Introduction to Linear Algebra for Statistics

Notation Used in These Notes

$\forall$	for all
$\exists$	there exists
$\ni$	such that
$\therefore$	therefore
$\because$	because
$\square$	end of proof

More Notation

$C_1 \implies C_2$ Condition C_1 implies condition C_2.

$C_1 \iff C_2$ C_1 is true if and only if (iff) C_2 is true.
(C_1 and C_2 are equivalent.)

$x \in \mathcal{S}$ x is an element of the set $\mathcal{S}$.

More Notation

$\mathcal{S}_1 \subset \mathcal{S}_2$	$\mathcal{S}_1$ is a proper subset of $\mathcal{S}_2$ (Every element of $\mathcal{S}_1$ is also in $\mathcal{S}_2$, but $\mathcal{S}_2$ has at least one element not in $\mathcal{S}_1$.)
$\mathcal{S}_1 \subseteq \mathcal{S}_2$	$\mathcal{S}_1$ is a subset of $\mathcal{S}_2$. (Every element in $\mathcal{S}_1$ is also in $\mathcal{S}_2$, and the sets may be exactly the same.)
$\mathbb{R}^n$	Euclidean n-space

Matrix Notation

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

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

- Throughout these slides, we consider the case where $a_{ij} \in \mathbb{R}$ $\forall\ i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Vectors

- A matrix with one column is called a <u>vector</u>. For example,

$$\begin{bmatrix} 4 \\ 3 \\ -9 \end{bmatrix} \text{ is a vector.}$$

- A matrix with one row is called a <u>row vector</u>. For example,

$$\begin{bmatrix} 2 & 4 & -2 & 8 & 1 \end{bmatrix} \text{ is a row vector.}$$

In these notes,

Matrices are represented with bold uppercase letters.

Vectors are represented with bold lowercase letters.

Some Special Vectors

- $\mathbf{0}$ (or $\mathbf{0}_n$) is a vector of (n) zeros.
- $\mathbf{1}$ (or $\mathbf{1}_n$) is a vector of (n) ones.
- For example,

$$\mathbf{0}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{1}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Square Matrices

- Matrix $\underset{m \times n}{A}$ is said to be square iff $m = n$.

- In other words, a matrix is square if and only if its number of rows is the same as its number of columns.

Special Types of Square Matrices

- A square matrix $\underset{n\times n}{A}$ is upper triangular if $a_{ij} = 0, \forall\ i > j$.
- A square matrix $\underset{n\times n}{A}$ is lower triangular if $a_{ij} = 0, \forall\ i < j$.
- A square matrix $\underset{n\times n}{A}$ 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 $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -13 \end{bmatrix}$.

Identity Matrices

- We use $\boldsymbol{I}$ (or $\boldsymbol{I}_n$ or $\underset{n\times n}{\boldsymbol{I}}$) to denote the $(n \times n)$ identity matrix, which is the diagonal matrix with all (n) ones on the diagonal.

- For example,

$$\boldsymbol{I}_3 = \boldsymbol{I}_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrix Transpose

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

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

A Symmetric Matrix

A square matrix $\underset{n \times n}{A}$ is <u>symmetric</u> if $A = A'$.

Transpose and Symmetric Matrix Examples

- Find the transpose of
$$\begin{bmatrix} 4 & -2 \\ 3 & 7 \end{bmatrix}.$$

- Provide an example of a symmetric matrix.

Transpose and 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 Addition

Suppose

$$\underset{m\times n}{\boldsymbol{A}} = \underset{m\times n}{[a_{ij}]} \text{ and } \underset{m\times n}{\boldsymbol{B}} = \underset{m\times n}{[b_{ij}]}.$$

