# Chapter 4: Vector Spaces


**San Jose State University**

Prof. Guangliang Chen


Fall 2022

# Outline

# Introduction

In this chapter we introduce vector spaces and the associated notions of

- Subspace
- Dimension
- Basis
- Coordinate system

Meanwhile, we cover the following matrix concepts

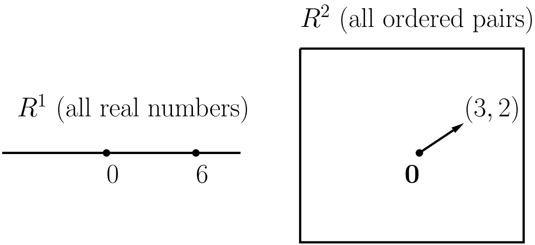
- Column/ null space
- Rank

## Euclidean spaces

For any integer $n \geq 1$, the $n$-dimensional Euclidean space is the set of all $n$-dimensional vectors

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \text{ are real numbers}\}$$

.



Euclidean spaces are endowed with two kinds of operations, vector addition and scalar multiplication, which satisfy the following properties:

(vector addition)

(1) For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the sum is in the same space: $\mathbf{u} + \mathbf{v} \in \mathbb{R}^n$.

(2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative law)

(3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (associative law)

(4) There is a zero vector: $\mathbf{u} + \mathbf{0} = \mathbf{u}$

(5) For each vector $\mathbf{u}$, there is a vector $-\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ (opposite vector must also be contained)

(scalar multiplication)

(6) Any scalar multiple of vector must also be in $\mathbb{R}^n$: $c\mathbf{u} \in \mathbb{R}^n$ (for any real number $c$ and vector $\mathbf{u} \in \mathbb{R}^n$)

(7) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (distributive law)

(8) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (distributive law)

(9) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(10) $1\mathbf{u} = \mathbf{u}$

# What are (abstract) vector spaces?

Formally, a vector space is a (nonempty) set $V$ of objects, called "vectors", that is endowed with two kinds of operations, addition and scalar multiplication, satisfying the same requirements (called axioms):

(1) For any $\mathbf{u}, \mathbf{v} \in V$, the sum is in the same space: $\mathbf{u} + \mathbf{v} \in V$.

(2) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative law)

(3) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (associative law)

(4) There is a zero vector $\mathbf{0}$ in $V$: $\mathbf{u} + \mathbf{0} = \mathbf{u}$

(5) For each vector $\mathbf{u} \in V$, there is a vector $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ (opposite vector must also be contained)

(6) Any scalar multiple of vector must also be in $V$: $c\mathbf{u} \in V$ (for any real number $c$ and vector $\mathbf{u} \in V$)

(7) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (distributive law)

(8) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (distributive law)

(9) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(10) $1\mathbf{u} = \mathbf{u}$

# Examples of (abstract) vector spaces

**Example 0.1.** The set of all functions $f : \mathbb{R} \mapsto \mathbb{R}$ is a vector space:

- Functions are (abstract) vectors;
- There is an addition defined between functions, e.g., for $f(x) = x^2 - 3x + 1$ and $g(x) = 3x + \sin x$, their sum is $f(x) + g(x) = x^2 + \sin x + 1$, and it satisfies all the requirements.
- Scalar multiplication (between a scalar and a function) is also defined: $5f(x) = 5x^2 - 15x + 5$, and it meets all the requirements.

**Example 0.2.** The set of all infinite sequences $(a_1, a_2, \ldots, a_n, \ldots)$ is a vector space:

- Sequences are (abstract) vectors;
- There is an addition defined between sequences

$$(a_1, a_2, \ldots, a_n, \ldots) + (b_1, b_2, \ldots, b_n, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n, \ldots)$$

  and it satisfies all the requirements.

- Scalar multiplication (between a scalar and a sequence) is also defined:

$$k(a_1, a_2, \ldots, a_n, \ldots) = (ka_1, ka_2, \ldots, ka_n, \ldots)$$

  and it meets all the requirements.

**Example 0.3.** The set of all matrices of a fixed size $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a vector space:

- Matrices are (abstract) vectors;
- There is an addition defined between matrices (of the same size)

$$\mathbf{A} + \mathbf{B}$$

  and it satisfies all the requirements.

- Scalar multiplication (between a scalar and a matrix) is also defined:

$$k\mathbf{A}$$

  and it meets all the requirements.

## Vector spaces are an algebraic system

These vector spaces, though consisting of very different objects (functions, sequences, matrices), are all equivalent to Euclidean spaces $\mathbb{R}^n$ in terms of algebraic properties.

- Concepts to be defined for $\mathbb{R}^n$, such as dimension, basis, and subspace, also apply to those vector spaces.

- Properties to be derived for $\mathbb{R}^n$ (based on the two operations) will also generalize to other vector spaces.

