# Introduction to Linear Algebra 

## George Voutsadakis ${ }^{1}$

${ }^{1}$ Mathematics and Computer Science

Lake Superior State University

## LSSU Math 305

(1) Vector Spaces

- Vector Spaces and Subspaces
- Null Spaces, Column Spaces and Linear Transformations
- Linearly Independent Sets; Bases
- Coordinate Systems
- The Dimension of a Vector Space
- Rank
- Change of Basis


## Subsection 1

## Vector Spaces and Subspaces

## Vector Spaces

## Definition

A vector space is a nonempty set $V$ of objects, called vectors, on which are defined two operations, called addition and multiplication by scalars (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ in $V$ and for all scalars c,d.

1. The sum of $\boldsymbol{u}$ and $\boldsymbol{v}$, denoted by $\boldsymbol{u}+\boldsymbol{v}$, is in $V$.
2. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
3. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
4. There is a zero vector $\mathbf{0}$ in $V$ such that $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$.
5. For each $\boldsymbol{u}$ in $V$, there is a vector $-\boldsymbol{u}$ in $V$ such that $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$.
6. The scalar multiple of $\boldsymbol{u}$ by $c$, denoted by $c \boldsymbol{u}$, is in $V$.
7. $c(\boldsymbol{u}+\boldsymbol{v})=c \boldsymbol{u}+c \boldsymbol{v}$.
8. $(c+d) \boldsymbol{u}=c \boldsymbol{u}+d \boldsymbol{u}$.
9. $c(d \boldsymbol{u})=(c d) \boldsymbol{u}$.
10. $1 \boldsymbol{u}=\boldsymbol{u}$.

## Additional Properties

- The zero vector is unique.
- The vector $-\boldsymbol{u}$, called the negative of $\boldsymbol{u}$, is unique for each $\boldsymbol{u}$ in $V$.
- For each $\boldsymbol{u}$ in $V$ and scalar $c$ :
- $0 u=0$;
- $c \mathbf{0}=\mathbf{0}$;
- $-\boldsymbol{u}=(-1) \boldsymbol{u}$.


## Examples

- The spaces $\mathbb{R}^{n}$, where $n \geq 1$, are the premier examples of vector spaces.
- Let $V$ be the set of all arrows (directed line segments) in three-dimensional space, with two arrows regarded as equal if they have the same length and point in the same direction.
Define addition by the parallelogram rule.
For each $\boldsymbol{v}$ in $V$, define $c \boldsymbol{v}$ to be the arrow whose length is $|c|$ times the length of $\boldsymbol{v}$, pointing in the same direction as $\boldsymbol{v}$ if $c \geq 0$ and otherwise pointing in the opposite direction. Then $V$ is a vector space.


## Examples (Cont'd)

- The definition of $V$ is geometric, using concepts of length and direction.
No xyz-coordinate system is involved.
An arrow of zero length is a single point and represents the zero vector.
The negative of $\boldsymbol{v}$ is $(-1) \boldsymbol{v}$.
So Axioms 1, 4, 5, 6, and 10 are evident.
The rest are verified by geometry:


$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.


## Example

- Let $\mathbb{S}$ be the space of all doubly infinite sequences of numbers (usually written in a row rather than a column):

$$
\left\{y_{k}\right\}=\left(\ldots, y_{-2}, y_{-1}, y_{0}, y_{1}, y_{2}, \ldots\right)
$$

If $\left\{z_{k}\right\}$ is another element of $\mathbb{S}$, then the sum $\left\{y_{k}\right\}+\left\{z_{k}\right\}$ is the sequence $\left\{y_{k}+z_{k}\right\}$ formed by adding corresponding terms of $\left\{y_{k}\right\}$ and $\left\{z_{k}\right\}$.
The scalar multiple $c\left\{y_{k}\right\}$ is the sequence $\left\{c y_{k}\right\}$.
The vector space axioms are verified in the same way as for $\mathbb{R}^{n}$.

- Elements of $\mathbb{S}$ arise in engineering, for example, whenever a signal is measured (or sampled) at discrete times.
- We will call $\mathbb{S}$ the space of (discrete-time) signals.


## Example

- For $n \geq 0$, the set $\mathbb{P}_{n}$ of polynomials of degree at most $n$ consists of all polynomials of the form

$$
\boldsymbol{p}(t)=a_{0}+a_{1} t+a_{2} t^{2}+\cdots+a_{n} t^{n}
$$

where the coefficients $a_{0}, \ldots, a_{n}$ and the variable $t$ are real numbers.

- The degree of $\boldsymbol{p}$ is the highest power of $t$ whose coefficient is not zero.
- If $\boldsymbol{p}(t)=a_{0} \neq 0$, the degree of $\boldsymbol{p}$ is zero.
- If all the coefficients are zero, $\boldsymbol{p}$ is called the zero polynomial.
- The zero polynomial is included in $\mathbb{P}_{n}$ even though its degree, for technical reasons, is not defined.


## Example (Cont'd)

- If $\boldsymbol{p}$ is given as above and if $\boldsymbol{q}(t)=b_{0}+b_{1} t+\cdots+b_{n} t^{n}$, then the sum $\boldsymbol{p}+\boldsymbol{q}$ is defined by

$$
\begin{aligned}
(\boldsymbol{p}+\boldsymbol{q})(t) & =\boldsymbol{p}(t)+\boldsymbol{q}(t) \\
& =\left(a_{0}+b_{0}\right)+\left(a_{1}+b_{1}\right) t+\cdots+\left(a_{n}+b_{n}\right) t^{n} .
\end{aligned}
$$

- The scalar multiple cp is the polynomial defined by

$$
(c \boldsymbol{p})(t)=c \boldsymbol{p}(t)=c a_{0}+\left(c a_{1}\right) t+\cdots+\left(c a_{n}\right) t^{n} .
$$

- These definitions satisfy Axioms 1 and 6 because $\boldsymbol{p}+\boldsymbol{q}$ and $c \boldsymbol{p}$ are polynomials of degree less than or equal to $n$.
- Axioms 2, 3, and 7-10 follow from properties of the real numbers.
- Clearly, the zero polynomial acts as the zero vector in Axiom 4.
- Finally, $(-1) \boldsymbol{p}$ acts as the negative of $\boldsymbol{p}$, so Axiom 5 is satisfied.
- Thus $\mathbb{P}_{n}$ is a vector space.


## Example

- Let $V$ be the set of all real valued functions defined on a set $\mathbb{D}$.
- Functions are added in the usual way: $\boldsymbol{f}+\boldsymbol{g}$ is the function whose value at $t$ in the domain $\mathbb{D}$ is $\boldsymbol{f}(t)+\boldsymbol{g}(t)$.
- Likewise, for a scalar $c$ and an $\boldsymbol{f}$ in $V$, the scalar multiple $c \boldsymbol{f}$ is the function whose value at $t$ is $c \boldsymbol{f}(t)$.
- For instance, if $\mathbb{D}=\mathbb{R}, \boldsymbol{f}(t)=1+\sin 2 t$, and $\boldsymbol{g}(t)=2+0.5 t$, then

$$
(\boldsymbol{f}+\boldsymbol{g})(t)=3+\sin 2 t+0.5 t \quad \text { and } \quad(2 \boldsymbol{g})(t)=4+t
$$

- Two functions in $V$ are equal if and only if their values are equal for every $t$ in $\mathbb{D}$.
- Hence the zero vector in $V$ is the function that is identically zero, $\boldsymbol{f}(t)=0$ for all $t$.
- The negative of $\boldsymbol{f}$ is $(-1) \boldsymbol{f}$.
- Axioms 1 and 6 are obviously true, and the other axioms follow from properties of the real numbers, so $V$ is a vector space.


## Subspaces

## Definition

A subspace of a vector space $V$ is a subset $H$ of $V$ that has three properties:
(a) The zero vector of $V$ is in $H$.
(b) $H$ is closed under vector addition. That is, for each $\boldsymbol{u}$ and $\boldsymbol{v}$ in $H$, the sum $\boldsymbol{u}+\boldsymbol{v}$ is in $H$.
(c) $H$ is closed under multiplication by scalars. That is, for each $\boldsymbol{u}$ in $H$ and each scalar $c$, the vector $c u$ is in $H$.

- Properties (a), (b) and (c) guarantee that a subspace $H$ of $V$ is itself a vector space, under the vector space operations already defined in $V$.


## Examples

- The set consisting of only the zero vector in a vector space $V$ is a subspace of $V$, called the zero subspace and written as $\{\mathbf{0}\}$.
- Let $\mathbb{P}$ be the set of all polynomials with real coefficients, with operations in $\mathbb{P}$ defined as for functions.
Then $\mathbb{P}$ is a subspace of the space of all real-valued functions defined on $\mathbb{R}$.
- For each $n \geq 0, \mathbb{P}_{n}$ is a subspace of $\mathbb{P}$, because $\mathbb{P}_{n}$ is a subset of $\mathbb{P}$ that contains the zero polynomial, the sum of two polynomials in $\mathbb{P}_{n}$ is also in $\mathbb{P}_{n}$, and a scalar multiple of a polynomial in $\mathbb{P}_{n}$ is also in $\mathbb{P}_{n}$.


## Example

- The vector space $\mathbb{R}^{2}$ is not a subspace of $\mathbb{R}^{3}$ because $\mathbb{R}^{2}$ is not even a subset of $\mathbb{R}^{3}$.

The set

$$
H=\left\{\left[\begin{array}{l}
s \\
t \\
0
\end{array}\right]: s \text { and } t \text { are real }\right\}
$$

is a subset of $\mathbb{R}^{3}$ that "emulates" $\mathbb{R}^{2}$, although it is logically distinct from $\mathbb{R}^{2}$.
$H$ is a subspace of $\mathbb{R}^{3}$.

- The zero vector is in $H$;
- $H$ is closed under vector addition and scalar multiplication because these operations on vectors in $H$ always produce vectors whose third entries are zero (and so belong to $H$ ).


## Example

- A plane in $\mathbb{R}^{3}$ not through the origin is not a subspace of $\mathbb{R}^{3}$, because the plane does not contain the zero vector of $\mathbb{R}^{3}$.
- A line in $\mathbb{R}^{2}$ not through the origin is not a subspace of $\mathbb{R}^{2}$.


## Subspace Spanned by a Set

- Recall that the term linear combination refers to any sum of scalar multiples of vectors.
- Moreover, $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ denotes the set of all vectors that can be written as linear combinations of $\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}$.
- Given $\boldsymbol{v}_{1}$ and $\boldsymbol{v}_{2}$ in a vector space $V$, let $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$. $H$ is a subspace of $V$.
- The zero vector is in $H$, since $\mathbf{0}=0 \boldsymbol{v}_{1}+0 \boldsymbol{v}_{2}$.


## Subspace Spanned by a Set (Cont'd)

- To show that $H$ is closed under vector addition, take two arbitrary vectors in $H$, say, $\boldsymbol{u}=s_{1} \boldsymbol{v}_{1}+s_{2} \boldsymbol{v}_{2}$ and $\boldsymbol{w}=t_{1} \boldsymbol{v}_{1}+t_{2} \boldsymbol{v}_{2}$. By Axioms 2,3 and 8 for the vector space $V$,

$$
\begin{aligned}
\boldsymbol{u}+\boldsymbol{w} & =\left(s_{1} \boldsymbol{v}_{1}+s_{2} \boldsymbol{v}_{2}\right)+\left(t_{1} \boldsymbol{v}_{1}+t_{2} \boldsymbol{v}_{2}\right) \\
& =\left(s_{1}+t_{1}\right) \boldsymbol{v}_{1}+\left(s_{2}+t_{2}\right) \boldsymbol{v}_{2} .
\end{aligned}
$$

So $\boldsymbol{u}+\boldsymbol{w}$ is in $H$.

- If $c$ is any scalar, then by Axioms 7 and 9 ,

$$
c \boldsymbol{u}=c\left(s_{1} \boldsymbol{v}_{1}+s_{2} \boldsymbol{v}_{2}\right)=\left(c s_{1}\right) \boldsymbol{v}_{1}+\left(c s_{2}\right) \boldsymbol{v}_{2} .
$$

This shows that $c u$ is in $H$ and $H$ is closed under scalar multiplication.

- Thus $H$ is a subspace of $V$.


## Spanned Subspace and Spanning Set of Vectors

- The argument of the preceding example can easily be generalized to prove the following theorem.


## Theorem

If $\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}$ are in a vector space $V$, then $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ is a subspace of $V$.

- We call $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ the subspace spanned (or generated) by $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$.
- Given any subspace $H$ of $V$, a spanning (or generating) set for $H$ is a set $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ in $H$ such that $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$.


## Example

- Let $H$ be the set of all vectors of the form $(a-3 b, b-a, a, b)$, where $a$ and $b$ are arbitrary scalars, i.e.,

$$
H=\{(a-3 b, b-a, a, b): a \text { and } b \text { in } \mathbb{R}\} .
$$

Show that $H$ is a subspace of $\mathbb{R}^{4}$.

- Write the vectors in $H$ as column vectors.