Then

$$\boldsymbol{A} + \boldsymbol{B} = [a_{ij} + b_{ij}].$$

That is, $\boldsymbol{A} + \boldsymbol{B}$ is the $m \times n$ matrix $\boldsymbol{C}$, whose i, jth entry c_{ij} is equal to $a_{ij} + b_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Matrix Addition Example

$$\begin{bmatrix} 8 & -2 \\ 5 & 1 \\ -2 & 4 \end{bmatrix} + \begin{bmatrix} -4 & 2 \\ -2 & 5 \\ 7 & -1 \end{bmatrix} = \begin{bmatrix} 8-4 & -2+2 \\ 5-2 & 1+5 \\ -2+7 & 4-1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 3 & 6 \\ 5 & 3 \end{bmatrix}$$

Scalar Multiplication of a Matrix

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

$$\underset{m\times n}{c\boldsymbol{A}} = c\underset{m\times n}{[a_{ij}]} = \underset{m\times n}{[ca_{ij}]} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.$$

Example Scalar Multiplication of a Matrix

$$3\begin{bmatrix} 8 & -2 \\ 5 & 1 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 24 & -6 \\ 15 & 3 \\ -6 & 12 \end{bmatrix}$$

Matrix Multiplication

Suppose

$$\underset{m\times n}{\boldsymbol{A}} = [a_{il}] = \begin{bmatrix} \boldsymbol{a}'_{(1)} \\ \vdots \\ \boldsymbol{a}'_{(m)} \end{bmatrix} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$$

and

$$\underset{n\times k}{\boldsymbol{B}} = [b_{lj}] = \begin{bmatrix} \boldsymbol{b}'_{(1)} \\ \vdots \\ \boldsymbol{b}'_{(n)} \end{bmatrix} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k].$$

In other words, the i, lth element of $\boldsymbol{A}$ is denoted a_{il}, the ith row of $\boldsymbol{A}$ is denoted $\boldsymbol{a}'_{(i)}$, and the lth column of $\boldsymbol{A}$ is denoted $\boldsymbol{a}_l$ (and analogously for the elements, rows, and columns of $\boldsymbol{B}$).

Matrix Multiplication

Then

$$\underset{m\times n}{\boldsymbol{A}}\ \underset{n\times k}{\boldsymbol{B}} = \underset{m\times k}{\boldsymbol{C}} = \left[c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}\right] = \left[c_{ij} = \boldsymbol{a}'_{(i)}\boldsymbol{b}_j\right]$$

$$= [\boldsymbol{A}\boldsymbol{b}_1, \ldots, \boldsymbol{A}\boldsymbol{b}_k] = \begin{bmatrix} \boldsymbol{a}'_{(1)}\boldsymbol{B} \\ \vdots \\ \boldsymbol{a}'_{(m)}\boldsymbol{B} \end{bmatrix} = \sum_{l=1}^{n} \boldsymbol{a}_l \boldsymbol{b}'_{(l)}.$$

(Note the many equivalent ways to think about and compute a matrix product.)

Matrix Multiplication

- Suppose
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$
and
$$B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$
- Work out AB using $AB = \sum_{l=1}^{n} a_l b'_{(l)}$.

Matrix Multiplication

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

Transpose of a Matrix Product

- $$(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$$

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

Linear Combination

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

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

is a <u>linear combination</u> (LC) of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

The <u>coefficients</u> of the LC are $c_1, \ldots, c_n$.

Linear Independence and Linear Dependence

- The vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent (LI) iff

$$\sum_{i=1}^{n} c_i \boldsymbol{x}_i = \mathbf{0} \text{ only when } c_1 = \cdots = c_n = 0.$$

- The vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly dependent (LD) iff

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

Prove or disprove: If one or more of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is the vector $\mathbf{0}$, the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are LD.

- Suppose $\boldsymbol{x}_j = \mathbf{0}$ for some $j \in \{1, \ldots, n\}$.

- If we take $c_j = 1$ and $c_k = 0$ for any $k \neq j$, then $\sum_{i=1}^{n} c_i \boldsymbol{x}_i = \mathbf{0}$ and $c_1, \ldots, c_n$ are not all zero.

- Thus, vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are LD if any of the vectors are $\mathbf{0}$. $\square$

Prove or disprove: The following vectors are LI.