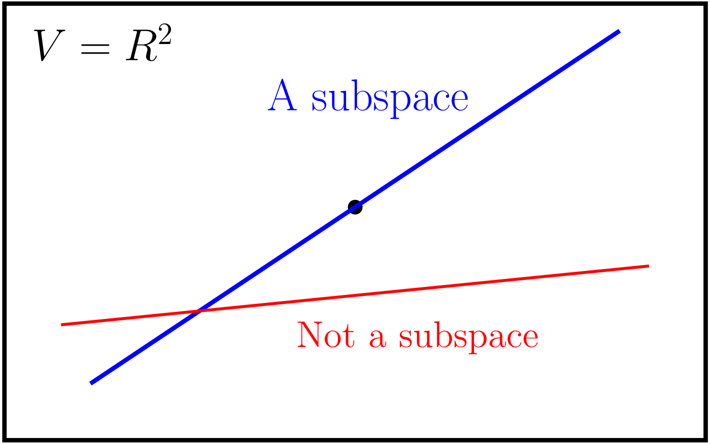
In this course we will focus on $\mathbb{R}^n$ (Math 129B deals with abstract vector spaces in more depth).

## Subspace (subset of a vector space)

A subset of a vector space is called a subspace, if the subset also resembles a vector space (such as a line in $\mathbb{R}^2$ through the origin).

**Def 0.1.** Let $V$ be a vector space (e.g., $\mathbb{R}^n$). A subspace of $V$ is a subset $H \subseteq V$ that is closed under addition and scalar multiplication:

- $H$ contains the zero vector: $\mathbf{0} \in H$
- $H$ is closed under scalar multiplication: For all real numbers $c$ and vectors $\mathbf{u} \in H$, we have $c\mathbf{u} \in H$
- $H$ is closed under addition: For all $\mathbf{u}, \mathbf{v} \in H$, we have $\mathbf{u} + \mathbf{v} \in H$


$V = R^2$
A subspace
Not a subspace

**Example 0.4.** Consider the vector space $V = \mathbb{R}^2$.

- Any line going through the origin in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$. In contrast, any line not passing through the origin is NOT a subspace.
- In fact, the single-element subset containing only the origin $\{\mathbf{0}\}$ is also a subspace of $\mathbb{R}^2$. It is called the zero subspace.
- The full vector space $\mathbb{R}^2$ is also a subspace of itself (though also a trivial one).

*Remark*. Lines going through the origin in $\mathbb{R}^2$ are called proper subspaces of $\mathbb{R}^2$.

**Example 0.5.** For the vector space $V = \mathbb{R}^3$,

- Lines and planes passing through the origin are proper subspaces.
- $\{\mathbf{0}\}$ and $\mathbb{R}^3$ are trivial subspaces.

Question: Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

**Example 0.6.** For the vector space $V = \mathbb{R}^3$,

- Lines and planes passing through the origin are proper subspaces.
- $\{\mathbf{0}\}$ and $\mathbb{R}^3$ are trivial subspaces.

Question: Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

**Answer**: It is not, because it is not even a subset of $\mathbb{R}^3$, as they contain vectors of different dimensions.

However, the following is a subspace of $\mathbb{R}^3$:

$$\{(x, y, 0)^T \in \mathbb{R}^3 \mid x, y \text{ are real numbers}\}$$

**Example 0.7.** Let $V$ be the vector space of all functions $f : \mathbb{R} \mapsto \mathbb{R}$. Then $H = \{\text{All polynomial functions}\}$ is a subspace.

To verify this, note that

- $0 \in H$ (just a trivial polynomial)
- Any scalar multiple of a polynomial is still a polynomial (closed under scalar multiplication)
- Sum of two polynomials is still a polynomial (closed under addition)

**A joke**

Q: What do you call it when a mathematician's parrot hasn't been fed?
A: Poly"no meal"

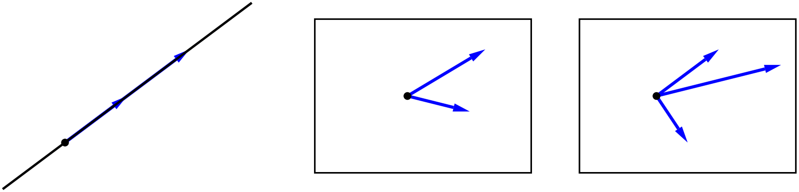

(Source: `https://mathwithbaddrawings.com/2018/03/07/matrix-jokes/`)

## Span of a set of vectors is always a subspace

*Theorem* 0.1. For any set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, their span

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

is a subspace of $V$.



*Proof*. We verify directly the three requirements:

- $\mathbf{0} \in \mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ (when $c_1 = \cdots = c_k = 0$);
- Let $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$. For any scalar $r$,

$$r\mathbf{v} = (rc_1)\mathbf{v}_1 + \cdots + (rc_k)\mathbf{v}_k \in \mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

- Let $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ and $\mathbf{w} = d_1\mathbf{v}_1 + \cdots + d_k\mathbf{w}_k$. Then

$$\mathbf{v} + \mathbf{w} = (c_1 + d_1)\mathbf{v}_1 + \cdots + (c_k + d_k)\mathbf{v}_k \in \mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

## Spaces defined over a matrix (overview)

Given a matrix $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \in \mathbb{R}^{m\times n}$, one can define the following spaces:

- **Column space**: Span of its column vectors

$$\mathrm{Col}(\mathbf{A}) = \mathrm{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subseteq \mathbb{R}^m$$

- **Null space**: Solution set of $\mathbf{A}\mathbf{x} = \mathbf{0}$:

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

## Column space of a matrix

**Def 0.2.** Let $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \in \mathbb{R}^{m\times n}$ be any matrix. Its column space is defined as

$$\mathrm{Col}(\mathbf{A}) = \mathrm{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}.$$

*Remark*. $\mathrm{Col}(\mathbf{A})$ must be a subspace of $\mathbb{R}^m$.