Then an arbitrary vector in $H$ has the form

$$
\left[\begin{array}{c}
a-3 b \\
b-a \\
a \\
b
\end{array}\right]=a\left[\begin{array}{r}
1 \\
-1 \\
1 \\
0
\end{array}\right]+b\left[\begin{array}{r}
-3 \\
1 \\
0 \\
1
\end{array}\right]=a \boldsymbol{v}_{1}+b \boldsymbol{v}_{2}
$$

This calculation shows that $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$, where $\boldsymbol{v}_{1}$ and $\boldsymbol{v}_{2}$ are the vectors indicated above.
Thus $H$ is a subspace of $\mathbb{R}^{4}$ by the theorem.

## Example

- For what value(s) of $h$ will $\boldsymbol{y}$ be in the subspace of $\mathbb{R}^{3}$ spanned by $\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{3}$, if

$$
\boldsymbol{v}_{1}=\left[\begin{array}{r}
1 \\
-1 \\
-2
\end{array}\right], \boldsymbol{v}_{2}=\left[\begin{array}{r}
5 \\
-4 \\
-7
\end{array}\right], \boldsymbol{v}_{3}=\left[\begin{array}{r}
-3 \\
1 \\
0
\end{array}\right], \boldsymbol{y}=\left[\begin{array}{r}
-4 \\
3 \\
h
\end{array}\right] .
$$

- We are asking for what value(s) of $h$ the equation $\boldsymbol{y}=x_{1} \boldsymbol{v}_{1}+x_{2} \boldsymbol{v}_{2}+x_{3} \boldsymbol{v}_{3}$ has solutions.
This happens if and only if the system

$$
\left[\begin{array}{rrr}
1 & 5 & -3 \\
-1 & -4 & 1 \\
-2 & -7 & 0
\end{array}\right]\left[\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3}
\end{array}\right]=\left[\begin{array}{r}
-4 \\
3 \\
h
\end{array}\right]
$$

is consistent.

## Example (Cont'd)

- We reduced the augmented matrix to row echelon form:

$$
\left.\left.\begin{array}{rlr|r}
{\left[\begin{array}{rrr}
1 & 5 & -3 \\
-1 & -4 & 1 \\
-4 & -7 & 0
\end{array}\right.} & h
\end{array}\right] \begin{array}{l}
\begin{array}{rlr|r}
R_{2} \leftarrow R_{2}+R_{1} \\
R_{3} \leftarrow R_{3}+2 R_{1}
\end{array}
\end{array} \begin{array}{rrr|r}
1 & 5 & -3 & -4 \\
0 & 1 & -2 & -1 \\
0 & 3 & -6 & h-8
\end{array}\right]
$$

Thus, $\boldsymbol{y}$ is in Span $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{3}\right\}$ if and only if $h=5$.

## Subsection 2

## Null Spaces, Column Spaces and Linear Transformations

## Example

- Consider the following system of homogeneous equations:

$$
\left\{\begin{aligned}
x_{1}-3 x_{2}-2 x_{3} & =0 \\
-5 x_{1}+9 x_{2}+x_{3} & =0
\end{aligned}\right.
$$

- In matrix form, this system is written as $\boldsymbol{A x}=\mathbf{0}$, where

$$
A=\left[\begin{array}{rrr}
1 & -3 & -2 \\
-5 & 9 & 1
\end{array}\right]
$$

- Recall that the set of all $\boldsymbol{x}$ that satisfy the system is called the solution set of the system.
- To relate this set directly to the matrix $A$ and the equation $A \boldsymbol{x}=\mathbf{0}$, we call it the null space of the matrix $A$.
- So the null space of $A$ is the set of $\boldsymbol{x}$ that satisfy $A \boldsymbol{x}=\mathbf{0}$.


## Null Space of a Matrix

## Definition

The null space of an $m \times n$ matrix $A$, written as $\operatorname{Nul} A$, is the set of all solutions of the homogeneous equation $A \boldsymbol{x}=\mathbf{0}$. In set notation,

$$
\operatorname{Nul} A=\left\{\boldsymbol{x}: \boldsymbol{x} \text { is in } \mathbb{R}^{n} \text { and } A \boldsymbol{x}=\mathbf{0}\right\} .
$$

- A more dynamic description of $\operatorname{Nul} A$ is the set of all $\boldsymbol{x}$ in $\mathbb{R}^{n}$ that are mapped into the zero vector of $\mathbb{R}^{m}$ via the linear transformation $x \rightarrow A x$.



## Example

- Let $A=\left[\begin{array}{rrr}1 & -3 & -2 \\ -5 & 9 & 1\end{array}\right]$ and let $\boldsymbol{u}=\left[\begin{array}{r}5 \\ 3 \\ -2\end{array}\right]$. Determine if $\boldsymbol{u}$ belongs to the null space of $A$.
- We need to test if $\boldsymbol{u}$ satisfies $A \boldsymbol{u}=\mathbf{0}$. We compute

$$
A \boldsymbol{u}=\left[\begin{array}{rrr}
1 & -3 & -2 \\
-5 & 9 & 1
\end{array}\right]\left[\begin{array}{r}
5 \\
3 \\
-2
\end{array}\right]=\left[\begin{array}{r}
5-9+4 \\
-25+27-2
\end{array}\right]=\left[\begin{array}{l}
0 \\
0
\end{array}\right]
$$

Thus $\boldsymbol{u}$ is in $\operatorname{Nul} A$.

## Vector Space Property of NulA

## Theorem

The null space of an $m \times n$ matrix $A$ is a subspace of $\mathbb{R}^{n}$. Equivalently, the set of all solutions to a system $A \boldsymbol{x}=\mathbf{0}$ of $m$ homogeneous linear equations in $n$ unknowns is a subspace of $\mathbb{R}^{n}$.

- Certainly Nul $A$ is a subset of $\mathbb{R}^{n}$ because $A$ has $n$ columns.

We must show that $\operatorname{Nul} A$ satisfies the three properties of a subspace. Of course, $\mathbf{0}$ is in NulA.
Next, let $\boldsymbol{u}$ and $\boldsymbol{v}$ represent any two vectors in Nul $A$. Then $A \boldsymbol{u}=\mathbf{0}$ and $A \boldsymbol{v}=\mathbf{0}$. To show that $\boldsymbol{u}+\boldsymbol{v}$ is in $\operatorname{Nul} A$, we must show that $A(\boldsymbol{u}+\boldsymbol{v})=\mathbf{0}$. Using a property of matrix multiplication, compute $A(\boldsymbol{u}+\boldsymbol{v})=A \boldsymbol{u}+A \boldsymbol{v}=\mathbf{0}+\mathbf{0}=\mathbf{0}$. Thus $\boldsymbol{u}+\boldsymbol{v}$ is in $\operatorname{Nul} A$, and $\operatorname{Nul} A$ is closed under vector addition.
If $c$ is any scalar, then $A(c \boldsymbol{u})=c(A \boldsymbol{u})=c \mathbf{0}=\mathbf{0}$. So $c \boldsymbol{u}$ is in NulA.
Thus Nul $A$ is a subspace of $\mathbb{R}^{n}$.

## Example

- Let $H$ be the set of all vectors in $\mathbb{R}^{4}$ whose coordinates $a, b, c, d$ satisfy the equations $a-2 b+5 c=d$ and $c-a=b$. Show that $H$ is a subspace of $\mathbb{R}^{4}$.
- Rearrange the equations that describe the elements of $H$, and note that $H$ is the set of all solutions of the following system of homogeneous linear equations:

$$
\left\{\begin{array}{ccc}
a-2 b+5 c-d & =0 \\
-a-b+c & =0
\end{array}\right.
$$

By the theorem, $H$ is a subspace of $\mathbb{R}^{4}$.

## An Explicit Description of NulA

- Find a spanning set for the null space of the matrix

$$
A=\left[\begin{array}{rrrrr}
-3 & 6 & -1 & 1 & -7 \\
1 & -2 & 2 & 3 & -1 \\
2 & -4 & 5 & 8 & -4
\end{array}\right]
$$

- We find the general solution of $A \boldsymbol{x}=\mathbf{0}$ in terms of free variables.
$\left.\begin{array}{c}{\left[\begin{array}{rrrrr|r}-3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0\end{array}\right]} \\ \longrightarrow\left[\begin{array}{rrrrr|r}1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 5 & 10 & -10 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0\end{array}\right]\end{array}{ }^{\left[\begin{array}{rrrrr|r}1 & -2 & 2 & 3 & -1 & 0 \\ -3 & 6 & -1 & 1 & -7 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0\end{array}\right]} \begin{array}{rlrrrr}1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0\end{array}\right]$.
The general solution is $x_{1}=2 x_{2}-2 x_{3}-3 x_{4}+x_{5}=$ $2 x_{2}-2\left(-2 x_{4}+2 x_{5}\right)-3 x_{4}+x_{5}=2 x_{2}+x_{4}-3 x_{5}, x_{3}=-2 x_{4}+2 x_{5}$, with $x_{2}, x_{4}$, and $x_{5}$ free.


## An Explicit Description of NulA (Cont'd)

- Next, decompose the vector giving the general solution into a linear combination of vectors where the weights are the free variables. That is,

$$
\left[\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3} \\
x_{4} \\
x_{5}
\end{array}\right]=\left[\begin{array}{c}
2 x_{2}+x_{4}-3 x_{5} \\
x_{2} \\
-2 x_{4}+2 x_{5} \\
x_{4} \\
x_{5}
\end{array}\right]=x_{2}\left[\begin{array}{l}
2 \\
1 \\
0 \\
0 \\
0
\end{array}\right]+x_{4}\left[\begin{array}{r}
1 \\
0 \\
-2 \\
1 \\
0
\end{array}\right]+x_{5}\left[\begin{array}{r}
-3 \\
0 \\
2 \\
0 \\
1
\end{array}\right]
$$

Every linear combination of $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ is an element of $\operatorname{Nul} A$. Thus $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a spanning set for NulA.

## The Column Space of a Matrix

## Definition

The column space of an $m \times n$ matrix $A$, written as $\operatorname{Col} A$, is the set of all linear combinations of the columns of $A$. If $A=\left[\begin{array}{lll}\boldsymbol{a}_{1} & \cdots & a_{n}\end{array}\right]$, then $\operatorname{Col} A=\operatorname{Span}\left\{\boldsymbol{a}_{1}, \ldots, \boldsymbol{a}_{n}\right\}$.

- Since $\operatorname{Span}\left\{\boldsymbol{a}_{1}, \ldots, \boldsymbol{a}_{n}\right\}$ is a subspace, by the preceding theorem, the next theorem follows from the definition of $\operatorname{Col} A$ and the fact that the columns of $A$ are in $\mathbb{R}^{m}$.


## Theorem

The column space of an $m \times n$ matrix $A$ is a subspace of $\mathbb{R}^{m}$.

## Column Spaces and Ranges of Linear Transformations

- Note that a typical vector in $\operatorname{Col} A$ can be written as $A \boldsymbol{x}$ for some $\boldsymbol{x}$ because the notation $A \boldsymbol{x}$ stands for a linear combination of the columns of $A$.
That is,

$$
\operatorname{Col} A=\left\{\boldsymbol{b}: \boldsymbol{b}=A \boldsymbol{x} \text { for some } \boldsymbol{x} \text { in } \mathbb{R}^{n}\right\} .
$$

- The notation $A \boldsymbol{x}$ for vectors in $\operatorname{Col} A$ also shows that $\operatorname{Col} A$ is the range of the linear transformation $x \mapsto A \boldsymbol{x}$.


## Example

- Find a matrix $A$ such that $W=\operatorname{Col} A$, where

$$
W=\left\{\left[\begin{array}{c}
6 a-b \\
a+b \\
-7 a
\end{array}\right]: a, b \text { in } \mathbb{R}\right\} .
$$

- Write $W$ as a set of linear combinations.

$$
\begin{aligned}
W & =\left\{a\left[\begin{array}{r}
6 \\
1 \\
-7
\end{array}\right]+b\left[\begin{array}{r}
-1 \\
1 \\
0
\end{array}\right]: a, b \text { in } \mathbb{R}\right\} \\
& =\operatorname{Span}\left\{\left[\begin{array}{r}
6 \\
1 \\
-7
\end{array}\right],\left[\begin{array}{r}
-1 \\
1 \\
0
\end{array}\right]\right\} .
\end{aligned}
$$

Now use the vectors in the spanning set as the columns of $A$. Let
$A=\left[\begin{array}{rr}6 & -1 \\ 1 & 1 \\ -7 & 0\end{array}\right]$. Then $W=\operatorname{Col} A$, as desired.

## Column Space as Range

- Recall from a previous theorem that
the columns of $A$ span $\mathbb{R}^{m}$
if and only if
the equation $A \boldsymbol{x}=\boldsymbol{b}$ has a solution for each $\boldsymbol{b}$.
- We can restate this fact as follows:

The column space of an $m \times n$ matrix $A$ is all of $\mathbb{R}^{m}$ if and only if the equation $A \boldsymbol{x}=\boldsymbol{b}$ has a solution for each $\boldsymbol{b}$ in $\mathbb{R}^{m}$.

## Example

- Let

$$
A=\left[\begin{array}{rrrr}
2 & 4 & -2 & 1 \\
-2 & -5 & 7 & 3 \\
3 & 7 & -8 & 6
\end{array}\right]
$$

(a) If the column space of $A$ is a subspace of $\mathbb{R}^{k}$, what is $k$ ?
(b) If the null space of $A$ is a subspace of $\mathbb{R}^{k}$, what is $k$ ?
(a) The columns of $A$ each have three entries. So Col $A$ is a subspace of $\mathbb{R}^{k}$, where $k=3$.
(b) A vector $\boldsymbol{x}$ such that $A \boldsymbol{x}$ is defined must have four entries. So $\operatorname{Nul} A$ is a subspace of $\mathbb{R}^{k}$, where $k=4$.

