Introduction to Vector Space Models

Vector span, Subspaces, and Basis Vectors

Part 1:
Vector Span and Subspaces

Linear Combinations

(Algebraically)

A **linear combination** is constructed from a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ by multiplying each vector by a constant and adding the result:

$$\mathbf{c} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_p\mathbf{v}_p = \sum_{i=1}^{n} a_i\mathbf{v}_i$$

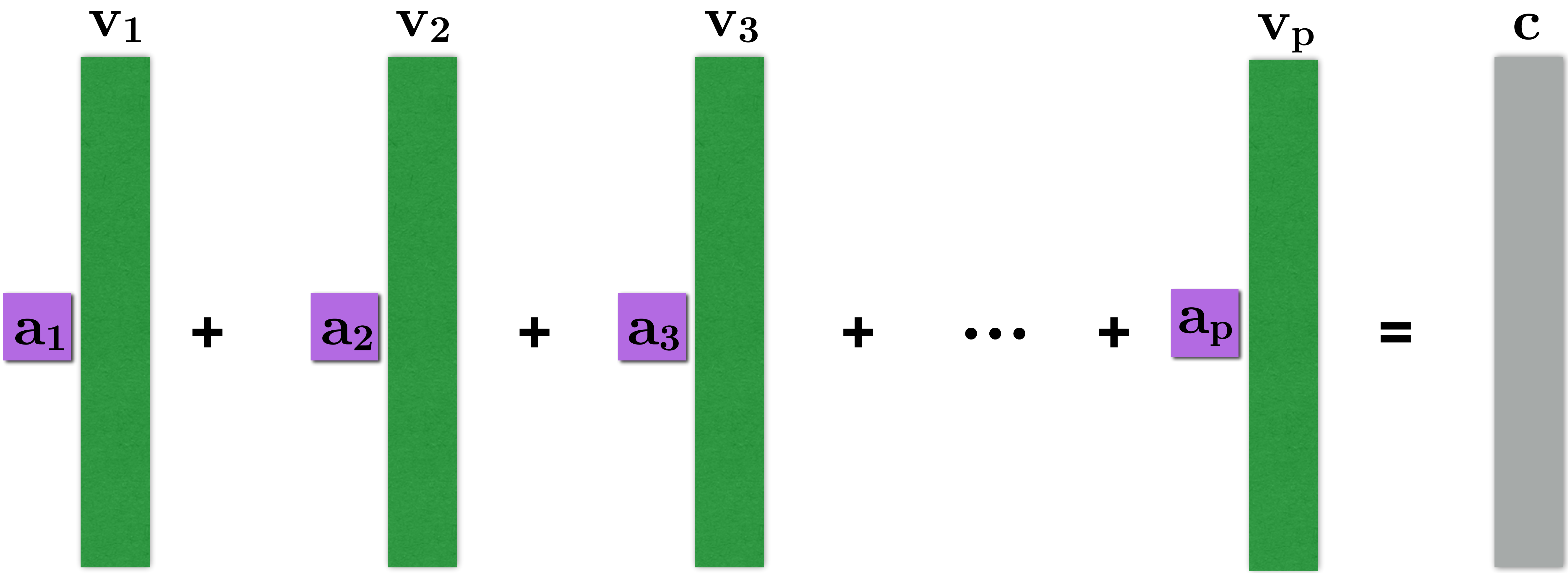
Linear Combinations (Algebraically)

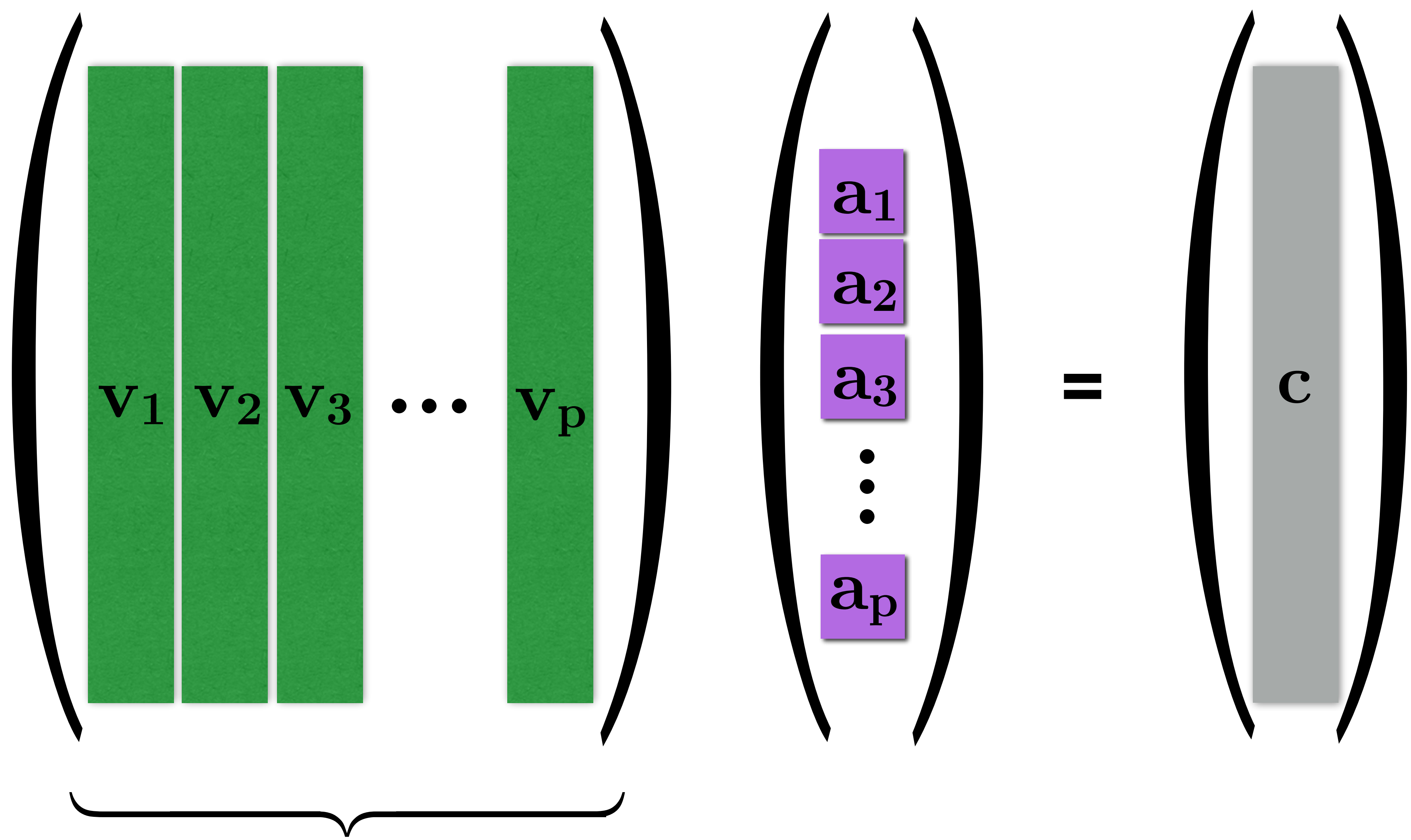
A **linear combination** is constructed from a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ by multiplying each vector by a constant and adding the result:

$$\mathbf{c} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_p\mathbf{v}_p = \sum_{i=1}^{n} a_i\mathbf{v}_i$$

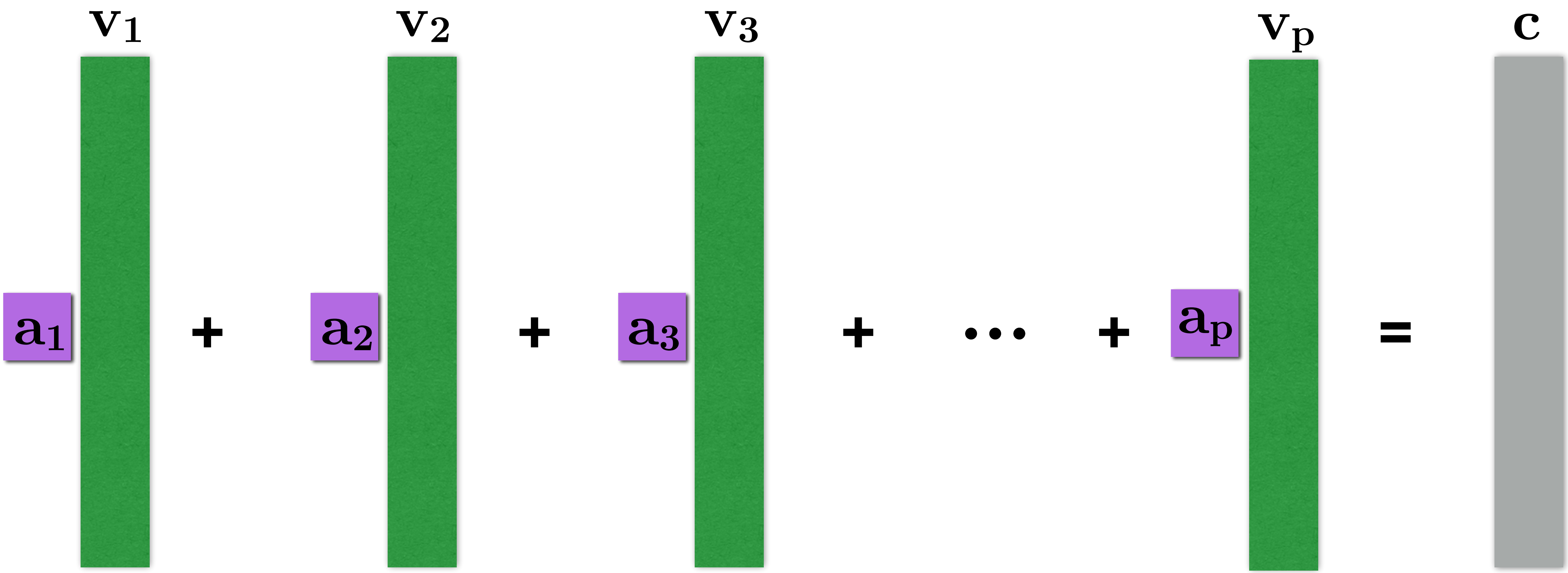
alternatively, we could write

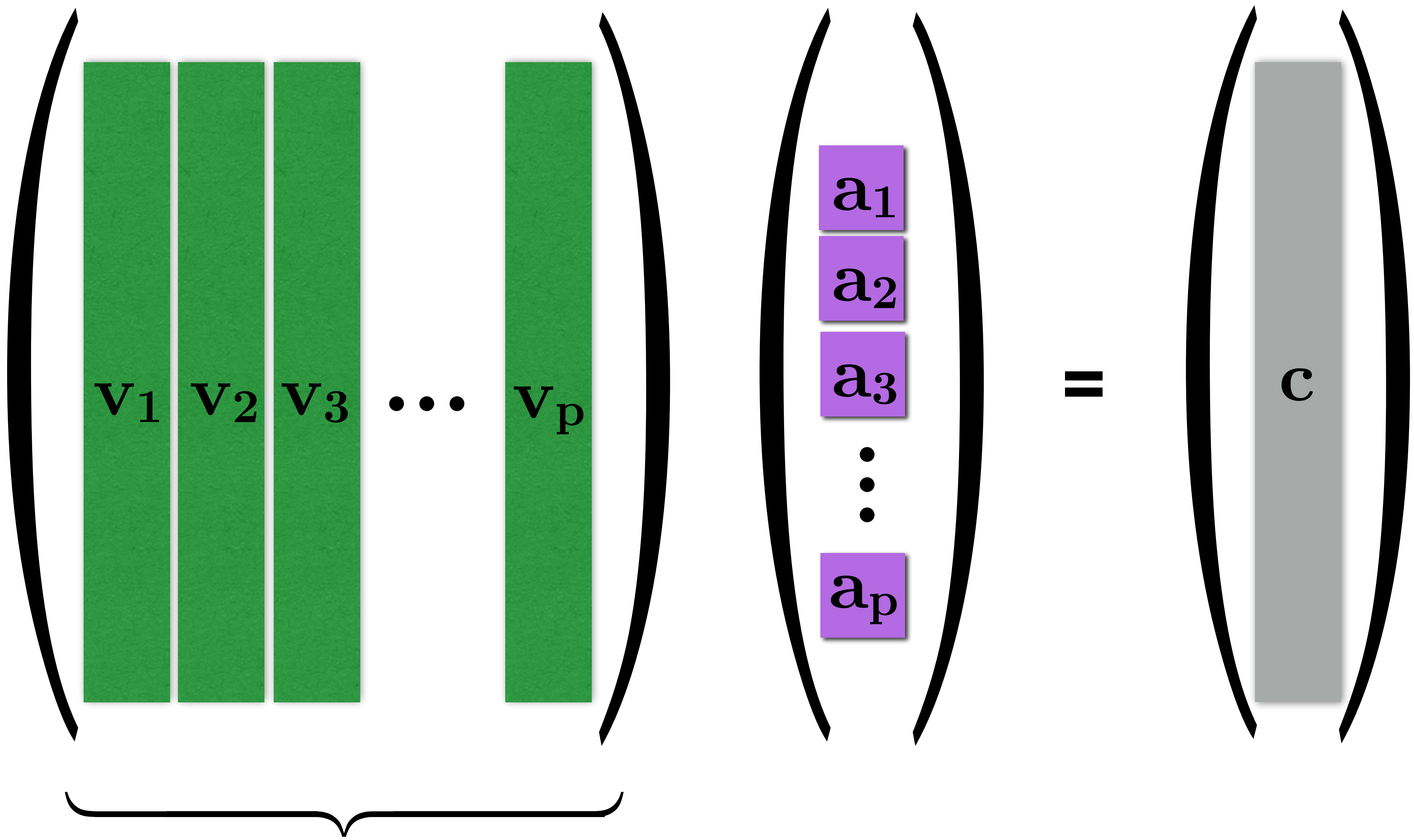
$$\mathbf{c} = \mathbf{V}\mathbf{a} \quad \text{where} \quad \mathbf{V} = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \ldots \mid \mathbf{v}_p] \quad \text{and} \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$


v1
v2
v3
vp
c
a1
+
a2
+
a3
+
• • •
+
ap
=



$$\mathbf{V} = [\mathbf{v}_1 \,|\, \mathbf{v}_2 \,|\, \ldots \,|\, \mathbf{v}_p]$$

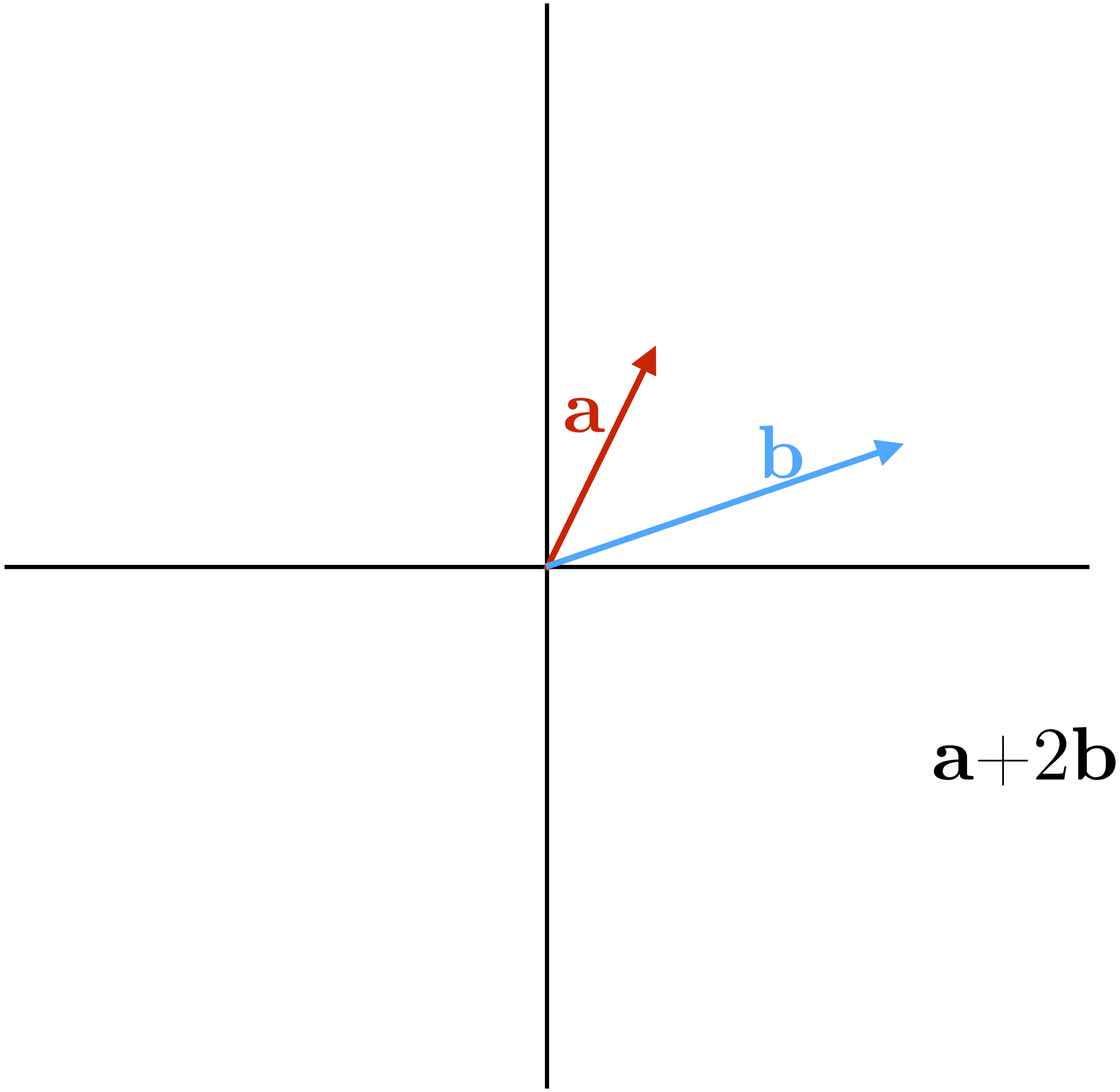

$\mathbf{v_1}$
$\mathbf{v_2}$
$\mathbf{v_3}$
$\mathbf{v_p}$
$\mathbf{c}$
$\mathbf{a_1}$
+
$\mathbf{a_2}$
+
$\mathbf{a_3}$
+
...
+
$\mathbf{a_p}$
=



$$\mathbf{V} = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \ldots \mid \mathbf{v}_p]$$

Linear Combinations

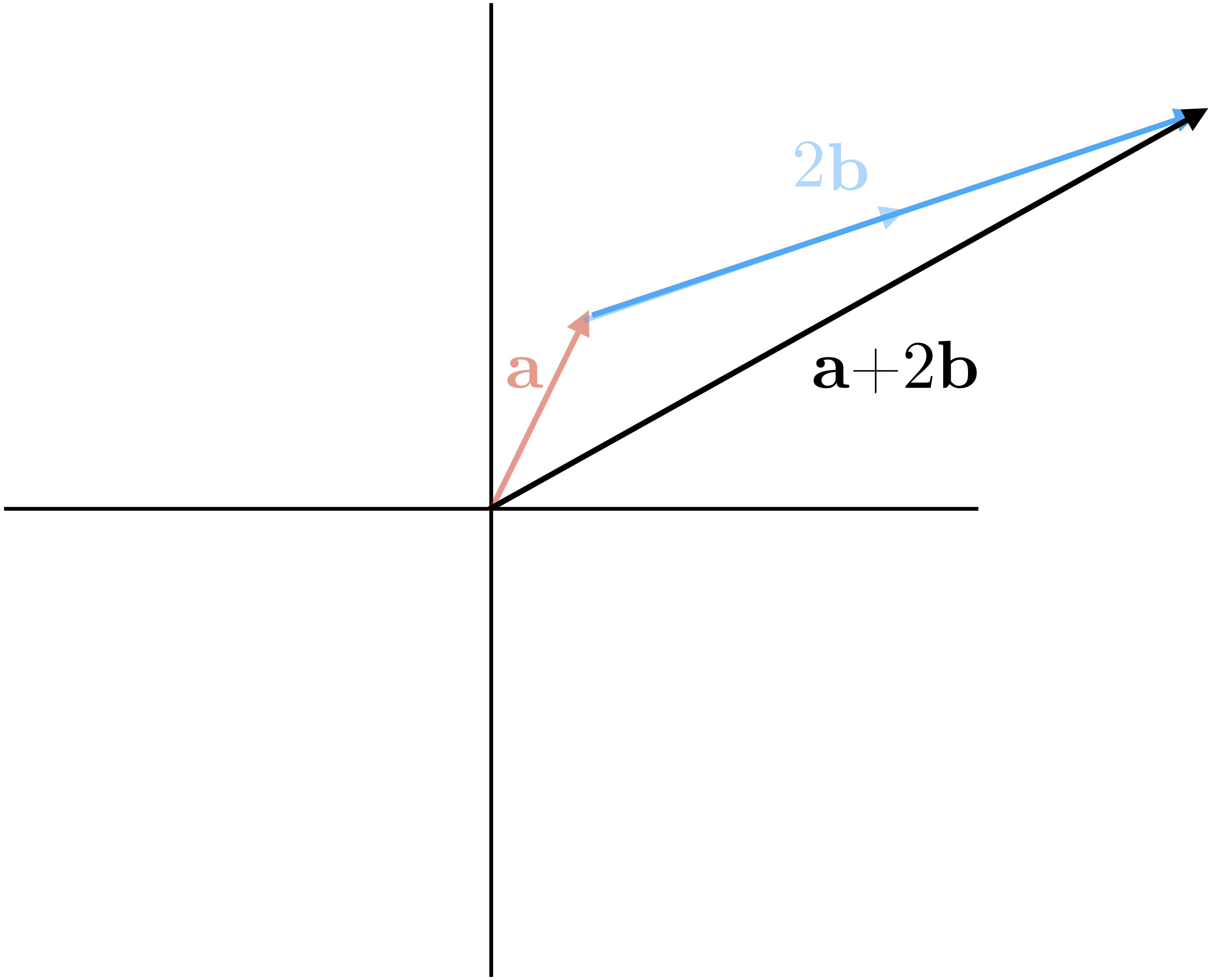
(Geometrically)



$\mathbf{a}+2\mathbf{b}$

Linear Combinations

(Geometrically)



Linear Dependence

(Algebraically)

A group of vectors, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are **linearly dependent** if there exists corresponding scalars, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *not all equal to zero* such that:

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n = 0$$

Linear Dependence

(Algebraically)

A group of vectors, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are **linearly dependent** if there exists corresponding scalars, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *not all equal to zero* such that:

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n = 0$$

For example, if $2\mathbf{v}_1 - 2\mathbf{v}_2 + 4\mathbf{v}_3 = 0$

Then, $\mathbf{v}_1 = \mathbf{v}_2 - 2\mathbf{v}_3$

#PerfectMulticollinearity

Linear Dependence
(Algebraically)

A group of vectors, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are **linearly dependent** if there exists corresponding scalars, $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *not all equal to zero* such that:

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n = 0$$

#PerfectMulticollinearity

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are **linearly independent** if the above equation has only the trivial solution (all $\alpha_i = 0$)

Linear Dependence

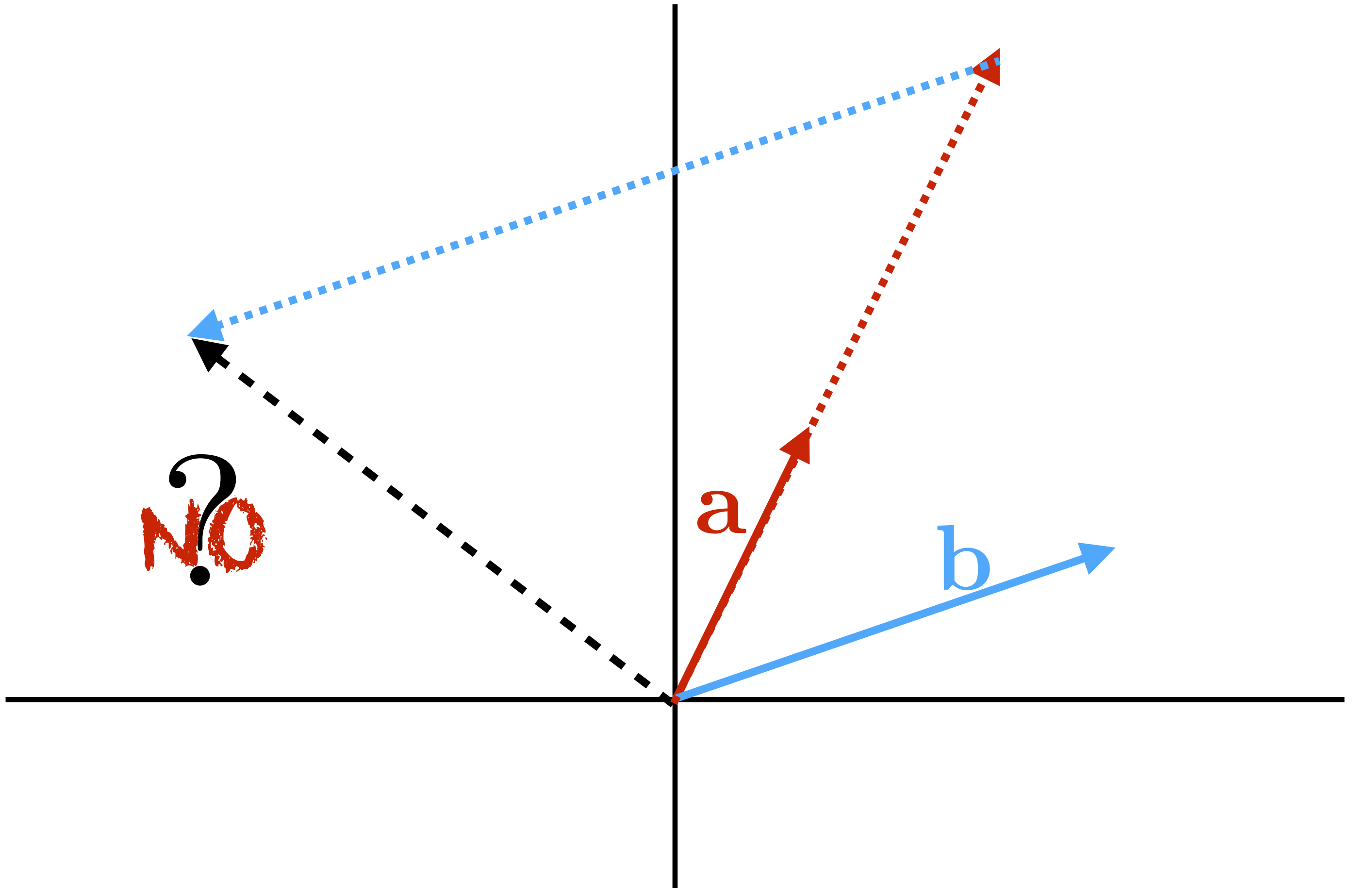
(Geometrically)

- ***Two*** vectors are **linearly dependent** if they are multiples of each other - point in same (or opposite) direction

- ***More than two*** vectors are **linearly dependent** if at least one is a linear combination of the others

Linear Dependence

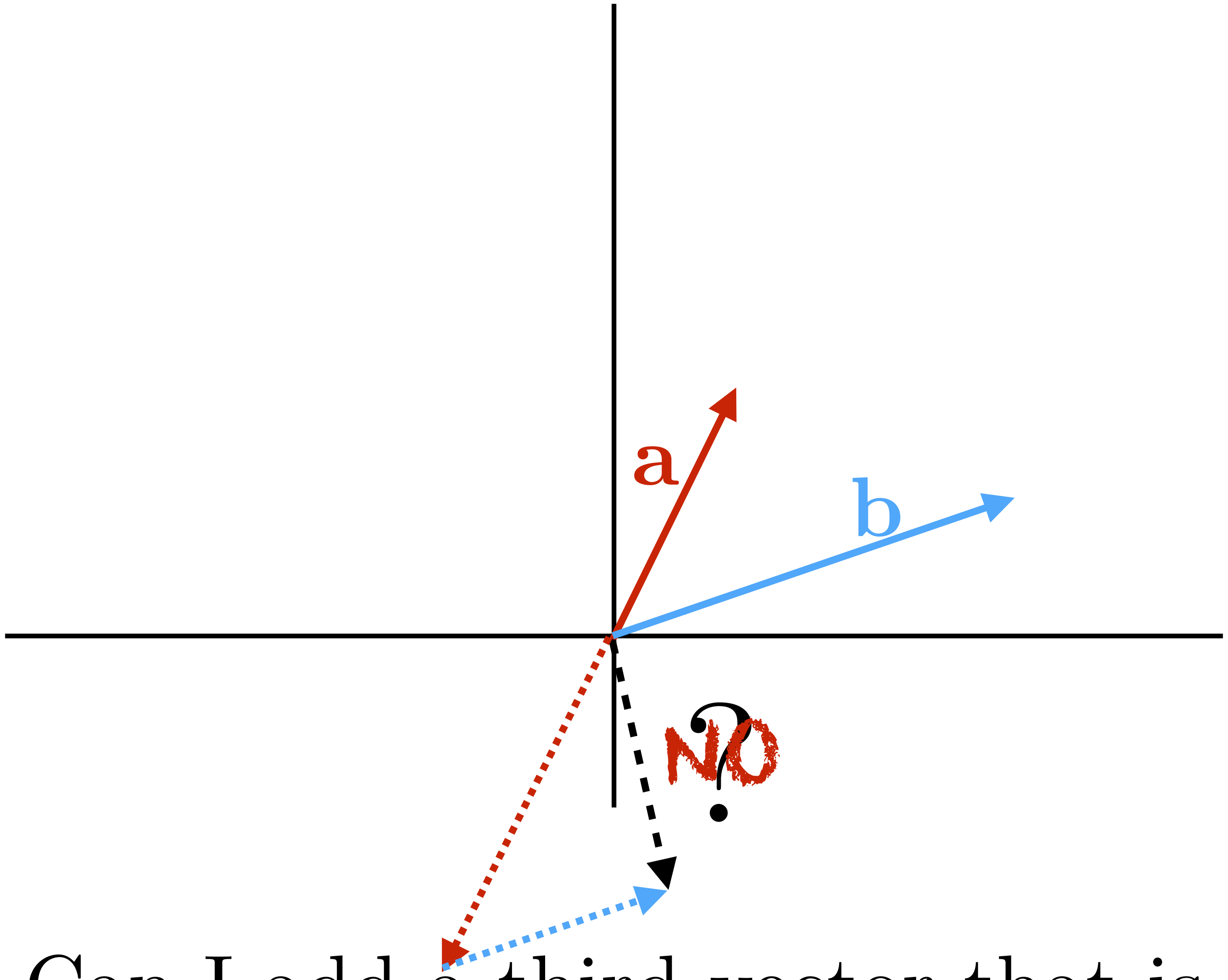
(Geometrically)



Can I add a third vector that is linearly ***in***dependent of **a** and **b**?