In terms of the linear transformation $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, the column space of $\mathbf{A}$ is exactly the range of $T$:

$$\mathrm{Range}(T) = \{\mathbf{b} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \in \mathbb{R}^n\}.$$

*Remark*. The linear transformation $T(\mathbf{x})$ is onto if and only if $\mathrm{Col}(\mathbf{A}) = \mathbb{R}^m$.

**Example 0.8.** Let

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 10 \\ 3 & 6 & 9 & 10 \end{bmatrix}$$

Do the following:

- Determine if $\mathbf{b} = [1 \ \ -1 \ 1]^T$ lies in the column space of $\mathbf{A}$
- Find $\mathrm{Col}(\mathbf{A})$. Is $f(\mathbf{x}) = \mathbf{Ax}$ onto?

## Null space of a matrix

**Def 0.3.** Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be any matrix. Its null space is defined as

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

In terms of the linear transformation $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, the null space of $\mathbf{A}$ is called the kernel of $T$:

$$\operatorname{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\} = \operatorname{Nul}(\mathbf{A})$$

*Remark*. The linear transformation $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ is one to one if and only if $\operatorname{Ker}(T) = \operatorname{Nul}(\mathbf{A}) = \{\mathbf{0}\}$.

**Example 0.9.** Let

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 10 \\ 3 & 6 & 9 & 10 \end{bmatrix}$$

Do the following:

- Determine if $\mathbf{x} = [1 \ \ -2 \ 1]^T$ and $\mathbf{y} = [-5 \ 0 \ 5 \ \ -3]^T$ lie in the null space of $\mathbf{A}$.
- Find $\mathrm{Nul}(\mathbf{A})$. Is $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ one to one?

*Theorem* 0.2. For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathrm{Nul}(\mathbf{A})$ is a subspace of $\mathbb{R}^n$.

*Proof.* We verify directly the three requirements:

- $\mathbf{0} \in \mathrm{Nul}(\mathbf{A})$ because $\mathbf{A}\mathbf{0} = \mathbf{0}$;
- Let $\mathbf{x} \in \mathrm{Nul}(\mathbf{A})$, i.e., $\mathbf{A}\mathbf{x} = \mathbf{0}$. For any scalar $c$,

  $$\mathbf{A}(c\mathbf{x}) = c(\mathbf{A}\mathbf{x}) = c\mathbf{0} = \mathbf{0}$$

  This shows that $c\mathbf{x} \in \mathrm{Nul}(\mathbf{A})$.
- Let $\mathbf{x}, \mathbf{y} \in \mathrm{Nul}(\mathbf{A})$. Then

  $$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

  This shows that $\mathbf{x} + \mathbf{y} \in \mathrm{Nul}(\mathbf{A})$.

**Example 0.10.** Find the null and column spaces of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -5 & 7 \\ 3 & 7 & -8 \end{bmatrix}$$

Of which Euclidean spaces are they each a subspace?

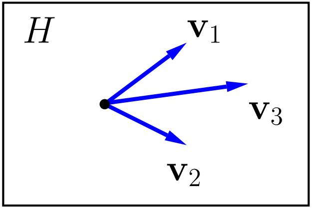
**Answer**: They are both subspaces of $\mathbb{R}^3$:

$$\mathrm{Col}(\mathbf{A}) = \{\mathbf{b} = (b_1, b_2, b_3)^T \in \mathbb{R}^3 \mid b_1 = b_2 + b_3\} = \mathrm{Span}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right)$$

$$\mathrm{Nul}(\mathbf{A}) = \mathrm{Span}\left(\begin{bmatrix}-9\\5\\1\end{bmatrix}\right)$$

# Basis of a subspace

Consider the span $H$ of the following three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$ (which are linearly dependent). We already know that it is a subspace of $V$.




Observe that we do not really need all three vectors to span $H$; in fact, any two of them (e.g., $\mathbf{v}_1, \mathbf{v}_2$) will be able to span $H$. $\longrightarrow$ simpler, and more efficient

**Question**: Why can we remove a vector, $\mathbf{v}_3$ in this case, from a set without changing the span of the set?

The reason is that $\mathbf{v}_3$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2$:

$$\mathbf{v}_3 = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 \quad \text{for some scalars } d_1, d_2$$

and makes no "new contribution" to the span:

$$\begin{aligned} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(d_1\mathbf{v}_1 + d_2\mathbf{v}_2) \\ &= (c_1 + c_3d_1)\mathbf{v}_1 + (c_2 + c_3d_2)\mathbf{v}_2 \end{aligned}$$

That is, any linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ can always be obtained from $\mathbf{v}_1, \mathbf{v}_2$ using a different set of coefficients.