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

- 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 vectors are LD. $\square$

- 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 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are LD $\Longleftrightarrow$ $\boldsymbol{x}_j$ is a LC of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{j-1}$ for some $j \in \{2, \ldots, n\}$.

Proof of Fact V1:

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

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

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

$$\sum_{i=1}^{j} c_i \boldsymbol{x}_i = \boldsymbol{0} \implies \sum_{i=1}^{j-1} c_i \boldsymbol{x}_i = -c_j \boldsymbol{x}_j.$$

$$\implies \sum_{i=1}^{j-1} \frac{-c_i}{c_j} \boldsymbol{x}_i = \boldsymbol{x}_j.$$

($\Longleftarrow$) Suppose $\boldsymbol{x}_j = \sum_{i=1}^{j-1} c_i \boldsymbol{x}_i$, then $\sum_{i=1}^n d_i \boldsymbol{x}_i = \boldsymbol{0}$, 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 $\boldsymbol{x}, \boldsymbol{y}$ are orthogonal to each other if their inner product is zero, i.e.,

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

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

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

- The vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are <u>mutually orthogonal</u> if

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

- The vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are <u>mutually orthonormal</u> if

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

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

- Write down a list 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 $\boldsymbol{Q}$ is orthogonal, then $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$.

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

- $\boldsymbol{Q}'\boldsymbol{Q} = [\boldsymbol{q}_i'\boldsymbol{q}_j]$, where $\boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n]$.
- By orthogonality of $\boldsymbol{Q}$, $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are mutually orthonormal.
- Thus,

$$\boldsymbol{q}_i'\boldsymbol{q}_j = 0 \quad \forall\, i \neq j$$

and

$$\|\boldsymbol{q}_i\| = 1 \quad \forall\, i = 1, \ldots, n.$$

$\therefore \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$.

$$
\begin{aligned}
\|\boldsymbol{Qx}\| &= \sqrt{(\boldsymbol{Qx})'\boldsymbol{Qx}} \\
&= \sqrt{\boldsymbol{x}'\boldsymbol{Q}'\boldsymbol{Qx}} \\
&= \sqrt{\boldsymbol{x}'\boldsymbol{Ix}} \\
&= \sqrt{\boldsymbol{x}'\boldsymbol{x}} \\
&= \|\boldsymbol{x}\|.
\end{aligned}
$$

An orthogonal matrix $\boldsymbol{Q}$ is sometimes called a rotation matrix because if a vector $\underset{n\times 1}{\boldsymbol{x}}$ is premultiplied by $\underset{n\times n}{\boldsymbol{Q}}$, the result ($\boldsymbol{Qx}$) is the vector $\boldsymbol{x}$ rotated to a new position in $\mathbb{R}^n$.

Vector Space in $\mathbb{R}^n$

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

In other words,

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

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

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

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

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

- Let

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

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

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

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

- Let

$$\mathcal{S} = \left\{ \underset{n\times m}{\boldsymbol{A}} \boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^m \right\}.$$

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

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

$$\boldsymbol{y}_1 = \boldsymbol{A}\boldsymbol{x}_1 \text{ and } \boldsymbol{y}_2 = \boldsymbol{A}\boldsymbol{x}_2.$$

- Thus,

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

  $\because c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 \in \mathbb{R}^m, c_1\boldsymbol{y}_1 + c_2\boldsymbol{y}_2 \in \mathcal{S}$. 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 the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ if

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

Span of Vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$

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

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

- $span\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is the vector space generated by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

Find vectors that generates the space

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

i.e., find vectors whose span is

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

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

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

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

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

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

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

Basis of a Vector Space

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

Fact V2:

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

Proof of Fact V2:

- Because $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ form a basis for $\mathcal{S}$ and $\boldsymbol{b}_1 \in \mathcal{S}$, $\boldsymbol{b}_1 = \sum_{i=1}^{n} c_i \boldsymbol{a}_i$ for some $c_1, \ldots, c_n \in \mathbb{R}$. Thus, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n, \boldsymbol{b}_1$ are LD by Fact V1.