Linear Dependence

(Geometrically)



Can I add a third vector that is linearly *in*dependent of **a** and **b**?

Vector Span

(Definition)

- The **span** of a single vector $\mathbf{v}$ is the set of all scalar multiples of $\mathbf{v}$:

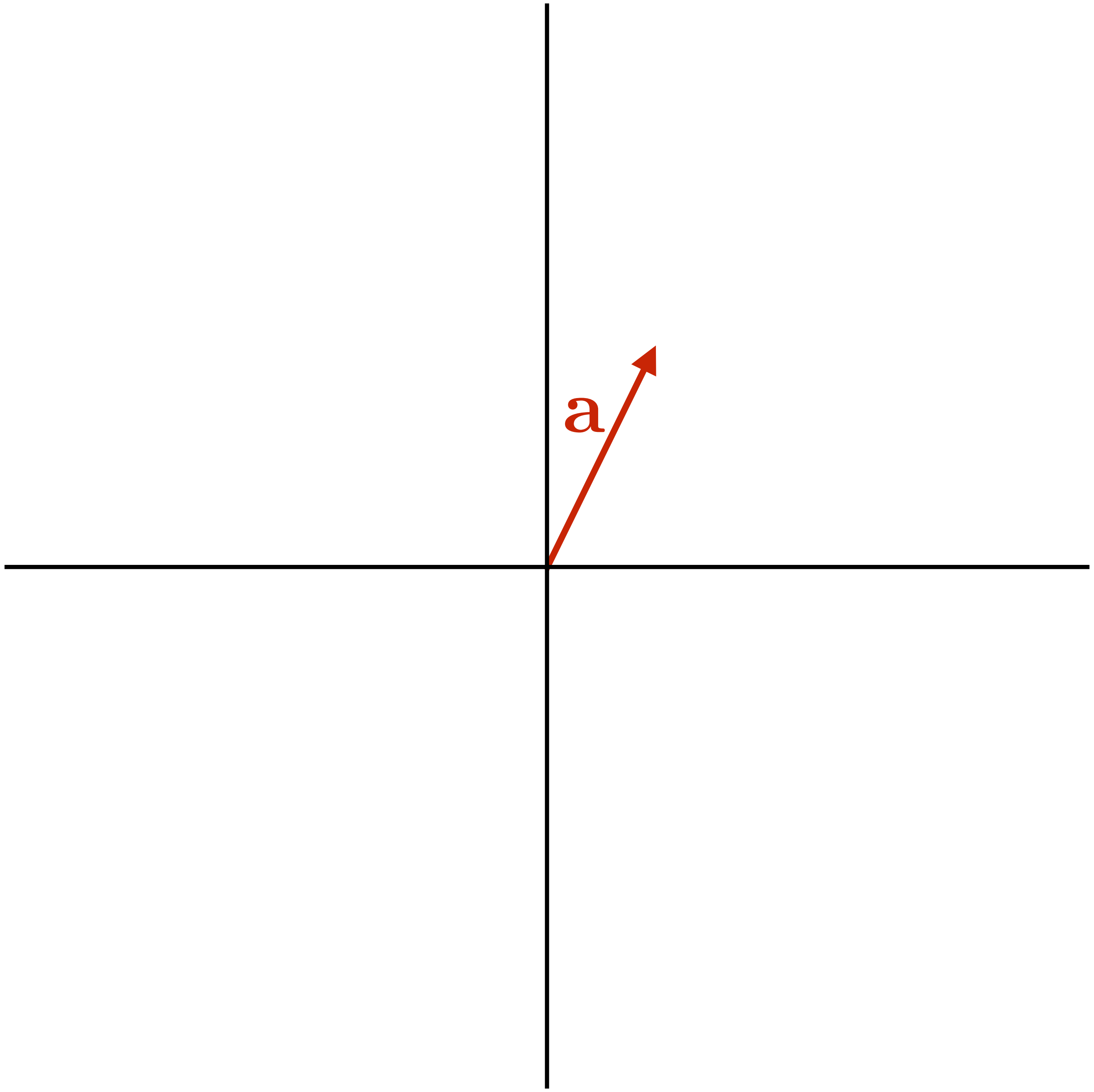
$$span(\mathbf{v}) = \{\alpha\mathbf{v} \quad \text{for all constants } \alpha\}$$

- The **span** of a collection of vectors $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is the set of *all* linear combinations of these vectors:

$$span(\mathbf{V}) = \{\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p \quad \text{for all constants } \alpha_1, \alpha_2, \ldots, \alpha_p\}$$

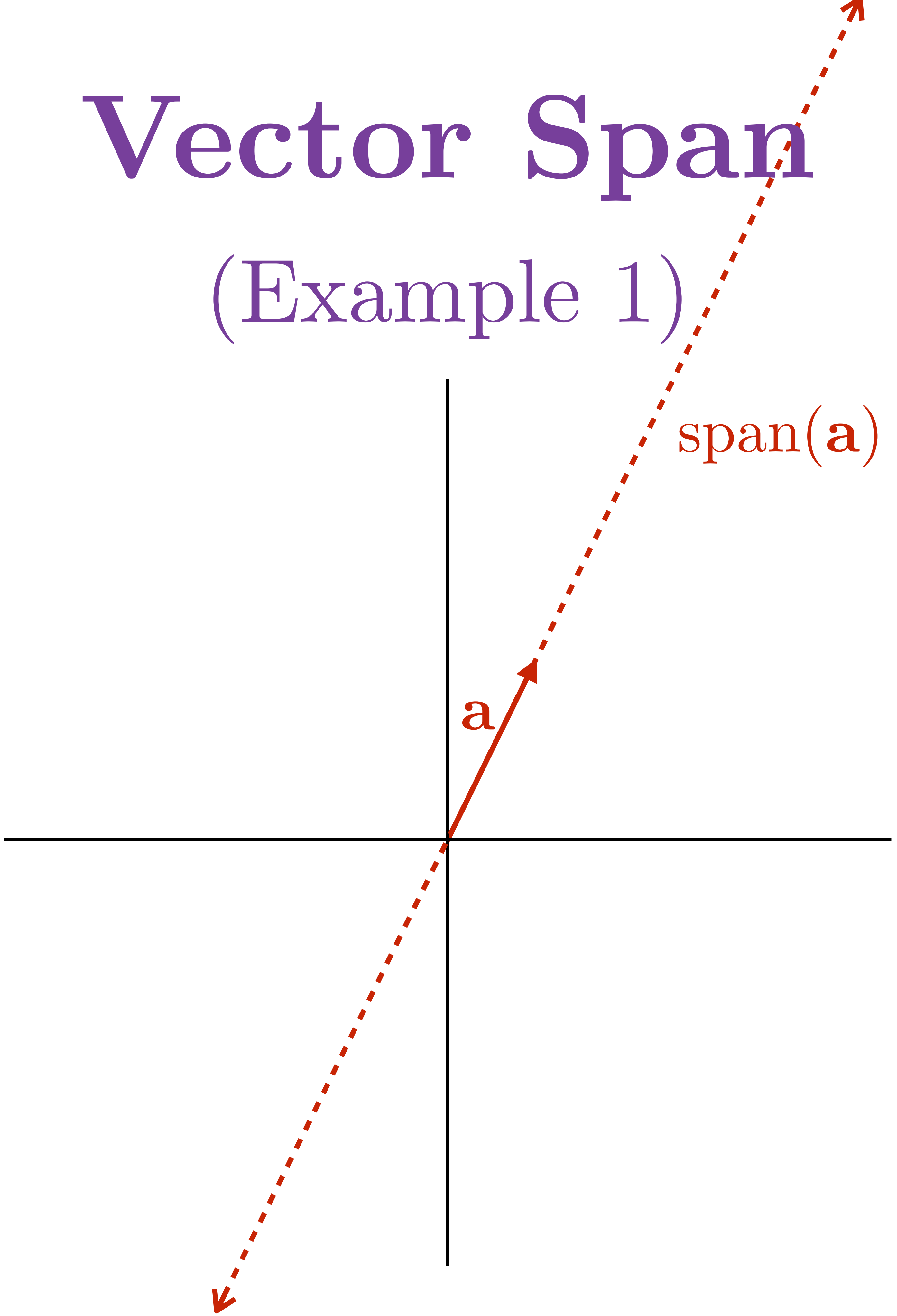
Vector Span

(Example 1)



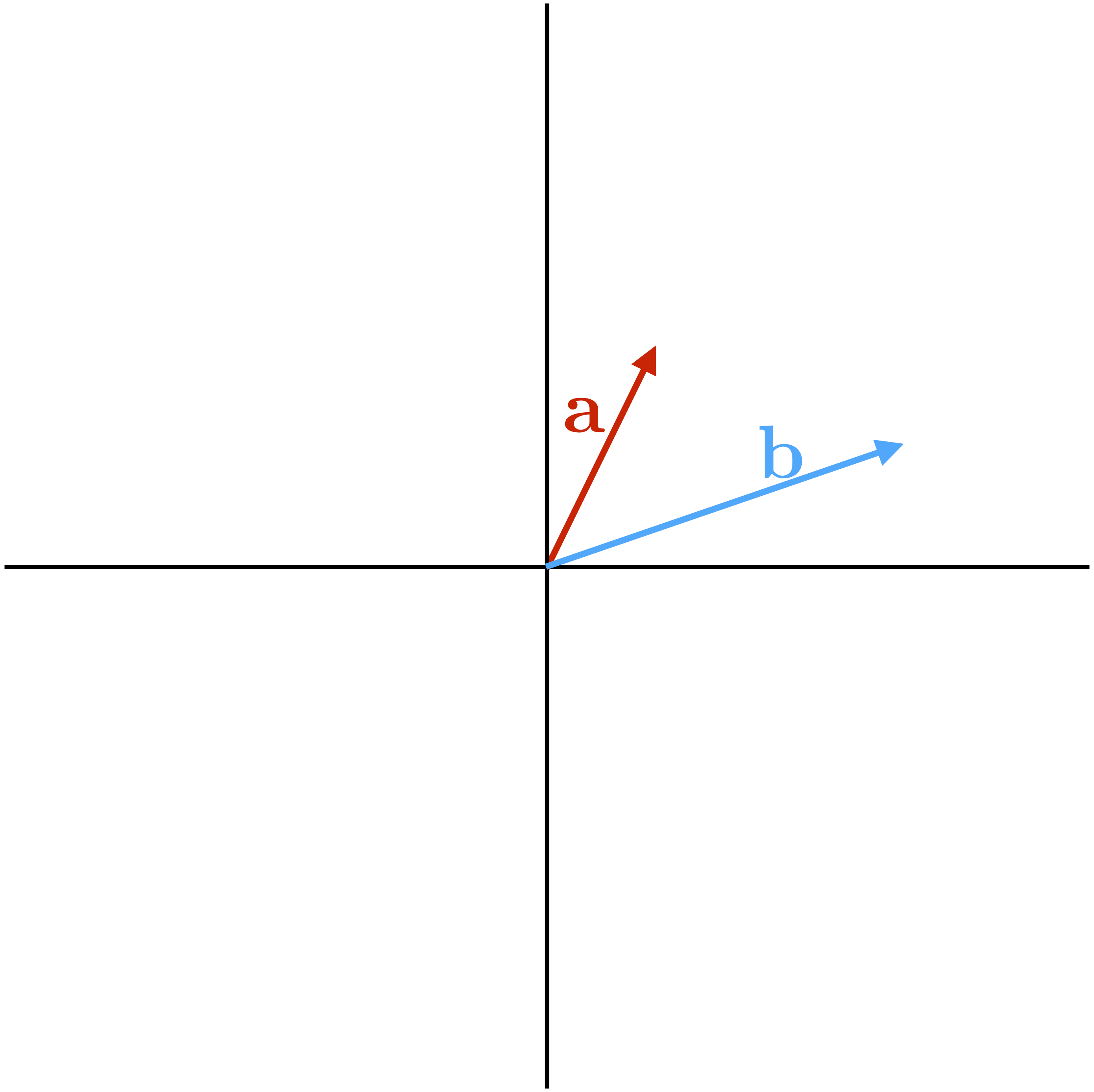
Vector Span

(Example 1)



Vector Span

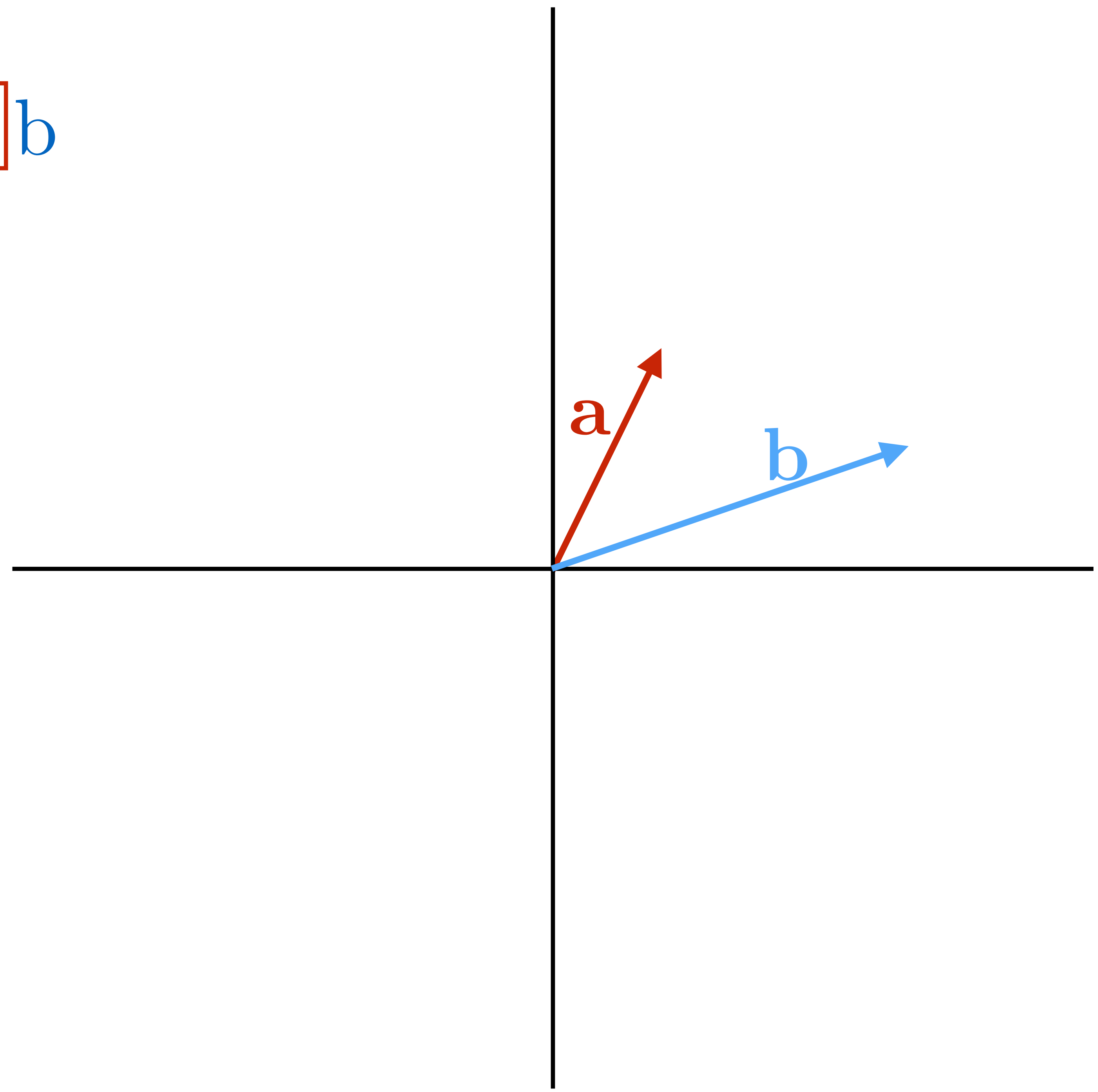
(Example 2)



Vector Span

(Example 2)

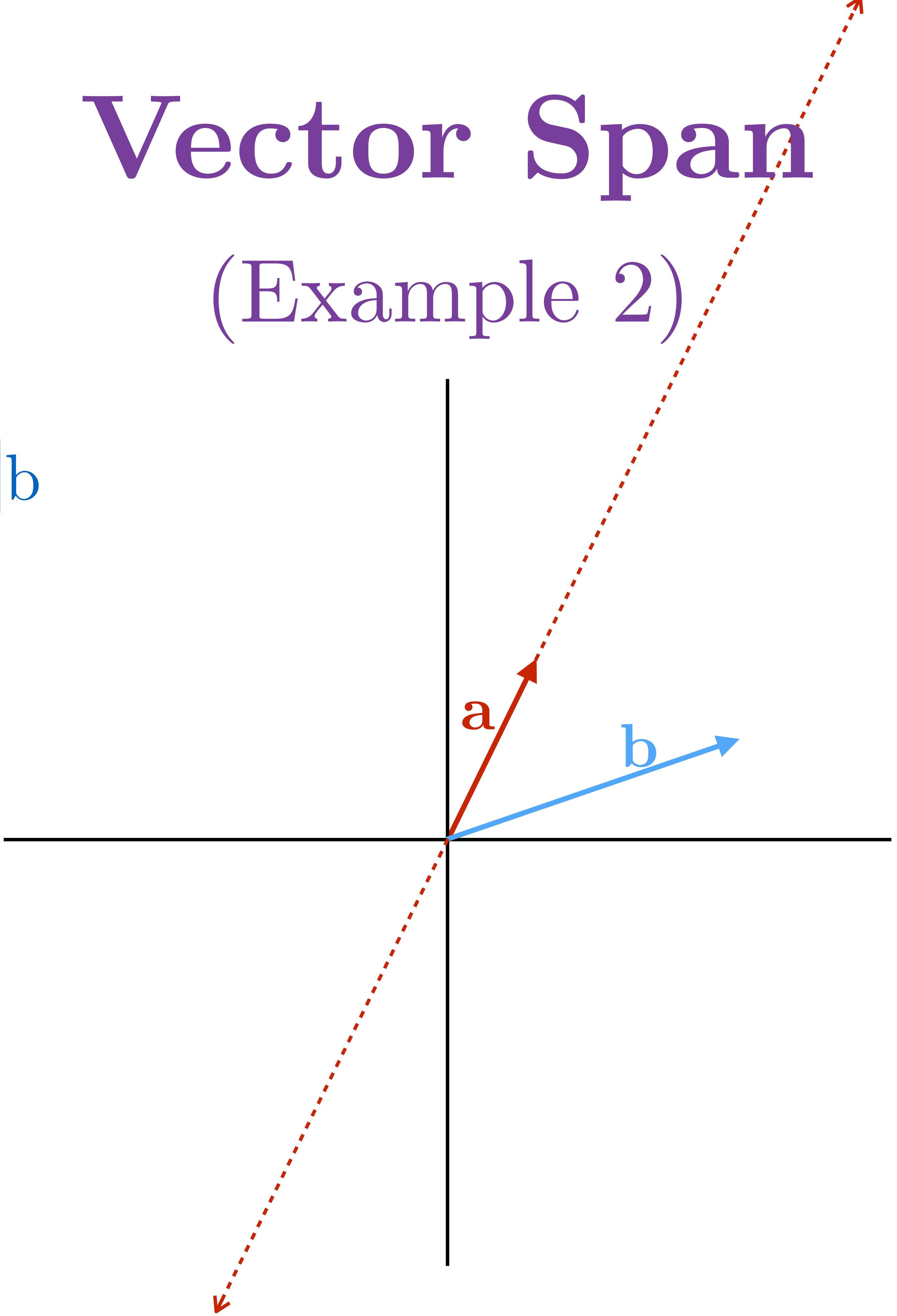
$\square\mathbf{a} + \square\mathbf{b}$



Vector Span

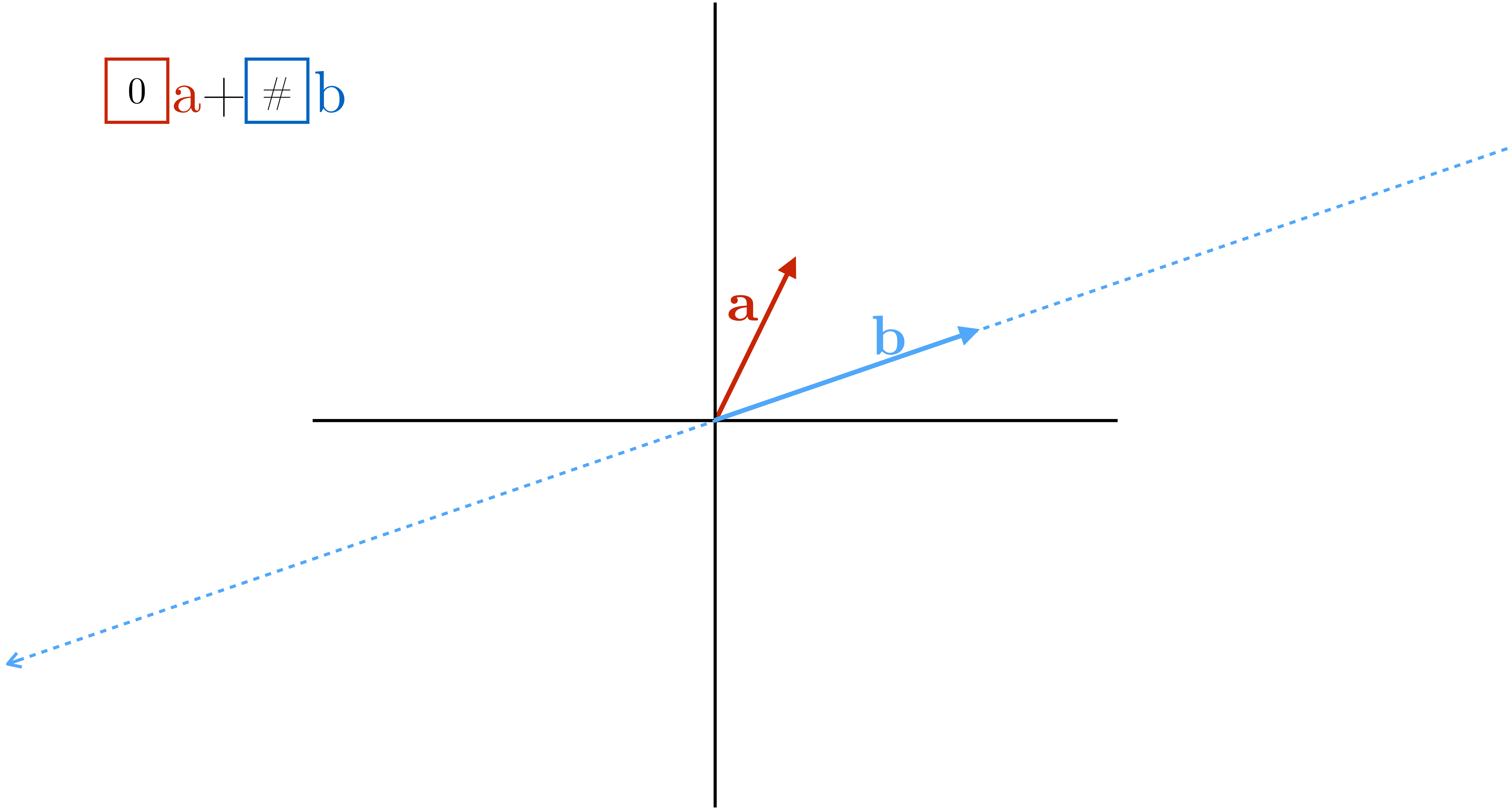
(Example 2)