## Example

- With $A=\left[\begin{array}{rrrr}2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6\end{array}\right]$, find a nonzero vector in $\operatorname{Col} A$ and a nonzero vector in $\operatorname{Nul} A$.
- For a vector in $\operatorname{Col} A$, any column of $A$ will do, say, $\left[\begin{array}{r}2 \\ -2 \\ 3\end{array}\right]$.

To find a nonzero vector in $\operatorname{Nul} A$, row reduce the augmented matrix:

## Example (Cont'd)

- We continue the reduction:

$$
\begin{gathered}
{\left[\begin{array}{rrrr|r}
1 & 2 & -1 & 9 & 0 \\
0 & -1 & 5 & 21 & 0 \\
0 & 0 & 0 & -17 & 0
\end{array}\right]} \\
\longrightarrow\left[\begin{array}{rrrr|r}
1 & 2 & -1 & 0 & 0 \\
0 & 1 & -5 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{array}\right] \rightarrow\left[\begin{array}{rrrr|r}
1 & 2 & -1 & 9 & 0 \\
0 & 1 & -5 & -21 & 0 \\
0 & 0 & 0 & 1 & 0
\end{array}\right]
\end{gathered}
$$

Thus, if $\boldsymbol{x}$ satisfies $A \boldsymbol{x}=\mathbf{0}$, then $x_{1}=-9 x_{3}, x_{2}=5 x_{3}, x_{4}=0$, and $x_{3}$ is free. Assigning a nonzero value to $x_{3}$ say, $x_{3}=1$ we obtain a
vector in Nul $A$, namely, $\boldsymbol{x}=\left[\begin{array}{r}-9 \\ 5 \\ 1 \\ 0\end{array}\right]$.

## Example

- Let $A=\left[\begin{array}{rrrr}2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6\end{array}\right], \boldsymbol{u}=\left[\begin{array}{r}3 \\ -2 \\ -1 \\ 0\end{array}\right]$ and $\boldsymbol{v}=\left[\begin{array}{r}3 \\ -1 \\ 3\end{array}\right]$.
(a) Determine if $\boldsymbol{u}$ is in $\operatorname{Nul} A$. Could $\boldsymbol{u}$ be in ColA?
(b) Determine if $\boldsymbol{v}$ is in ColA. Could $\boldsymbol{v}$ be in NulA?
(a) We compute the product

$$
A \boldsymbol{u}=\left[\begin{array}{rrrr}
2 & 4 & -2 & 1 \\
-2 & -5 & 7 & 3 \\
3 & 7 & -8 & 6
\end{array}\right]\left[\begin{array}{r}
3 \\
-2 \\
-1 \\
0
\end{array}\right]=\left[\begin{array}{r}
0 \\
-3 \\
3
\end{array}\right] \neq\left[\begin{array}{l}
0 \\
0 \\
0
\end{array}\right]
$$

Obviously, $\boldsymbol{u}$ is not a solution of $A \boldsymbol{x}=\mathbf{0}$. So $\boldsymbol{u}$ is not in $\operatorname{Nul} A$.
Also, with four entries, $\boldsymbol{u}$ could not possibly be in $\operatorname{Col} A$, since $\operatorname{Col} A$ is a subspace of $\mathbb{R}^{3}$.

## Example (Cont'd)

(b) Reduce $[A \boldsymbol{v}]$ to an echelon form.

$$
\begin{aligned}
& {[A \boldsymbol{v}]=} \\
& {\left[\begin{array}{rrrr|r}
2 & 4 & -2 & 1 & 3 \\
-2 & -5 & 7 & 3 & -1 \\
3 & 7 & -8 & 6 & 3
\end{array}\right] \rightarrow\left[\begin{array}{rrrr|r}
2 & 4 & -2 & 1 & 3 \\
0 & -1 & 5 & 4 & 2 \\
0 & 1 & -5 & \frac{9}{2} & -\frac{3}{2}
\end{array}\right]} \\
& \longrightarrow\left[\begin{array}{rrrrr}
2 & 4 & -2 & 1 & 3 \\
0 & 1 & -5 & -4 & -2 \\
0 & 0 & 0 & \frac{17}{2} & \frac{1}{2}
\end{array}\right] .
\end{aligned}
$$

At this point, it is clear that the equation $A \boldsymbol{x}=\boldsymbol{v}$ is consistent. So $\boldsymbol{v}$ is in $\operatorname{Col} A$.

With only three entries, $\boldsymbol{v}$ could not possibly be in $\operatorname{Nul} A$, since $\operatorname{Nul} A$ is a subspace of $\mathbb{R}^{4}$.

## Linear Transformations; Kernel; Range

## Definition

A linear transformation $T$ from a vector space $V$ into a vector space $W$ is a rule that assigns to each vector $\boldsymbol{x}$ in $V$ a unique vector $T(\boldsymbol{x})$ in $W$, such that:
(i) $T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$, for all $\boldsymbol{u}, \boldsymbol{v}$ in $V$;
(ii) $T(c \boldsymbol{u})=c T(\boldsymbol{u})$, for all $\boldsymbol{u}$ in $V$ and all scalars $c$.

- The kernel (or null space) of such a $T$ is the set of all $\boldsymbol{u}$ in $V$ such that $T(\boldsymbol{u})=\mathbf{0}$ (the zero vector in $W$ ).
- The range of $T$ is the set of all vectors in $W$ of the form $T(\boldsymbol{x})$, for some $\boldsymbol{x}$ in $V$.
- If $T$ happens to arise as a matrix transformation say, $T(\boldsymbol{x})=A \boldsymbol{x}$ for some matrix $A$ then the kernel and the range of $T$ are just the null space and the column space of $A$.


## Properties of Kernel and Range

- It is not difficult to show that the kernel of $T$ is a subspace of $V$.
- Also, the range of $T$ is a subspace of $W$.



## Example

- Let $V$ be the vector space of all real-valued functions $f$ defined on an interval $[a, b]$ with the property that they are differentiable and their derivatives are continuous functions on $[a, b]$.
- Let $W$ be the vector space $C[a, b]$ of all continuous functions on $[a, b]$.
- Let $D: V \rightarrow W$ be the transformation that changes $f$ in $V$ into its derivative $f^{\prime}$.
- In calculus, two simple differentiation rules are

$$
D(f+g)=D(f)+D(g) \quad \text { and } \quad D(c f)=c D(f)
$$

- That is, $D$ is a linear transformation.
- It can be shown that the kernel of $D$ is the set of constant functions on $[a, b]$.
- Moreover, and the range of $D$ is the set $W$ of all continuous functions on $[a, b]$.


## Example

- Consider the differential equation

$$
y^{\prime \prime}+\omega^{2} y=0
$$

where $\omega$ is a constant, used to describe a variety of physical systems.

- The set of its solutions is precisely the kernel of the linear transformation that maps a function $y=f(t)$ into the function $f^{\prime \prime}(t)+\omega^{2} f(t)$.
- Finding an explicit description of this vector space is a problem in differential equations.


## Subsection 3

## Linearly Independent Sets; Bases

## Linearly Dependence and Independence

- An indexed set of vectors $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ in $V$ is said to be linearly independent if the vector equation

$$
c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}+\cdots c_{p} \boldsymbol{v}_{p}=\mathbf{0}
$$

has only the trivial solution, $c_{1}=0, \ldots, c_{p}=0$.

- The set $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ is said to be linearly dependent if the equation has a nontrivial solution, that is, if there are some weights, $c_{1}, \ldots, c_{p}$, not all zero, such that the equation holds.
- In such a case, the equation is called a linear dependence relation among $\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}$.
- Just as in $\mathbb{R}^{n}$, a set containing a single vector $\boldsymbol{v}$ is linearly independent if and only if $\boldsymbol{v} \neq \mathbf{0}$.
- Also, a set of two vectors is linearly dependent if and only if one of the vectors is a multiple of the other.
- And any set containing the zero vector is linearly dependent.


## Characterization of Linearly Dependence

## Theorem

An indexed set $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ of two or more vectors, with $\boldsymbol{v}_{1} \neq \mathbf{0}$, is linearly dependent if and only if some $\boldsymbol{v}_{j}$ (with $j>1$ ) is a linear combination of the preceding vectors, $\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{j-1}$.

- The main difference between linear dependence in $\mathbb{R}^{n}$ and in a general vector space is that when the vectors are not $n$-tuples, the homogeneous equation

$$
c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}+\cdots c_{p} \boldsymbol{v}_{p}=\mathbf{0}
$$

usually cannot be written as a system of $n$ linear equations.
That is, the vectors cannot be made into the columns of a matrix $A$ in order to study the equation $A \boldsymbol{x}=\mathbf{0}$.

- We must rely instead on the definition of linear dependence and the theorem above.


## Example

- Let $\boldsymbol{p}_{1}(t)=1, \boldsymbol{p}_{2}(t)=t$ and $\boldsymbol{p}_{3}(t)=4-t$. Is $\left\{\boldsymbol{p}_{1}(t), \boldsymbol{p}_{2}(t), \boldsymbol{p}_{3}(t)\right\}$ linearly dependent?
- We set

$$
x_{1} \boldsymbol{p}_{1}(t)+x_{2} \boldsymbol{p}_{2}(t)+x_{3} \boldsymbol{p}_{3}(t)=0
$$

If we can find $x_{1}, x_{2}, x_{3}$ not all zero satisfying the equation, then the set $\left\{\boldsymbol{p}_{1}(t), \boldsymbol{p}_{2}(t), \boldsymbol{p}_{3}(t)\right\}$ is linearly dependent. We have:

$$
\begin{aligned}
& x_{1}+x_{2} t+x_{3}(4-t)=0 \\
& \left(x_{1}+4 x_{3}\right)+\left(x_{2}-x_{3}\right) t=0 \\
& \left\{\begin{array}{r}
x_{1}+4 x_{3}=0 \\
x_{2}-x_{3}=
\end{array}\right\} \Rightarrow\left\{\begin{array}{l}
x_{1}=-4 x_{3} \\
x_{2}=x_{3}
\end{array}\right.
\end{aligned}
$$

So $x_{1}=-4, x_{2}=1, x_{3}=1$ satisfy the equation. We conclude $\left\{\boldsymbol{p}_{1}(t), \boldsymbol{p}_{2}(t), \boldsymbol{p}_{3}(t)\right\}$ is linearly dependent.

## Example

- The set $\{\sin t, \cos t\}$ is linearly independent in $C[0,1]$, the space of all continuous functions on $0 \leq t \leq 1$, because $\sin t$ and $\cos t$ are not multiples of one another as vectors in $C[0,1]$.
That is, there is no scalar $c$ such that $\cos t=c \sin t$ for all $t$ in $[0,1]$.
- However, $\{\sin t \cos t, \sin 2 t\}$ is linearly dependent because of the identity:

$$
\sin 2 t=2 \sin t \cos t, \text { for all } t
$$

## Basis of a Vector Space

## Definition

Let $H$ be a subspace of a vector space $V$. An indexed set of vectors $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{p}\right\}$ in $V$ is a basis for $H$ if:
(i) $\mathcal{B}$ is a linearly independent set;
(ii) the subspace spanned by $\mathcal{B}$ coincides with $H$; that is,

$$
H=\operatorname{Span}\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{p}\right\}
$$

- The definition of a basis applies to the case when $H=V$, because any vector space is a subspace of itself.
- Thus a basis of $V$ is a linearly independent set that spans $V$.
- Observe that when $H \neq V$, condition (ii) includes the requirement that each of the vectors $\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{p}$ must belong to $H$, because $\operatorname{Span}\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{p}\right\}$ contains $\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{p}$.


## Example

- Let $A$ be an invertible $n \times n$ matrix - say, $A=\left[\begin{array}{lll}\boldsymbol{a}_{1} & \cdots & \boldsymbol{a}_{n}\end{array}\right]$.
- Then the columns of $A$ form a basis for $\mathbb{R}^{n}$ because they are linearly independent and they span $\mathbb{R}^{n}$, by the Invertible Matrix Theorem.


## Example: Standard Basis

- Let $\boldsymbol{e}_{1}, \ldots, \boldsymbol{e}_{n}$ be the columns of the $n \times n$ identity matrix, $I_{n}$.
- That is

$$
\boldsymbol{e}_{1}=\left[\begin{array}{c}
1 \\
0 \\
\vdots \\
0
\end{array}\right], \boldsymbol{e}_{2}=\left[\begin{array}{c}
0 \\
1 \\
\vdots \\
0
\end{array}\right], \ldots, \boldsymbol{e}_{n}=\left[\begin{array}{c}
0 \\
0 \\
\vdots \\
1
\end{array}\right]
$$

- The set $\left\{\boldsymbol{e}_{1}, \ldots, \boldsymbol{e}_{n}\right\}$ is called the standard basis for $\mathbb{R}^{n}$.