- Again, using V1, we have $\boldsymbol{a}_j$ a LC of $\boldsymbol{b}_1, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}$ for some $j \in \{1, 2, \ldots, n\}$.

- Thus, $\boldsymbol{b}_1, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n$ generate $\mathcal{S}$. It follows that $\boldsymbol{b}_1, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n, \boldsymbol{b}_2$ are LD vectors by V1.

- Again by V1, one of the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n$ is a LC of the preceding vectors. It is not $\boldsymbol{b}_2 \because \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ are LI.

- Thus $\boldsymbol{b}_1, \boldsymbol{b}_2$ and $n-2$ of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ generate $\mathcal{S}$.

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

Fact V3:

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

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

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

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

- As demonstrated previously,

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

- Because $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ are LI vectors, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ 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 $\{\underset{n\times 1}{\mathbf{0}}\}$. Is this a vector space? If so, what is its dimension?

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

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

- The vector $\mathbf{0}$ generates the vector space $\{\mathbf{0}\}$. However, $\mathbf{0}$ is not a LI list of vectors and thus not a basis. By convention, we say $dim(\{\mathbf{0}\}) = 0$.

Fact V4:

Suppose $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are LI vectors in a vector space $\mathcal{S}$ with dimension n.

Then $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ form a basis for $\mathcal{S}$.

Proof of Fact V4:

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

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

and the result follows. $\square$

Fact V5:

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

Proof of Fact V5:

- $k \leq n$ by V2.
- If $k = n$, the result follows from V4.
- Suppose $k < n$. Then, there exist $\boldsymbol{a}_{k+1} \in \mathcal{S}$ such that $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k+1}$ are LI. Because if not, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ 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 $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k+1}\}$ until we have $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ LI vectors. The result follows from V4. $\square$

Fact V6:

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

Proof: Try to come up with it on your own.

Rank of a Matrix

It can be shown that

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

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

Inverse of a Matrix

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

- Prove that $r(\underset{n\times n}{\boldsymbol{A}}) = n \iff \exists \underset{n\times n}{\boldsymbol{B}} \ni \underset{n\times n}{\boldsymbol{A}}\underset{n\times n}{\boldsymbol{B}} = \underset{n\times n}{\boldsymbol{I}}$.

- Prove that $\underset{n\times n}{\boldsymbol{A}}\underset{n\times n}{\boldsymbol{B}} = \underset{n\times n}{\boldsymbol{I}} \Longrightarrow \underset{n\times n}{\boldsymbol{B}}\underset{n\times n}{\boldsymbol{A}} = \underset{n\times n}{\boldsymbol{I}}$

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

Proof:

($\Longrightarrow$):

- The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$ by V4. Thus, there exists a LC of columns of $\underset{n\times n}{\boldsymbol{A}}$ that equals $\boldsymbol{e}_i$ for all $i = 1, \ldots, n$, where $\boldsymbol{e}_i$ is the i^{th} column of the identity matrix $\underset{n\times n}{\boldsymbol{I}}$.

- Let $\boldsymbol{b}_i$ denote the coefficients of the LC of the columns of $\boldsymbol{A}$ that yields $\boldsymbol{e}_i$. Then, with $\boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n]$, we have $\boldsymbol{AB} = [\boldsymbol{Ab}_1, \ldots, \boldsymbol{Ab}_n] = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n] = \boldsymbol{I}$.

($\Longleftarrow$):

- If $\exists \underset{n\times n}{\boldsymbol{B}} \ni \underset{n\times n}{\boldsymbol{A}}\,\underset{n\times n}{\boldsymbol{B}} = \underset{n\times n}{\boldsymbol{I}}$, then the columns of $\boldsymbol{A}$ generate $\mathbb{R}^n$
  $\because \forall\, \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{ABx} = \boldsymbol{Ix} = \boldsymbol{x}$.