$\boxed{\#}\,\mathbf{a} + \boxed{0}\,\mathbf{b}$



Vector Span

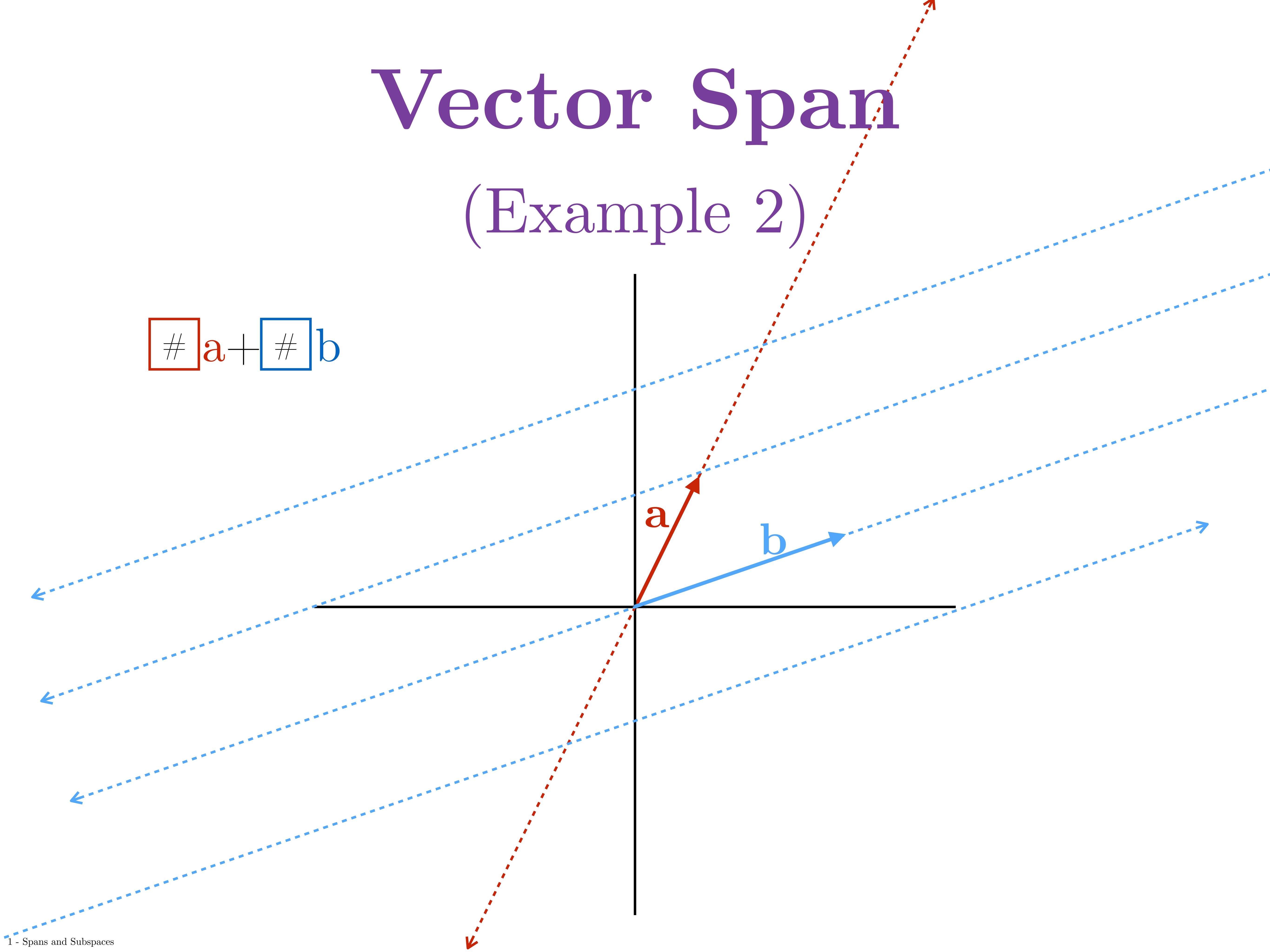
(Example 2)



Vector Span

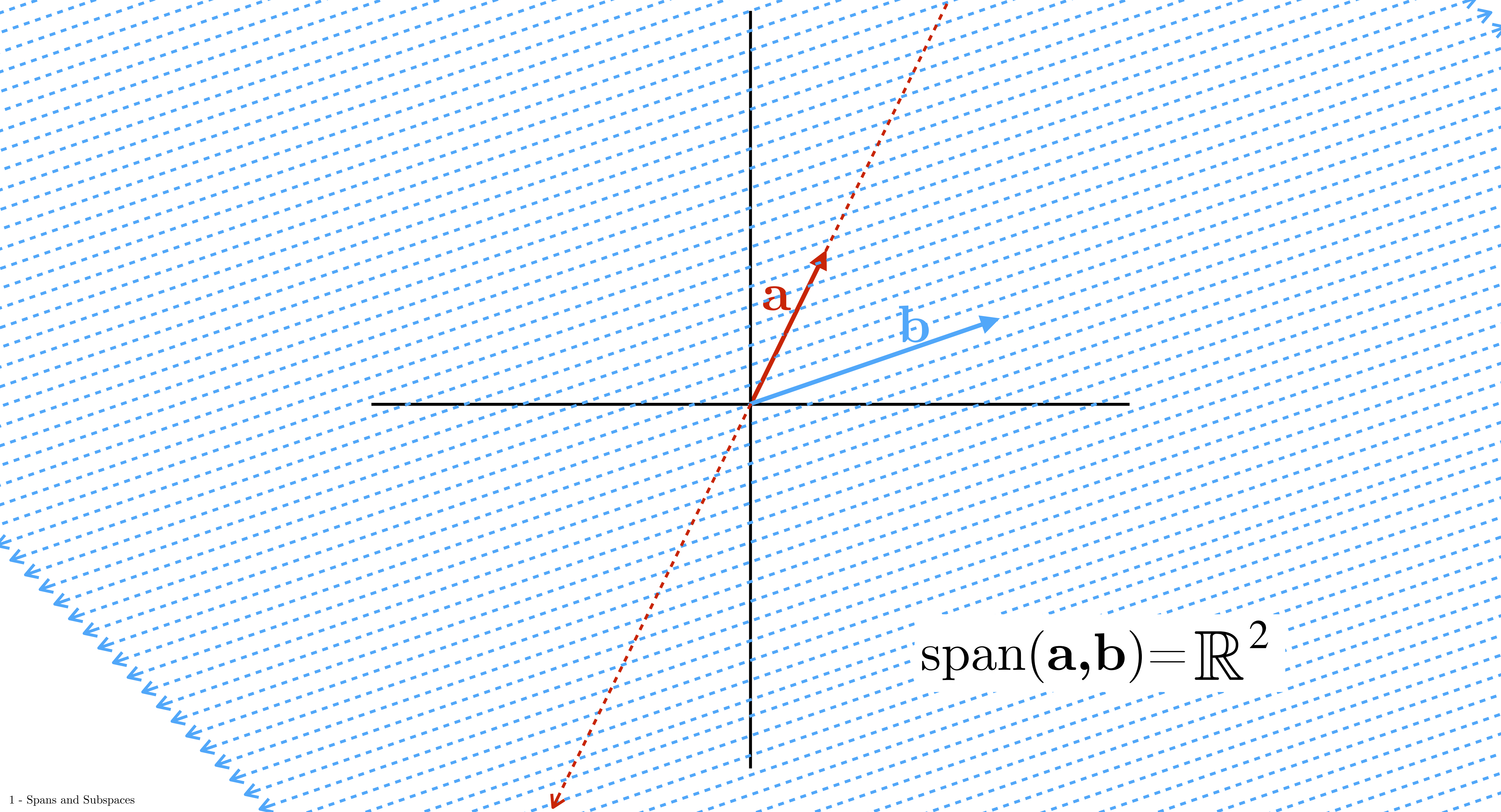
(Example 2)

$\boxed{\#}\,\mathbf{a} + \boxed{\#}\,\mathbf{b}$



Vector Span

(Example 2)

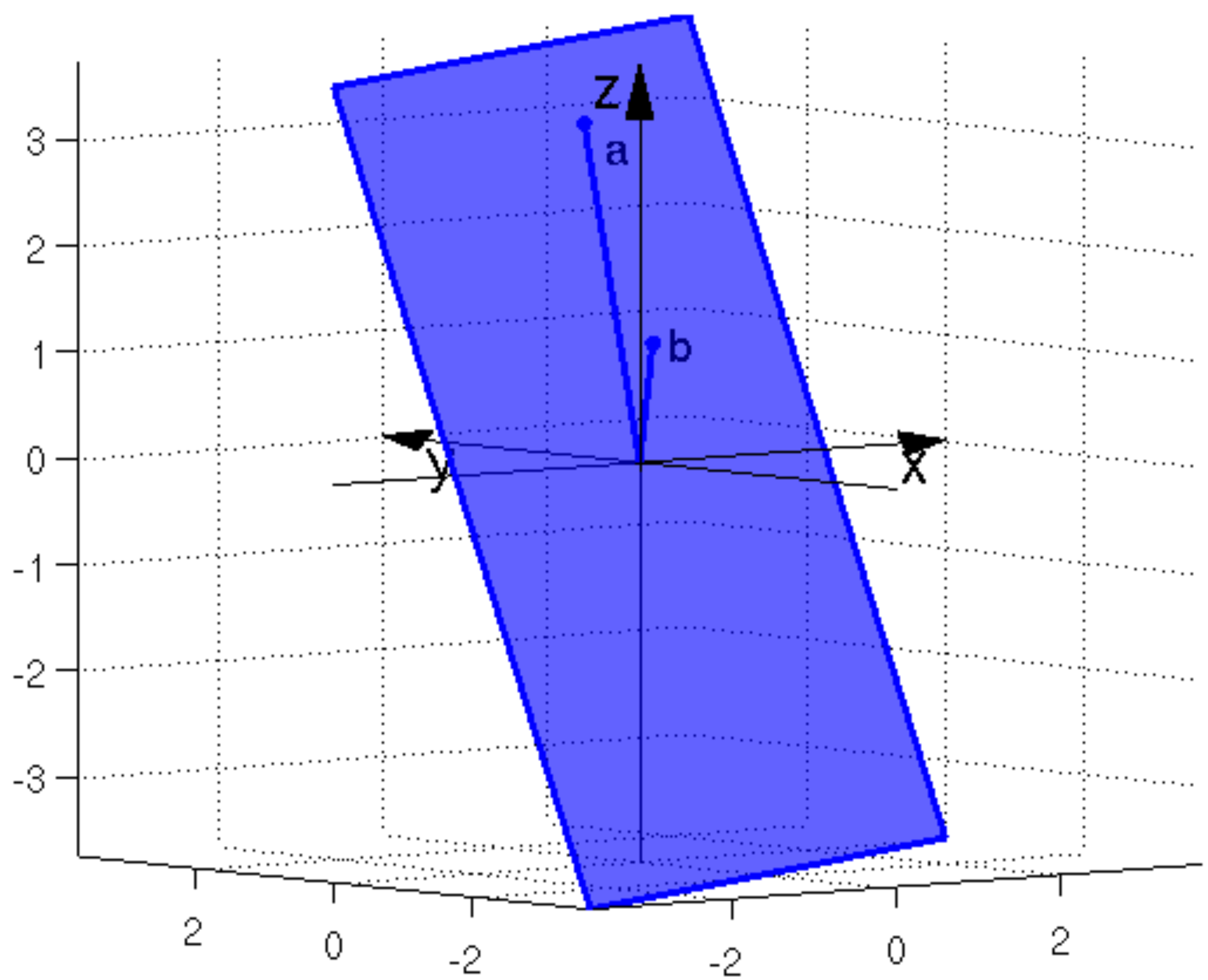


$$\text{span}(\mathbf{a},\mathbf{b}) = \mathbb{R}^2$$

Vector Span

(Example 3)

What is the span of two linearly independent vectors in $\mathbb{R}^3$?



The plane (hyperplane) that contains both vectors.

(A 2-dimensional **subspace** of $\mathbb{R}^3$)

Subspace

(Definition)

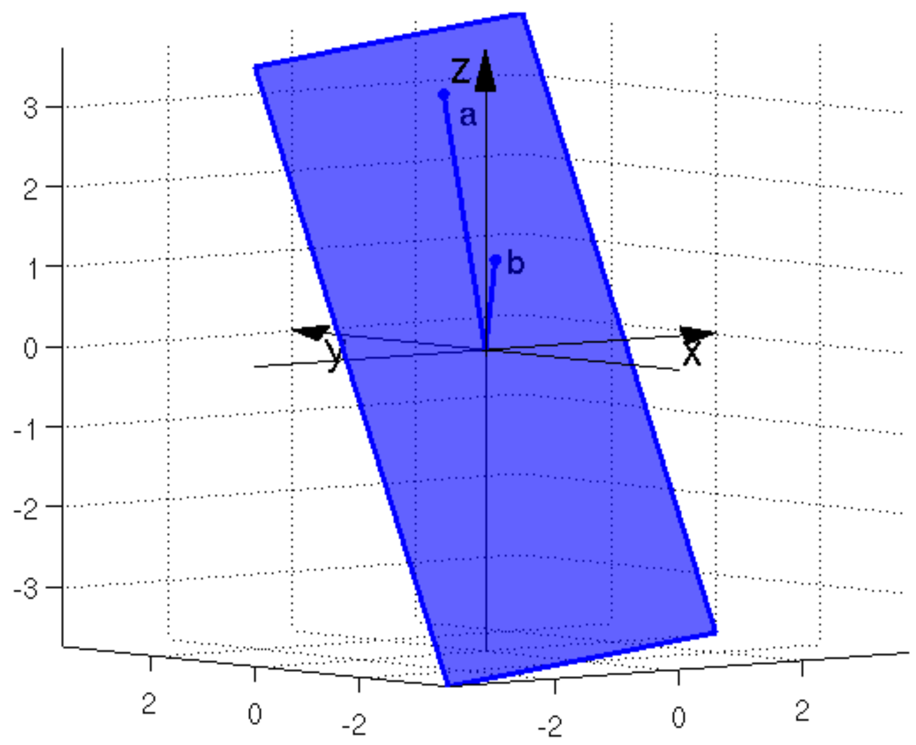
- A **subspace** S of $\mathbb{R}^n$ is thought of as a "flat" (having no curvature) surface within $\mathbb{R}^n$. It is a collection of vectors which satisfies the following conditions:

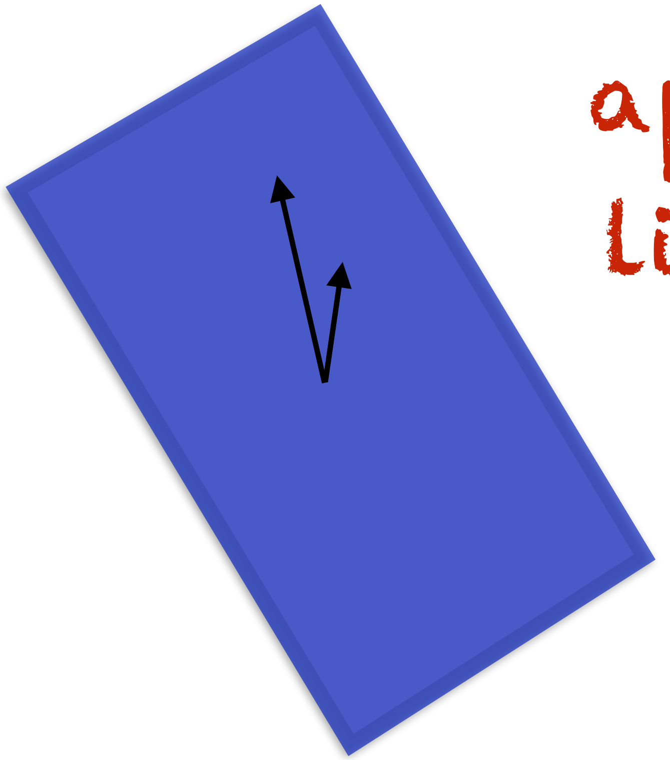
  - The origin ($\mathbf{0}$ vector) is contained in S.
  - If $\mathbf{x}$ and $\mathbf{y}$ are in S then $\mathbf{x}+\mathbf{y}$ also in S.
  - If $\mathbf{x}$ is in S then $\alpha\mathbf{x}$ is in S for any scalar α.

Subspace

(Definition)

- In other words, it is an infinite subset of vectors (points) from a larger space ($\mathbb{R}^n$) that when taken alone, appears like $\mathbb{R}^p$, $p<n$





- The **dimension** of the subspace is the minimum number of vectors it takes to span the space. (Think: # of axes)

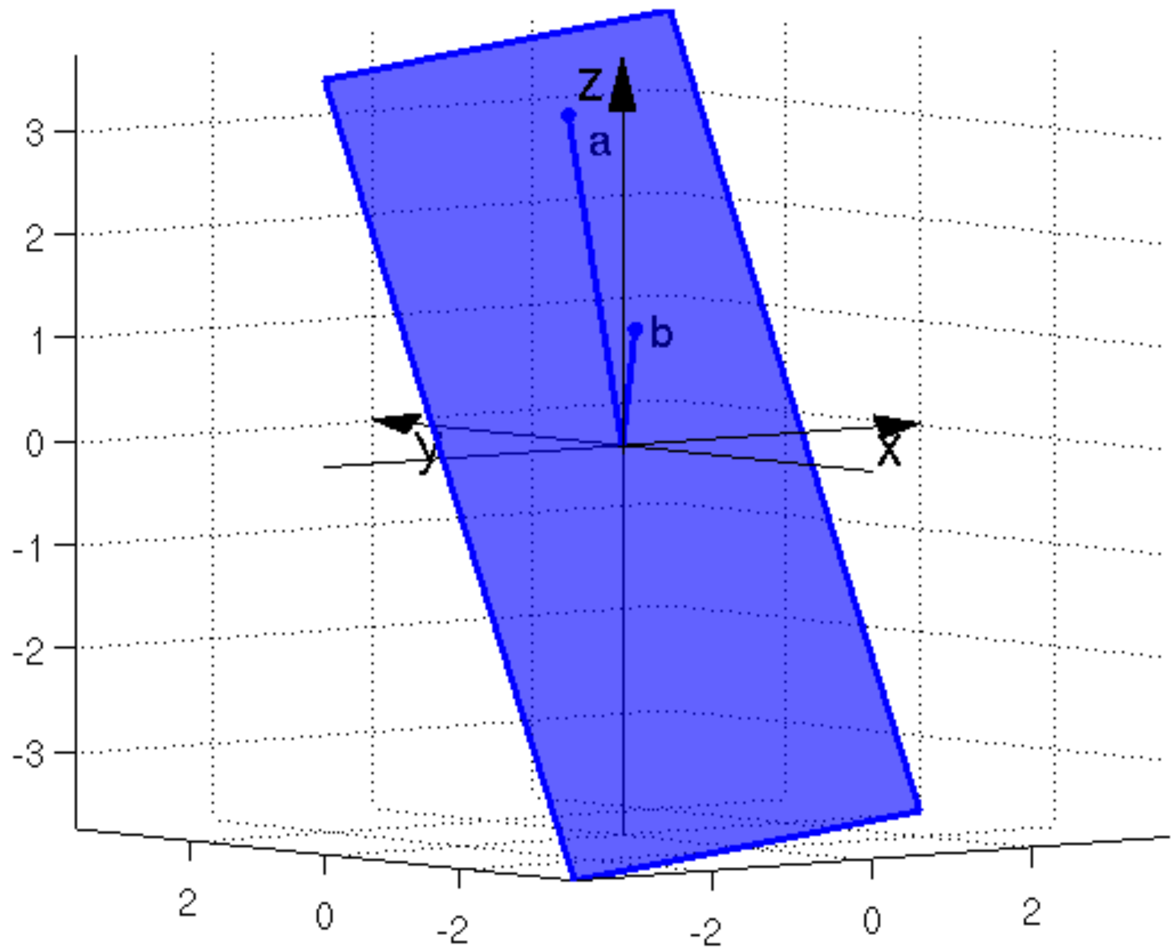
Hyperplane

(Definition)

- A **hyperplane** is a subspace that has one less dimension than its ambient space.
- In 3-dimensional space a hyperplane would be a 2-dimensional plane.
- In 4-dimensional space, a hyperplane would be a 3-dimensional plane (helps to keep same picture in mind: a "flat" subspace in 4D!)

Hyperplane
(Definition)

- A **hyperplane** cuts the ambient space into two parts, one ‘above’ it and one ‘below it’



Practice

1 Is the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the $span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$?

2 Describe the span of one vector in $\mathbb{R}^3$

3 Describe the span of two linearly dependent vectors in $\mathbb{R}^3$

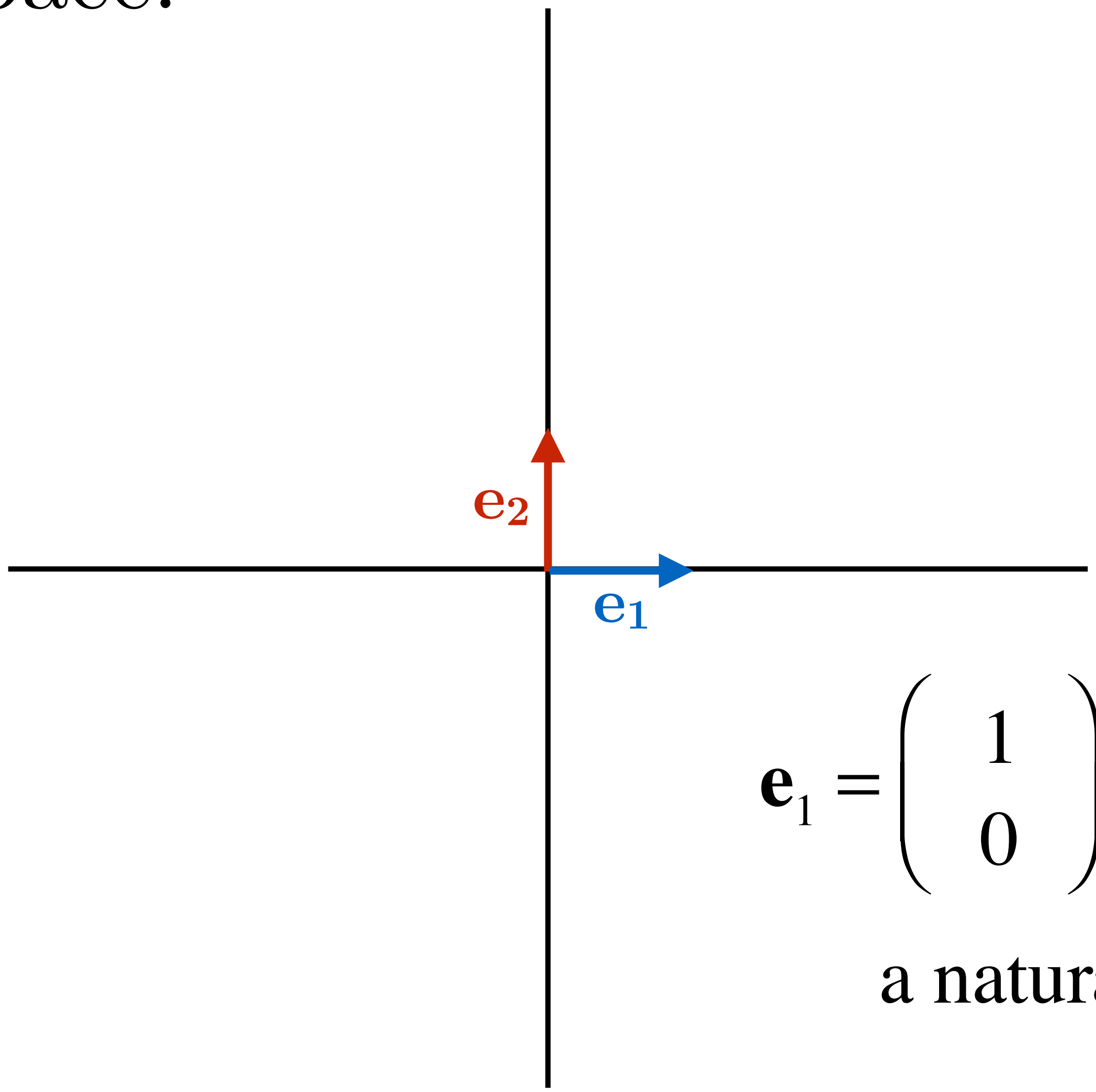
4 Compare the $span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ to the $span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right\}$

5 What is the dimension of the $span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right\}$

Part 2:
Basis and Coordinates

Basis and Coordinates

- A collection of vectors makes a **basis** for a space (or a subspace) if they are linearly independent and span the space.