## Example

- Let $\boldsymbol{v}_{1}=\left[\begin{array}{r}3 \\ 0 \\ -6\end{array}\right], \boldsymbol{v}_{2}=\left[\begin{array}{r}-4 \\ 1 \\ 7\end{array}\right]$ and $\boldsymbol{v}_{3}=\left[\begin{array}{r}-2 \\ 1 \\ 5\end{array}\right]$. Determine if $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{3}\right\}$ is a basis for $\mathbb{R}^{3}$.
- Since there are exactly three vectors in $\mathbb{R}^{3}$, we can use any of several methods to determine if the matrix $A=\left[\begin{array}{lll}\boldsymbol{v}_{1} & \boldsymbol{v}_{2} & \boldsymbol{v}_{3}\end{array}\right]$ is invertible.
Reduce to row echelon form:

$$
\left[\begin{array}{rrr}
3 & -4 & -2 \\
0 & 1 & 1 \\
-6 & 7 & 5
\end{array}\right] \rightarrow\left[\begin{array}{rrr}
3 & -4 & -2 \\
0 & 1 & 1 \\
0 & -1 & 1
\end{array}\right] \rightarrow\left[\begin{array}{rrr}
3 & -4 & -2 \\
0 & 1 & 1 \\
0 & 0 & 2
\end{array}\right]
$$

Thus, $A$ has three pivot positions. Hence $A$ is invertible. It follows the columns of $A$ form a basis for $\mathbb{R}^{3}$.

## Example

- Let $S=\left\{1, t, t^{2}, \ldots, t^{n}\right\}$. Verify that $S$ is a basis for $\mathbb{P}_{n}$. This basis is called the standard basis for $\mathbb{P}_{n}$.
- Certainly $S$ spans $\mathbb{P}_{n}$.

To show that $S$ is linearly independent, suppose that $c_{0}, \ldots, c_{n}$ satisfy

$$
c_{0} \cdot 1+c_{1} t+c_{2} t^{2}+\cdots+c_{n} t^{n}=\mathbf{0}(t)
$$

This equality means that the polynomial on the left has the same values as the zero polynomial on the right. A fundamental theorem in algebra says that the only polynomial in $\mathbb{P}_{n}$ with more than $n$ zeros is the zero polynomial. That is, the equation above holds for all $t$ only if $c_{0}=\cdots=c_{n}=0$. This proves that $S$ is linearly independent and hence is a basis for $\mathbb{P}_{n}$.

## Example: The Spanning Set Theorem

- Let

$$
\boldsymbol{v}_{1}=\left[\begin{array}{r}
0 \\
2 \\
-1
\end{array}\right], \boldsymbol{v}_{2}=\left[\begin{array}{l}
2 \\
2 \\
0
\end{array}\right], \boldsymbol{v}_{3}=\left[\begin{array}{r}
6 \\
16 \\
-5
\end{array}\right], H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{3}\right\}
$$

Note that $\boldsymbol{v}_{3}=5 \boldsymbol{v}_{1}+3 \boldsymbol{v}_{2}$, and show that
$\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{3}\right\}=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$.
Then find a basis for the subspace $H$.

- Every vector in $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$ belongs to $H$ because $c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}=c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}+0 \boldsymbol{v}_{3}$.
Now let $\boldsymbol{x}$ be any vector in $H$ say, $\boldsymbol{x}=c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}+c_{3} \boldsymbol{v}_{3}$. Since $\boldsymbol{v}_{3}=5 \boldsymbol{v}_{1}+3 \boldsymbol{v}_{2}$, we may substitute

$$
\begin{aligned}
\boldsymbol{x} & =c_{1} \boldsymbol{v}_{1}+c_{2} \boldsymbol{v}_{2}+c_{3}\left(5 \boldsymbol{v}_{1}+3 \mathbf{v}_{2}\right) \\
& =\left(c_{1}+5 c_{3}\right) \boldsymbol{v}_{1}+\left(c_{2}+3 c_{3}\right) \boldsymbol{v}_{2} .
\end{aligned}
$$

Thus $\boldsymbol{x}$ is in $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$.

## Example: The Spanning Set Theorem (Cont'd)

- We showed that every vector in $H$ already belongs to
$\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}=\operatorname{Span}\left\{\left[\begin{array}{r}0 \\ 2 \\ -1\end{array}\right],\left[\begin{array}{l}2 \\ 2 \\ 0\end{array}\right]\right\}$.
So $H$ and $\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$ are actually the same set of vectors. Now, notice that $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$ is obviously linearly independent.
- It follows that $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$ both spans $H$ and is a linearly independent set, whence it is is a basis of $H$


## The Spanning Set Theorem

## Theorem (The Spanning Set Theorem)

Let $S=\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$ be a set in $V$, and let $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p}\right\}$.
(a) If one of the vectors in $S$ say, $\boldsymbol{v}_{k}$ - is a linear combination of the remaining vectors in $S$, then the set formed from $S$ by removing $\boldsymbol{v}_{k}$ still spans $H$.
(b) If $H \neq\{\mathbf{0}\}$, some subset of $S$ is a basis for $H$.
(a) By rearranging the list of vectors in $S$, if necessary, we may suppose that $\boldsymbol{v}_{p}$ is a linear combination of $\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p-1}$. say,

$$
\mathbf{v}_{p}=a_{1} \mathbf{v}_{1}+\cdots+a_{p-1} \mathbf{v}_{p-1} .
$$

Given any $\boldsymbol{x}$ in $H$, we may write

$$
\boldsymbol{x}=c_{1} \boldsymbol{v}_{1}+\cdots+c_{p-1} \mathbf{v}_{p-1}+c_{p} \boldsymbol{v}_{p}
$$

for suitable scalars $c_{1}, \ldots, c_{p}$.

## The Spanning Set Theorem (Cont'd)

- Substituting the expression for $\boldsymbol{v}_{p}$ from the first equation, we get

$$
\begin{aligned}
\boldsymbol{x} & =c_{1} \mathbf{v}_{1}+\cdots+c_{p-1} \mathbf{v}_{p-1}+c_{p}\left(a_{1} \mathbf{v}_{1}+\cdots+a_{p-1} \mathbf{v}_{p-1}\right) \\
& =\left(c_{1}+c_{p} a_{1}\right) \boldsymbol{v}_{1}+\cdots+\left(c_{p-1}+c_{p} a_{p-1}\right) \boldsymbol{v}_{p-1} .
\end{aligned}
$$

Thus $\left\{\boldsymbol{v}_{1}, \ldots, \boldsymbol{v}_{p-1}\right\}$ spans $H$, since $\boldsymbol{x}$ was an arbitrary element of $H$.
(b) If the original spanning set $S$ is linearly independent, then it is already a basis for $H$.
Otherwise, one of the vectors in $S$ depends on the others and can be deleted, by Part (a).
So long as there are two or more vectors in the spanning set, we can repeat this process until the spanning set is linearly independent and hence is a basis for $H$.
If the spanning set is eventually reduced to one vector, that vector will be nonzero (and hence linearly independent) because $H \neq\{\mathbf{0}\}$.

## Example: Bases for $\operatorname{Col} A$

- Find a basis for $\operatorname{Col} B$, where

$$
B=\left[\begin{array}{llll}
\boldsymbol{b}_{1} & \boldsymbol{b}_{2} & \cdots & \boldsymbol{b}_{5}
\end{array}\right]=\left[\begin{array}{rrrrr}
1 & 4 & 0 & 2 & 0 \\
0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

- Each nonpivot column of $B$ is a linear combination of the pivot columns. In fact, $\boldsymbol{b}_{2}=4 \boldsymbol{b}_{1}$ and $\boldsymbol{b}_{4}=2 \boldsymbol{b}_{1}-\boldsymbol{b}_{3}$.
By the Spanning Set Theorem, we may discard $\boldsymbol{b}_{2}$ and $\boldsymbol{b}_{4}$, and $\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{3}, \boldsymbol{b}_{5}\right\}$ will still span Col $B$. Let

$$
S=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{3}, \boldsymbol{b}_{5}\right\}=\left\{\left[\begin{array}{l}
1 \\
0 \\
0 \\
0
\end{array}\right],\left[\begin{array}{l}
0 \\
1 \\
0 \\
0
\end{array}\right],\left[\begin{array}{l}
0 \\
0 \\
0 \\
1
\end{array}\right]\right\}
$$

Since $\boldsymbol{b}_{1} \neq \mathbf{0}$ and no vector in $S$ is a linear combination of the vectors that precede it, $S$ is linearly independent. Thus $S$ is a basis for Col $B$.

## The Case of a Non-Reduced Matrix

- Suppose a matrix $A$ is not in reduced echelon form.
- We know that any linear dependence relationship among the columns of $A$ can be expressed in the form $A \boldsymbol{x}=\mathbf{0}$, where $\boldsymbol{x}$ is a column of weights.
- When $A$ is row reduced to a matrix $B$, the columns of $B$ are often totally different from the columns of $A$.
- However, the equations $A \boldsymbol{x}=\mathbf{0}$ and $B \boldsymbol{x}=\mathbf{0}$ have exactly the same set of solutions.
- If $A=\left[\begin{array}{lll}\boldsymbol{a}_{1} & \cdots & \boldsymbol{a}_{n}\end{array}\right]$ and $B=\left[\begin{array}{lll}\boldsymbol{b}_{1} & \cdots & \boldsymbol{b}_{n}\end{array}\right]$, then the vector equations

$$
x_{1} \boldsymbol{a}_{1}+\cdots+x_{n} \boldsymbol{a}_{n}=\mathbf{0} \quad \text { and } \quad x_{1} \boldsymbol{b}_{1}+\cdots+x_{n} \boldsymbol{b}_{n}=\mathbf{0}
$$

also have the same set of solutions.

- That is, the columns of $A$ have exactly the same linear dependence relationships as the columns of $B$.


## Example: Bases for NulA and ColA

- The matrix $A$ below is row equivalent to the matrix $B$ :

$$
A=\left[\begin{array}{llll}
a_{1} & a_{2} & \cdots & a_{5}
\end{array}\right]
$$

$$
=\left[\begin{array}{rrrrr}
1 & 4 & 0 & 2 & -1 \\
3 & 12 & 1 & 5 & 5 \\
2 & 8 & 1 & 3 & 2 \\
5 & 20 & 2 & 8 & 8
\end{array}\right], \quad B=\left[\begin{array}{rrrrr}
1 & 4 & 0 & 2 & 0 \\
0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{array}\right] .
$$

Find a basis for $\operatorname{Col} A$.

- We saw that $\boldsymbol{b}_{2}=4 \boldsymbol{b}_{1}$ and $\boldsymbol{b}_{4}=2 \boldsymbol{b}_{1}-\boldsymbol{b}_{3}$.

So we have $\boldsymbol{a}_{2}=4 \boldsymbol{a}_{1}$ and $\boldsymbol{a}_{4}=2 \boldsymbol{a}_{1}-\boldsymbol{a}_{3}$.
Thus we may discard $\boldsymbol{a}_{2}$ and $\boldsymbol{a}_{4}$ when selecting a minimal spanning set for ColA.
$\left\{\boldsymbol{a}_{1}, \boldsymbol{a}_{3}, \boldsymbol{a}_{5}\right\}$ must be linearly independent because any linear dependence relationship among $\boldsymbol{a}_{1}, \boldsymbol{a}_{3}, \boldsymbol{a}_{5}$ would imply a linear dependence relationship among $\boldsymbol{b}_{1}, \boldsymbol{b}_{3}, \boldsymbol{b}_{5}$. But we know that $\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{3}, \boldsymbol{b}_{5}\right\}$ is a linearly independent set.
Thus $\left\{\boldsymbol{a}_{1}, \boldsymbol{a}_{3}, \boldsymbol{a}_{5}\right\}$ is a basis for $\operatorname{Col} A$.

## Basis for ColA and Pivot Columns

## Theorem

The pivot columns of a matrix $A$ form a basis for $\operatorname{Col} A$.

- Let $B$ be the reduced echelon form of $A$. The set of pivot columns of $B$ is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since $A$ is row equivalent to $B$, the pivot columns of $A$ are linearly independent as well, because any linear dependence relation among the columns of $A$ corresponds to a linear dependence relation among the columns of $B$. For this same reason, every non-pivot column of $A$ is a linear combination of the pivot columns of $A$. Thus the non-pivot columns of $A$ may be discarded from the spanning set for $\operatorname{Col} A$, by the Spanning Set Theorem. This leaves the pivot columns of $A$ as a basis for $\operatorname{Col} A$.


## Two Views of a Basis

- When the Spanning Set Theorem is used, the deletion of vectors from a spanning set must stop when the set becomes linearly independent. If an additional vector is deleted, it will not be a linear combination of the remaining vectors, and hence the smaller set will no longer span $V$.

Thus a basis is a spanning set that is as small as possible.

- A basis is also a linearly independent set that is as large as possible. If $S$ is a basis for $V$, and if $S$ is enlarged by one vector say, $\boldsymbol{w}$ from $V$, then the new set cannot be linearly independent, because $S$ spans $V$, and $\boldsymbol{w}$ is therefore a linear combination of the elements in $S$.