**Next question**: Can we remove one of $\mathbf{v}_1, \mathbf{v}_2$ while still preserving the span?

The answer is obviously no:

- We can only remove a vector that is a linear combination of the others (according to previous reasoning).
- We have to stop removing vectors from a set (if we want to preserve the span) when there is no more vector that is a linear combination of the rest.

This just means that the remaining vectors are linearly independent. This is the smallest set you can use to span the same subspace $H$, and it is nonempty as long as $H \neq \{\mathbf{0}\}$.

In the book this is called the Spanning Set Theorem.

# Basis for a subspace

Briefly speaking, a **basis** for a subspace of a vector space, $H \subseteq V$, is a set of linearly independent vectors that can span $H$.

**Def 0.4.** Let $H \subseteq V$ be a subspace of the vector space $V$. We say that a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ form a basis for $H$ if

- Span$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = H$; $\longleftarrow$ This implies that every $\mathbf{v}_i$ must be in $H$.
- The set $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linear independent.

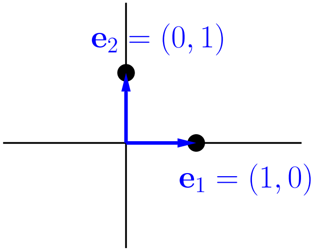
*Remark*. The definition covers the case of $H = V$, so we can talk about basis for the vector space $V$.

It is easy to see that the set of vectors $\mathbf{v}_1 = [1,0]^T, \mathbf{v}_2 = [0,1]^T$ is a basis for $\mathbb{R}^2$.

In fact, for any positive integer $n$, the following set of vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \ddots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

is always a basis for $\mathbb{R}^n$.




It is called the standard basis for $\mathbb{R}^n$.

## Columns of any square, invertible matrix are a basis

*Theorem* 0.3. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be an invertible matrix. Then the columns of $\mathbf{A}$ form a basis for $\mathbb{R}^n$ because

- The columns of $\mathbf{A}$ are linearly independent, and
- They span $\mathbb{R}^n$

both by the Invertible Matrix Theorem.

**Example 0.11.** Show that the columns of $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ form a basis for $\mathbb{R}^2$.

**Example 0.12.** Determine if the columns of the matrix

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

form a basis for $\mathbb{R}^3$.

## Finding a basis for Col($\mathbf{A}$)

We first consider a matrix in the RREF and explain how to find a basis for its column space by direct observation.

**Example 0.13.** Find a basis for the column space of

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

**Example 0.14.** Find a basis for the column space of

$$\mathbf{B} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

A couple things to note first:

- $\mathbf{A}$ on the preceding slide is actually the RREF of $\mathbf{B}$.
- $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ are pivot columns. For the other columns of $\mathbf{A}$,

$$\mathbf{a}_2 = 4\mathbf{a}_1, \quad \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3.$$

Do exactly the same dependence relationships hold true for the columns of $\mathbf{B}$?

The answer is yes, which implies that we can remove $\{\mathbf{b}_2, \mathbf{b}_4\}$ and use the remaining columns $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ as a basis for Col$(\mathbf{B})$.

To see this, we first point out that there exist a sequence of elementary matrices such that

$$\underbrace{\mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1}_{=\mathbf{E}\,\text{(invertible)}} \mathbf{B} = \underbrace{\mathbf{A}}_{\text{RREF}}, \qquad \text{or} \quad \mathbf{B} = \mathbf{E}^{-1}\mathbf{A}$$

Using this equation and the columnwise multiplication

$$\mathbf{b}_1 = \mathbf{E}^{-1}\mathbf{a}_1, \quad \ldots, \quad \mathbf{b}_5 = \mathbf{E}^{-1}\mathbf{a}_5$$

we can show that any dependence relation among the columns of $\mathbf{A}$, such as $\mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3$, also holds true for $\mathbf{B}$:

$$\mathbf{E}^{-1}\mathbf{a}_4 = \mathbf{E}^{-1}(2\mathbf{a}_1 - \mathbf{a}_3) \qquad \longrightarrow \quad \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3.$$

We have effectively proved the following result.

*Theorem* 0.4. The pivot columns of any matrix $\mathbf{A}$ form a basis for its column space Col($\mathbf{A}$).

*Remark*. To identify the pivot columns of $\mathbf{A}$, a REF would suffice. There is no need to obtain the RREF (which requires more work).

*Remark*. Do not use the pivot columns from any REF of $\mathbf{A}$ toward a basis for $\mathrm{Col}(\mathbf{A})$. Instead, always use the pivot columns of $\mathbf{A}$ to create a basis.

The reason is that row operations actually change the column space, but preserve the dependence relationship among the columns.

# Basis for Nul($\mathbf{A}$)

We use an example to explain how to find a basis for the null space of a matrix.

**Example 0.15.** Find a basis for the null space of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -5 & 7 & 5 \\ 3 & 7 & -8 & -5 \end{bmatrix}$$

**Answer**. By direct calculation, the solution of $\mathbf{A}\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = x_3 \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -10 \\ 5 \\ 0 \\ 1 \end{bmatrix}, \quad x_3, x_4 \text{ are free variables}$$

## The Unique Representation Theorem

One nice thing about the basis of a vector space $V$ is that it can uniquely span any vector in $V$.