- If the columns of $\boldsymbol{A}$ were LD, then a subset of the columns of $\boldsymbol{A}$ 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 $\boldsymbol{A}$ must be LI. Hence, $r(\underset{n\times n}{\boldsymbol{A}}) = n$. $\square$
-

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

- Thus,

$$\begin{aligned} AB = I &\implies CAB = CI \\ &\implies IB = C \\ &\implies B = C. \end{aligned}$$

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

Singular / Nonsingular Matrix

- If $r(\underset{n\times n}{\boldsymbol{A}}) = n$, $\underset{n\times n}{\boldsymbol{A}}$ is said to be <u>nonsingular</u>.
- If $r(\underset{n\times n}{\boldsymbol{A}}) < n$, $\underset{n\times n}{\boldsymbol{A}}$ is said to be <u>singular</u>.

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 $\underset{m\times n}{\boldsymbol{A}}$, denoted by $\mathcal{C}(\boldsymbol{A})$, is the vector space generated by the columns of $\underset{m\times n}{\boldsymbol{A}}$; i.e.,

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

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

- Let

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

- Find $r(\boldsymbol{A})$.

- Give a basis for $\mathcal{C}(\boldsymbol{A})$.

- Characterize $\mathcal{C}(\boldsymbol{A})$.

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

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

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

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

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

- Note

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

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

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

Result A.1:

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

Proof of Result A.1:

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

- $dim(\mathcal{C}(\boldsymbol{A}))$ is $rank(\boldsymbol{A})$.

- There does not exist a list of LI vectors in $\mathcal{C}(\boldsymbol{A})$ with more than $rank(\boldsymbol{A})$ vectors by Fact V2.

- It follows that $rank(\boldsymbol{AB}) \leq rank(\boldsymbol{A})$.

- It remains to show that

$$rank(\boldsymbol{AB}) \leq rank(\boldsymbol{B}).$$

- Note that $rank(\boldsymbol{AB})$ is the same as $rank((\boldsymbol{AB})') = rank(\boldsymbol{B}'\boldsymbol{A}')$.

- Our previous argument shows that

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

□

Provide an example where

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

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

- Then
$$\boldsymbol{AB} = \boldsymbol{0}.$$

- Therefore,
$$rank(\boldsymbol{AB}) = 0, \quad rank(\boldsymbol{A}) = 1, \quad rank(\boldsymbol{B}) = 1.$$

Result A.2:

(a) If $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$, then $\mathcal{C}(\boldsymbol{A}) \subseteq \mathcal{C}(\boldsymbol{B})$.

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

Proof of A.2(a):

- Suppose $\boldsymbol{x} \in \mathcal{C}(\boldsymbol{A})$. Then $\exists\, \boldsymbol{y} \ni \boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}$.
- Now $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C} \implies \boldsymbol{x} = \boldsymbol{B}\boldsymbol{C}\boldsymbol{y}$.
- Thus, $\exists\, \boldsymbol{z} \ni \boldsymbol{x} = \boldsymbol{B}\boldsymbol{z}$ (namely, $\boldsymbol{z} = \boldsymbol{C}\boldsymbol{y}$). $\therefore \boldsymbol{x} \in \mathcal{C}(\boldsymbol{B})$.
- We have shown $\boldsymbol{x} \in \mathcal{C}(\boldsymbol{A}) \implies \boldsymbol{x} \in \mathcal{C}(\boldsymbol{B})$. $\therefore \mathcal{C}(\boldsymbol{A}) \subseteq \mathcal{C}(\boldsymbol{B})$. $\square$

Proof of A.2(b):

- Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ denote the columns of $\boldsymbol{A}$.
  $\mathcal{C}(\boldsymbol{A}) \subseteq \mathcal{C}(\boldsymbol{B}) \Longrightarrow \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathcal{C}(\boldsymbol{B})$.

- Let $\boldsymbol{c}_i$ be such that $\boldsymbol{B}\boldsymbol{c}_i = \boldsymbol{a}_i \; \forall \; i = 1, \ldots, n$. Then denote $\boldsymbol{C} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n]$.