## Example

- The following three sets in $\mathbb{R}^{3}$ show how a linearly independent set can be enlarged to a basis and how further enlargement destroys the linear independence of the set.
- Also, a spanning set can be shrunk to a basis, but further shrinking destroys the spanning property.

$$
\left\{\underset{\text { Linearly Independent }}{\left\{\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right]\right.}\left[\begin{array}{l}
2 \\
3 \\
0
\end{array}\right]\right\},\left\{\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right], \underset{A}{[ } \underset{A}{\left[\begin{array}{l}
2 \\
3 \\
0
\end{array}\right]},\left[\begin{array}{l}
4 \\
5 \\
6
\end{array}\right]\right\},
$$

Not Spanning $\mathbb{R}^{3}$

$$
\left\{\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right],\left[\begin{array}{l}
2 \\
3 \\
0
\end{array}\right],\left[\begin{array}{l}
4 \\
5 \\
6
\end{array}\right],\left[\begin{array}{l}
7 \\
8 \\
9
\end{array}\right]\right\}
$$

## Subsection 4

## Coordinate Systems

## The Unique Representation Theorem

## Theorem (The Unique Representation Theorem)

Let $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ be a basis for a vector space $V$. Then for each $\boldsymbol{x}$ in $V$, there exists a unique set of scalars $c_{1}, \ldots, c_{n}$ such that

$$
\boldsymbol{x}=c_{1} \boldsymbol{b}_{1}+\cdots+c_{n} \boldsymbol{b}_{n} .
$$

- Since $\mathcal{B}$ spans $V$, there exist scalars such that the equation holds. Suppose $\boldsymbol{x}$ also has the representation

$$
\boldsymbol{x}=d_{1} \boldsymbol{b}_{1}+\cdots+d_{n} \boldsymbol{b}_{n}
$$

for scalars $d_{1}, \ldots, d_{n}$. Then, subtracting, we have

$$
\mathbf{0}=\boldsymbol{x}-\boldsymbol{x}=\left(c_{1}-d_{1}\right) \boldsymbol{b}_{1}+\cdots+\left(c_{n}-d_{n}\right) \boldsymbol{b}_{n} .
$$

Since $\mathcal{B}$ is linearly independent, the weights must all be zero. That is, $c_{j}=d_{j}$ for $1 \leq j \leq n$.

## Coordinates Relative to a Basis

## Definition

Suppose $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ is a basis for $V$ and $\boldsymbol{x}$ is in $V$. The coordinates of $\boldsymbol{x}$ relative to the basis $\mathcal{B}$ (or the $\mathcal{B}$-coordinates of $\boldsymbol{x}$ ) are the weights $c_{1}, \ldots, c_{n}$ such that $\boldsymbol{x}=c_{1} \boldsymbol{b}_{1}+\cdots+c_{n} \boldsymbol{b}_{n}$.

- If $c_{1}, \ldots, c_{n}$ are the $\mathcal{B}$-coordinates of $\boldsymbol{x}$, then the vector in $\mathbb{R}^{n}$

$$
[\boldsymbol{x}]_{\mathcal{B}}=\left[\begin{array}{c}
c_{1} \\
\vdots \\
c_{n}
\end{array}\right]
$$

is the coordinate vector of $\boldsymbol{x}($ relative to $\mathcal{B})$, or the $\mathcal{B}$-coordinate vector of $x$.

- The mapping $x \mapsto[x]_{\mathcal{B}}$ is the coordinate mapping (determined by $\mathcal{B})$.


## Example

- Consider a basis $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}$ for $\mathbb{R}^{2}$, where $\boldsymbol{b}_{1}=\left[\begin{array}{l}1 \\ 0\end{array}\right]$ and $\boldsymbol{b}_{2}=\left[\begin{array}{l}1 \\ 2\end{array}\right]$.
- Suppose an $\boldsymbol{x}$ in $\mathbb{R}^{2}$ has the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}=\left[\begin{array}{r}-2 \\ 3\end{array}\right]$.
- The $\mathcal{B}$-coordinates of $\boldsymbol{x}$ tell how to build $\boldsymbol{x}$ from the vectors in $\mathcal{B}$. That is,

$$
\boldsymbol{x}=(-2) \boldsymbol{b}_{1}+3 \boldsymbol{b}_{2}=(-2)\left[\begin{array}{l}
1 \\
0
\end{array}\right]+3\left[\begin{array}{l}
1 \\
2
\end{array}\right]=\left[\begin{array}{l}
1 \\
6
\end{array}\right]
$$

## Example

- The entries in the vector $\boldsymbol{x}=\left[\begin{array}{l}1 \\ 6\end{array}\right]$ are the coordinates of $\boldsymbol{x}$ relative to the standard basis $\mathcal{E}=\left\{\boldsymbol{e}_{1}, \boldsymbol{e}_{2}\right\}$.
- Indeed, we have

$$
\left[\begin{array}{l}
1 \\
6
\end{array}\right]=1 \cdot\left[\begin{array}{l}
1 \\
0
\end{array}\right]+6 \cdot\left[\begin{array}{l}
0 \\
1
\end{array}\right]=1 \cdot \boldsymbol{e}_{1}+6 \cdot \boldsymbol{e}_{2} .
$$

- If $\mathcal{E}=\left\{\boldsymbol{e}_{1}, \boldsymbol{e}_{2}\right\}$, then $[\boldsymbol{x}]_{\mathcal{E}}=\boldsymbol{x}$.


## Example

- Let $\boldsymbol{b}_{1}=\left[\begin{array}{l}2 \\ 1\end{array}\right], \boldsymbol{b}_{2}=\left[\begin{array}{r}-1 \\ 1\end{array}\right], \boldsymbol{x}=\left[\begin{array}{l}4 \\ 5\end{array}\right]$ and $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}$. Find the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ of $\boldsymbol{x}$ relative to $\mathcal{B}$.
- The $\mathcal{B}$-coordinates $c_{1}, c_{2}$ of $\boldsymbol{x}$ satisfy

$$
c_{1}\left[\begin{array}{l}
2 \\
1
\end{array}\right]+c_{2}\left[\begin{array}{r}
-1 \\
1
\end{array}\right]=\left[\begin{array}{l}
4 \\
5
\end{array}\right] \quad \text { or } \quad\left[\begin{array}{rr}
2 & -1 \\
1 & 1
\end{array}\right]\left[\begin{array}{l}
c_{1} \\
c_{2}
\end{array}\right]=\left[\begin{array}{l}
4 \\
5
\end{array}\right] .
$$

The solution is

$$
c_{1}=\frac{\left|\begin{array}{rr}
4 & -1 \\
5 & 1
\end{array}\right|}{\left|\begin{array}{rr}
2 & -1 \\
1 & 1
\end{array}\right|}=\frac{9}{3}=3, \quad c_{2}=\frac{\left|\begin{array}{ll}
2 & 4 \\
1 & 5
\end{array}\right|}{\left|\begin{array}{rr}
2 & -1 \\
1 & 1
\end{array}\right|}=\frac{6}{3}=2
$$

Thus $\boldsymbol{x}=3 \boldsymbol{b}_{1}+2 \boldsymbol{b}_{2}$, and $[\boldsymbol{x}]_{\mathcal{B}}=\left[\begin{array}{l}3 \\ 2\end{array}\right]$.

## Change-of-Coordinates Matrix

- Let $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ be a basis for $\mathbb{R}^{n}$.
- Let $P_{\mathcal{B}}=\left[\begin{array}{llll}\boldsymbol{b}_{1} & \boldsymbol{b}_{2} & \cdots & \boldsymbol{b}_{n}\end{array}\right]$.
- Then the vector equation $\boldsymbol{x}=c_{1} \boldsymbol{b}_{1}+c_{2} \boldsymbol{b}_{2}+\cdots+c_{n} \boldsymbol{b}_{n}$ is equivalent to

$$
\boldsymbol{x}=P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}
$$

- We call $P_{\mathcal{B}}$ the change-of-coordinates matrix from $\mathcal{B}$ to the standard basis in $\mathbb{R}^{n}$.
- Left-multiplication by $P_{\mathcal{B}}$ transforms the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ into $x$.


## Properties of the Change-of-Coordinates Matrix

- Since the columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^{n}, P_{\mathcal{B}}$ is invertible (by the Invertible Matrix Theorem).
- Left-multiplication by $P_{\mathcal{B}}^{-1}$ converts $\boldsymbol{x}$ into its $\mathcal{B}$-coordinate vector:

$$
P_{\mathcal{B}}^{-1} x=[\boldsymbol{x}]_{\mathcal{B}}
$$

- The correspondence $\boldsymbol{x} \mapsto[\boldsymbol{x}]_{\mathcal{B}}$, produced here by $P_{\mathcal{B}}^{-1}$, is the coordinate mapping mentioned earlier.
- Since $P_{\mathcal{B}}^{-1}$ is an invertible matrix, the coordinate mapping is a one-to-one linear transformation from $\mathbb{R}^{n}$ onto $\mathbb{R}^{n}$, by the Invertible Matrix Theorem.


## Coordinatization

- Choosing a basis $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ for a vector space $V$ introduces a coordinate system in $V$.
- The coordinate mapping $x \mapsto[\boldsymbol{x}]_{\mathcal{B}}$ connects the possibly unfamiliar space $V$ to the familiar space $\mathbb{R}^{n}$.

- Points in $V$ can now be identified by their new "names".


## Properties of the Coordinate Mapping

## Theorem

Let $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ be a basis for a vector space $V$. Then the coordinate mapping $\boldsymbol{x} \mapsto[\boldsymbol{x}]_{\mathcal{B}}$ is a one-to-one linear transformation from $V$ onto $\mathbb{R}^{n}$.

- Take two typical vectors in $V$, say,

$$
\begin{aligned}
\boldsymbol{u} & =c_{1} \boldsymbol{b}_{1}+\cdots+c_{n} \boldsymbol{b}_{n} \\
\boldsymbol{w} & =d_{1} \boldsymbol{b}_{1}+\cdots+d_{n} \boldsymbol{b}_{n} .
\end{aligned}
$$

Then, using vector operations,

$$
\boldsymbol{u}+\boldsymbol{w}=\left(c_{1}+d_{1}\right) \boldsymbol{b}_{1}+\cdots+\left(c_{n}+d_{n}\right) \boldsymbol{b}_{n} .
$$

It follows that

$$
[\boldsymbol{u}+\boldsymbol{w}]_{\mathcal{B}}=\left[\begin{array}{c}
c_{1}+d_{1} \\
\vdots \\
c_{n}+d_{n}
\end{array}\right]=\left[\begin{array}{c}
c_{1} \\
\vdots \\
c_{n}
\end{array}\right]+\left[\begin{array}{c}
d_{1} \\
\vdots \\
d_{n}
\end{array}\right]=[\boldsymbol{u}]_{\mathcal{B}}+[\boldsymbol{w}]_{\mathcal{B}}
$$

So the coordinate mapping preserves addition.

## Properties of the Coordinate Mapping (Cont'd)

- If $r$ is any scalar, then

$$
r \boldsymbol{u}=r\left(c_{1} \boldsymbol{b}_{1}+\cdots+c_{n} \boldsymbol{b}_{n}\right)=\left(r c_{1}\right) \boldsymbol{b}_{1}+\cdots+\left(r c_{n}\right) \boldsymbol{b}_{n} .
$$

So

$$
[r \boldsymbol{u}]_{\mathcal{B}}=\left[\begin{array}{c}
r c_{1} \\
\vdots \\
r c_{n}
\end{array}\right]=r\left[\begin{array}{c}
c_{1} \\
\vdots \\
c_{n}
\end{array}\right]=r[\boldsymbol{u}]_{\mathcal{B}} .
$$

Thus the coordinate mapping also preserves scalar multiplication and hence is a linear transformation.

The coordinate mapping is one-to-one: If $[\boldsymbol{u}]_{\mathcal{B}}=[\boldsymbol{w}]_{\mathcal{B}}$, then $\boldsymbol{u}=P_{\mathcal{B}}[\boldsymbol{u}]_{\mathcal{B}}=P_{\mathcal{B}}[\boldsymbol{w}]_{\mathcal{B}}=\boldsymbol{w}$.
It is also onto: If $\boldsymbol{y}$ in $\mathbb{R}^{n}$, then letting $\boldsymbol{u}=P_{\mathcal{B}} \boldsymbol{y}$, we get
$[\boldsymbol{u}]_{\mathcal{B}}=P_{\mathcal{B}}^{-1} \boldsymbol{u}=P_{\mathcal{B}}^{-1}\left(P_{\mathcal{B}} \boldsymbol{y}\right)=\boldsymbol{y}$.

## Isomorphisms of Vector Spaces

- The linearity of the coordinate mapping extends to linear combinations: If $\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{p}$ are in $V$ and if $c_{1}, \ldots, c_{p}$ are scalars, then

$$
\left[c_{1} \boldsymbol{u}_{1}+\cdots+c_{p} \boldsymbol{u}_{p}\right]_{\mathcal{B}}=c_{1}\left[\boldsymbol{u}_{1}\right]_{\mathcal{B}}+\cdots+c_{p}\left[\boldsymbol{u}_{p}\right]_{\mathcal{B}}
$$

- A one-to-one linear transformation from a vector space $V$ onto a vector space $W$ is called an isomorphism from $V$ onto $W$.
- The notation and terminology for $V$ and $W$ may differ, but the two spaces are indistinguishable as vector spaces.
- Every vector space calculation in $V$ is accurately reproduced in $W$, and vice versa.
- In particular, any real vector space with a basis of $n$ vectors is indistinguishable from $\mathbb{R}^{n}$.