*Theorem* 0.5. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for a vector space $V$. Then for any vector $\mathbf{v} \in V$, there exists a unique set of scalars $c_1, \ldots, c_k$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

*Proof.* For any given vector $\mathbf{v}$, the existence of the scalars is due to Span$(\mathcal{B}) = V$.

To prove the uniqueness, suppose there are two sets of scalars such that

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{v} = d_1\mathbf{v}_1 + \cdots + d_k\mathbf{v}_k.$$

Merging terms gives that

$$(c_1 - d_1)\mathbf{v}_1 + \cdots + (c_k - d_k)\mathbf{v}_k = \mathbf{0}$$

Because the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent, we conclude that

$$c_1 - d_1 = \cdots = c_k - d_k = 0, \qquad \text{i.e.,} \quad c_1 = d_1, \ldots, c_k = d_k.$$

Thus, the set of scalars must be unique. □

**Example 0.16.** Consider the Euclidean space $\mathbb{R}^n$. Every vector $\mathbf{b} = (b_1, \ldots, b_n)^T$ in it has a unique representation under the standard basis:

$$\mathbf{b} = b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n$$

**Example 0.17.** We have previously showed that the columns of the matrix form a basis for $\mathbb{R}^3$:

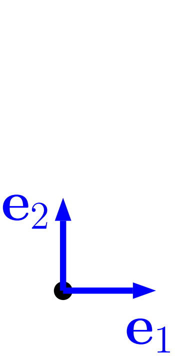
$$\mathbf{A} = \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix}$$

Let $\mathbf{b} = [1\ 0\ 2]^T$. Find the unique set of scalars $c_1, c_2, c_3$ such that

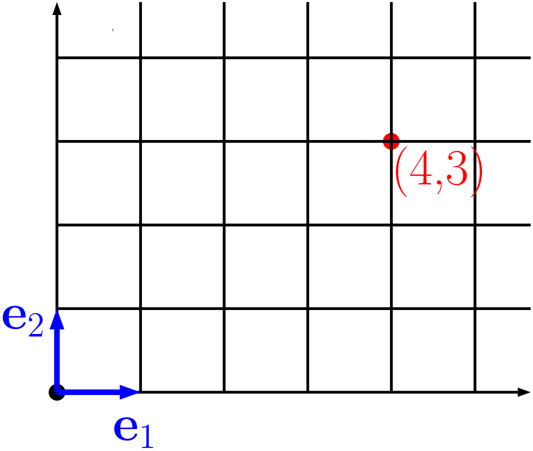
$$\mathbf{b} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3. \longleftarrow \textit{Answer}: c_1 = -1, c_2 = -2, c_3 = 2$$

## Coordinate system

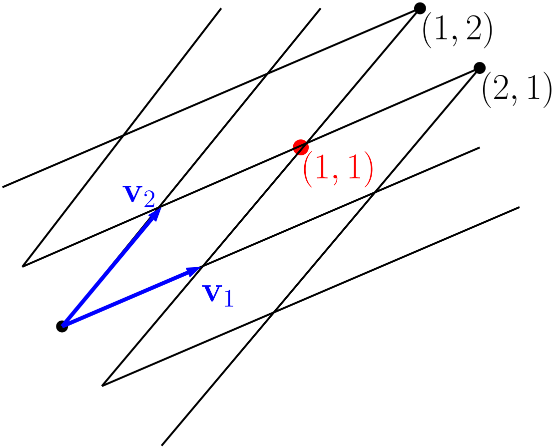
The describe the location of a point in the plane, we need to specify a reference point (origin) and two direction vectors (e.g., east and north).



The red point is 4 units to the east and 3 units to the north, relative to the origin.



Here is a new but weird way of describing the location of the red point.



# Coordinates of a vector relative to a basis

In fact, any basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of a vector space $V$ can be used as a coordinate system to describe the locations of all vectors in the vector space.

For any $\mathbf{v} \in V$, due to the Unique Representation Theorem, there exist a unique set of scalars $c_1, \ldots, c_k$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

**Def 0.5.** The (unique) coefficients $c_1, \ldots, c_k$ are called **coordinates of the vector $\mathbf{v}$ relative to the basis $\mathcal{B}$**, or in short $\mathcal{B}$-**coordinates**.

We collect them to form a (coordinate) vector and denote it by $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$.