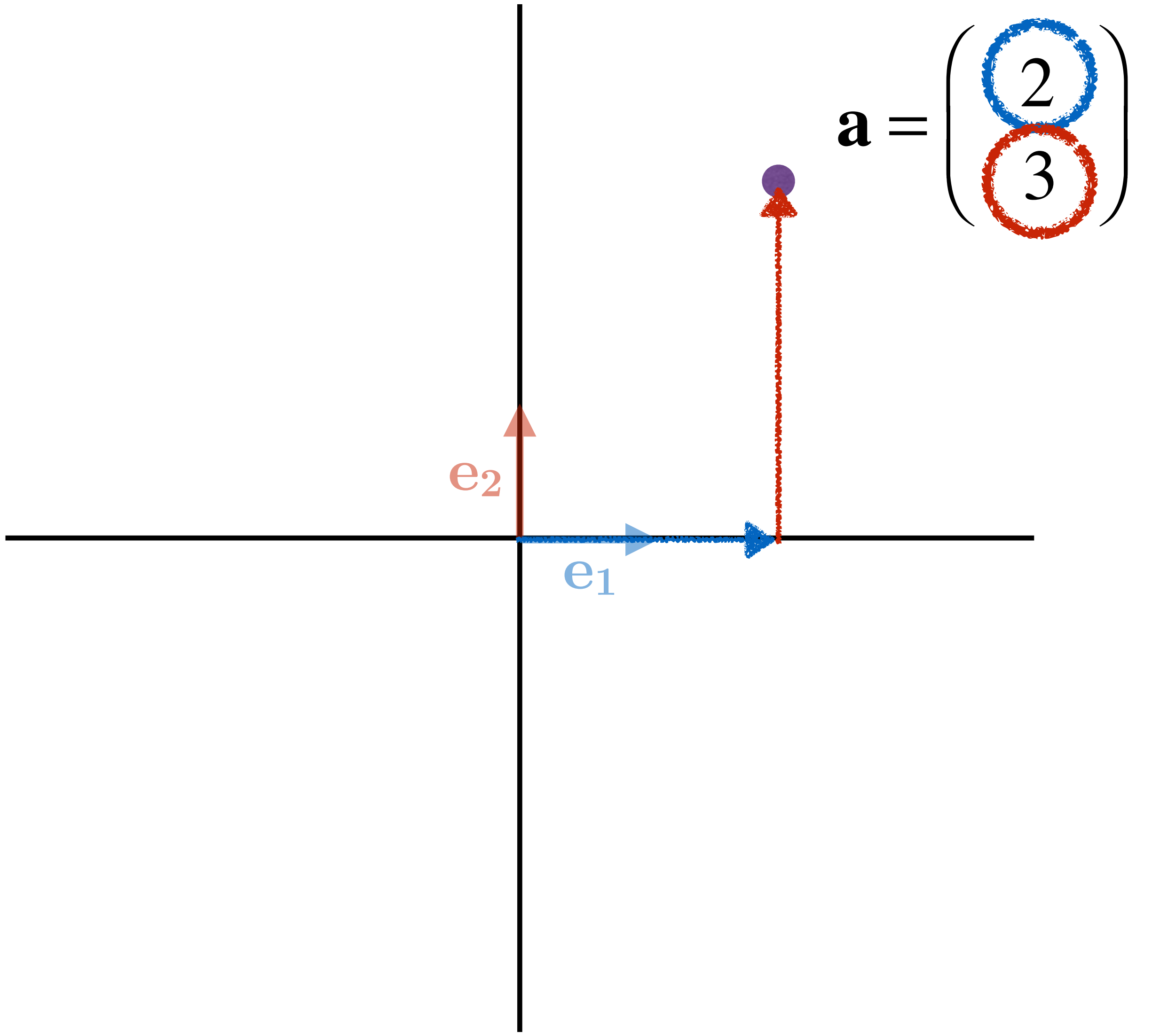


$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ are}$$

a natural basis for $\mathbb{R}^2$

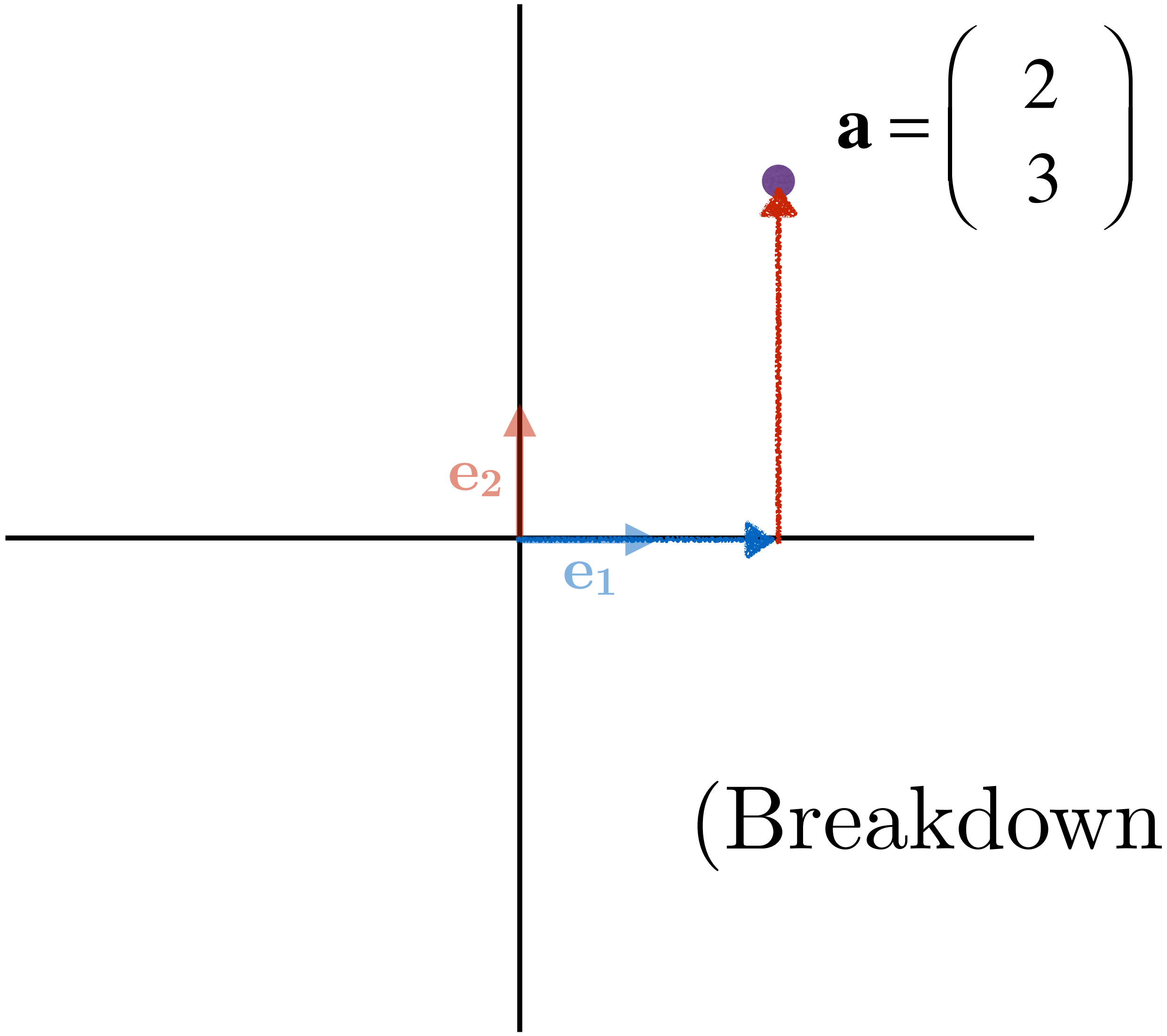
Basis and Coordinates

- Coordinate pairs are represented in a basis. Each **coordinate** tells you how far to move along each basis direction.



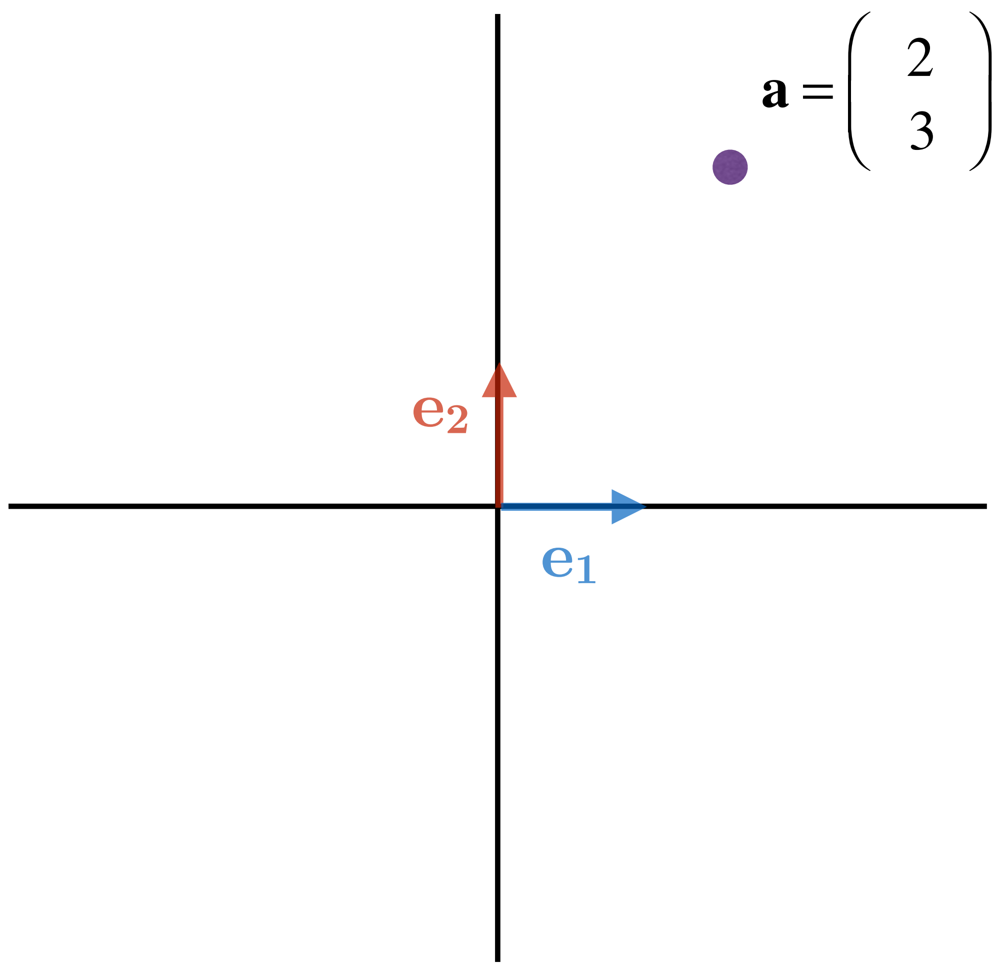
Basis and Coordinates

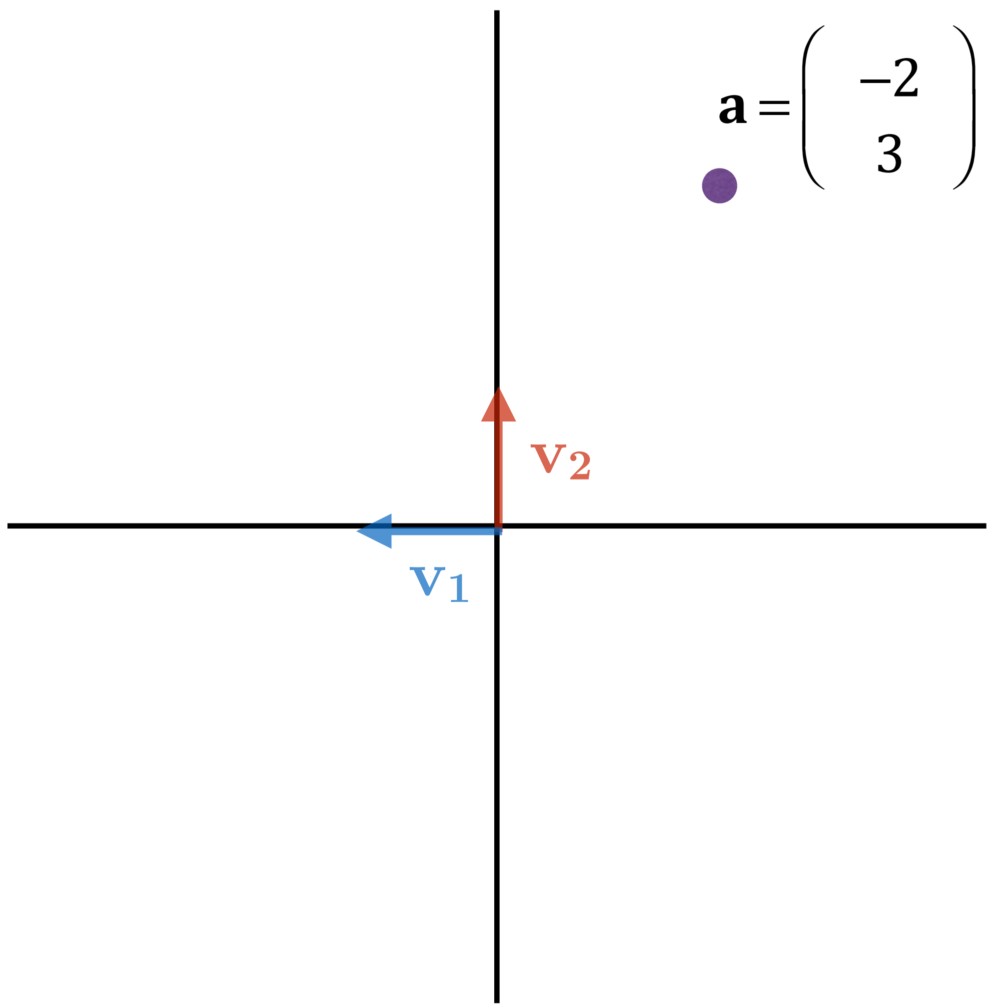
- Coordinate pairs are represented in a basis. Each **coordinate** tells you how far to move along each basis direction.



(Breakdown into parts)

Change of Basis

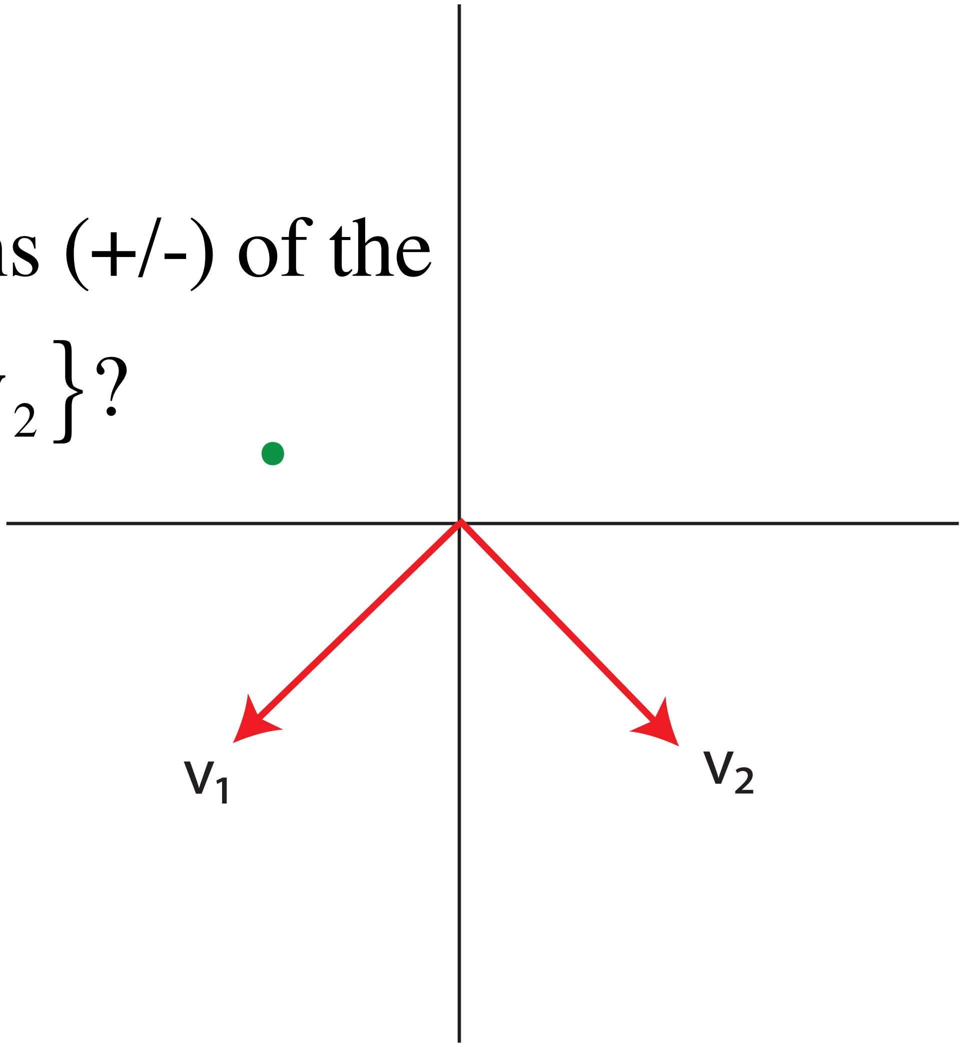




$$2\mathbf{e}_1 + 3\mathbf{e}_2 = -2\mathbf{v}_1 + 3\mathbf{v}_2$$

Practice

1 In the following picture, what would be the signs (+/-) of the coordinates of the green point in the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$?



2 Find the coordinates of the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the basis $\left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

Draw a picture to confirm your answer matches your intuition.

Part 3:
A Change of Basis

Basis and Coordinates

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

Basis and Coordinates

$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

Coordinates

Bases and Coordinates

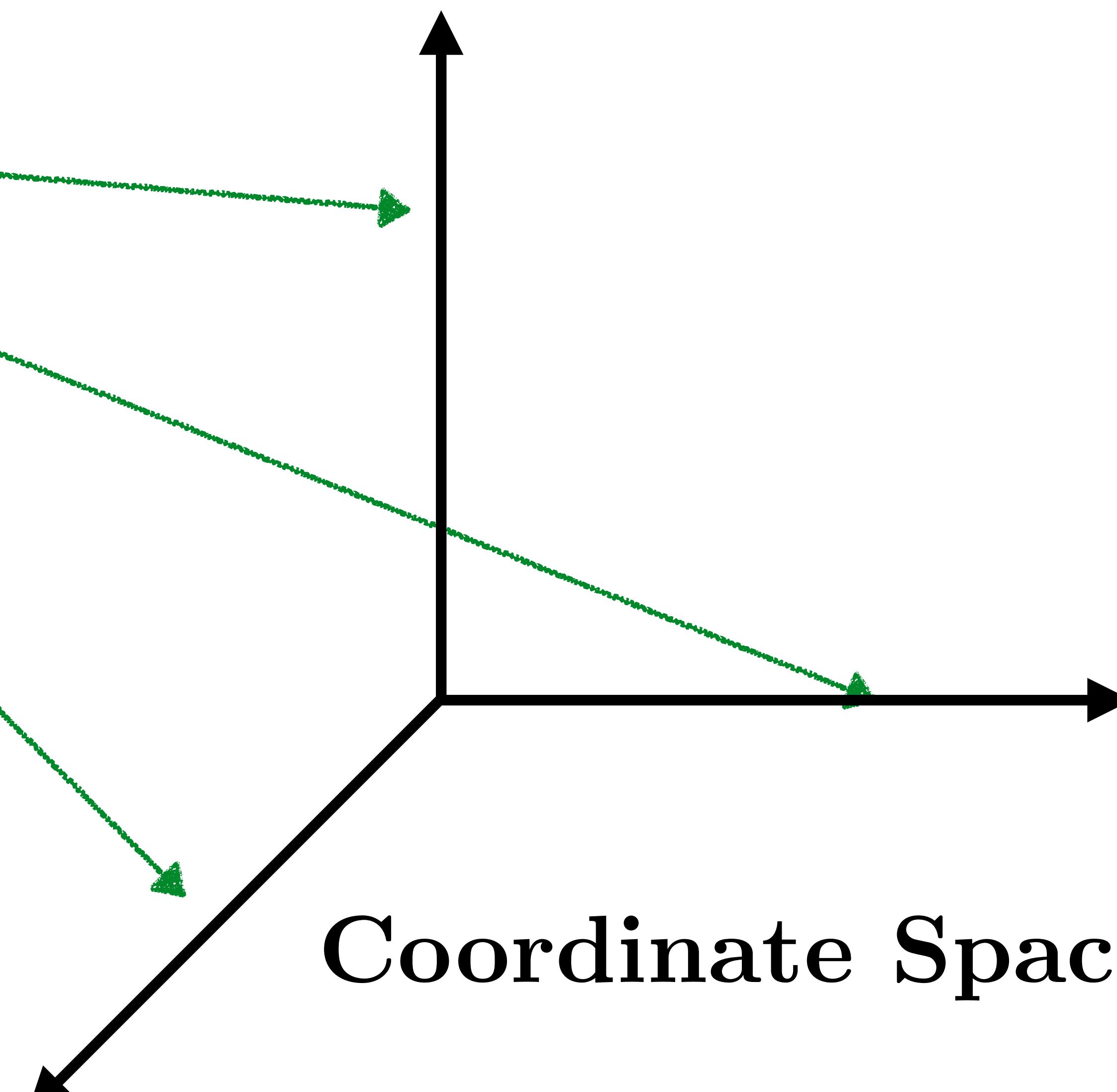
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

Basis Vectors

$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

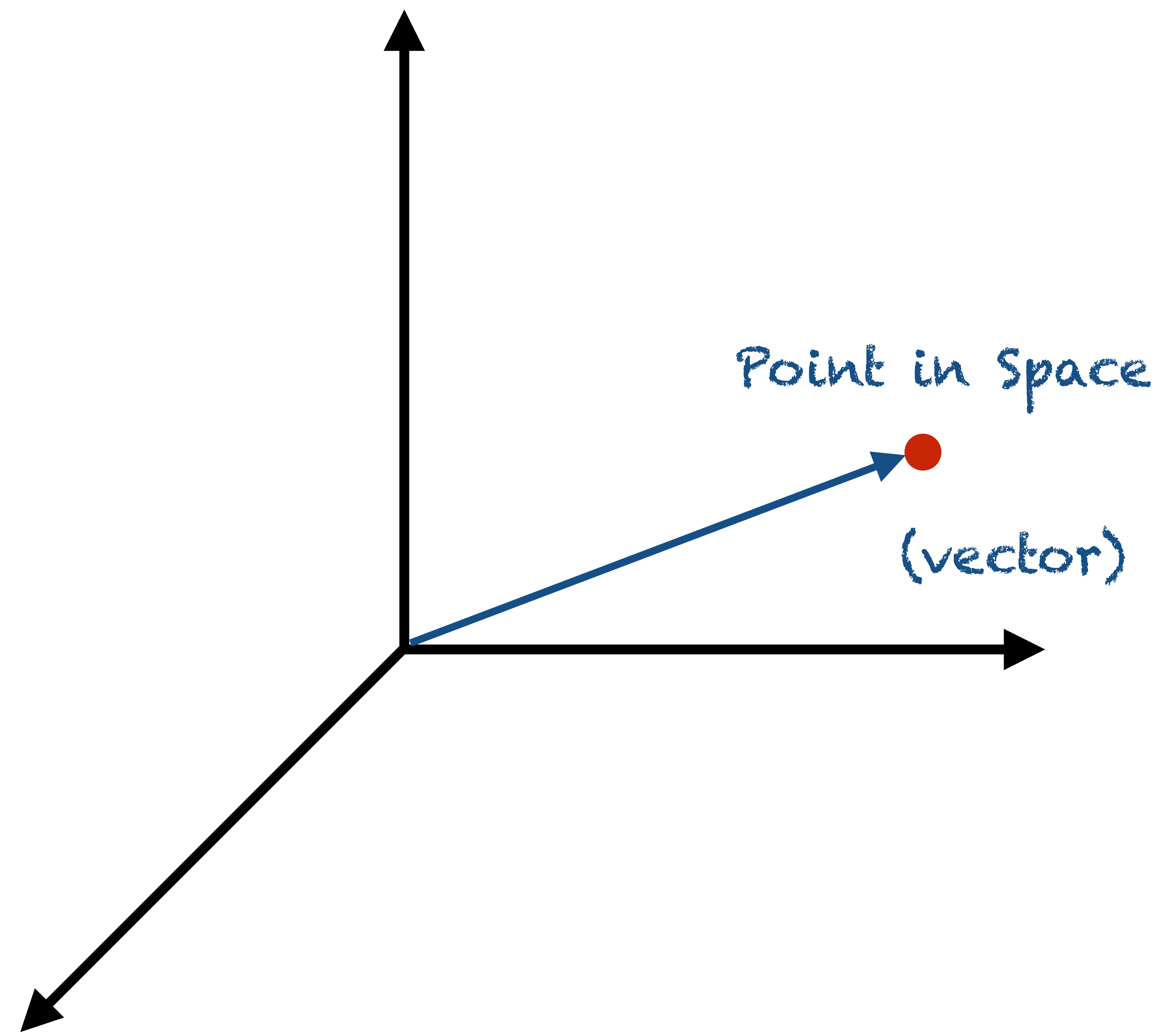
Basis Vectors

(Think: Axes)



Coordinate Space

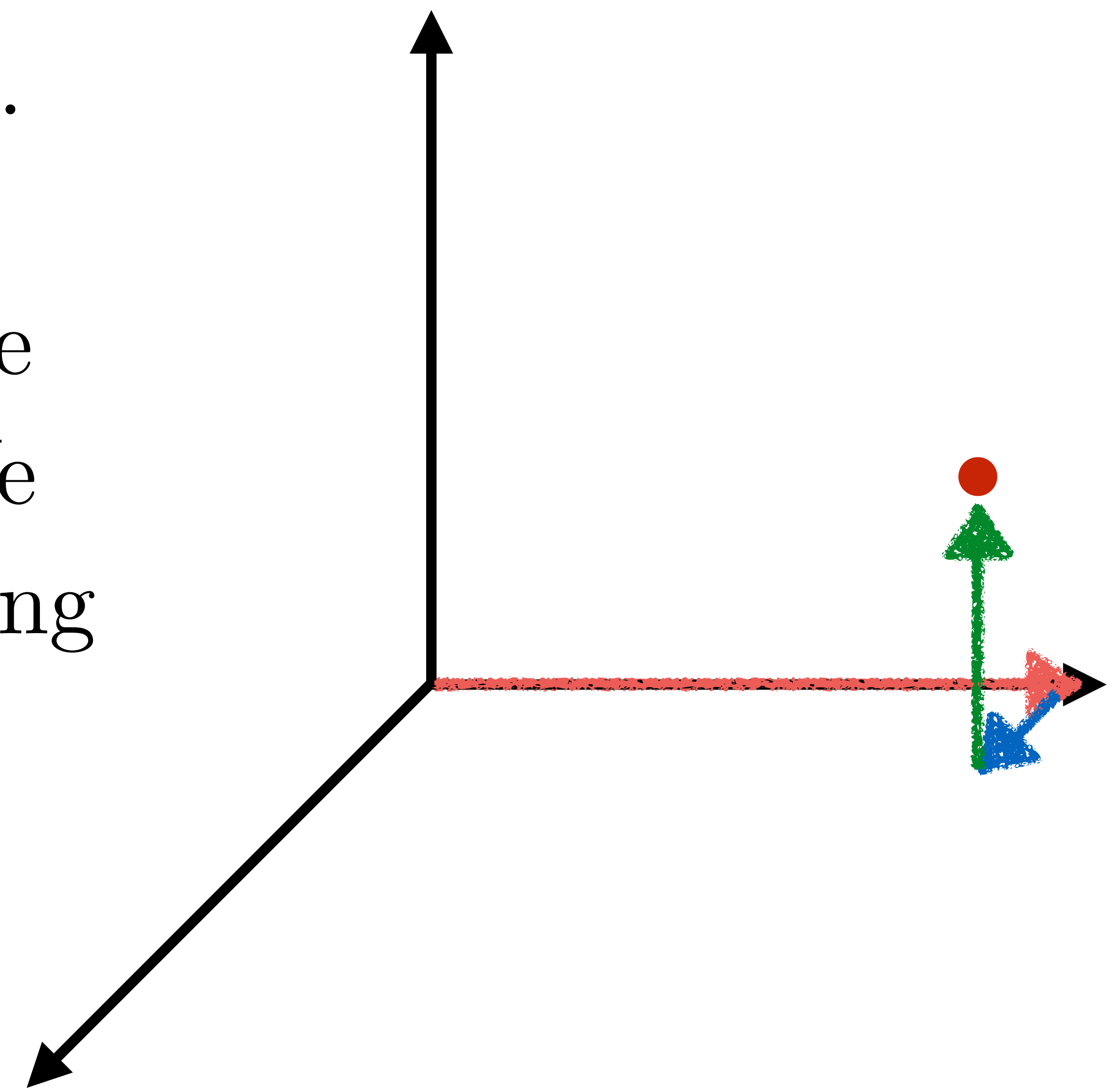
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$



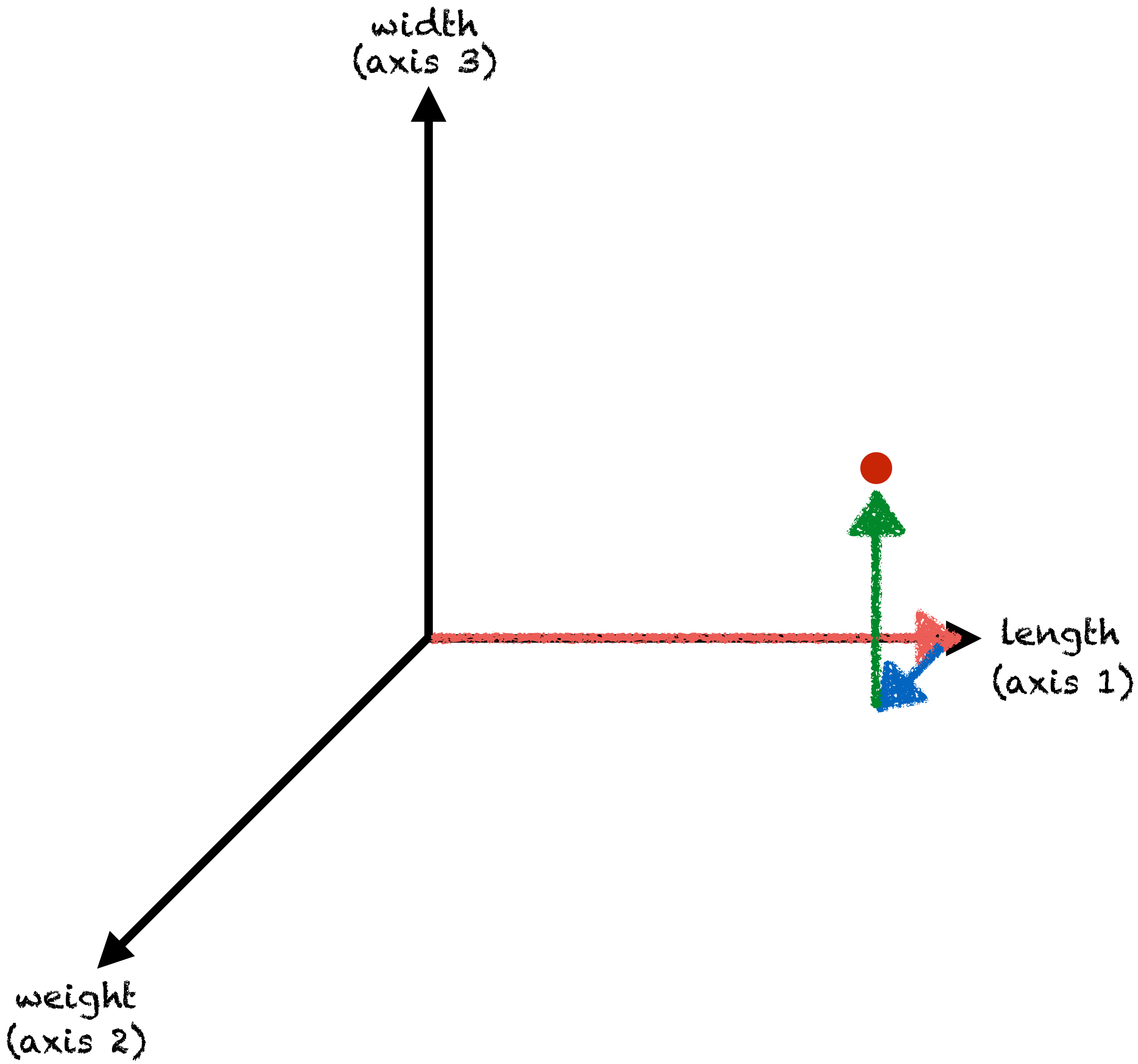
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

Coordinates give directions to a point along basis vectors.