## Example

- Let $\mathcal{B}$ be the standard basis of the space $\mathbb{P}_{3}$ of polynomials; that is, let $\mathcal{B}=\left\{1, t, t^{2}, t^{3}\right\}$.
- A typical element $\boldsymbol{p}$ of $\mathbb{P}_{3}$ has the form

$$
\boldsymbol{p}(t)=a_{0}+a_{1} t+a_{2} t^{2}+a_{3} t^{3} .
$$

- Since $\boldsymbol{p}$ is already displayed as a linear combination of the standard basis vectors, we conclude that

$$
[\boldsymbol{p}]_{\mathcal{B}}=\left[\begin{array}{l}
a_{0} \\
a_{1} \\
a_{2} \\
a_{3}
\end{array}\right]
$$

- Thus the coordinate mapping $\boldsymbol{p} \mapsto[\boldsymbol{p}]_{\mathcal{B}}$ is an isomorphism from $\mathbb{P}_{3}$ onto $\mathbb{R}^{4}$.
- All vector space operations in $\mathbb{P}_{3}$ correspond to operations in $\mathbb{R}^{4}$.


## Example

- Use coordinate vectors to verify that the polynomials $1+2 t^{2}$, $4+t+5 t^{2}$, and $3+2 t$ are linearly dependent in $\mathbb{P}_{2}$.
- The coordinate mapping from the preceding example produces the coordinate vectors $(1,0,2),(4,1,5)$ and $(3,2,0)$, respectively. Writing these vectors as the columns of a matrix $A$, we can determine their independence by row reducing the augmented matrix for $\boldsymbol{A x}=\mathbf{0}$ :

$$
\left[\begin{array}{lll|l}
1 & 4 & 3 & 0 \\
0 & 1 & 2 & 0 \\
2 & 5 & 0 & 0
\end{array}\right] \rightarrow\left[\begin{array}{rrr|r}
1 & 4 & 3 & 0 \\
0 & 1 & 2 & 0 \\
0 & -3 & -6 & 0
\end{array}\right] \rightarrow\left[\begin{array}{llll}
1 & 4 & 3 & 0 \\
0 & 1 & 2 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]
$$

The columns of $A$ are linearly dependent. So the corresponding polynomials are linearly dependent.
In fact, it is easy to check that column 3 of $A$ is 2 times column 2 minus 5 times column 1.
The corresponding relation for the polynomials is

$$
3+2 t=2\left(4+t+5 t^{2}\right)-5\left(1+2 t^{2}\right)
$$

## Example

- Let $\boldsymbol{v}_{1}=\left[\begin{array}{l}3 \\ 6 \\ 2\end{array}\right], \boldsymbol{v}_{2}=\left[\begin{array}{r}-1 \\ 0 \\ 1\end{array}\right], \boldsymbol{x}=\left[\begin{array}{r}3 \\ 12 \\ 7\end{array}\right]$ and $\mathcal{B}=\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$.

Then $\mathcal{B}$ is a basis for $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$. Determine if $\boldsymbol{x}$ is in $H$, and if it is, find the coordinate vector of $\boldsymbol{x}$ relative to $\mathcal{B}$.

- If $\boldsymbol{x}$ is in $H$, then the following vector equation is consistent:

$$
c_{1}\left[\begin{array}{l}
3 \\
6 \\
2
\end{array}\right]+c_{2}\left[\begin{array}{r}
-1 \\
0 \\
1
\end{array}\right]=\left[\begin{array}{r}
3 \\
12 \\
7
\end{array}\right]
$$

The scalars $c_{1}$ and $c_{2}$, if they exist, are the $\mathcal{B}$-coordinates of $\boldsymbol{x}$. In matrix form

$$
\left[\begin{array}{rr}
3 & -1 \\
6 & 0 \\
2 & 1
\end{array}\right]\left[\begin{array}{l}
c_{1} \\
c_{2}
\end{array}\right]=\left[\begin{array}{r}
3 \\
12 \\
7
\end{array}\right] .
$$

## Example (Cont'd)

- Using row operations, we obtain

$$
\begin{aligned}
& {\left[\begin{array}{rrr}
3 & -1 & 3 \\
6 & 0 & 12 \\
2 & 1 & 7
\end{array}\right] \underset{\substack{ \\
R_{1} \leftrightarrow R_{2}}}{\substack{\leftarrow \frac{1}{6} R_{2}}}\left[\begin{array}{rrr}
1 & 0 & 2 \\
3 & -1 & 3 \\
2 & 1 & 7
\end{array}\right]} \\
& \underset{R_{2} \leftarrow R_{2}-3 R_{1}}{\substack{R_{3}-2 R_{1}}}\left[\begin{array}{rrr}
1 & 0 & 2 \\
0 & -1 & -3 \\
0 & 1 & 3
\end{array}\right] \xrightarrow{R_{3} \leftarrow R_{3}+R_{2}} \begin{array}{rlll}
R_{2} \leftarrow(-1) R_{2}
\end{array}\left[\begin{array}{lll}
1 & 0 & 2 \\
0 & 1 & 3 \\
0 & 0 & 0
\end{array}\right] .
\end{aligned}
$$

Thus $c_{1}=2, c_{2}=3$, and $[x]_{\mathcal{B}}=\left[\begin{array}{l}2 \\ 3\end{array}\right]$.

## Subsection 5

## The Dimension of a Vector Space

## Dependent Vectors in a Vector Space

## Theorem

If a vector space $V$ has a basis $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$, then any set in $V$ containing more than $n$ vectors must be linearly dependent.

- Let $\left\{\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{p}\right\}$ be a set in $V$ with more than $n$ vectors. The coordinate vectors $\left[\boldsymbol{u}_{1}\right]_{\mathcal{B}}, \ldots,\left[\boldsymbol{u}_{p}\right]_{\mathcal{B}}$ form a linearly dependent set in $\mathbb{R}^{n}$, because there are more vectors $(p)$ than entries $(n)$ in each vector. So there exist scalars $c_{1}, \ldots, c_{p}$, not all zero, such that $c_{1}\left[\boldsymbol{u}_{1}\right]_{\mathcal{B}}+\cdots+c_{p}\left[\boldsymbol{u}_{p}\right]_{\mathcal{B}}=\mathbf{0}$, the zero vector in $\mathbb{R}^{n}$. Since the coordinate mapping is a linear transformation, $\left[c_{1} \boldsymbol{u}_{1}+\cdots+c_{p} \boldsymbol{u}_{p}\right]_{\mathcal{B}}=\mathbf{0}$. The zero vector on the right displays the n weights needed to build the vector $c_{1} \boldsymbol{u}_{1}+\cdots+c_{p} \boldsymbol{u}_{p}$ from the basis vectors in $\mathcal{B}$. That is, $c_{1} \boldsymbol{u}_{1}+\cdots+c_{p} \boldsymbol{u}_{p}=0 \cdot \boldsymbol{b}_{1}+\cdots+0 \cdot \boldsymbol{b}_{n}=\mathbf{0}$. Since the $c_{i}$ are not all zero, $\left\{\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{p}\right\}$ is linearly dependent.


## Number of Vectors in Bases

## Theorem

If a vector space $V$ has a basis of $n$ vectors, then every basis of $V$ must consist of exactly $n$ vectors.

- Let $\mathcal{B}_{1}$ be a basis of $n$ vectors and $\mathcal{B}_{2}$ be any other basis (of $V$ ). Since $\mathcal{B}_{1}$ is a basis and $\mathcal{B}_{2}$ is linearly independent, $\mathcal{B}_{2}$ has no more than $n$ vectors, by the preceding theorem.
Also, since $\mathcal{B}_{2}$ is a basis and $\mathcal{B}_{1}$ is linearly independent, $\mathcal{B}_{2}$ has at least $n$ vectors.
Thus $\mathcal{B}_{2}$ consists of exactly $n$ vectors.


## Dimension of a Vector Space

- If a nonzero vector space $V$ is spanned by a finite set $S$, then a subset of $S$ is a basis for $V$, by the Spanning Set Theorem.
- In this case, the preceding theorem ensures that the following definition makes sense:


## Definition

If $V$ is spanned by a finite set, then $V$ is said to be finite-dimensional, and the dimension of $V$, written as $\operatorname{dim} V$, is the number of vectors in a basis for $V$.
The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. If $V$ is not spanned by a finite set, then $V$ is said to be infinite-dimensional.

## Examples

- The standard basis for $\mathbb{R}^{n}$ contains $n$ vectors. So $\operatorname{dim} \mathbb{R}^{n}=n$.
- The standard polynomial basis $\left\{1, t, t^{2}\right\}$ shows that $\operatorname{dim} \mathbb{P}_{2}=3$.
- In general, $\operatorname{dim} \mathbb{P}_{n}=n+1$.
- The space $\mathbb{P}$ of all polynomials is infinite-dimensional.


## Example

- Let $H=\operatorname{Span}\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$, where $\boldsymbol{v}_{1}=\left[\begin{array}{l}3 \\ 6 \\ 2\end{array}\right], \boldsymbol{v}_{2}=\left[\begin{array}{r}-1 \\ 0 \\ 1\end{array}\right]$.
- Then a basis for $H$ is $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}\right\}$, since $\boldsymbol{v}_{1}$ and $\boldsymbol{v}_{2}$ are not multiples and hence are linearly independent.
- Thus $\operatorname{dim} H=2$.


## Example

- Find the dimension of the subspace

$$
H=\left\{\left[\begin{array}{c}
a-3 b+6 c \\
5 a+4 d \\
b-2 c-d \\
5 d
\end{array}\right]: a, b, c, d \text { in } \mathbb{R}\right\}
$$

- Note that

$$
\left[\begin{array}{c}
a-3 b+6 c \\
5 a+4 d \\
b-2 c-d \\
5 d
\end{array}\right]=a\left[\begin{array}{l}
1 \\
5 \\
0 \\
0
\end{array}\right]+b\left[\begin{array}{r}
-3 \\
0 \\
1 \\
0
\end{array}\right]+c\left[\begin{array}{r}
6 \\
0 \\
-2 \\
0
\end{array}\right]+d\left[\begin{array}{r}
0 \\
4 \\
-1 \\
5
\end{array}\right]
$$

So $H$ is the set of all linear combinations of the vectors

$$
\boldsymbol{v}_{1}=\left[\begin{array}{l}
1 \\
5 \\
0 \\
0
\end{array}\right], \quad \boldsymbol{v}_{2}=\left[\begin{array}{r}
-3 \\
0 \\
1 \\
0
\end{array}\right], \quad \boldsymbol{v}_{3}=\left[\begin{array}{r}
6 \\
0 \\
-2 \\
0
\end{array}\right], \quad \boldsymbol{v}_{4}=\left[\begin{array}{r}
0 \\
4 \\
-1 \\
5
\end{array}\right] .
$$

## Example (Cont'd)

- We saw $H=$ Span $\left\{\left[\begin{array}{l}1 \\ 5 \\ 0 \\ 0\end{array}\right],\left[\begin{array}{r}-3 \\ 0 \\ 1 \\ 0\end{array}\right],\left[\begin{array}{r}6 \\ 0 \\ -2 \\ 0\end{array}\right],\left[\begin{array}{r}0 \\ 4 \\ -1 \\ 5\end{array}\right]\right\}$.

Clearly, $\boldsymbol{v}_{1} \neq \mathbf{0}, \boldsymbol{v}_{2}$ is not a multiple of $\boldsymbol{v}_{1}$, but $\boldsymbol{v}_{3}$ is a multiple of $\boldsymbol{v}_{2}$. By the Spanning Set Theorem, we may discard $\boldsymbol{v}_{3}$ and still have a set that spans $H$.
Finally, $\boldsymbol{v}_{4}$ is not a linear combination of $\boldsymbol{v}_{1}$ and $\boldsymbol{v}_{2}$.
So $\left\{\boldsymbol{v}_{1}, \boldsymbol{v}_{2}, \boldsymbol{v}_{4}\right\}$ is linearly independent and hence is a basis for $H$.
Thus $\operatorname{dim} H=3$.

## Classification of Subspaces of $\mathbb{R}^{3}$

- The subspaces of $\mathbb{R}^{3}$ can be classified by dimension:
- 0-dimensional subspaces: Only the zero subspace.
- 1-dimensional subspaces: Any subspace spanned by a single nonzero vector. Such subspaces are lines through the origin.
- 2-dimensional subspaces: Any subspace spanned by two linearly independent vectors. Such subspaces are planes through the origin.
- 3-dimensional subspaces: Only $\mathbb{R}^{3}$ itself. Any three linearly independent vectors in $\mathbb{R}^{3}$ span all of $\mathbb{R}^{3}$, by the Invertible Matrix Theorem.



## Subspaces of a Finite-Dimensional Space

## Theorem

Let $H$ be a subspace of a finite-dimensional vector space $V$. Any linearly independent set in $H$ can be expanded, if necessary, to a basis for $H$. Also, $H$ is finite-dimensional and $\operatorname{dim} H \leq \operatorname{dim} V$.

- If $H=\{\mathbf{0}\}$, then certainly $\operatorname{dim}=0 \leq \operatorname{dim} V$.