Always label your axes
HEAVY
DOUBLE
GraphJam.com

**Example 0.18.** Find the coordinate vector of $\mathbf{x} = [2, 5]^T \in \mathbb{R}^2$ relative to the basis given by the columns of $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

**Example 0.19.** We have previously showed that the columns of the matrix form a basis for $\mathbb{R}^3$:

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

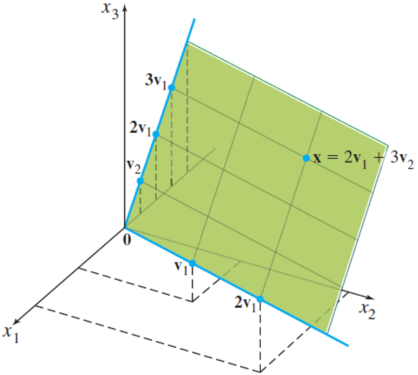
and for $\mathbf{b} = [1\ 0\ 2]^T \in \mathbb{R}^3$, we obtained that

$$\mathbf{b} = (-1)\mathbf{a}_1 + (-2)\mathbf{a}_2 + 2\mathbf{a}_3.$$

Therefore, the coordinates of $\mathbf{b}$ relative to the basis (columns of $\mathbf{A}$) are $[-1, -2, 2]^T$.

## Coordinate axes for a subspace

It is also possible to select a coordinate system for a subspace $H \subset V$.



**Example 0.20.** Let $\mathbf{v}_1 = [1, 1]^T$. Then $\mathcal{B} = \{\mathbf{v}_1\}$ is a basis for $H = \text{Span}\{\mathbf{v}_1\} \subset \mathbb{R}^2$. Determine if $\mathbf{x} = [5, 5]^T$ is in $H$, and if yes, find its coordinate vector relative to $\mathcal{B}$.

**Example 0.21.** Let $\mathbf{v}_1 = [1, 1, 0]^T, \mathbf{v}_2 = [1, 0, 1]^T$. Then $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\} \subset \mathbb{R}^3$. Determine if $\mathbf{x} = [3, 2, 1]^T$ is in $H$, and if yes, find its coordinate vector relative to $\mathcal{B}$.

## What is a dimension?

We have seen that vector spaces have infinitely many vectors inside, yet all of them can be uniquely spanned by a basis (which is often a small, finite set).

The cardinality of the basis (as a set) is an intrinsic property of a vector space. We will use it to define the dimension of the vector space.

Before that we need to address the following question: Could different bases have different sizes?

The answer is no, according to the following theorem.

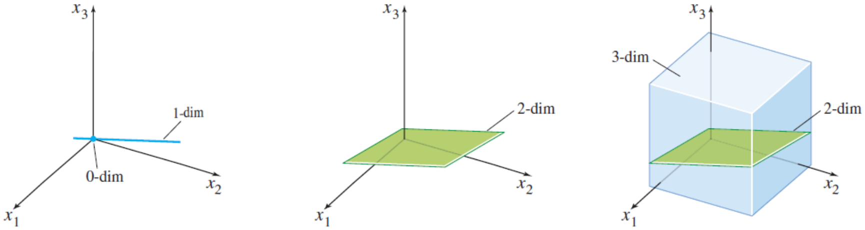
*Theorem* 0.6. **Any two bases of a vector space must have the same size.**

**Def 0.6.** Let $V$ be a vector space with basis $\mathcal{B}$. The size (or cardinality) of $\mathcal{B}$ is called the **dimension** of $V$, and written as $\dim(V)$.

- The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be 0.
- If $V$ has a finite basis, then it is said to be finite dimensional.
- If $V$ cannot be spanned by a finite set, then it is said to be infinite dimensional.

*Remark*. In a $k$-dimension vector space $V$, any set of $k+1$ or more vectors must be linearly dependent.

Dimensions of various subspaces of $\mathbb{R}^3$:




*Remark*. An example of infinite dimensional vector spaces is the space of all polynomials. However, the subspace of all polynomials of degree no more than a fixed number, say $n$, has a dimension $n+1$, thus it is finite-dimensional.

## Dimension of a subspace

**Example 0.22.** Let $\mathbf{v}_1 = [1, 1, 0]^T, \mathbf{v}_2 = [1, 0, 1]^T$. Then $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\} \subset \mathbb{R}^3$. It follows that the dimension of $H$ is 2, i.e., dim$(H) = 2$.

*Theorem* 0.7. If $H \subseteq V$, then $\dim(H) \leq \dim(V)$.

*Proof.* Suppose $H \neq \{\mathbf{0}\}$ (otherwise it is trivially true). Let $\mathcal{B}$ be a basis for $H$. Because $\mathcal{B}$ is a linearly independent subset of $V$, its size cannot exceed the dimension of $V$. That is, $\dim(H) \leq \dim(V)$. □

## The Basis Theorem

Recall that a set of vectors is a basis for a vector space if they are linearly independent and span the vector space.

However, with a correct size ($=$ dimension of the vector space), any one of linear independence or spanning must imply the other and thus makes a basis.

*Theorem* 0.8. Let $V$ be a $k$-dimensional vector space.

- Any $k$ linearly independent vectors in $V$ are a basis for $V$;
- Any $k$ vectors that span $V$ must also be a basis for $V$.

# Dimensions of null and column spaces

*Theorem* 0.9. Let $\mathbf{A}$ be any matrix. Then

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

**Example 0.23.** For the following matrix, $\dim(\text{Col}(\mathbf{A})) = 3$ (pivot columns), and $\dim(\text{Nul}(\mathbf{A})) = 2$ (free variables).

$$\mathbf{A} = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 3 & -9 & 12 & -9 & 6 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & 3 & 0 \\ 0 & 1 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

## Rank of a matrix?

Briefly speaking, the rank of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the maximal number of linearly independent columns (or rows) of $\mathbf{A}$.

