}_n \text{ are LI} \\ &\iff \mathbf{A}\mathbf{y} = \mathbf{0} \text{ only if } \mathbf{y} = \mathbf{0} \\ &\iff \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}.\end{aligned}$$



Theorem A.1:

If the matrix \mathbf{A} is $m \times n$ with rank r , then

$$\dim(\mathcal{N}(\mathbf{A})) = n - r,$$

or more elegantly,

$$\dim(\mathcal{N}(\mathbf{A})) + \dim(\mathcal{C}(\mathbf{A})) = n.$$

Proof of Theorem A.1:

- Let $k = \dim(\mathcal{N}(A))$. Results A.3 covers the case where $k = 0$.
Suppose now that $k > 0$.

- Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ form a basis for $\mathcal{N}(A)$. Then

$$\underset{m \times n}{A} \mathbf{u}_i = \mathbf{0} \quad \forall i = 1, \dots, k.$$

- By Fact V5, there exist $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ such that $\mathbf{u}_1, \dots, \mathbf{u}_n$ form a basis for \mathbb{R}^n .

- We will now argue that the $n - k$ vectors $\mathbf{A}\mathbf{u}_{k+1}, \dots, \mathbf{A}\mathbf{u}_n$ form a basis for $\mathcal{C}(\mathbf{A})$.
- If so, then

$$\dim(\mathcal{N}(\mathbf{A})) + \dim(\mathcal{C}(\mathbf{A})) = k + n - k = n,$$

i.e.,

$$\dim(\mathcal{N}(\mathbf{A})) = n - \dim(\mathcal{C}(\mathbf{A})) = n - r.$$

- First note that $\mathbf{A}\mathbf{u}_i \in \mathcal{C}(\mathbf{A}) \quad \forall i = k + 1, \dots, n.$
- Now note that

$$c_{k+1}\mathbf{A}\mathbf{u}_{k+1} + \cdots + c_n\mathbf{A}\mathbf{u}_n = \mathbf{0}$$

$$\implies \mathbf{A}(c_{k+1}\mathbf{u}_{k+1} + \cdots + c_n\mathbf{u}_n) = \mathbf{0}$$

$$\implies c_{k+1}\mathbf{u}_{k+1} + \cdots + c_n\mathbf{u}_n \in \mathcal{N}(\mathbf{A})$$

$$\implies \exists c_1, \dots, c_k \in \mathbb{R} \ni c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \sum_{j=k+1}^n c_j\mathbf{u}_j$$

$$\implies c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k - c_{k+1}\mathbf{u}_{k+1} - \cdots - c_n\mathbf{u}_n = \mathbf{0}$$

$$\implies c_1 = \cdots = c_n = 0 \text{ by LI of } \mathbf{u}_1, \dots, \mathbf{u}_n.$$

- Therefore, $A\mathbf{u}_{k+1}, \dots, A\mathbf{u}_n$ are LI.
- Now let $U = [\mathbf{u}_1, \dots, \mathbf{u}_n]$.
- Because $\mathbf{u}_1, \dots, \mathbf{u}_n$ are LI and a basis for \mathbb{R}^n , $\exists U^{-1} \ni UU^{-1} = I$.
- Let $\mathbf{x} \in \mathbb{R}^n$ be arbitrary and define $\mathbf{z} = U^{-1}\mathbf{x}$.

- Then

$$\begin{aligned}\mathbf{Ax} &= \mathbf{A} \mathbf{U} \mathbf{U}^{-1} \mathbf{x} = \mathbf{A} \mathbf{U} \mathbf{z} \\ &= [\mathbf{A} \mathbf{u}_1, \dots, \mathbf{A} \mathbf{u}_k, \mathbf{A} \mathbf{u}_{k+1}, \dots, \mathbf{A} \mathbf{u}_n] \mathbf{z} \\ &= [\mathbf{0}, \dots, \mathbf{0}, \mathbf{A} \mathbf{u}_{k+1}, \dots, \mathbf{A} \mathbf{u}_n] \mathbf{z} \\ &= z_{k+1} \mathbf{A} \mathbf{u}_{k+1} + \dots + z_n \mathbf{A} \mathbf{u}_n.\end{aligned}$$

- Therefore, any vector in $\mathcal{C}(\mathbf{A})$ can be written as a LC of $\mathbf{A} \mathbf{u}_{k+1}, \dots, \mathbf{A} \mathbf{u}_n$.

- It follows that

$\mathbf{A}u_{k+1}, \dots, \mathbf{A}u_n$ is a basis for $\mathcal{C}(\mathbf{A})$.

- $\therefore n - k = r$ and $k + r = n$.