Otherwise, let $S=\left\{\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{k}\right\}$ be any linearly independent set in $H$. If $S$ spans $H$, then $S$ is a basis for $H$. Otherwise, there is some $\boldsymbol{u}_{k+1}$ in $H$ that is not in SpanS. But then $\left\{\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{k}, \boldsymbol{u}_{k+1}\right\}$ will be linearly independent, because no vector in the set can be a linear combination of vectors that precede it (by a previous theorem). So long as the new set does not span $H$, we can continue this process of expanding $S$ to a larger linearly independent set in $H$. But the number of vectors in a linearly independent expansion of $S$ can never exceed the dimension of $V$, by a previous theorem. So eventually the expansion of $S$ will span $H$ and hence will be a basis for $H$.

## The Basis Theorem

## Theorem (The Basis Theorem)

Let $V$ be a $p$-dimensional vector space, $p \geq 1$.

- Any linearly independent set of exactly $p$ elements in $V$ is automatically a basis for $V$.
- Any set of exactly $p$ elements that spans $V$ is automatically a basis for $V$.
- By the preceding theorem, a linearly independent set $S$ of $p$ elements can be extended to a basis for $V$. But that basis must contain exactly $p$ elements, since $\operatorname{dim} V=p$. So $S$ must already be a basis for $V$.
Now suppose that $S$ has $p$ elements and spans $V$. Since $V$ is nonzero, the Spanning Set Theorem implies that a subset $S_{0}$ of $S$ is a basis of $V$. Since $\operatorname{dim} V=p, S_{0}$ must contain $p$ vectors. Hence $S=S_{0}$.


## The Dimensions of Nul $A$ and $\operatorname{Col} A$

- Since the pivot columns of a matrix $A$ form a basis for $\operatorname{Col} A$, we know the dimension of $\operatorname{Col} A$ as soon as we know the pivot columns.
- Let $A$ be an $m \times n$ matrix, and suppose the equation $A \boldsymbol{x}=\mathbf{0}$ has $k$ free variables. We know that the standard method of finding a spanning set for $\mathrm{Nul} A$ will produce exactly $k$ linearly independent vectors say, $\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{k}$ one for each free variable. So $\left\{\boldsymbol{u}_{1}, \ldots, \boldsymbol{u}_{k}\right\}$ is a basis for $\operatorname{Nul} A$, and the number of free variables determines the size of the basis.

> The dimension of Nul $A$ is the number of free variables in the equation $A \boldsymbol{x}=\mathbf{0}$, and the dimension of $\operatorname{Col} A$ is the number of pivot columns in $A$.

## Example

- Find the dimensions of the null space and the column space of

$$
A=\left[\begin{array}{rrrrr}
-3 & 6 & -1 & 1 & -7 \\
1 & -2 & 2 & 3 & -1 \\
2 & -4 & 5 & 8 & -4
\end{array}\right]
$$

- Row reduce the augmented matrix $\left[\begin{array}{ll}A & \mathbf{0}\end{array}\right]$ to echelon form:

$$
\begin{aligned}
& {\left[\begin{array}{rrrrr|r}
-3 & 6 & -1 & 1 & -7 & 0 \\
1 & -2 & 2 & 3 & -1 & 0 \\
2 & -4 & 5 & 8 & -4 & 0
\end{array}\right]} \\
& \longrightarrow\left[\begin{array}{rrrrr|r}
1 & -2 & 2 & 3 & -1 & 0 \\
0 & 0 & 5 & 10 & -10 & 0 \\
0 & 0 & 1 & 2 & -2 & 0
\end{array}\right] \longrightarrow\left[\begin{array}{rrrrr|r}
1 & -2 & 2 & 3 & -1 & 0 \\
-3 & 6 & -1 & 1 & -7 & 0 \\
2 & -4 & 5 & 8 & -4 & 0
\end{array}\right]
\end{aligned}
$$

There are three free variables $x_{2}, x_{4}$ and $x_{5}$.
Hence the dimension of $\operatorname{Nul} A$ is 3 .
Also, $\operatorname{dim} \operatorname{Col} A=2$ because $A$ has two pivot columns.

## Subsection 6

## Rank

## The Row Space

- If $A$ is an $m \times n$ matrix, each row of $A$ has $n$ entries and thus can be identified with a vector in $\mathbb{R}^{n}$.
- The set of all linear combinations of the row vectors is called the row space of $A$ and is denoted by Row $A$.
- Each row has $n$ entries, so $\operatorname{Row} A$ is a subspace of $\mathbb{R}^{n}$.
- Since the rows of $A$ are identified with the columns of $A^{T}$, we could also write $\operatorname{Col} A^{T}$ in place of $\operatorname{Row} A$.


## Example

- Let

$$
A=\left[\begin{array}{rrrrr}
-2 & -5 & 8 & 0 & -17 \\
1 & 3 & -5 & 1 & 5 \\
3 & 11 & -19 & 7 & 1 \\
1 & 7 & -13 & 5 & -3
\end{array}\right], \begin{aligned}
& \boldsymbol{r}_{1}=(-2,-5,8,0,-17) \\
& \boldsymbol{r}_{2}=(1,3,-5,1,5) \\
& \boldsymbol{r}_{3}=(3,11,-19,7,1) \\
& \boldsymbol{r}_{4}=(1,7,-13,5,-3) .
\end{aligned}
$$

- The row space of $A$ is the subspace of $\mathbb{R}^{5}$ spanned by $\left\{\boldsymbol{r}_{1}, \boldsymbol{r}_{2}, \boldsymbol{r}_{3}, \boldsymbol{r}_{4}\right\}$.
- That is, $\operatorname{Row} A=\operatorname{Span}\left\{\boldsymbol{r}_{1}, \boldsymbol{r}_{2}, \boldsymbol{r}_{3}, \boldsymbol{r}_{4}\right\}$.
- It is natural to write row vectors horizontally;
- However, they may also be written as column vectors if that is more convenient.


## Row Space and Elementary Operations

## Theorem

If two matrices $A$ and $B$ are row equivalent, then their row spaces are the same. If $B$ is in echelon form, the nonzero rows of $B$ form a basis for the row space of $A$ as well as for that of $B$.

- If $B$ is obtained from $A$ by row operations, the rows of $B$ are linear combinations of the rows of $A$. It follows that any linear combination of the rows of $B$ is automatically a linear combination of the rows of $A$. Thus the row space of $B$ is contained in the row space of $A$. Since row operations are reversible, the same argument shows that the row space of $A$ is a subset of the row space of $B$.
So the two row spaces are the same.
If $B$ is in echelon form, its nonzero rows are linearly independent because no nonzero row is a linear combination of the nonzero rows below it. Thus the nonzero rows of $B$ form a basis of the (common) row space of $B$ and $A$.


## The Spaces RowA, ColA and NulA

- Find bases for the row space, the column space and the null space of
the matrix $A=\left[\begin{array}{rrrrr}-2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3\end{array}\right]$.
- To find bases for the row space and the column space, row reduce $A$ to an echelon form:

$$
\begin{gathered}
A \longrightarrow\left[\begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
-2 & -5 & 8 & 0 & -17 \\
3 & 11 & -19 & 7 & 1 \\
1 & 7 & -13 & 5 & -3
\end{array}\right] \\
\longrightarrow\left[\begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
0 & 1 & -2 & 2 & -7 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -4 & 20
\end{array}\right] \longrightarrow\left[\begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
0 & 1 & -2 & 2 & -7 \\
0 & 2 & -4 & 4 & -14 \\
0 & 4 & -8 & 4 & -8
\end{array}\right] \\
\left.\hline \begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
0 & 1 & -2 & 2 & -7 \\
0 & 0 & 0 & -4 & 20 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{gathered}
$$

## The Spaces RowA (Cont'd)

- We found

$$
A \sim\left[\begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
0 & 1 & -2 & 2 & -7 \\
0 & 0 & 0 & -4 & 20 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

By the theorem, the first three rows of $B$ form a basis for the row space of $A$ (as well as for the row space of $B$ ):

Basis for Row $A:\{(1,3,-5,1,5),(0,1,-2,2,-7),(0,0,0,-4,20)\}$.

## The Spaces ColA (Cont'd)

- We found

$$
A \sim\left[\begin{array}{rrrrr}
1 & 3 & -5 & 1 & 5 \\
0 & 1 & -2 & 2 & -7 \\
0 & 0 & 0 & -4 & 20 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

For the column space, observe from $B$ that the pivots are in columns 1,2 , and 4.
Hence columns 1,2 , and 4 of $A($ not $B)$ form a basis for $\operatorname{Col} A$ :

$$
\text { Basis for Col } A:\left\{\left[\begin{array}{r}
-2 \\
1 \\
3 \\
1
\end{array}\right],\left[\begin{array}{r}
-5 \\
3 \\
11 \\
7
\end{array}\right],\left[\begin{array}{l}
0 \\
1 \\
7 \\
5
\end{array}\right]\right\} .
$$

## The Space NulA (Cont'd)

- For $\operatorname{Nul} A$, we need the reduced echelon form:
$A \sim\left[\begin{array}{rrrrr}1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0\end{array}\right] \sim\left[\begin{array}{rrrrr}1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0\end{array}\right]$.
The equation $A \boldsymbol{x}=\mathbf{0}$ is equivalent to

$$
\left\{\begin{array}{rlrl}
x_{1}+x_{3} & +x_{5}=0 \\
x_{2}-2 x_{3} & & +3 x_{5}=0 \\
& & x_{4}-5 x_{5}=0
\end{array} .\right.
$$

So $x_{1}=-x_{3}-x_{5}, x_{2}=2 x_{3}-3 x_{5}, x_{4}=5 x_{5}$, with $x_{3}$ and $x_{5}$ free:

$$
\left[\begin{array}{l}
x_{1} \\
x_{2} \\
x_{3} \\
x_{4} \\
x_{5}
\end{array}\right]=x_{3}\left[\begin{array}{r}
-1 \\
2 \\
1 \\
0 \\
0
\end{array}\right]+x_{5}\left[\begin{array}{r}
-1 \\
-3 \\
0 \\
5 \\
1
\end{array}\right] \text {; Basis for NulA: }\left\{\left[\begin{array}{r}
-1 \\
2 \\
1 \\
0 \\
0
\end{array}\right],\left[\begin{array}{r}
-1 \\
-3 \\
0 \\
5 \\
1
\end{array}\right]\right\} .
$$

## Rank and Nullity of a Matrix

## Definition

The rank of $A$ is the dimension of the column space of $A$.

- Since Row $A$ is the same as $\operatorname{Col} A^{T}$, the dimension of the row space of $A$ is the rank of $A^{T}$.
- The dimension of the null space is sometimes called the nullity of $A$.


## The Rank Theorem

## Theorem (The Rank Theorem)

The dimensions of the column space and the row space of an $m \times n$ matrix $A$ are equal. This common dimension, the rank of $A$, also equals the number of pivot positions in $A$ and satisfies the equation

$$
\operatorname{rank} A+\operatorname{dimNul} A=n
$$

- By a previous theorem, rank $A$ is the number of pivot columns in $A$. Equivalently, rank $A$ is the number of pivot positions in an echelon form $B$ of $A$. Furthermore, since $B$ has a nonzero row for each pivot, and since these rows form a basis for the row space of $A$, the rank of $A$ is also the dimension of the row space.
We know that the dimension of $\operatorname{Nul} A$ equals the number of free variables in the equation $A \boldsymbol{x}=\mathbf{0}$. Expressed another way, the dimension of $\operatorname{Nul} A$ is the number of columns of $A$ that are not pivot columns. Obviously, the sum of these two numbers in $n$.


## Example

(a) If $A$ is a $7 \times 9$ matrix with a two-dimensional null space, what is the rank of $A$ ?
(b) Could a $6 \times 9$ matrix have a two-dimensional null space?
(a) Since $A$ has 9 columns,

$$
\begin{aligned}
\operatorname{rank} A+\operatorname{dimNul} A=9 & \Rightarrow \quad \operatorname{rank} A+2=9 \\
& \Rightarrow \quad \operatorname{rank} A=7 .
\end{aligned}
$$

(b) If a $6 \times 9$ matrix, call it $B$, had a two-dimensional null space, it would have to have rank 7 .
But the columns of $B$ are vectors in $\mathbb{R}^{6}$.
So the dimension of $\operatorname{Col} B$, i.e., rank $B$ cannot exceed 6 .
So a $6 \times 9$ matrix cannot have a two-dimensional null space.

## The Invertible Matrix Theorem

## Theorem (The Invertible Matrix Theorem Cont'd)

Let $A$ be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that $A$ is an invertible matrix:
(m) The columns of $A$ form a basis of $\mathbb{R}^{n}$.
(n) $\operatorname{Col} A=\mathbb{R}^{n}$.
(o) $\operatorname{dim} \operatorname{Col} A=n$.
(p) $\operatorname{rank} A=n$.
(q) $\operatorname{Nul} A=\{\mathbf{0}\}$.
(r) $\operatorname{dimNul} A=0$.

- Obvious statements about the row space of $A$ may be added to the Invertible Matrix Theorem, because the row space is the column space of $A^{T}$ and $A$ is invertible if and only if $A^{T}$ is invertible.