It is one of the most fundamental characteristics of a matrix.

A lot of properties of a matrix can be determined by its rank. For example, "An $n \times n$ matrix is invertible if and only if the rank is $n$".

Formally, we define the matrix rank as follows.

**Def 0.7.** The rank of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is defined as the dimension of the column space of $\mathbf{A}$, i.e.,

$$\operatorname{rank}(\mathbf{A}) = \dim(\operatorname{Col}(\mathbf{A}))$$

**Example 0.24.** For the following matrix,

$$\mathbf{A} = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 3 & -9 & 12 & -9 & 6 \end{bmatrix}$$

its rank is $3$ (because we already know that $\dim(\operatorname{Col}(\mathbf{A})) = 3$). Thus, the maximal number of linearly independent columns is also 3.

## The Rank Theorem

*Theorem* 0.10. For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$,

$$\operatorname{rank}(\mathbf{A}) + \dim(\operatorname{Nul}(\mathbf{A})) = n.$$

That is, the column and null spaces of a matrix have a combined dimension that is equal to the number of columns.

*Proof.* Because

- $\operatorname{rank}(\mathbf{A}) = \dim(\operatorname{Col}(\mathbf{A}))$ is equal to the number of pivot columns
- $\dim(\operatorname{Nul}(\mathbf{A}))$ is equal to the number of free variables in $\mathbf{A}\mathbf{x} = \mathbf{0}$

their sum must be equal to $n$ (the number of columns). □

**Example 0.25.** Consider the matrix again:

$$\mathbf{A} = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 3 & -9 & 12 & -9 & 6 \end{bmatrix}$$

Because its rank is $3$, we must have

$$\dim(\mathrm{Nul}(\mathbf{A})) = n - \mathrm{rank}(\mathbf{A}) = 5 - 3 = 2.$$

You may want to verify this by finding a basis for the null space.

# The Invertible Matrix Theorem (cont'd)

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ (square matrix). Then each of the following statements is equivalent to "$\mathbf{A}$ is invertible".

- The columns of $\mathbf{A}$ form a basis for $\mathbb{R}^n$.
- $\mathrm{Col}(\mathbf{A}) = \mathbb{R}^n$
- $\dim(\mathrm{Col}(\mathbf{A})) = n$
- $\mathrm{rank}(\mathbf{A}) = n$
- $\mathrm{Nul}(\mathbf{A}) = \{\mathbf{0}\}$
- $\dim(\mathrm{Nul}(\mathbf{A})) = 0$

# The row space of a matrix

Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we have defined its column space as the span of the column vectors (which are in $\mathbb{R}^m$). It is a subspace of $\mathbb{R}^m$.

Similarly, we can consider the span of the rows of $\mathbf{A}$ (treated as vectors in $\mathbb{R}^n$), which is called the row space and denoted $\mathrm{Row}(\mathbf{A})$. It is a subspace of $\mathbb{R}^n$.

Clearly, $\mathrm{Row}(\mathbf{A}) = \mathrm{Col}(\mathbf{A}^T)$.

**Example 0.26.** Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. The column space is the span of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3$, while the row space is the span of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2$.

## Row operations preserve row space (but not column space)

*Theorem* 0.11. If two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ are row equivalent, then their row spaces are the same.

**Example 0.27.** The following two matrices have the same row space, but not the same column space:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \quad \longrightarrow \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The reason is that linear combinations of rows of $\mathbf{B}$, which are linear combinations of rows of $\mathbf{A}$, are always linear combinations of rows of $\mathbf{A}$ (and vice versa).

*Proof.* Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ be two row equivalent matrices. Then there exists an invertible matrix $\mathbf{E} = \mathbf{E}_\ell \cdots \mathbf{E}_2 \mathbf{E}_1$ (which is a product of elementary matrices) such that

$$\mathbf{E}\mathbf{A} = \mathbf{B}$$

Any linear combination of the rows of $\mathbf{B}$ must be a linear combination of the rows of $\mathbf{A}$:

$$\mathbf{c}^T \mathbf{B} = \mathbf{c}^T (\mathbf{E}\mathbf{A}) = (\mathbf{c}^T \mathbf{E})\mathbf{A}$$

This shows that $\mathrm{Row}(\mathbf{B}) \subseteq \mathrm{Row}(\mathbf{A})$.

Similarly, by using $\mathbf{E}^{-1}\mathbf{B} = \mathbf{A}$ we can show that $\mathrm{Row}(\mathbf{A}) \subseteq \mathrm{Row}(\mathbf{B})$. Therefore, we must have $\mathrm{Row}(\mathbf{A}) = \mathrm{Row}(\mathbf{B})$. $\square$

The previous theorem implies the following result.

*Corollary* 0.12. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be any matrix and $\mathbf{R}$ its echelon form. Then

$$\mathrm{Row}(\mathbf{A}) = \mathrm{Row}(\mathbf{R})$$

**Example 0.28.** Find a basis for the row space of $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 0 & 3 & -6 & 6 \\ 3 & -7 & 8 & -5 \\ 3 & -9 & 12 & -9 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note that the first two rows of $\mathbf{A}$ are not necessarily a basis of its row space!