For any set of data points, the basis vectors are the same. We compare the points by comparing their coordinates.

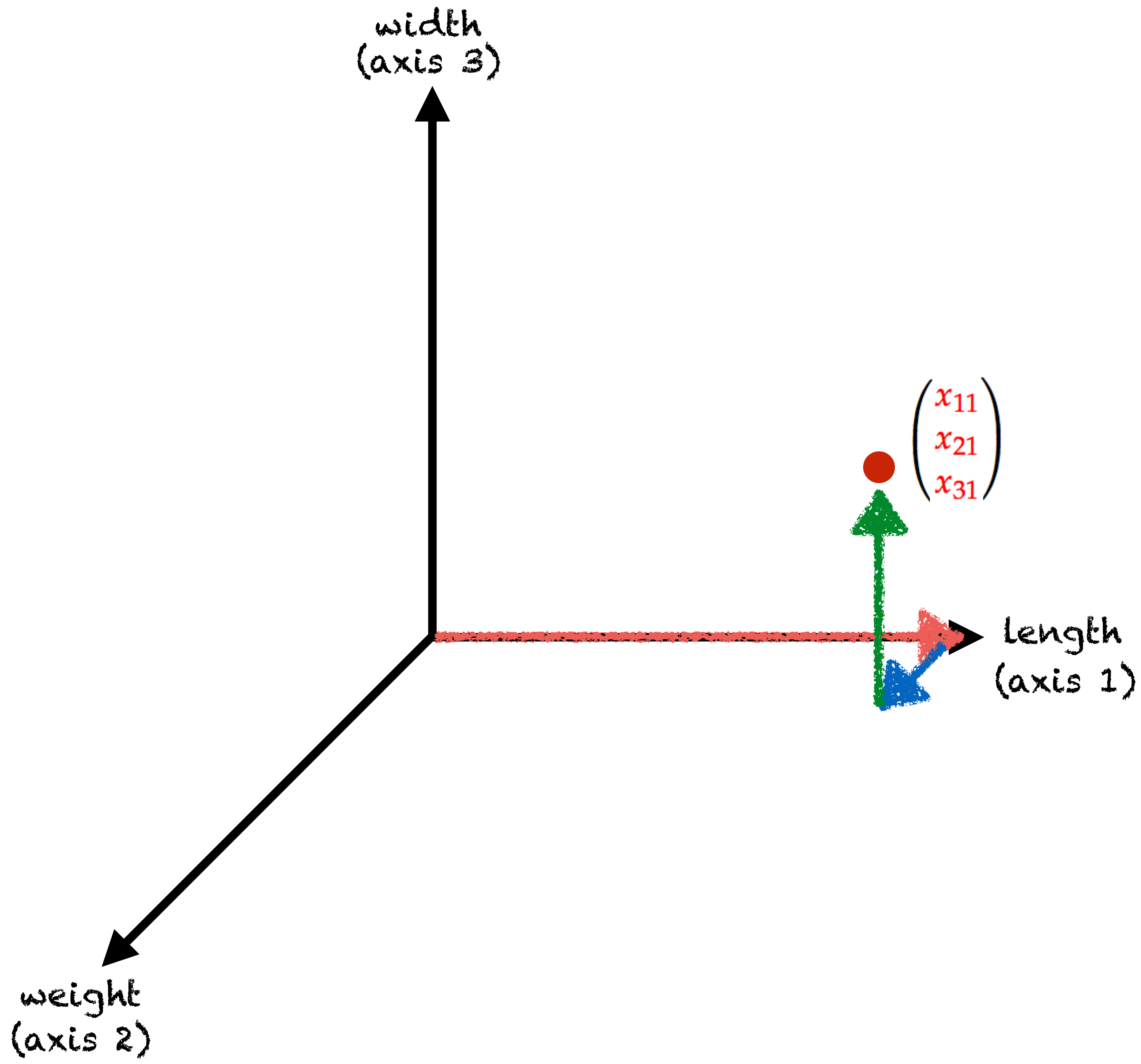


$$x_{11}\underset{\text{length}}{\begin{pmatrix}1\\0\\0\end{pmatrix}} + x_{21}\underset{\text{weight}}{\begin{pmatrix}0\\1\\0\end{pmatrix}} + x_{31}\underset{\text{width}}{\begin{pmatrix}0\\0\\1\end{pmatrix}} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

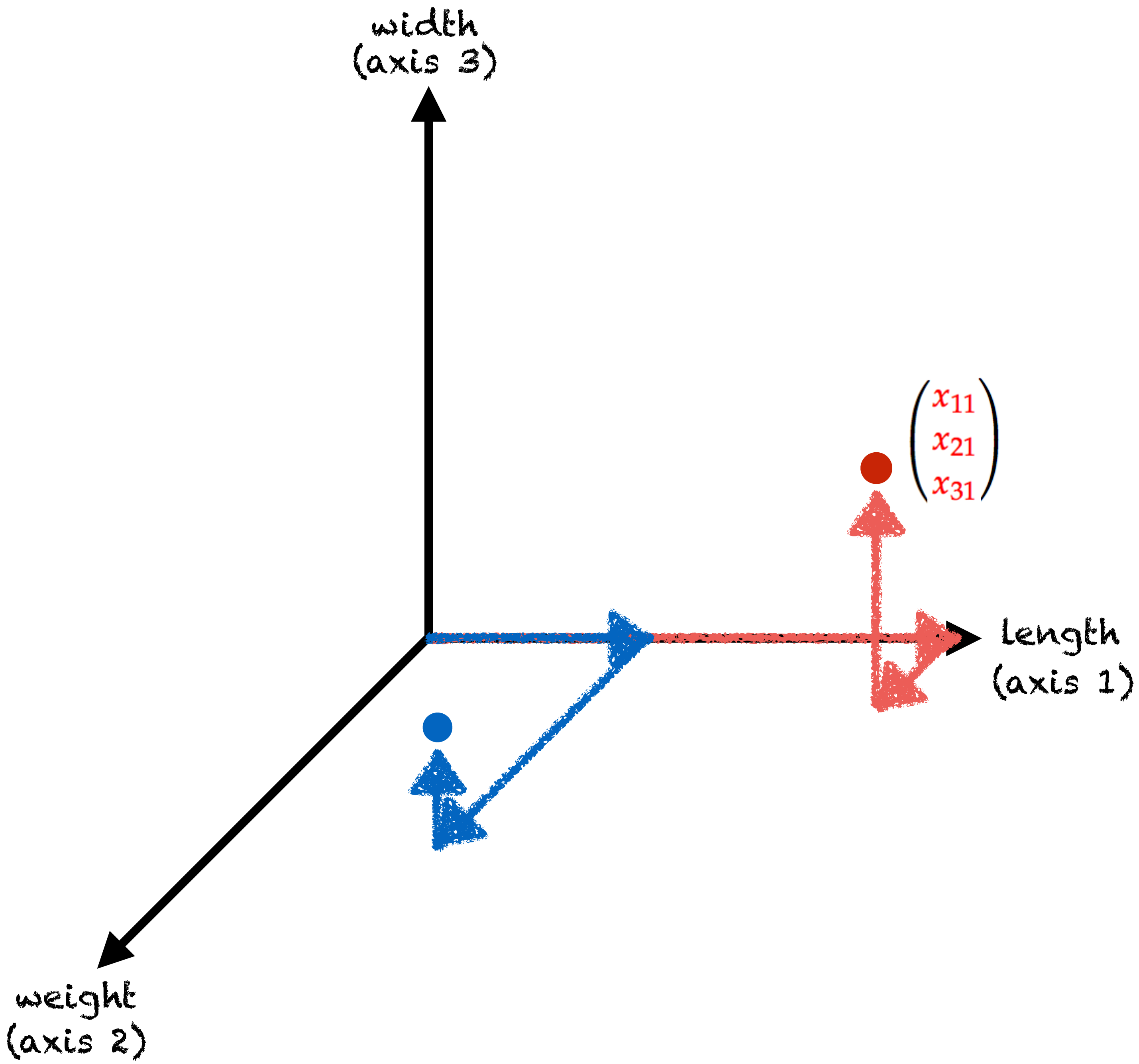


$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} =$$

length weight width

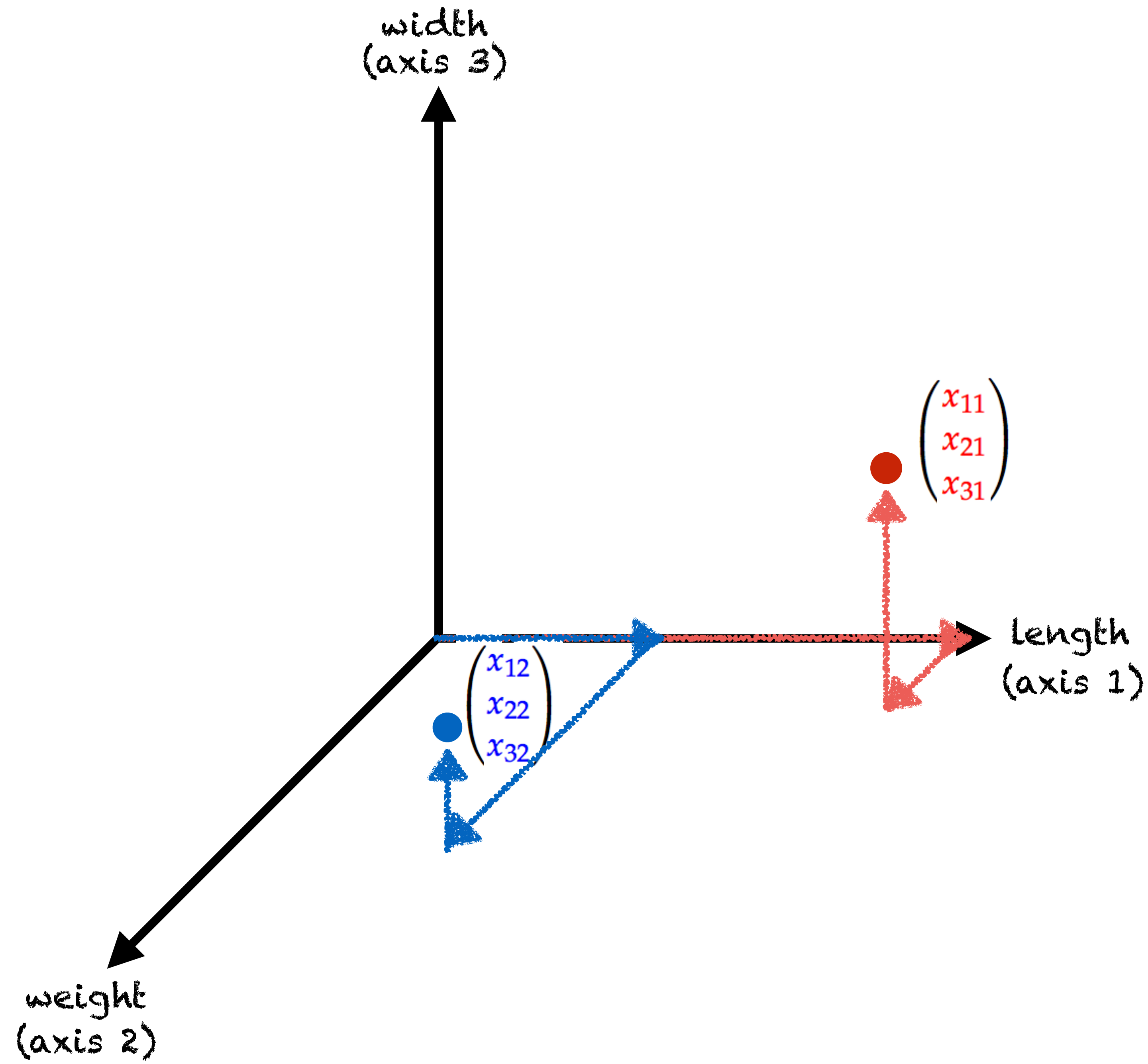


$$x_{12}\underbrace{\begin{pmatrix}1\\0\\0\end{pmatrix}}_{\text{length}} + x_{22}\underbrace{\begin{pmatrix}0\\1\\0\end{pmatrix}}_{\text{weight}} + x_{32}\underbrace{\begin{pmatrix}0\\0\\1\end{pmatrix}}_{\text{width}} = \begin{pmatrix}x_{12}\\x_{22}\\x_{32}\end{pmatrix}$$



$$x_{12}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{22}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{32}\begin{pmatrix}0\\0\\1\end{pmatrix} =$$

length weight width

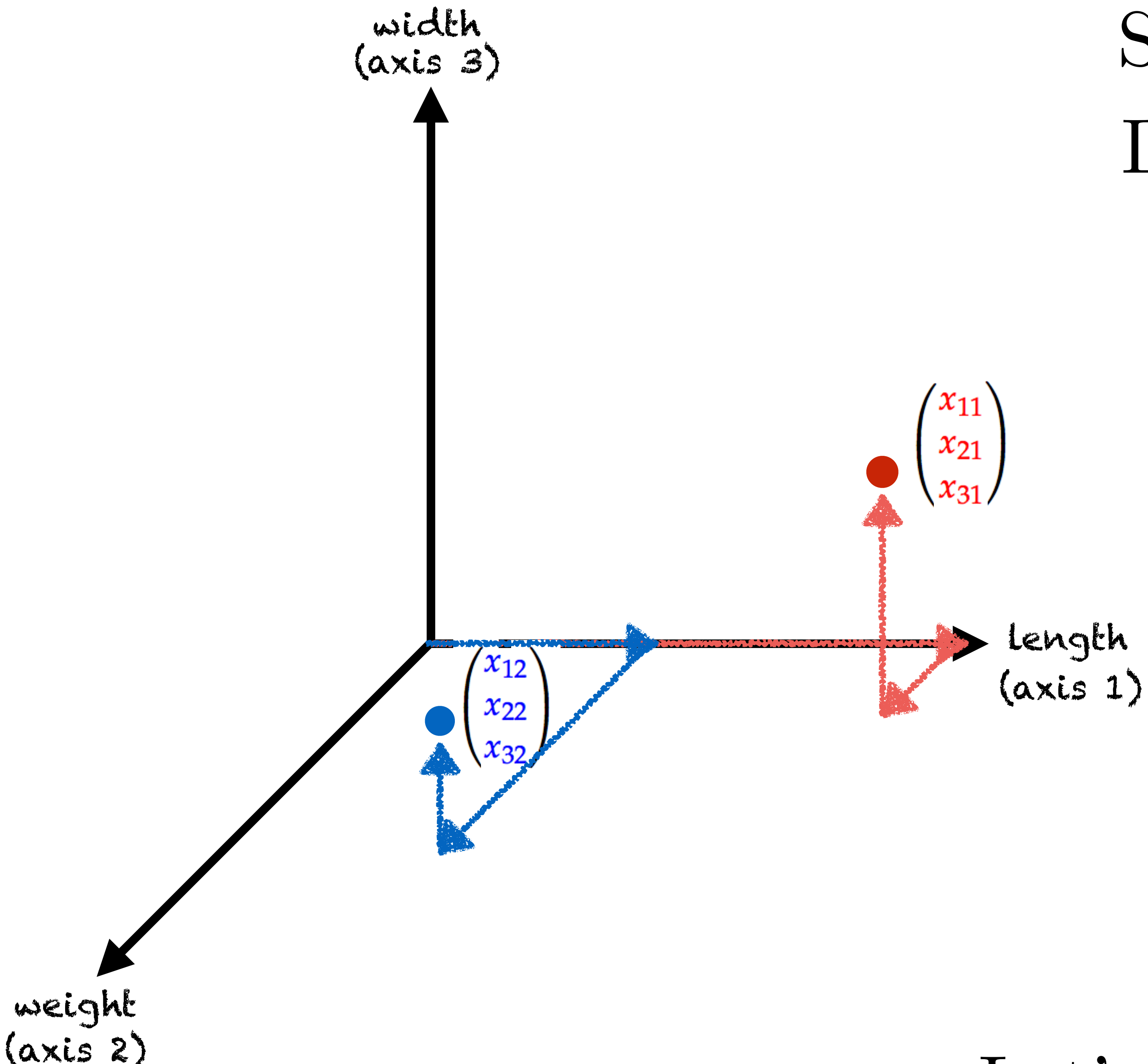


Compared to red point, blue point has:

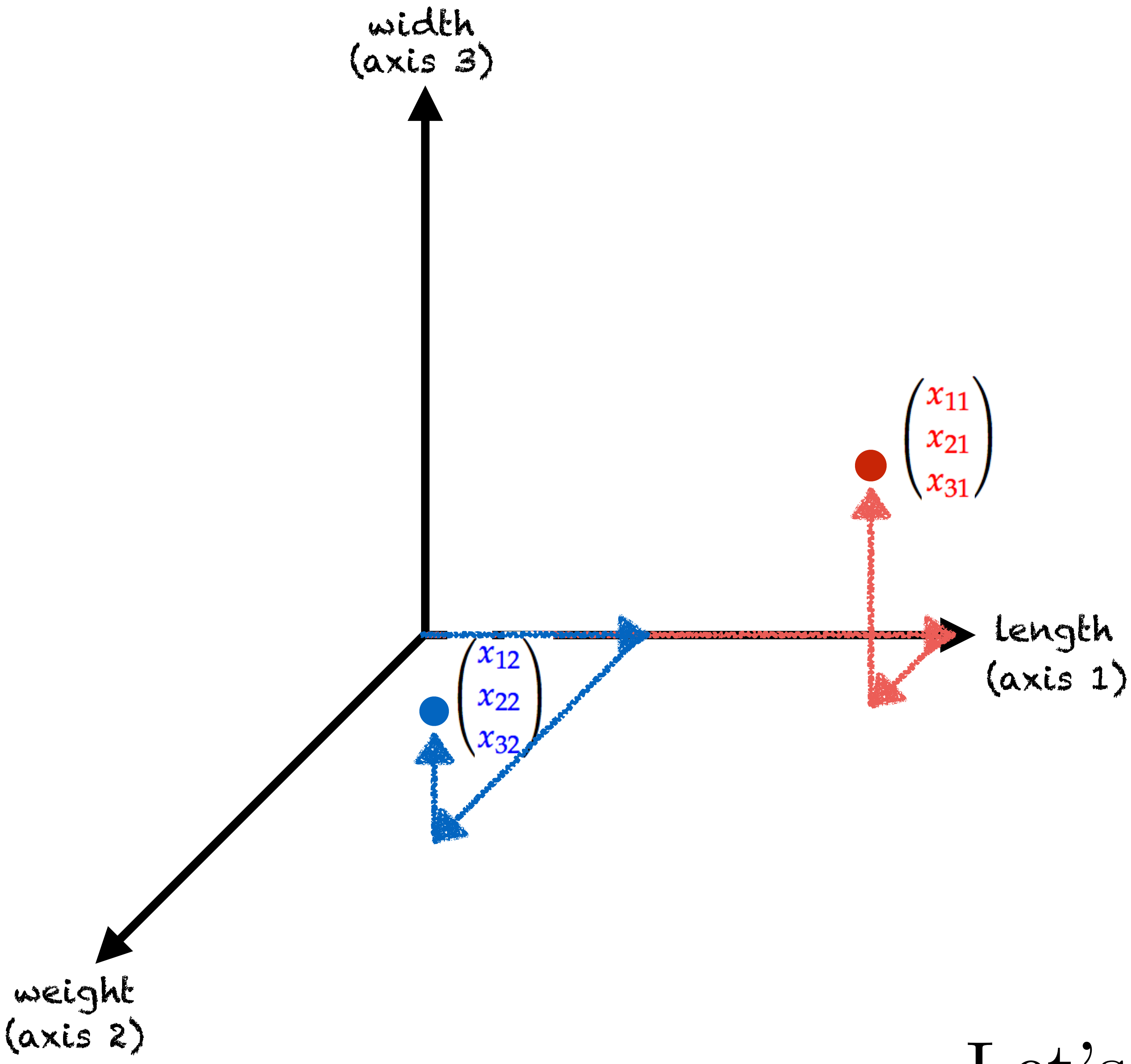
Smaller length

Smaller width

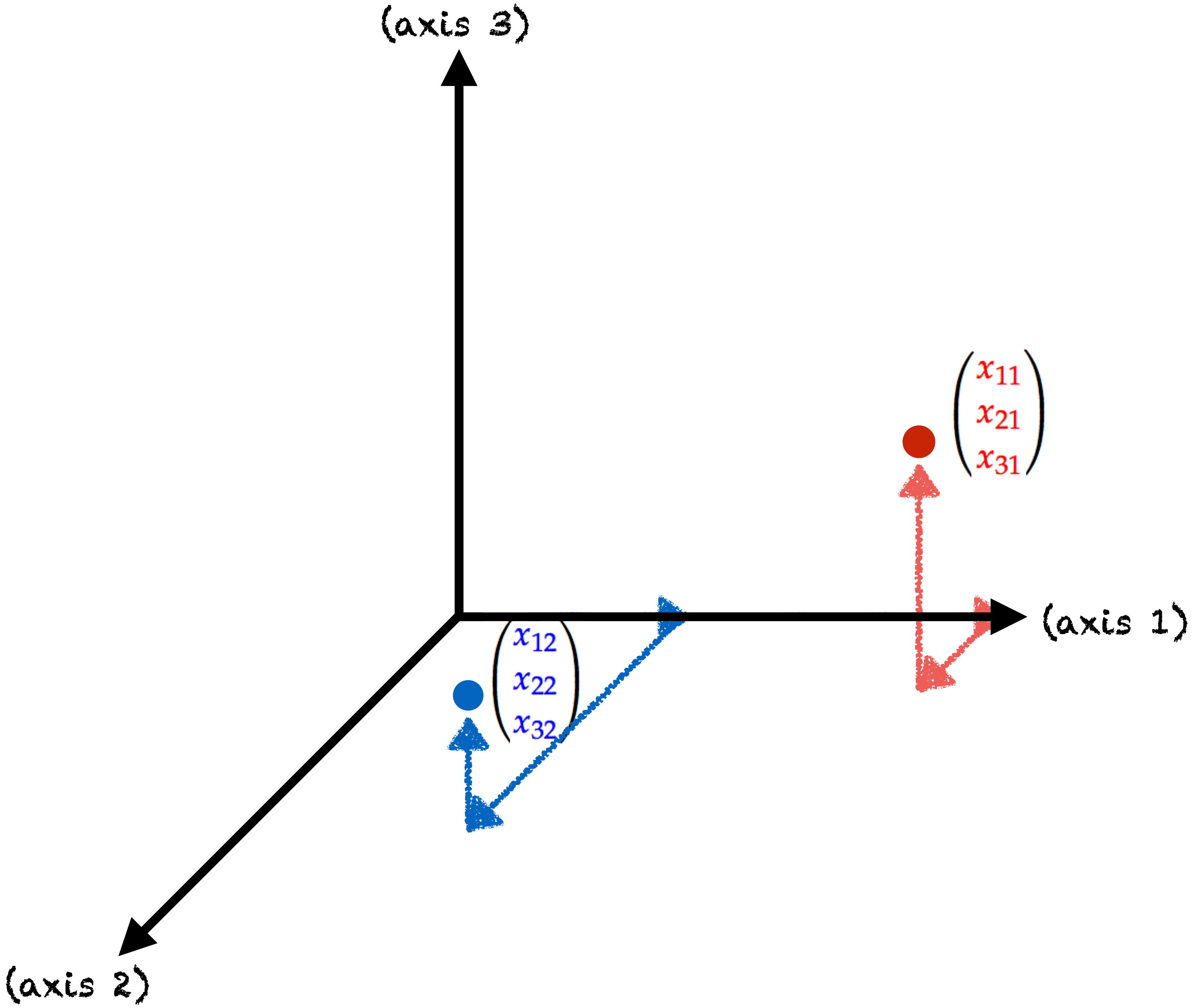
Larger weight



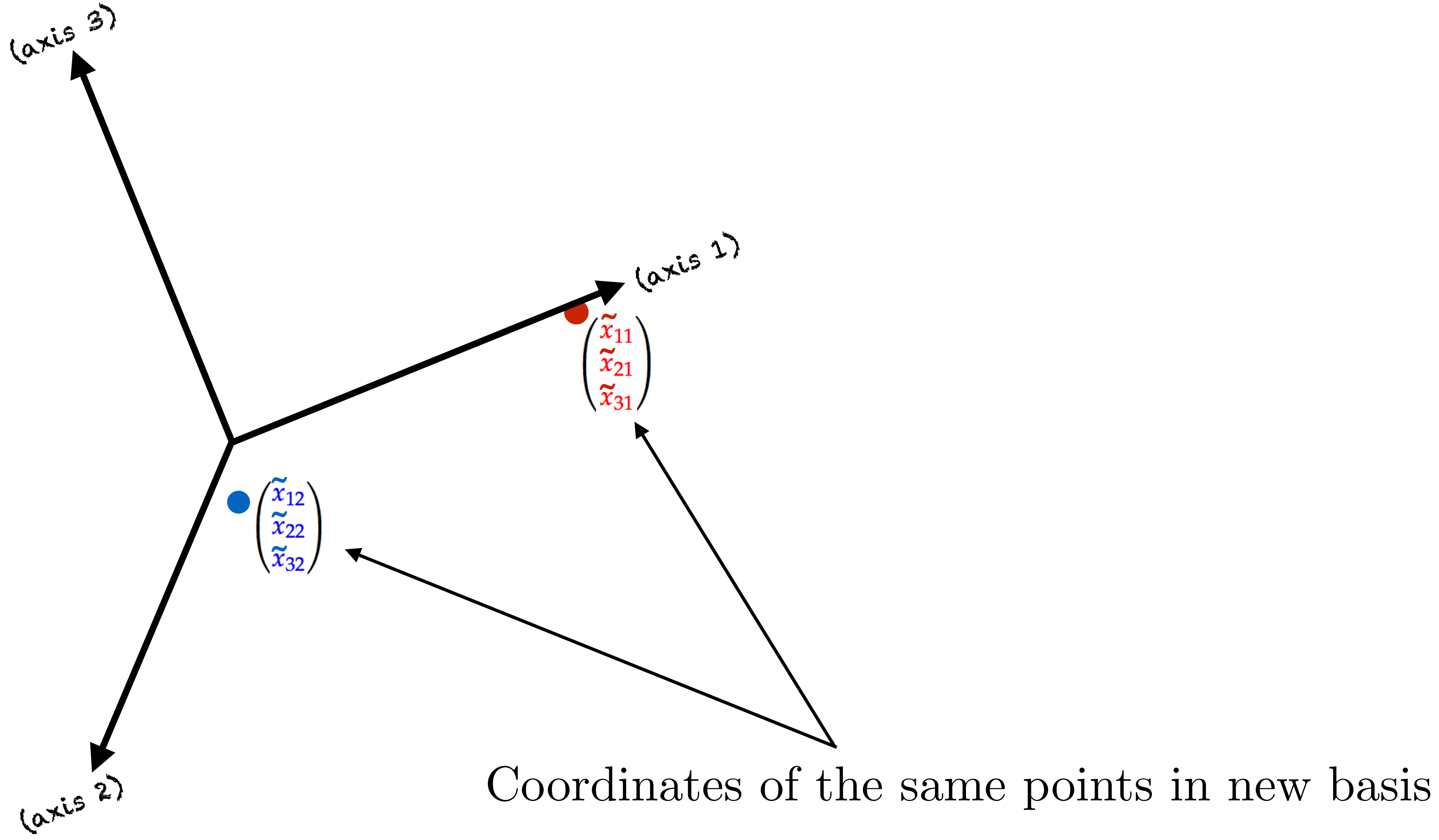
Let's change the basis...