## Subsection 7

## Change of Basis

## Changing Bases

- When a basis $\mathcal{B}$ is chosen for an $n$-dimensional vector space $V$, the associated coordinate mapping onto $\mathbb{R}^{n}$ provides a coordinate system for $V$.
- Each $\boldsymbol{x}$ in $V$ is identified uniquely by its $\mathcal{B}$-coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$.
- In some applications, a problem is described initially using a basis $\mathcal{B}$, but the problem's solution is aided by changing $\mathcal{B}$ to a new basis $\mathcal{C}$.
- Each vector is assigned a new $\mathcal{C}$-coordinate vector.
- We study how $[\boldsymbol{x}]_{\mathcal{C}}$ and $[\boldsymbol{x}]_{\mathcal{B}}$ are related for each $\boldsymbol{x}$ in $V$.


## Example

- Consider two bases $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}$ and $\mathcal{C}=\left\{\boldsymbol{c}_{1}, \boldsymbol{c}_{2}\right\}$ for a vector space $V$, such that $\boldsymbol{b}_{1}=4 \boldsymbol{c}_{1}+\boldsymbol{c}_{2}$ and $\boldsymbol{b}_{2}=-6 \boldsymbol{c}_{1}+\boldsymbol{c}_{2}$. Suppose $\boldsymbol{x}=3 \boldsymbol{b}_{1}+\boldsymbol{b}_{2}$. That is, suppose $[\boldsymbol{x}]_{\mathcal{B}}=\left[\begin{array}{l}3 \\ 1\end{array}\right]$. Find $[\boldsymbol{x}]_{\mathcal{C}}$.
- Apply the coordinate mapping determined by $\mathcal{C}$ to $\boldsymbol{x}$, taking into account that it is a linear transformation,

$$
[\boldsymbol{x}]_{\mathcal{C}}=\left[3 \boldsymbol{b}_{1}+\boldsymbol{b}_{2}\right]_{\mathcal{C}}=3\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}+\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}
$$

We can write this vector equation as a matrix equation, using the vectors in the linear combination as the columns of a matrix:

$$
[\boldsymbol{x}]_{\mathcal{C}}=\left[\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}\right]\left[\begin{array}{l}
3 \\
1
\end{array}\right]
$$

## Example (Cont'd)

- This formula gives $[\boldsymbol{x}]_{\mathcal{C}}$, once we know the columns of the matrix. We have $\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}=\left[\begin{array}{l}4 \\ 1\end{array}\right],\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}=\left[\begin{array}{r}-6 \\ 1\end{array}\right]$.
Thus we get the solution:

$$
\begin{aligned}
{[\boldsymbol{x}]_{\mathcal{C}} } & =\left[\begin{array}{ll}
{\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}} & {\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}}
\end{array}\right]\left[\begin{array}{l}
3 \\
1
\end{array}\right] \\
& =\left[\begin{array}{rr}
4 & -6 \\
1 & 1
\end{array}\right]\left[\begin{array}{l}
3 \\
1
\end{array}\right]=\left[\begin{array}{l}
6 \\
4
\end{array}\right]
\end{aligned}
$$

## The Change-of-Coordinates Matrix

## Theorem

Let $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ and $\mathcal{C}=\left\{\boldsymbol{c}_{1}, \ldots, \boldsymbol{c}_{n}\right\}$ be bases of a vector space $V$. Then there is a unique $n \times n$ matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ such that $[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}}$. The columns of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ are the $\mathcal{C}$-coordinate vectors of the vectors in the basis $\mathcal{B}$. That is,

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}=\left[[ \boldsymbol { b } _ { 1 } ] _ { \mathcal { C } } \left[\begin{array}{lll}
\left.\boldsymbol{b}_{2}\right]_{\mathcal{C}} & \cdots & \left.\left[\boldsymbol{b}_{n}\right]_{\mathcal{C}}\right] .
\end{array}\right.\right.
$$

- $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ is called the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.



## Properties of the Change-of-Coordinates Matrix

- The columns of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ are linearly independent because they are the coordinate vectors of the linearly independent set $\mathcal{B}$.
- Since $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ is square, it must be invertible, by the Invertible Matrix Theorem.
- Left-multiplying both sides of equation $[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}}$ by $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P-1}$ yields

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}{ }^{-1}[x]_{\mathcal{C}}=[\boldsymbol{x}]_{\mathcal{B}}
$$

- Thus $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ is the matrix that converts $\mathcal{C}$-coordinates into $\mathcal{B}$-coordinates, i.e.,

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P-1}=\underset{\mathcal{B} \leftarrow \mathcal{C}}{P}
$$

## The Case of the Standard Basis

- If $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ and $\mathcal{E}$ is the standard basis $\left\{\boldsymbol{e}_{1}, \ldots, \boldsymbol{e}_{n}\right\}$ in $\mathbb{R}^{n}$, then $\left[\boldsymbol{b}_{1}\right]_{\mathcal{E}}=\boldsymbol{b}_{1}$;
- Likewise for the other vectors in $\mathcal{B}$.
- In this case, $\underset{\mathcal{E} \leftarrow \mathcal{B}}{P}$ is the same as the change-of coordinates matrix $P_{\mathcal{B}}$ introduced previously, namely,

$$
\underset{\mathcal{E} \leftarrow \mathcal{B}}{P}=P_{\mathcal{B}}=\left[\begin{array}{llll}
\boldsymbol{b}_{1} & \boldsymbol{b}_{2} & \cdots & \boldsymbol{b}_{n}
\end{array}\right] .
$$

- To change coordinates between two nonstandard bases in $\mathbb{R}^{n}$, we need the theorem.
- The theorem shows that to solve the change-of-basis problem, we need the coordinate vectors of the old basis relative to the new basis.


## Example

- Let $\boldsymbol{b}_{1}=\left[\begin{array}{r}-9 \\ 1\end{array}\right], \boldsymbol{b}_{2}=\left[\begin{array}{l}-5 \\ -1\end{array}\right], \boldsymbol{c}_{1}=\left[\begin{array}{r}1 \\ -4\end{array}\right], \boldsymbol{c}_{2}=\left[\begin{array}{r}3 \\ -5\end{array}\right]$.

Consider the bases for $\mathbb{R}^{2}$ given by $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}$ and $\mathcal{C}=\left\{\boldsymbol{c}_{1}, \boldsymbol{c}_{2}\right\}$.
Find the change-of coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.

- The matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ involves the $\mathcal{C}$-coordinate vectors of $\boldsymbol{b}_{1}$ and $\boldsymbol{b}_{2}$.

Let $\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}=\left[\begin{array}{l}x_{1} \\ x_{2}\end{array}\right],\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}=\left[\begin{array}{l}y_{1} \\ y_{2}\end{array}\right]$.
Then, by definition,

$$
x_{1} \boldsymbol{c}_{1}+x_{2} \boldsymbol{c}_{2}=\boldsymbol{b}_{1}, \quad y_{1} \boldsymbol{c}_{1}+y_{2} \boldsymbol{c}_{2}=\boldsymbol{b}_{2} .
$$

Equivalently in matrix form,

$$
\left[\begin{array}{ll}
\boldsymbol{c}_{1} & \boldsymbol{c}_{2}
\end{array}\right]\left[\begin{array}{l}
x_{1} \\
x_{2}
\end{array}\right]=\boldsymbol{b}_{1}, \quad\left[\begin{array}{ll}
\boldsymbol{c}_{1} & \boldsymbol{c}_{2}
\end{array}\right]\left[\begin{array}{l}
y_{1} \\
y_{2}
\end{array}\right]=\boldsymbol{b}_{2}
$$

## Example (Cont'd)

- To solve both systems simultaneously, augment the coefficient matrix with $\boldsymbol{b}_{1}$ and $\boldsymbol{b}_{2},\left[\begin{array}{cc|cc}\boldsymbol{c}_{1} & \boldsymbol{c}_{2} & \boldsymbol{b}_{1} & \boldsymbol{b}_{2}\end{array}\right]$, and row reduce:

$$
\begin{aligned}
& {\left[\begin{array}{rr|rr}
1 & 3 & -9 & -5 \\
-4 & -5 & 1 & -1
\end{array}\right] \rightarrow\left[\begin{array}{ll|rr}
1 & 3 & -9 & -5 \\
0 & 7 & -35 & -21
\end{array}\right]} \\
& \longrightarrow\left[\begin{array}{ll|ll}
1 & 3 & -9 & -5 \\
0 & 1 & -5 & -3
\end{array}\right] \rightarrow\left[\begin{array}{ll|rr}
1 & 0 & 6 & 4 \\
0 & 1 & -5 & -3
\end{array}\right] .
\end{aligned}
$$

Thus

$$
\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}=\left[\begin{array}{r}
6 \\
-5
\end{array}\right], \quad\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}=\left[\begin{array}{r}
4 \\
-3
\end{array}\right] .
$$

The desired change-of-coordinates matrix is therefore

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}=\left[\left[\boldsymbol{b}_{1}\right]_{\mathcal{C}}\left[\boldsymbol{b}_{2}\right]_{\mathcal{C}}\right]=\left[\begin{array}{rr}
6 & 4 \\
-5 & -3
\end{array}\right] .
$$

## Summary of the Process

- Observe that the matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ in the preceding example was the one appearing on the right of the reduced echelon form.
- This is not surprising because the first column of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ results from row reducing $\left[\boldsymbol{c}_{1} \boldsymbol{c}_{2} \mid \boldsymbol{b}_{1}\right]$ to $\left[I \mid[\boldsymbol{b}]_{\mathcal{C}}\right]$, and similarly for the second column of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$.
- Thus

$$
\left[\begin{array}{ll|ll}
\boldsymbol{c}_{1} & \boldsymbol{c}_{2} & \boldsymbol{b}_{1} & \boldsymbol{b}_{2}
\end{array}\right] \sim\left[\begin{array}{l|l}
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}
\end{array}\right] .
$$

- An analogous procedure works for finding the change-of-coordinates matrix between any two bases in $\mathbb{R}^{n}$.


## Example

- Let $\boldsymbol{b}_{1}=\left[\begin{array}{r}1 \\ -3\end{array}\right], \boldsymbol{b}_{2}=\left[\begin{array}{r}-2 \\ 4\end{array}\right], \boldsymbol{c}_{1}=\left[\begin{array}{r}-7 \\ 9\end{array}\right], \boldsymbol{c}_{2}=\left[\begin{array}{r}-5 \\ 7\end{array}\right]$.

Consider the bases for $\mathbb{R}^{2}$ given by $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}$ and $\mathcal{C}=\left\{\boldsymbol{c}_{1}, \boldsymbol{c}_{2}\right\}$.
(a) Find the change-of-coordinates matrix from $\mathcal{C}$ to $\mathcal{B}$.
(b) Find the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.
(a) We compute $\left[\begin{array}{ll|ll}\boldsymbol{b}_{1} & \boldsymbol{b}_{2} & \boldsymbol{c}_{1} & \boldsymbol{c}_{2}\end{array}\right]=$

$$
\begin{aligned}
& {\left[\begin{array}{rr|rr}
1 & -2 & -7 & -5 \\
-3 & 4 & 9 & 7
\end{array}\right] \rightarrow\left[\begin{array}{ll|ll}
1 & -2 & -7 & -5 \\
0 & -2 & -12 & -8
\end{array}\right]} \\
& \longrightarrow\left[\begin{array}{rr|rr}
1 & -2 & -7 & -5 \\
0 & 1 & 6 & 4
\end{array}\right] \longrightarrow\left[\begin{array}{ll|rr}
1 & 0 & 5 & 3 \\
0 & 1 & 6 & 4
\end{array}\right]
\end{aligned}
$$

So $\underset{\mathcal{B} \leftarrow \mathcal{C}}{P}=\left[\begin{array}{ll}5 & 3 \\ 6 & 4\end{array}\right]$.
(b) Now we get

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}=\underset{\mathcal{B} \leftarrow \mathcal{C}}{P}=\frac{1}{2}\left[\begin{array}{rr}
4 & -3 \\
-6 & 5
\end{array}\right]=\left[\begin{array}{rr}
2 & -\frac{3}{2} \\
-3 & \frac{5}{2}
\end{array}\right] .
$$

## Alternative Description of Change-of-Coordinate Matrix

- Another description of the change-of-coordinates matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ uses the change-ofcoordinate matrices $P_{\mathcal{B}}$ and $P_{\mathcal{C}}$ that convert $\mathcal{B}$-coordinates and $\mathcal{C}$-coordinates, respectively, into standard coordinates.
- Recall that for each $\boldsymbol{x}$ in $\mathbb{R}^{n}$,

$$
P_{\mathcal{B}}[x]_{\mathcal{B}}=x, \quad P_{\mathcal{C}}[x]_{\mathcal{C}}=x, \quad[x]_{\mathcal{C}}=P_{\mathcal{C}}^{-1} x
$$

- Thus

$$
[x]_{\mathcal{C}}=P_{\mathcal{C}}^{-1} x=P_{\mathcal{C}}^{-1} P_{\mathcal{B}}[x]_{\mathcal{B}}
$$

- In $\mathbb{R}^{n}$, the change-of-coordinates matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ may be computed as

$$
\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}=P_{\mathcal{C}}^{-1} P_{\mathcal{B}} \quad \text { or, more suggestively, } \quad[\boldsymbol{x}]_{\mathcal{C}}=\underset{\mathcal{E} \leftarrow \mathcal{C} \mathcal{E} \leftarrow \mathcal{B}}{P}[\boldsymbol{x}]_{\mathcal{B}} .
$$