*Theorem* 0.13. For any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we have

$$\dim(\text{Row}(\mathbf{A})) = \dim(\text{Col}(\mathbf{A})) = \text{rank}(\mathbf{A}).$$

*Proof.* This is because

- $\dim(\text{Row}(\mathbf{A}))$ = number of pivot rows (nonzero rows);
- $\dim(\text{Col}(\mathbf{A}))$ = number of pivot columns

which must be the same. □

## $\mathbf{A}, \mathbf{A}^T$ must have the same rank

*Corollary* 0.14. For any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we have

$$\operatorname{rank}(\mathbf{A}^T) = \operatorname{rank}(\mathbf{A}).$$

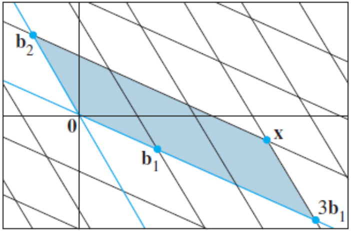
*Proof.* $\operatorname{rank}(\mathbf{A}^T) = \dim(\operatorname{Col}(\mathbf{A}^T)) = \dim(\operatorname{Row}(\mathbf{A})) = \operatorname{rank}(\mathbf{A}).$ □

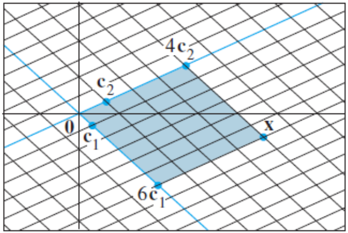
**Example 0.29** (p233). Find a basis for each of the row/column/null spaces of the following matrix

$$\mathbf{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \longrightarrow \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

## The change of basis problem

Assume two bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ for $\mathbb{R}^2$. For a fixed point $\mathbf{x} \in \mathbb{R}^2$, suppose we know its coordinates with respect to $\mathcal{B}$: $[\mathbf{x}]_{\mathcal{B}} = [3, 1]^T$. What is its coordinate vector $[\mathbf{x}]_{\mathcal{C}}$ with respect to $\mathcal{C}$?



**Solution**. Define two basis matrices

$$\mathbf{P}_{\mathcal{B}} = [\mathbf{b}_1, \mathbf{b}_2], \quad \mathbf{P}_{\mathcal{C}} = [\mathbf{c}_1, \mathbf{c}_2].$$

Then we have

$$\mathbf{P}_{\mathcal{B}} \cdot [\mathbf{x}]_{\mathcal{B}} = \mathbf{x} = \mathbf{P}_{\mathcal{C}} \cdot [\mathbf{x}]_{\mathcal{C}}$$

Since $\mathbf{P}_{\mathcal{C}}$ is invertible, we obtain

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}} \cdot [\mathbf{x}]_{\mathcal{B}}$$

*Remark*.

- $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}}$ is called the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.
- Similarly, $[\mathbf{x}]_{\mathcal{B}} = \mathbf{P}_{\mathcal{B}}^{-1}\mathbf{P}_{\mathcal{C}} \cdot [\mathbf{x}]_{\mathcal{C}}$. The change-of-coordinates matrix from $\mathcal{C}$ to $\mathcal{B}$ is
$$\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} = \mathbf{P}_{\mathcal{B}}^{-1}\mathbf{P}_{\mathcal{C}} = (\mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}})^{-1} = (\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$$

**Example 0.30.** Suppose

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}, \quad \mathbf{b}_1 = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

$$\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}, \quad \mathbf{c}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and for some vector $\mathbf{x} \in \mathbb{R}^2$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Find $[\mathbf{x}]_{\mathcal{C}}$.

## How to compute $\mathrm{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathrm{P}_{\mathcal{C}}^{-1}\mathrm{P}_{\mathcal{B}}$ efficiently in general

Recall how to compute the inverse of a matrix

$$[\mathbf{P}_{\mathcal{C}} \mid \mathbf{I}] \longrightarrow [\mathbf{I} \mid \mathbf{P}_{\mathcal{C}}^{-1}] \quad \text{(via elementary row operations)}$$

which is equivalent to the following matrix equation:

$$\mathbf{P}_{\mathcal{C}}^{-1} \cdot [\mathbf{P}_{\mathcal{C}} \mid \mathbf{I}] = [\mathbf{I} \mid \mathbf{P}_{\mathcal{C}}^{-1}]$$

Similarly, we can compute $\mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}}$ as follows:

$$[\mathbf{P}_{\mathcal{C}} \mid \mathbf{P}_{\mathcal{B}}] \longrightarrow [\mathbf{I} \mid \mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}}] \quad \text{(via elementary row operations)}$$

which is equivalent to the following matrix equation:

$$\mathbf{P}_{\mathcal{C}}^{-1} \cdot [\mathbf{P}_{\mathcal{C}} \mid \mathbf{P}_{\mathcal{B}}] = [\mathbf{I} \mid \mathbf{P}_{\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{B}}]$$