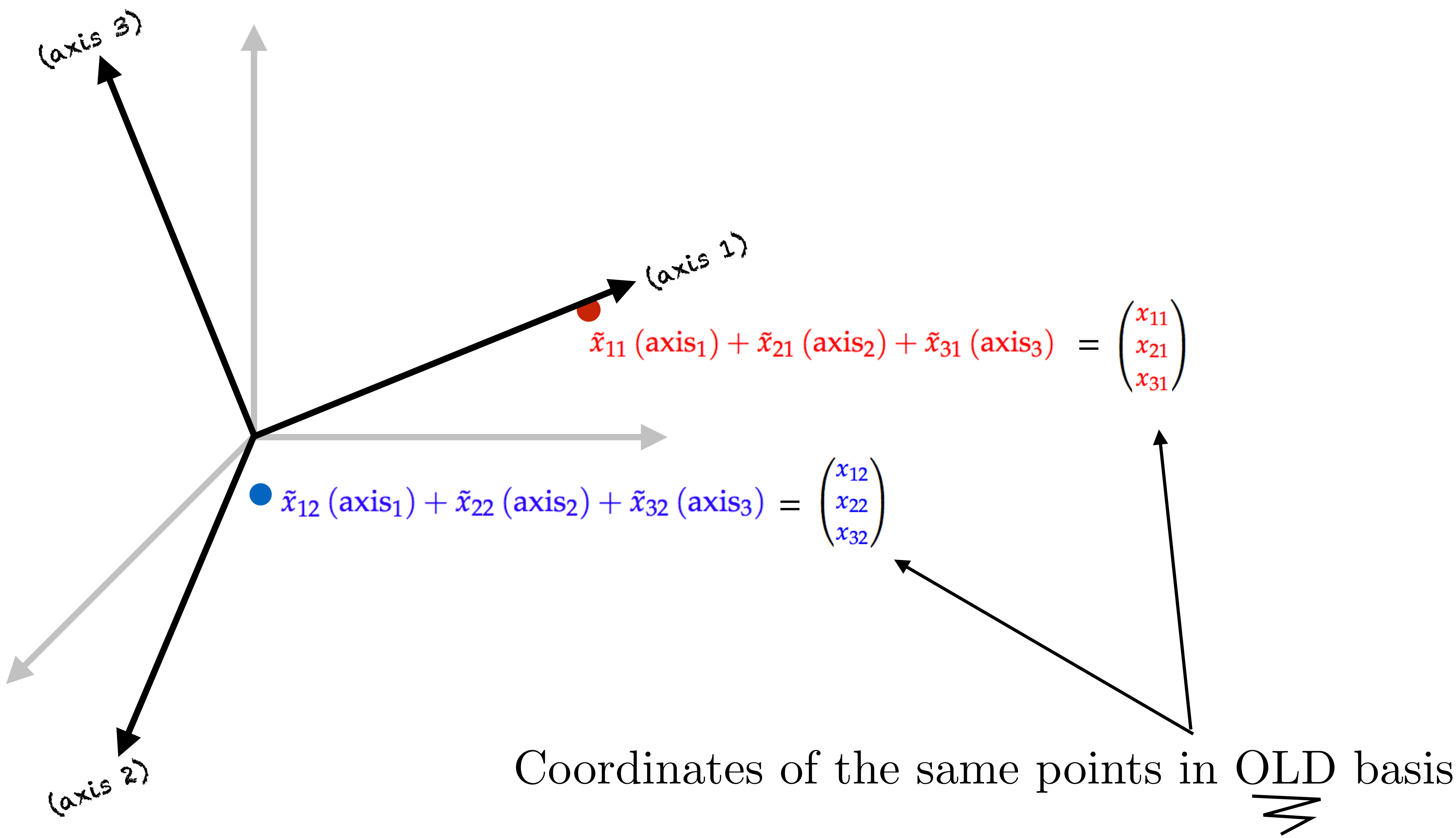


Let's change the basis...

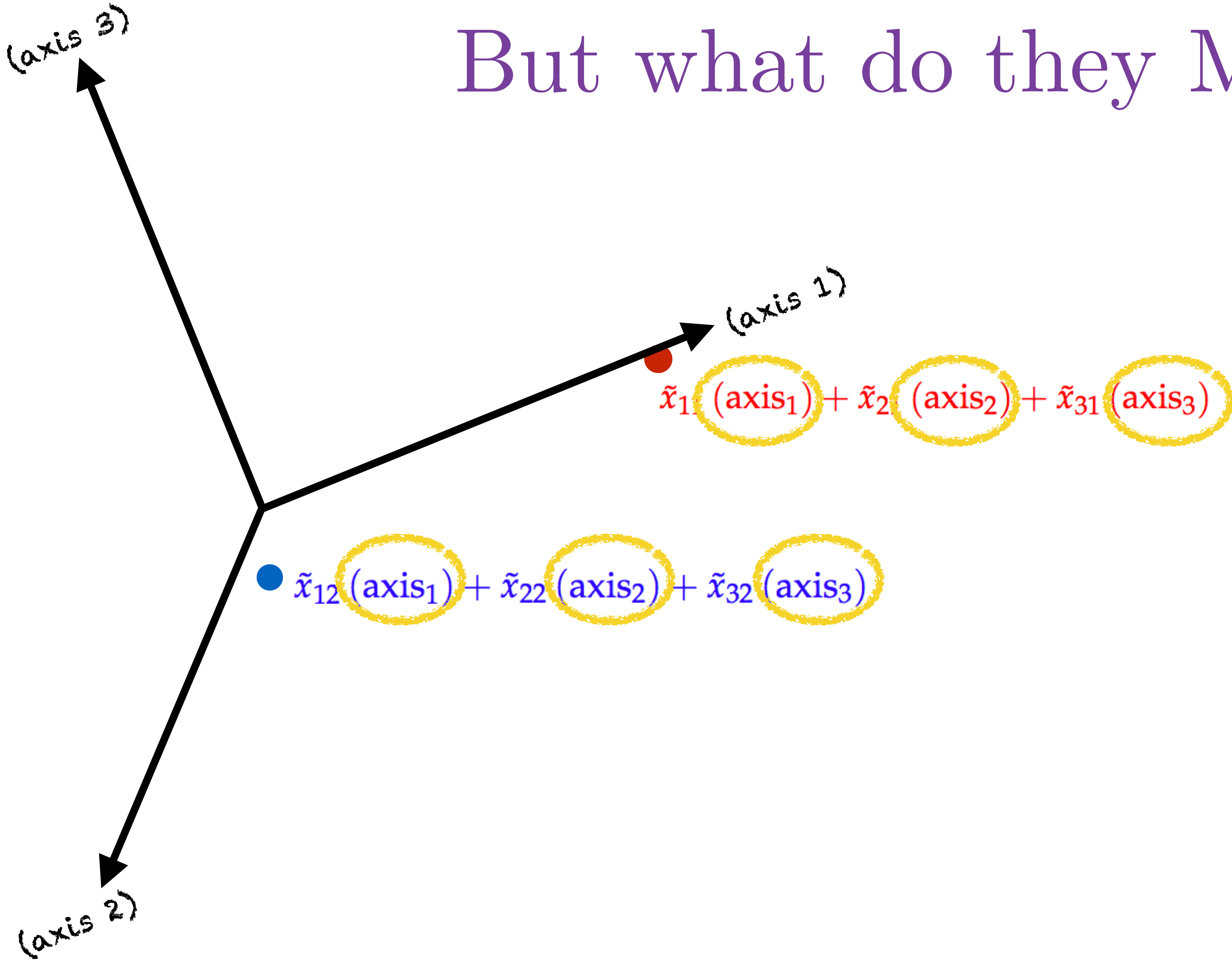


Let's change the basis...


(axis 3)
(axis 1)
(axis 2)
$\begin{pmatrix}\tilde{x}_{11}\\ \tilde{x}_{21}\\ \tilde{x}_{31}\end{pmatrix}$
$\begin{pmatrix}\tilde{x}_{12}\\ \tilde{x}_{22}\\ \tilde{x}_{32}\end{pmatrix}$
Coordinates of the same points in new basis


(axis 3)
(axis 1)
(axis 2)
$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$
$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$
Coordinates of the same points in OLD basis

These are the new basis vectors... But what do they MEAN?

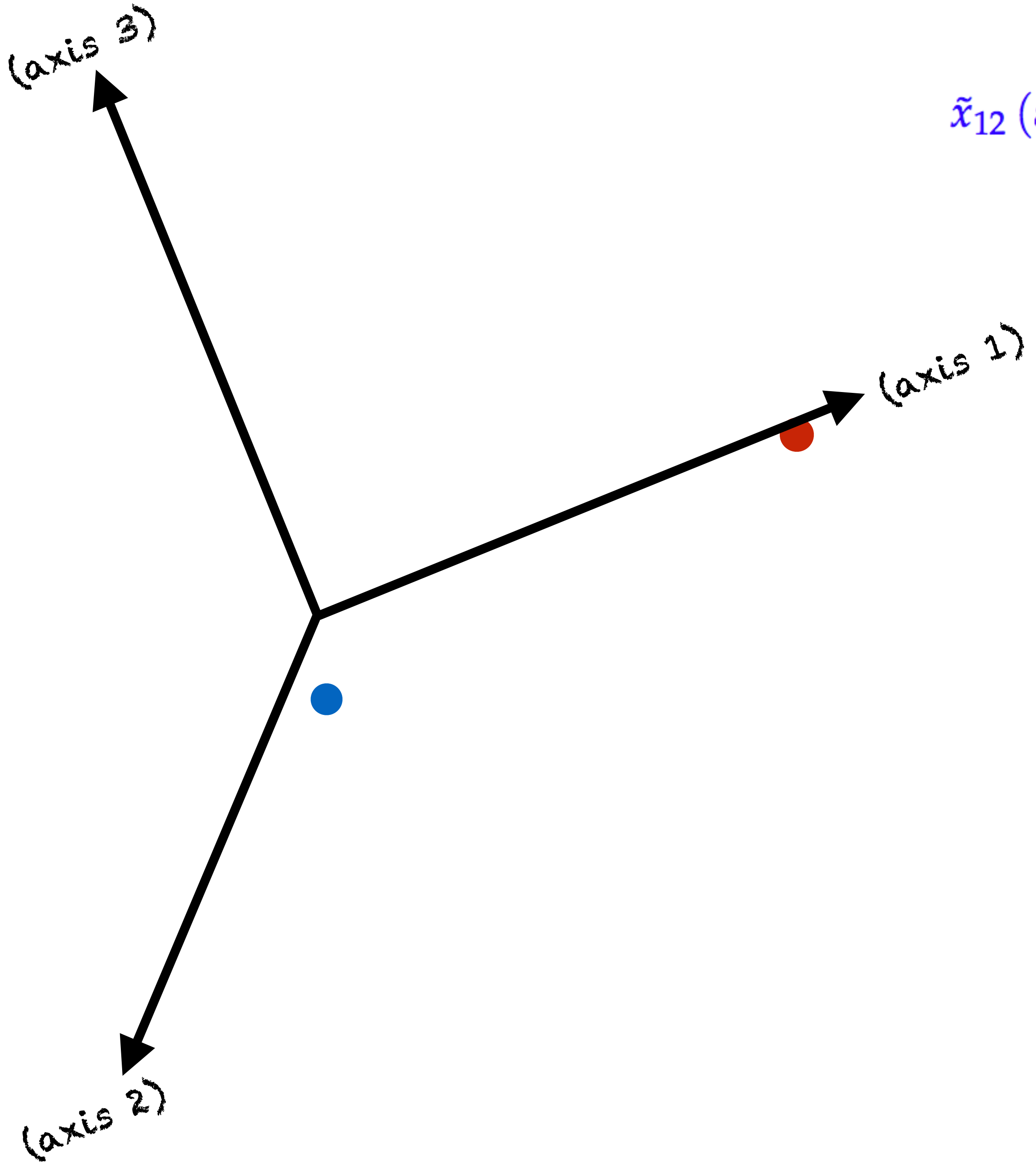


These are the new basis vectors...

But what do they MEAN?

$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

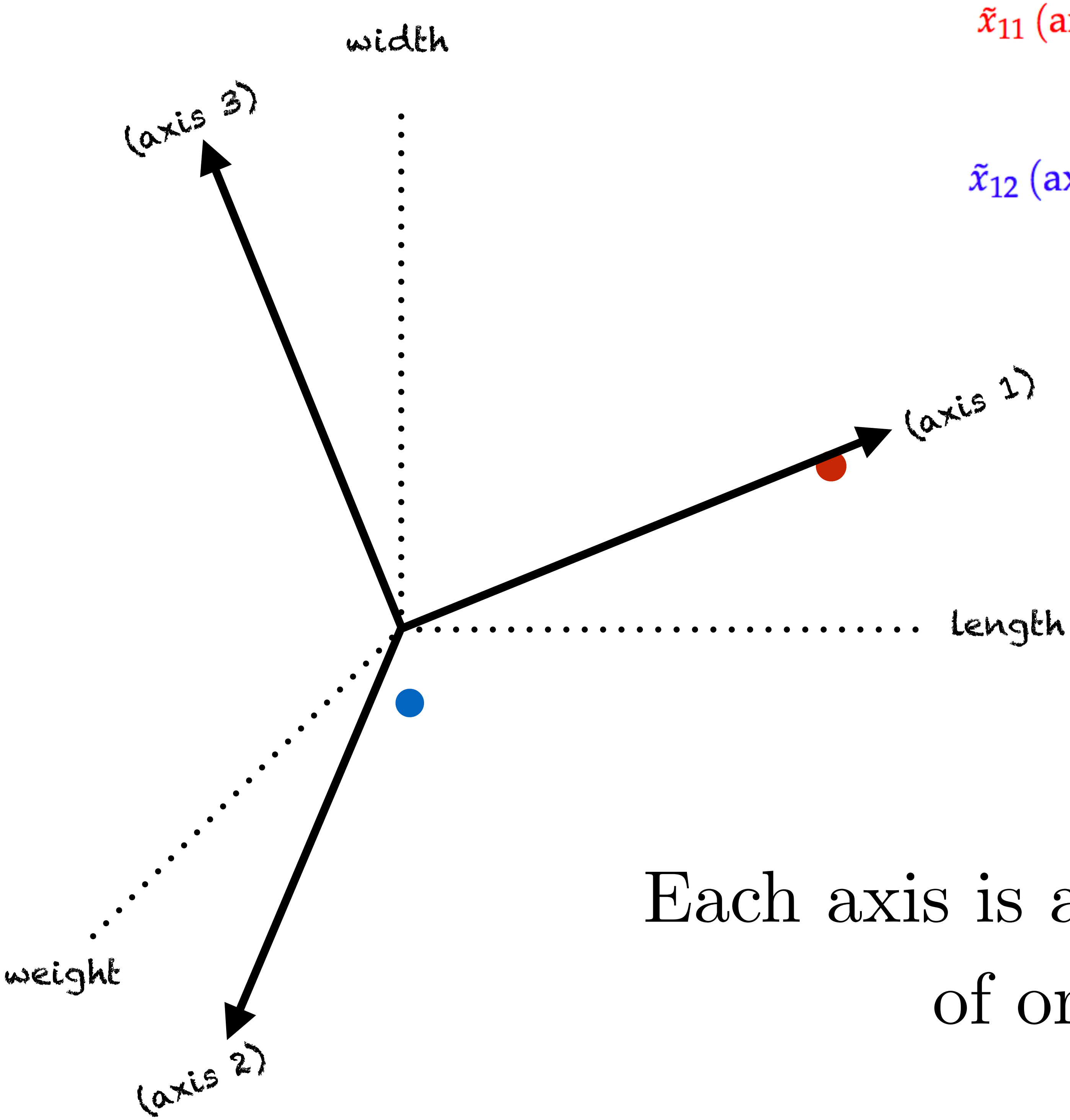
$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$



These are the new basis vectors... But what do they MEAN?

$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$



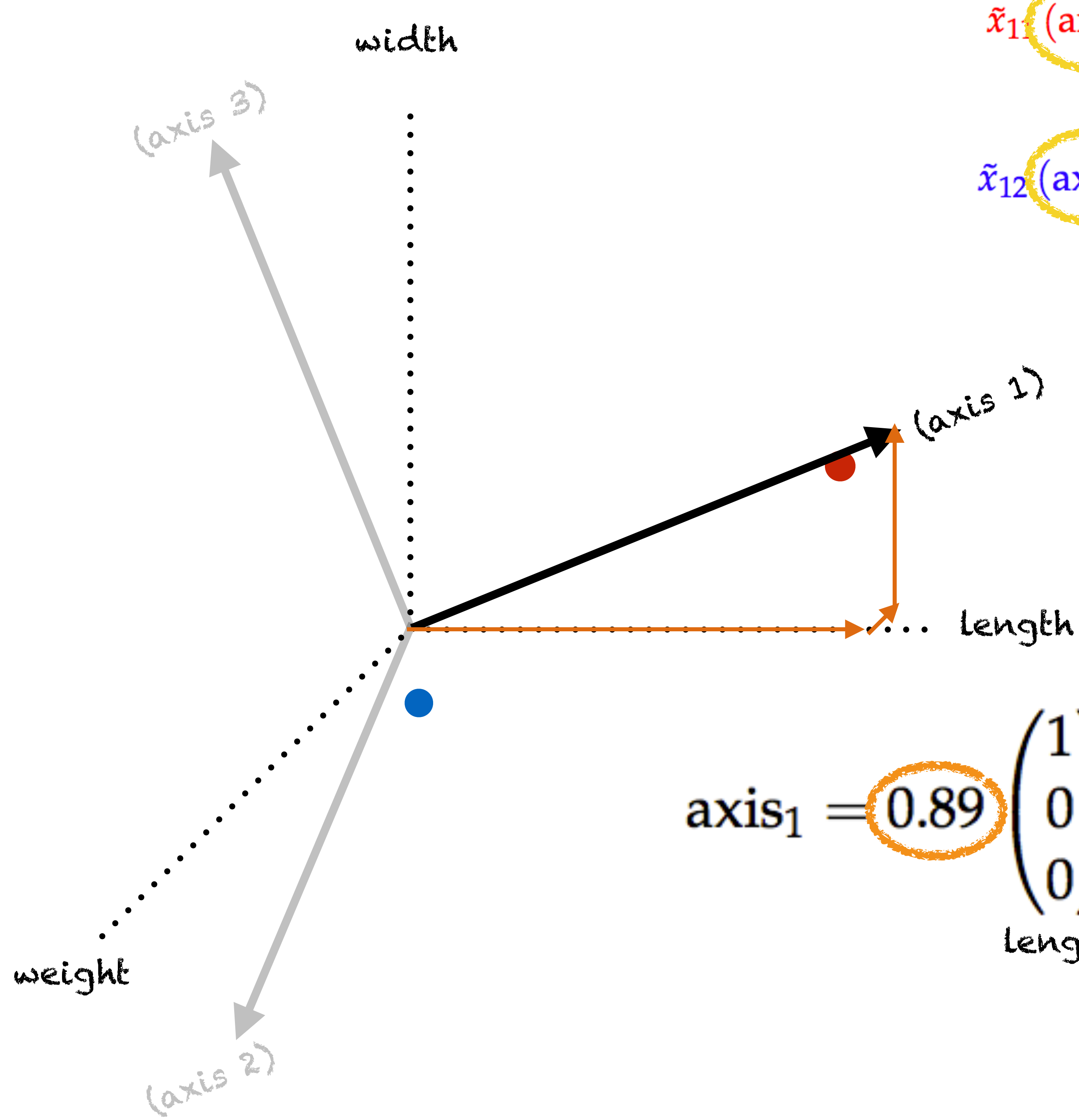
Each axis is a linear combination of original axes

These are the new basis vectors...

But what do they MEAN?

$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

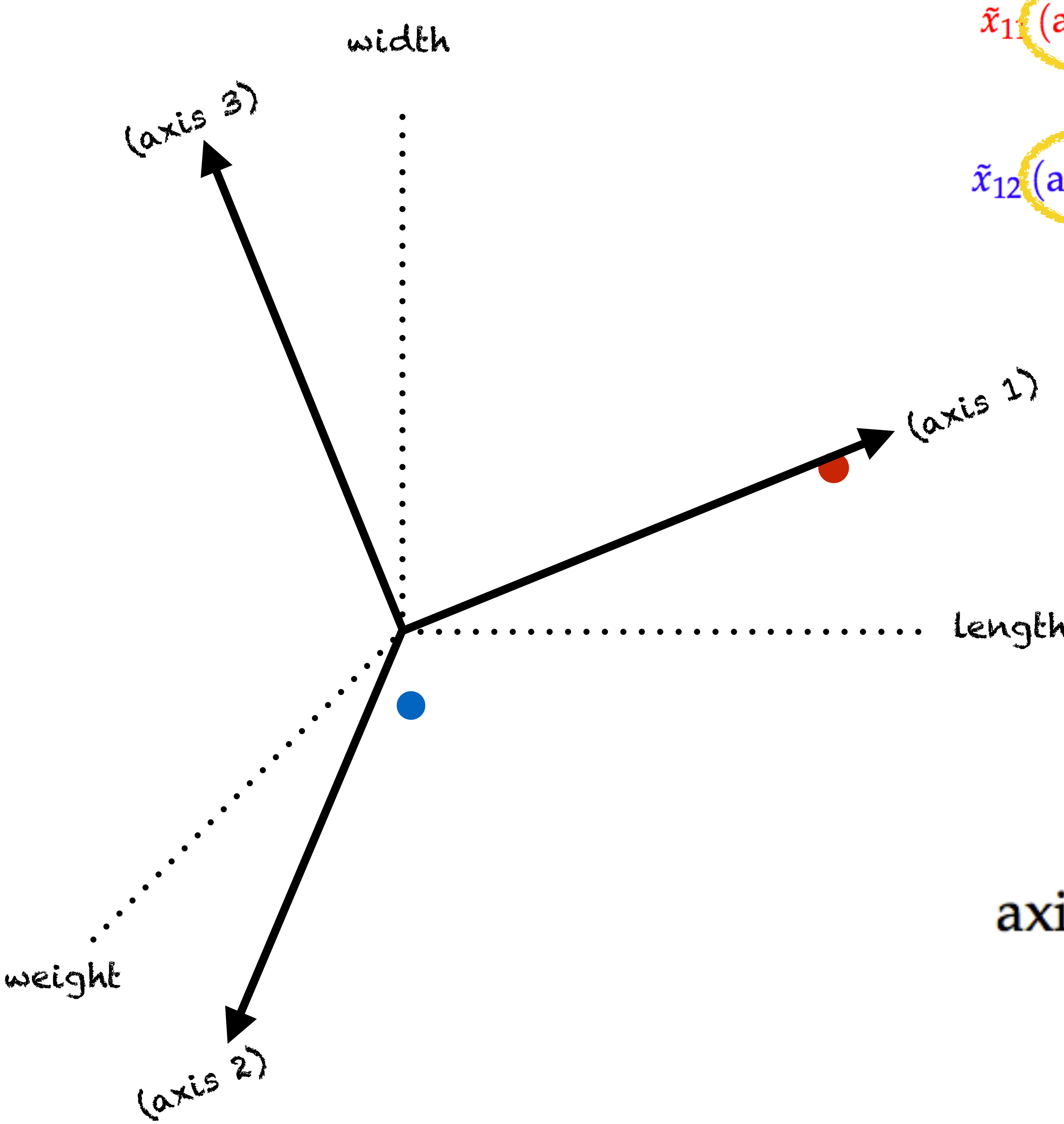


$$\text{axis}_1 = 0.89 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 0.09 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0.45 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

length weight width

These are the new basis vectors...

But what do they MEAN?