- It follows that

$$
\begin{aligned}
\boldsymbol{B}\boldsymbol{C} &= \boldsymbol{B}[\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n] \\
&= [\boldsymbol{B}\boldsymbol{c}_1, \ldots, \boldsymbol{B}\boldsymbol{c}_n] \\
&= [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] = \boldsymbol{A}.
\end{aligned}
$$

□

Null Space of a Matrix

- The <u>null space</u> of a matrix $\boldsymbol{A}$, denoted $\mathcal{N}(\boldsymbol{A})$ is defined as

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

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

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

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

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

- Yes.

- Suppose $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A})$. Then $\forall\, c \in \mathbb{R}, \boldsymbol{A}(c\boldsymbol{x}) = c\boldsymbol{A}\boldsymbol{x} = c\boldsymbol{0} = \boldsymbol{0}$. Thus $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A}) \Longrightarrow c\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A}) \quad \forall\, c \in \mathbb{R}$.

- Suppose $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{N}(\boldsymbol{A})$. Then $\boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$. Thus, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{N}(\boldsymbol{A}) \Longrightarrow \boldsymbol{x}_1 + \boldsymbol{x}_2 \in \mathcal{N}(\boldsymbol{A})$.

Find the null space of

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

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

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

Result A.3:

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

Proof of Result A.3:

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

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

-

$$\begin{aligned} r(\underset{m\times n}{\boldsymbol{A}}) = n &\iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \quad \text{are LI} \\ &\iff \boldsymbol{A}\boldsymbol{y} = \boldsymbol{0} \text{ only if } \boldsymbol{y} = \boldsymbol{0} \\ &\iff \mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{0}\}. \quad \square \end{aligned}$$

Theorem A.1:

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

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

or more elegantly,

$$dim(\mathcal{N}(\underset{m\times n}{\boldsymbol{A}})) + dim(\mathcal{C}(\underset{m\times n}{\boldsymbol{A}})) = n.$$

Proof of Theorem A.1:

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

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

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

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

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

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

i.e.,

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

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

- Now note that

$$
\begin{aligned}
& c_{k+1}\boldsymbol{A}\boldsymbol{u}_{k+1} + \cdots + c_n\boldsymbol{A}\boldsymbol{u}_n = \boldsymbol{0} \\
\Longrightarrow\ & \boldsymbol{A}(c_{k+1}\boldsymbol{u}_{k+1} + \cdots + c_n\boldsymbol{u}_n) = \boldsymbol{0} \\
\Longrightarrow\ & c_{k+1}\boldsymbol{u}_{k+1} + \cdots + c_n\boldsymbol{u}_n \in \mathcal{N}(\boldsymbol{A}) \\
\Longrightarrow\ & \exists\ c_1, \ldots, c_k \in \mathbb{R} \ni c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k = \sum_{j=k+1}^{n} c_j\boldsymbol{u}_j \\
\Longrightarrow\ & c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k - c_{k+1}\boldsymbol{u}_{k+1} - \cdots - c_n\boldsymbol{u}_n = \boldsymbol{0} \\
\Longrightarrow\ & c_1 = \cdots = c_n = 0 \text{ by LI of } \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n.
\end{aligned}
$$

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

- Then

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{AUU}^{-1}\boldsymbol{x} = \boldsymbol{AUz} \\ &= [\boldsymbol{Au}_1, \ldots, \boldsymbol{Au}_k, \boldsymbol{Au}_{k+1}, \ldots, \boldsymbol{Au}_n]\boldsymbol{z} \\ &= [\boldsymbol{0}, \ldots, \boldsymbol{0}, \boldsymbol{Au}_{k+1}, \ldots, \boldsymbol{Au}_n]\boldsymbol{z} \\ &= z_{k+1}\boldsymbol{Au}_{k+1} + \cdots + z_n\boldsymbol{Au}_n. \end{aligned}$$

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

- It follows that

$$\boldsymbol{A}\boldsymbol{u}_{k+1}, \ldots, \boldsymbol{A}\boldsymbol{u}_n \text{ is a basis for } \mathcal{C}(\boldsymbol{A}).$$

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