$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

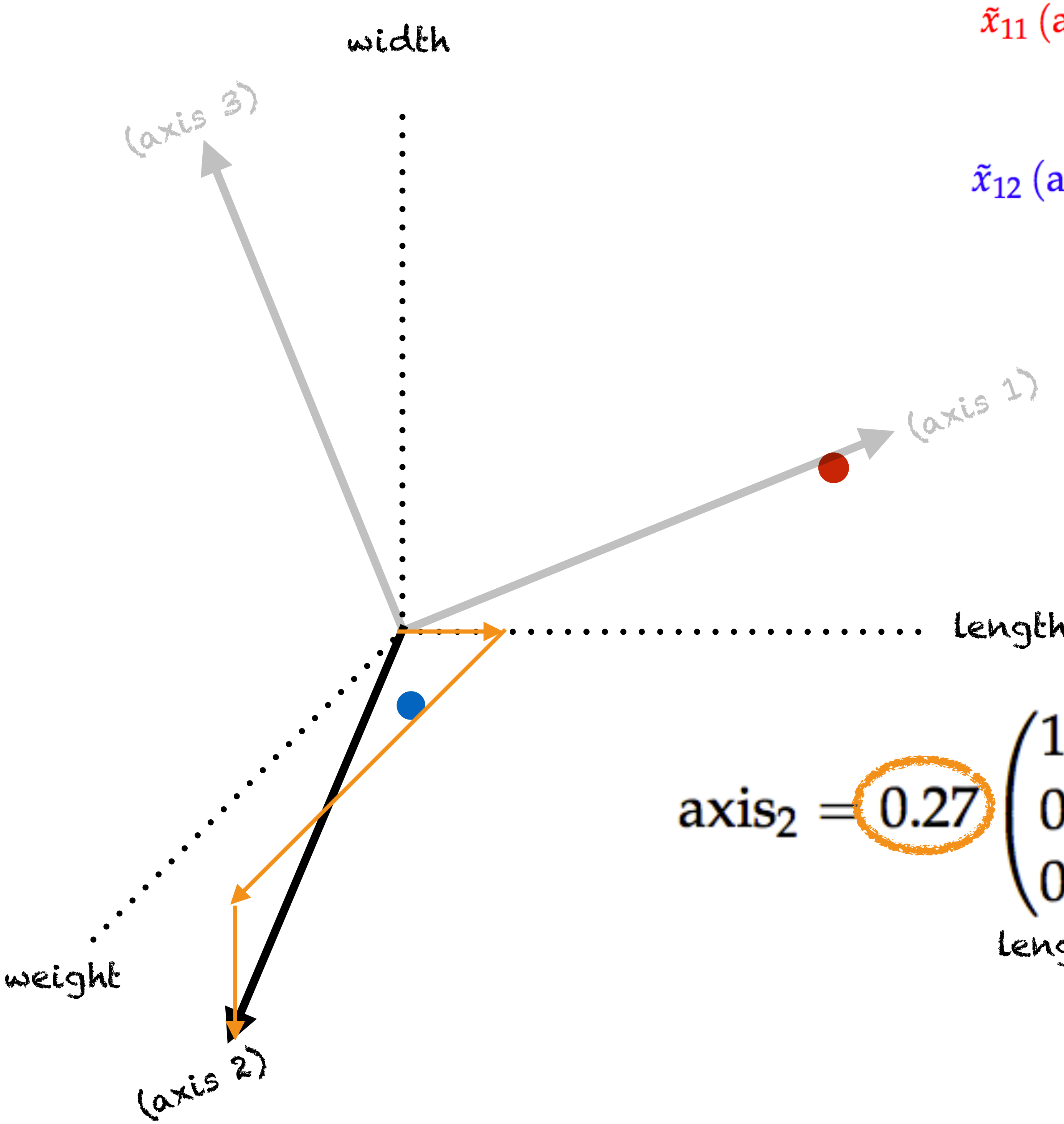
represents ...size?

These are the new basis vectors...

But what do they MEAN?

$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$



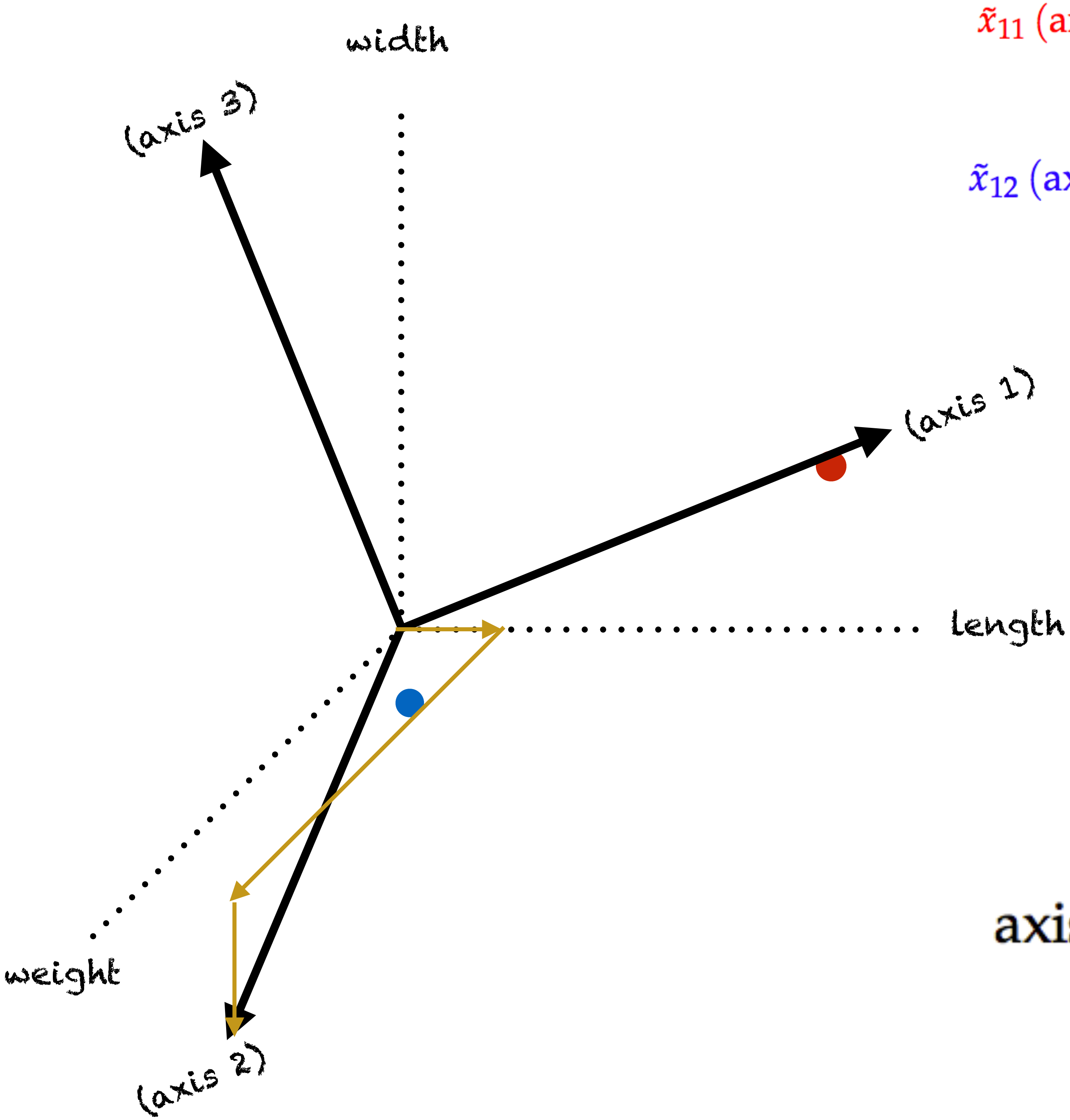
$$\text{axis}_2 = 0.27\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0.89\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 0.35\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

length weight width

These are the new basis vectors... But what do they MEAN?

$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$



$$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

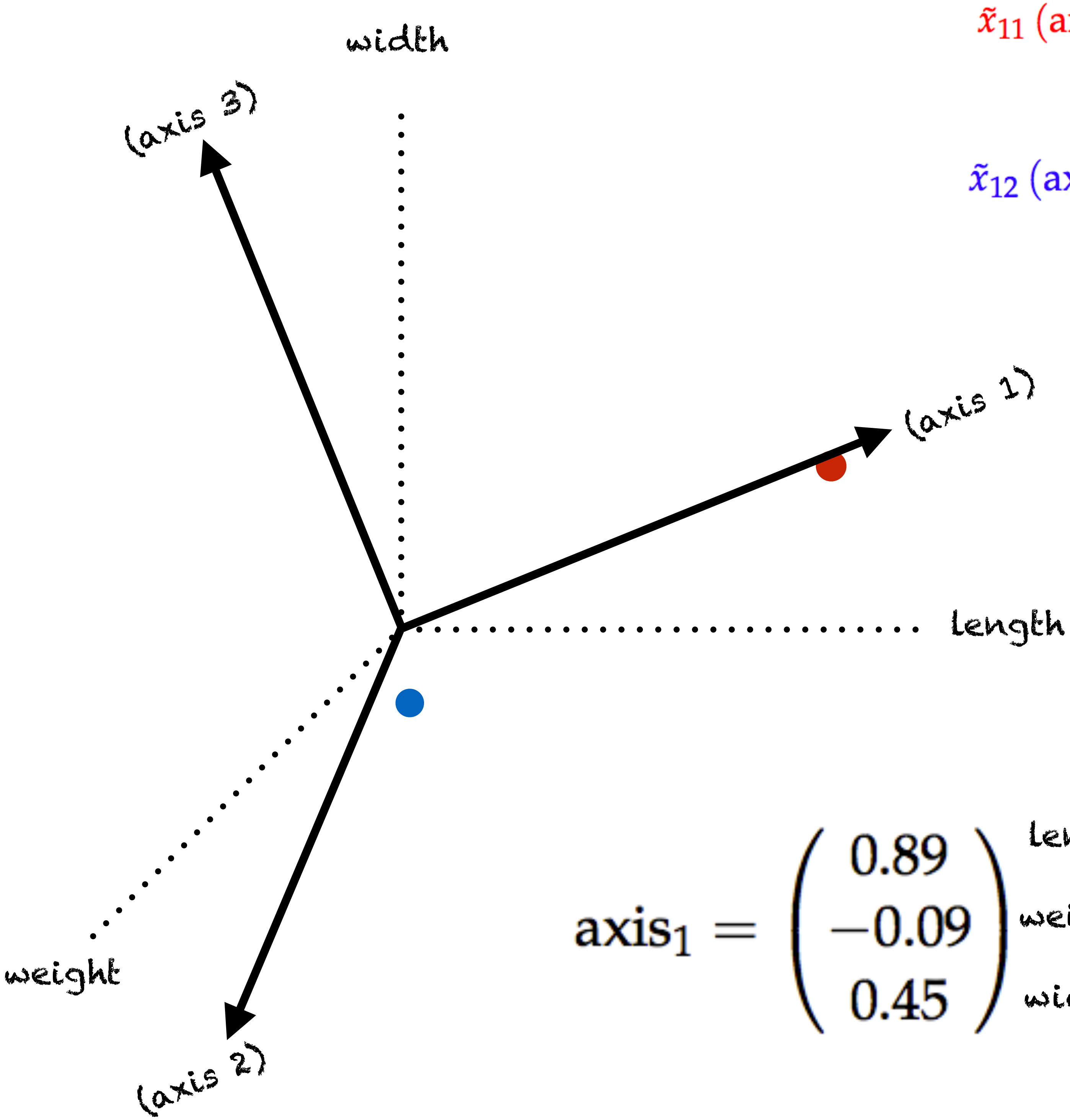
represents ...weight?

These are the new basis vectors... But what do they MEAN?

$$\color{red}{\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}}$$

$$\color{blue}{\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}}$$

=0



$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

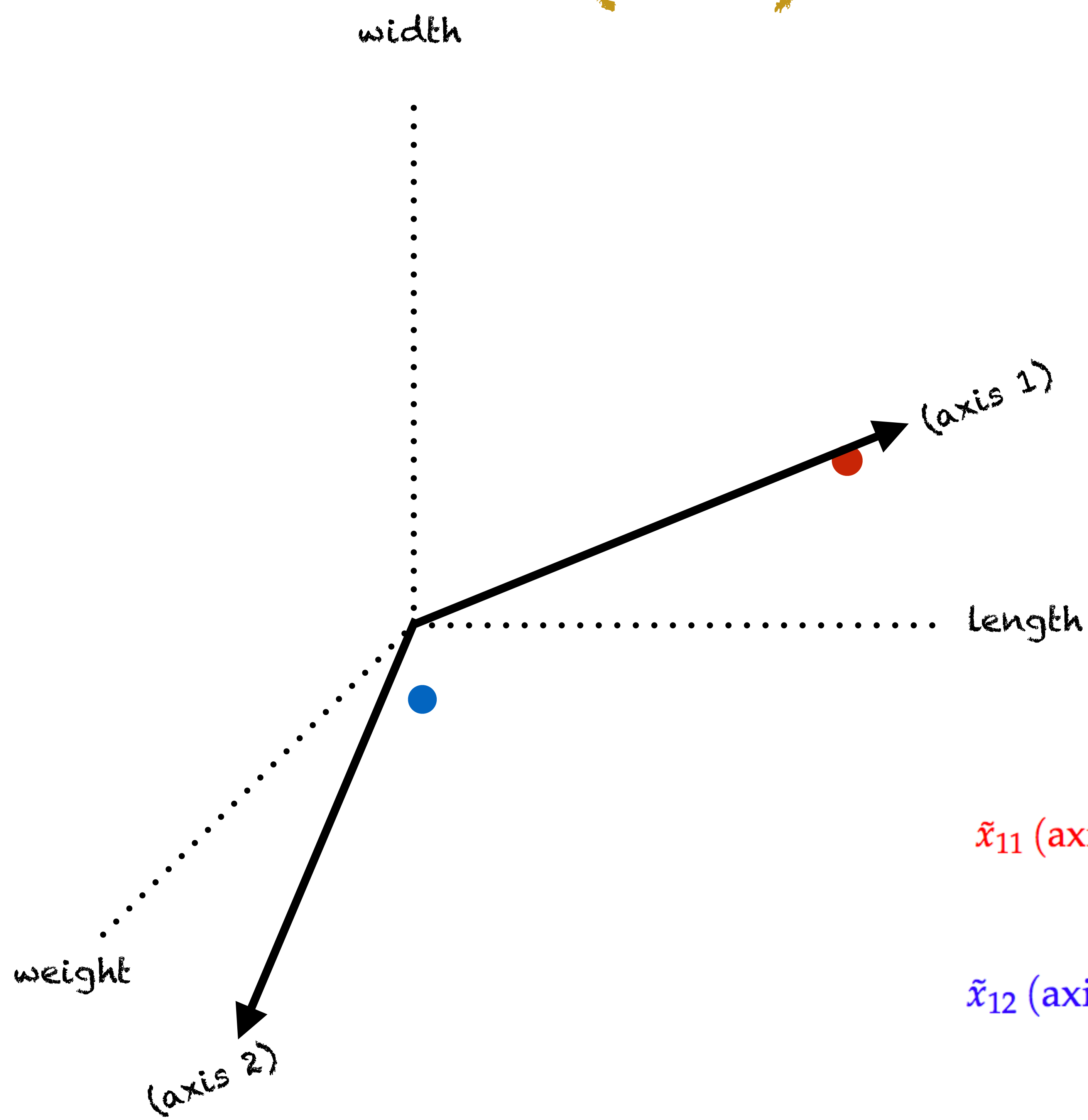
Let's ignore axis 3 for now...

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(size)

$$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(weight)

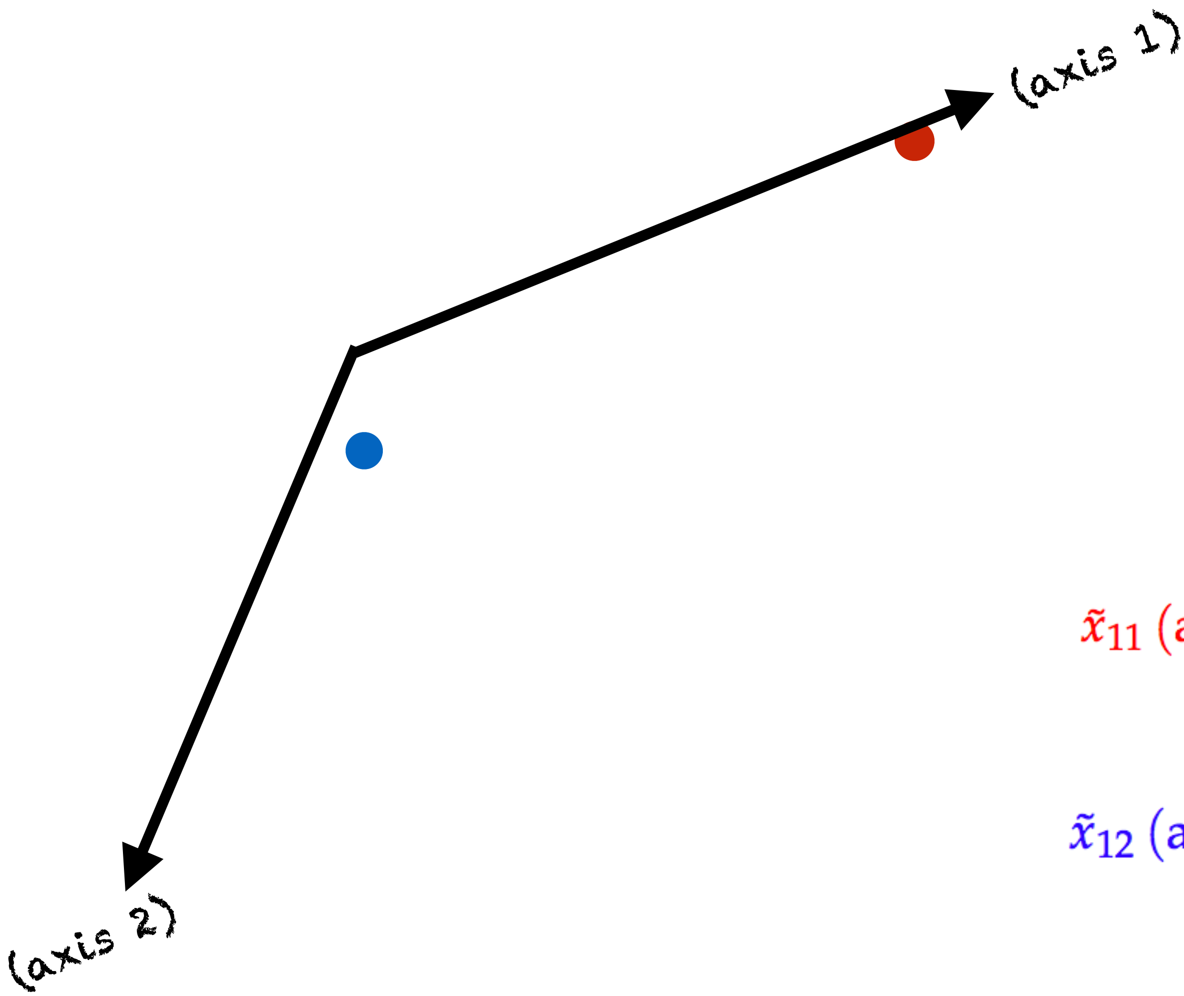


$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(size) (weight)

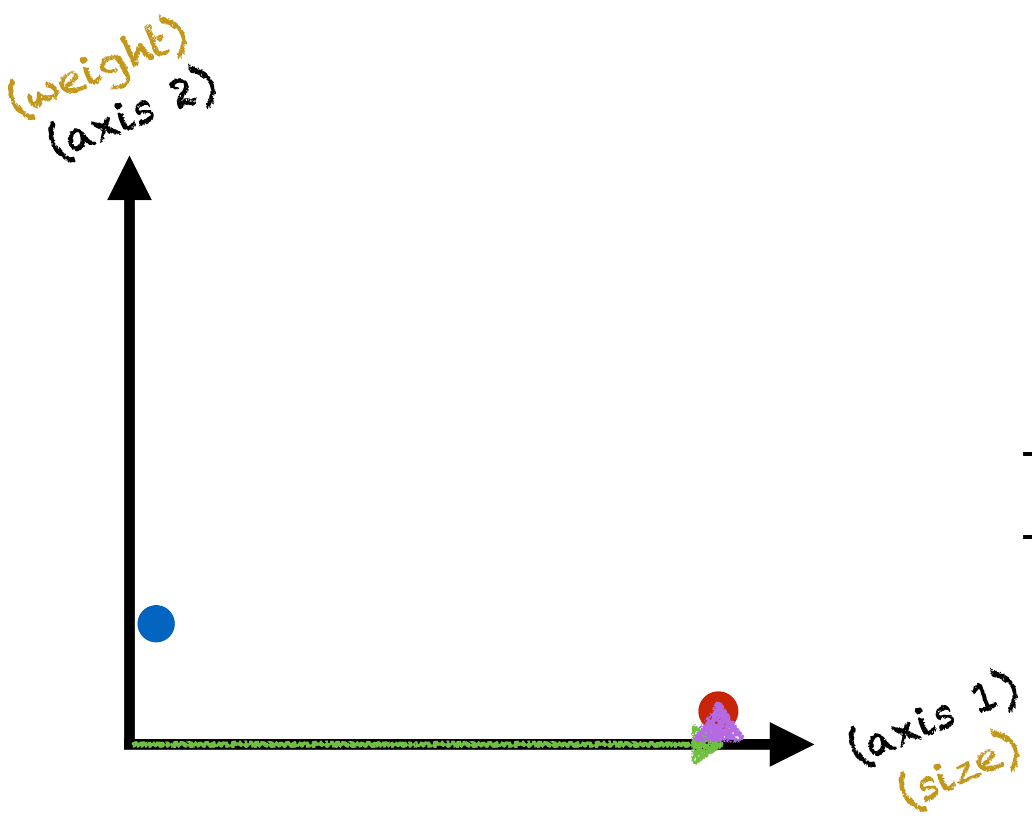


$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(size) (weight)



$$\tilde{x}_{11}(\text{axis}_1) + \tilde{x}_{21}(\text{axis}_2) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

Infer that the red point has:

large size

and

small weight

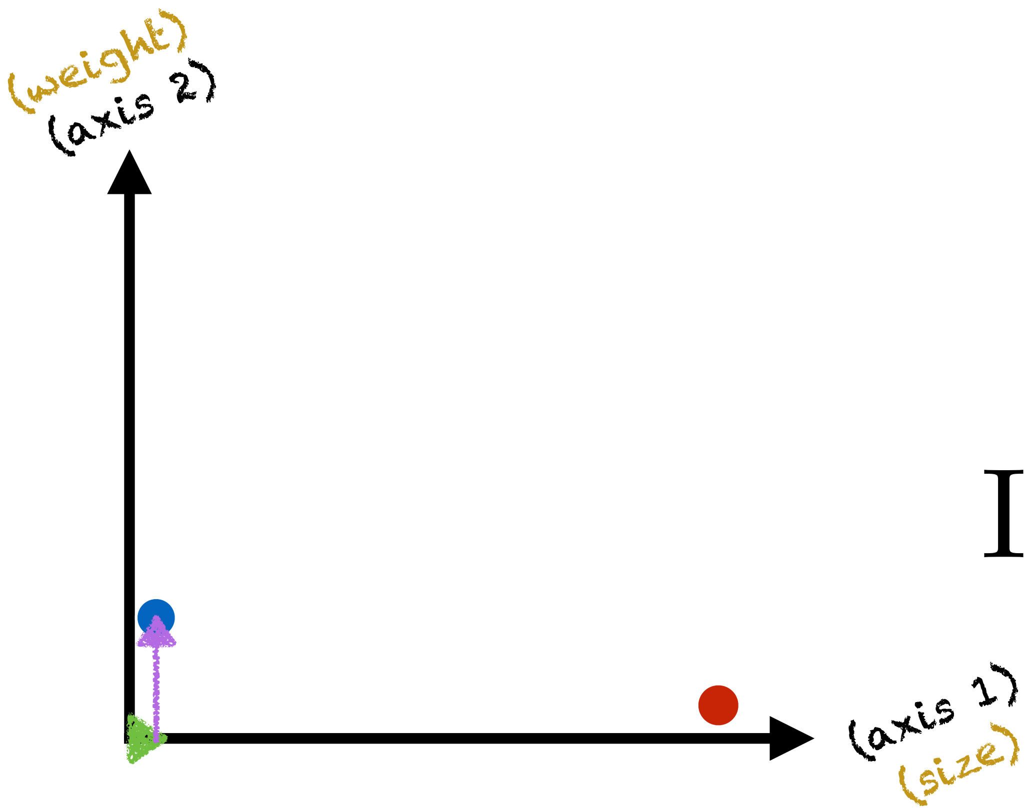
relative to blue point

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(size)

$$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

(weight)



$$\tilde{x}_{12}(\text{axis}_1) + \tilde{x}_{22}(\text{axis}_2) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

Infer that the blue point has:

small size

and

large weight

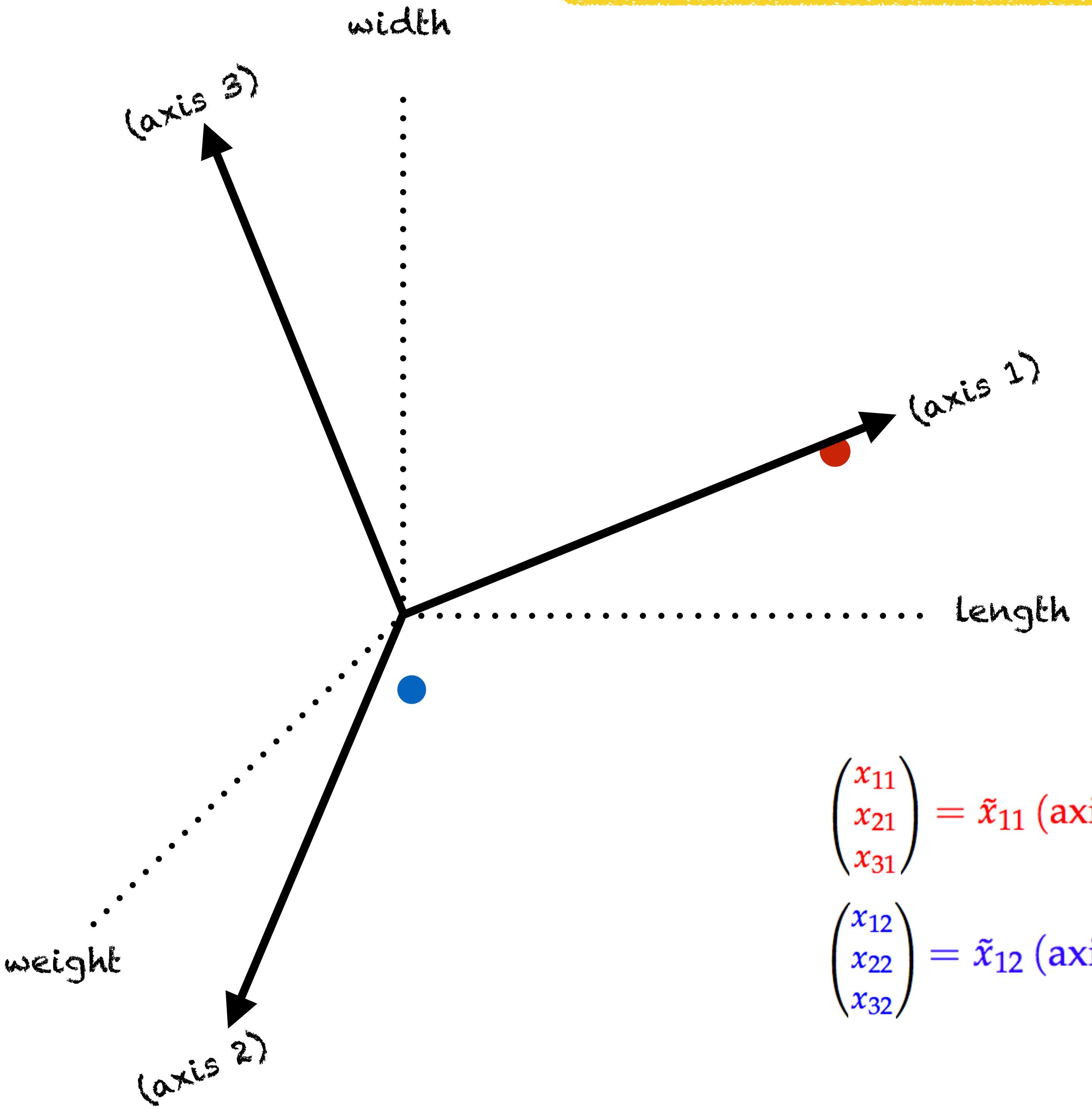
relative to red point

Some Terminology

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

$$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

Factors

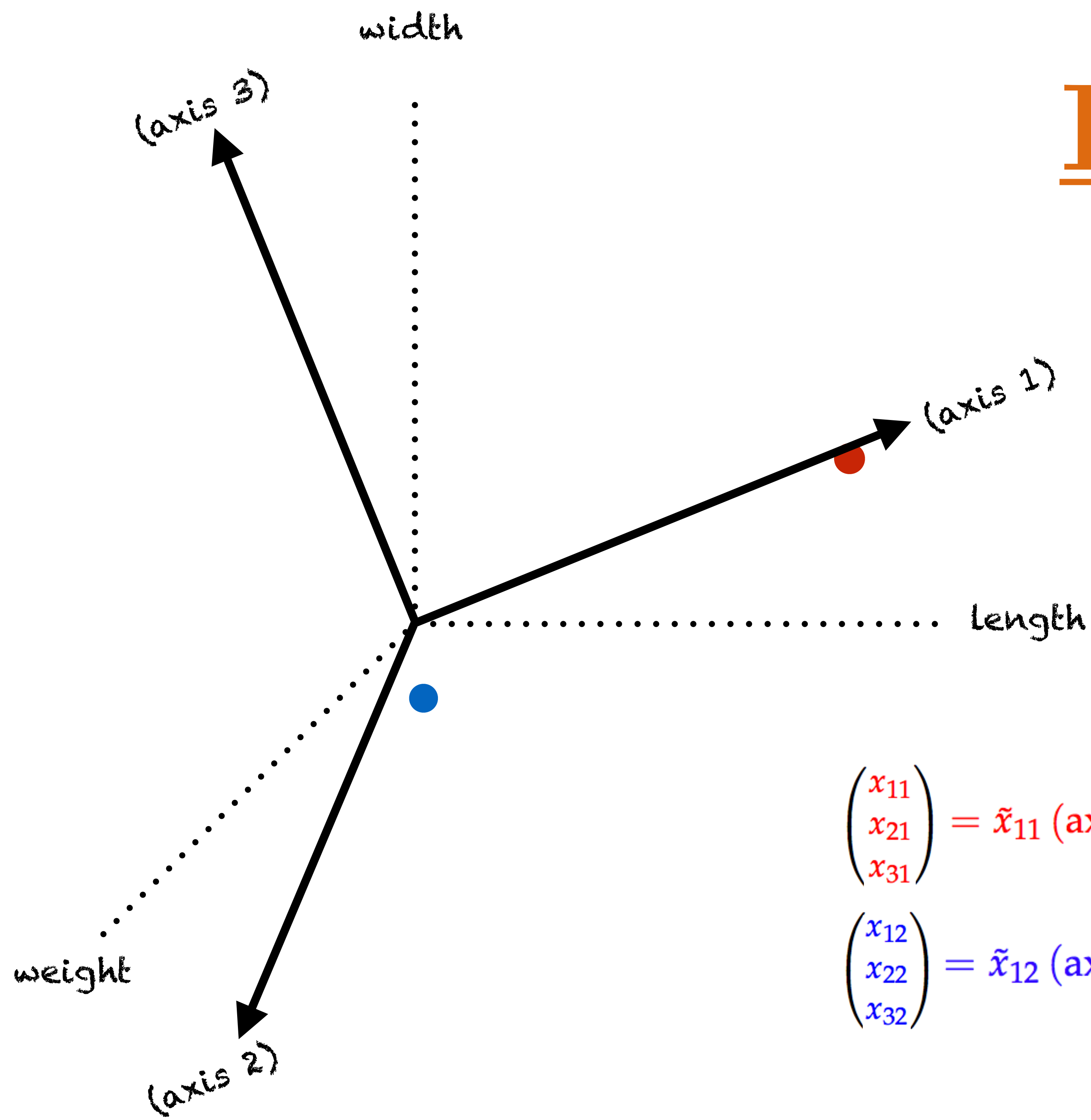


$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3)$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

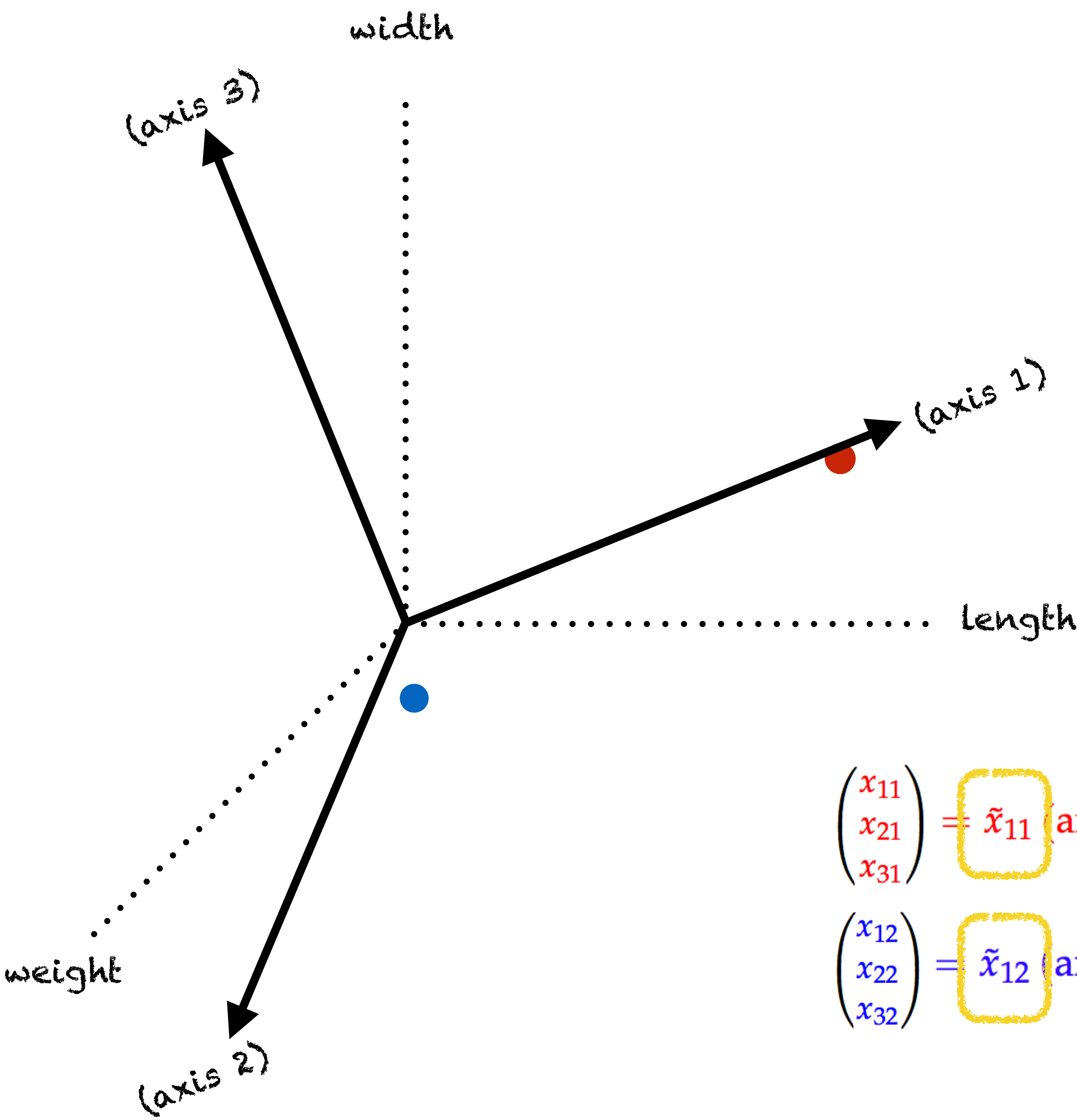
Loadings



$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3)$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



Scores

or

Coordinates

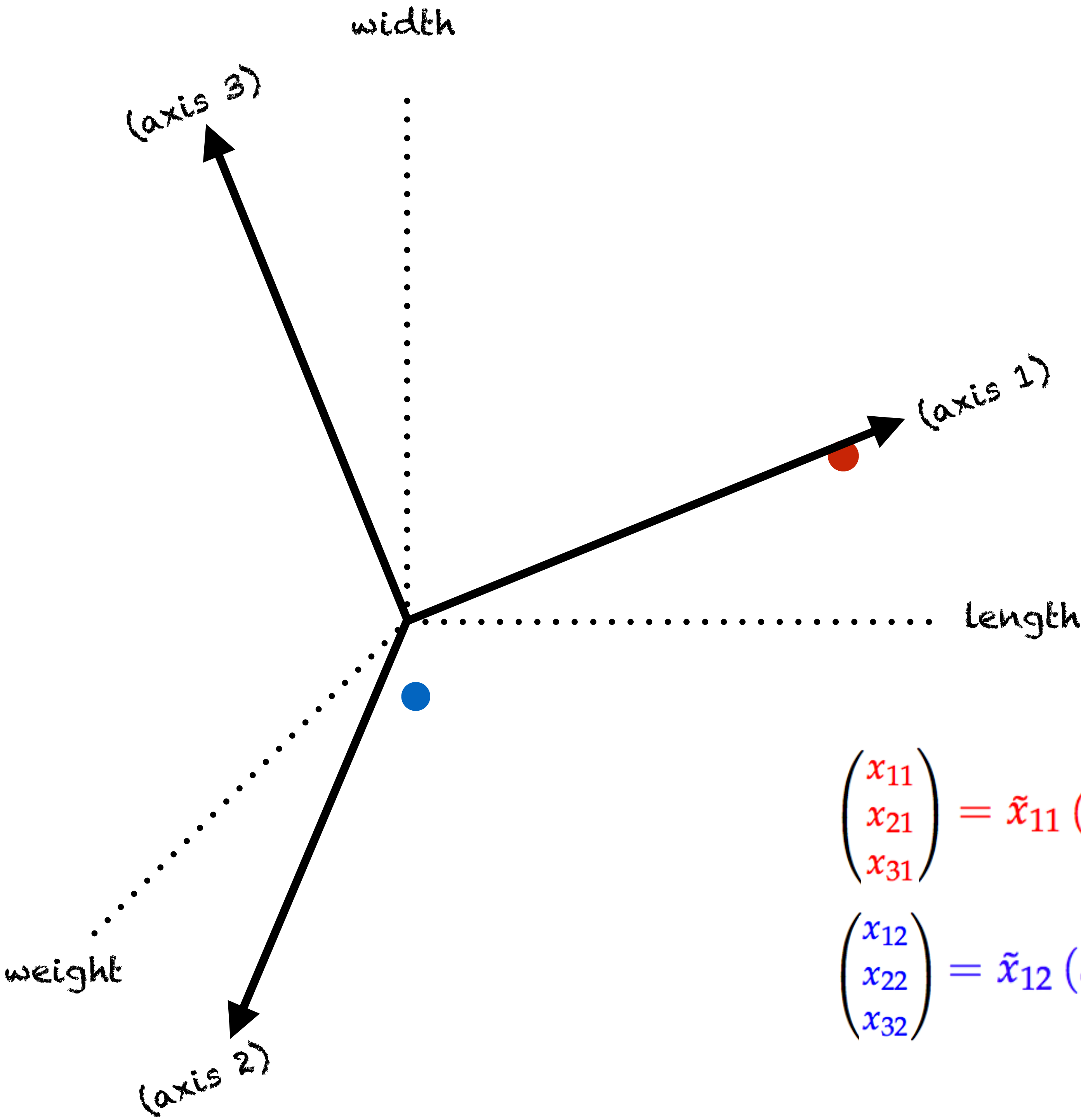
$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}(\text{axis}_1) + \tilde{x}_{21}(\text{axis}_2) + \tilde{x}_{31}(\text{axis}_3)$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}(\text{axis}_1) + \tilde{x}_{22}(\text{axis}_2) + \tilde{x}_{32}(\text{axis}_3)$$

Part 4:
One Step Further

Back to Matrix Multiplication

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

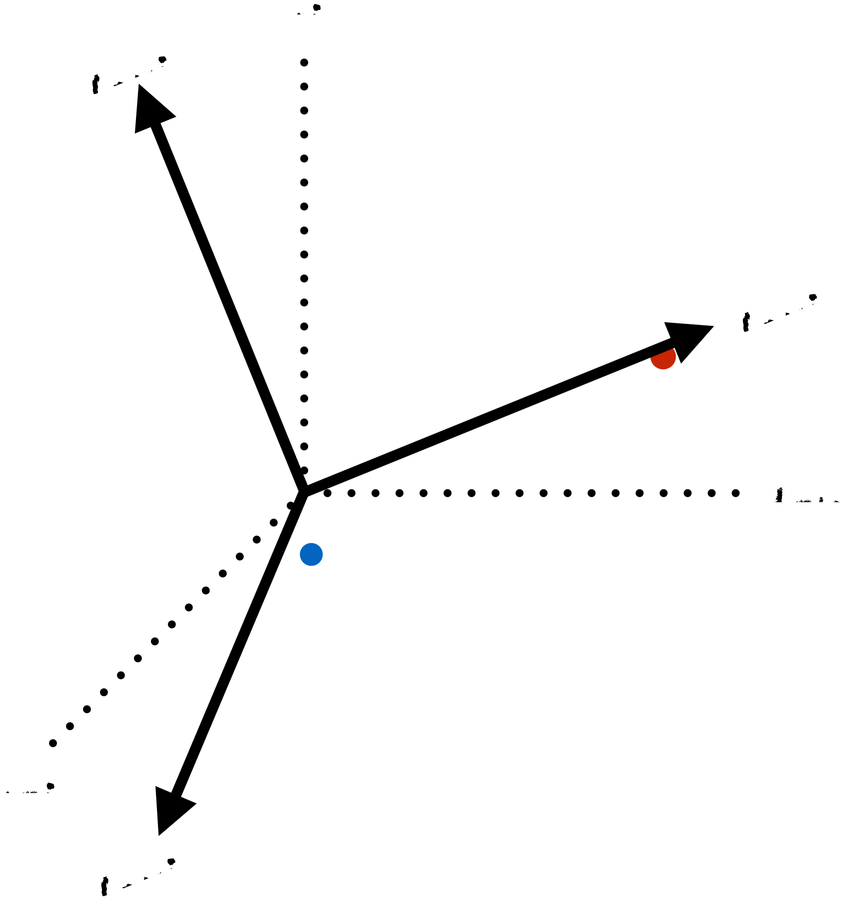


$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3)$$

$$\approx 0$$

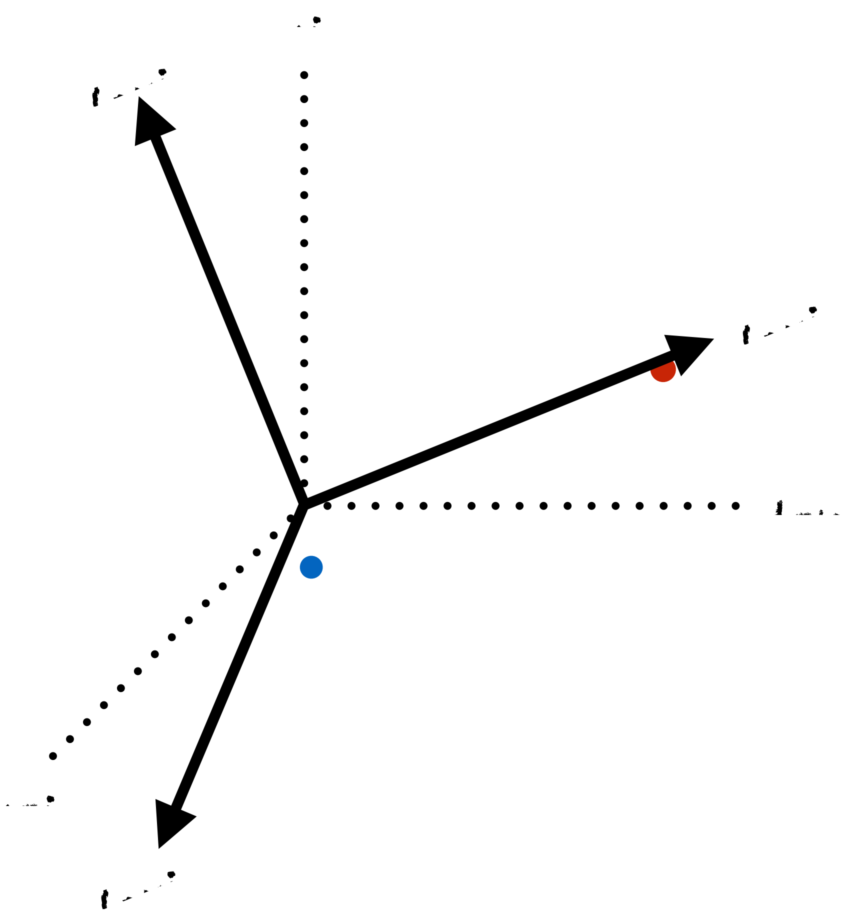
$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} \approx \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2)$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} \approx \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2)$$

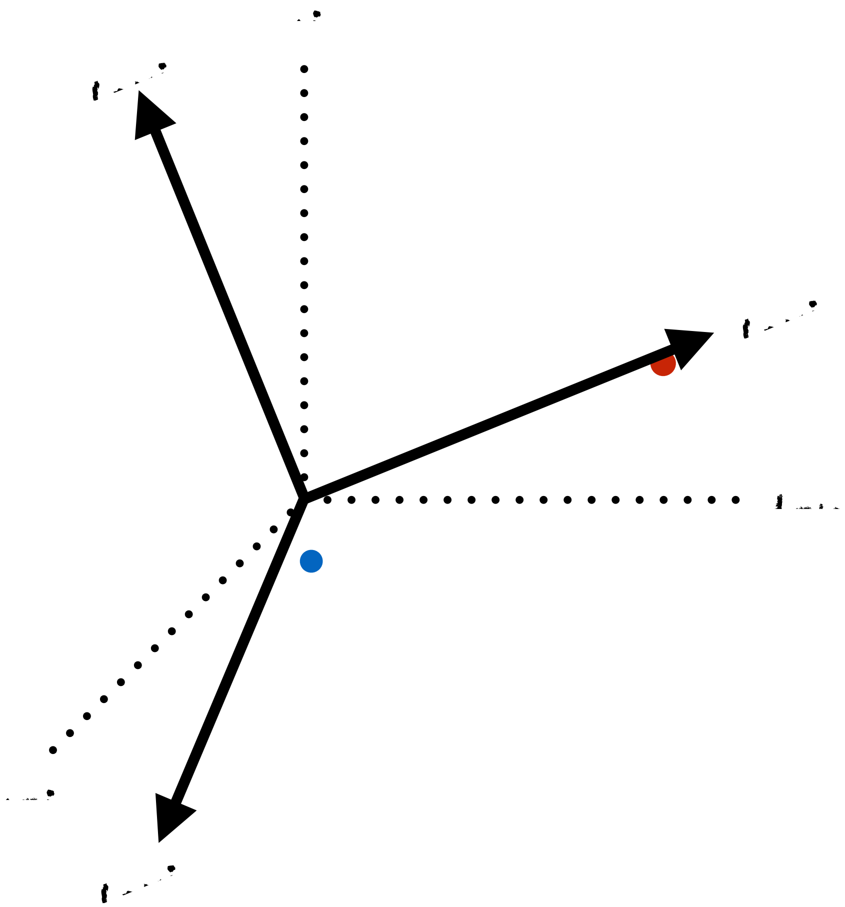
$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} \approx \tilde{x}_{11} \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} + \tilde{x}_{21} \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} \approx \tilde{x}_{12} \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} + \tilde{x}_{22} \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$$

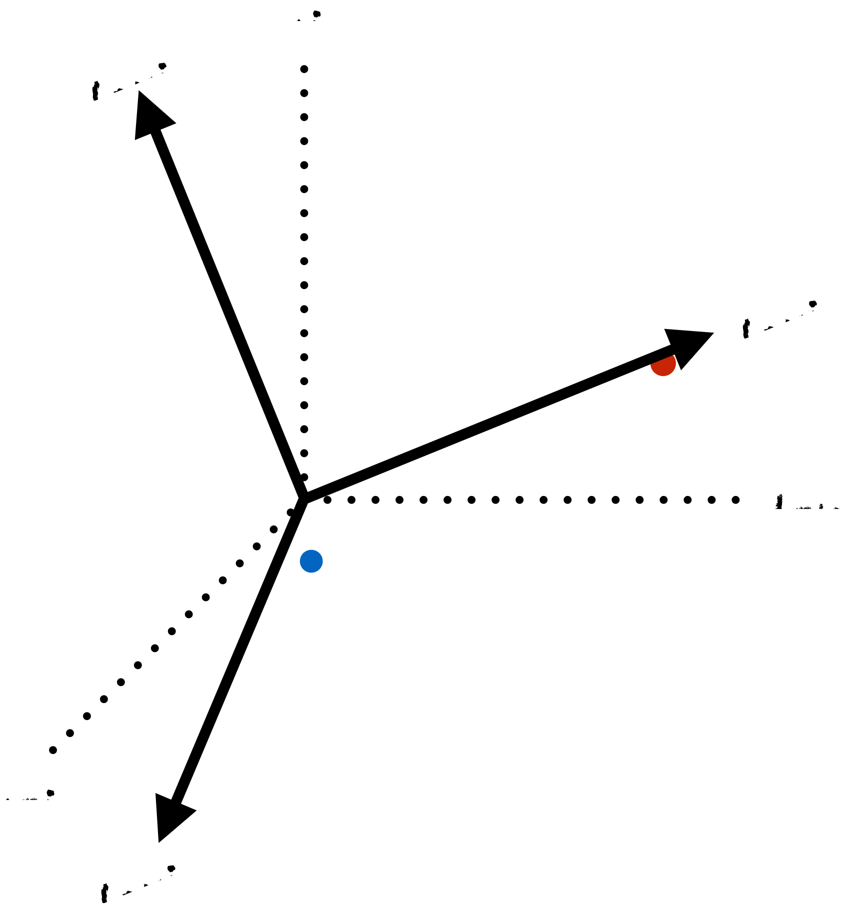
$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{11} \\ \tilde{x}_{21} \end{matrix}$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{12} \\ \tilde{x}_{22} \end{matrix}$$

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



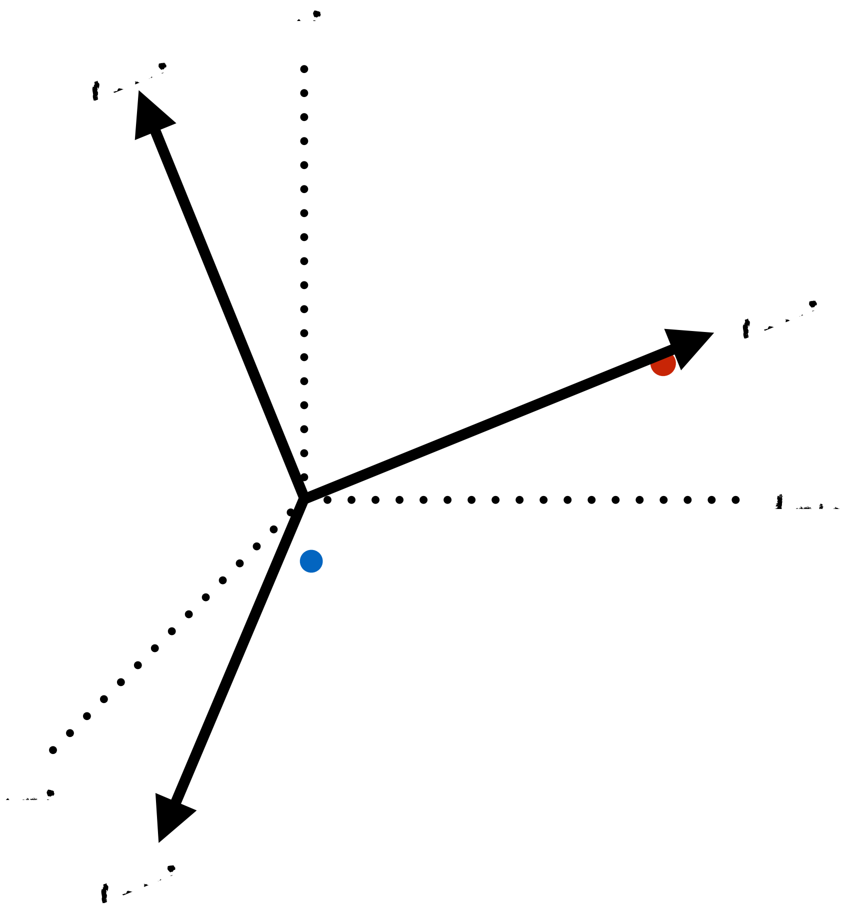
$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{11} \\ \tilde{x}_{21} \end{matrix}$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{12} \\ \tilde{x}_{22} \end{matrix}$$

$$\Longrightarrow \underbrace{\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix}}_{\mathbf{X}} \approx \underbrace{\begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix}}_{\mathbf{F}} \underbrace{\begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} \\ \tilde{x}_{21} & \tilde{x}_{22} \end{pmatrix}}_{\mathbf{C}}$$

Data Matrix — "Latent" Factors — Scores

$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$



Interpretability of latent factors is a little subjective, but soon you will be more comfortable with the idea!

$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{11} \\ \tilde{x}_{21} \end{matrix}$$

$$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} \approx \begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix} \begin{matrix} \tilde{x}_{12} \\ \tilde{x}_{22} \end{matrix}$$

$$\Longrightarrow \underbrace{\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix}}_{\mathbf{X}} \approx \underbrace{\begin{pmatrix} 0.89 & 0.27 \\ -0.09 & 0.89 \\ 0.45 & -0.35 \end{pmatrix}}_{\mathbf{F}} \underbrace{\begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} \\ \tilde{x}_{21} & \tilde{x}_{22} \end{pmatrix}}_{\mathbf{C}}$$

Data Matrix (X) — "Latent" Factors (F: size, weight) — Scores (C)

Part 5:
A More Complete Example

(Nonnegative) Matrix Factorization for Text

More Complete Example

(Factors in Text)

Document 1

My **cat** likes to eat **dog** food. It's insane. He won't eat tuna, but **dog** food? He's all over it.

Document 2

Check out this video of my **dog** chasing my **cat** around the house! He never gets **tired**! Simon! The **cat** is not a **dog** toy! Dumb **dog**.

Document 3

I **injured** my **ankle** playing football yesterday. It is bruised and swollen. Maybe **sprained**?

Document 4

So **tired** of being **injured**. My **ankle** just won't get better! I **sprained** it 2 months ago!

More Complete Example

(Factors in Text)

$$\mathbf{B} = \begin{array}{c} \\ \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & \boxed{3} & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

In this example, our *observations are the documents* and the *words are the variables*

More Complete Example

(Factors in Text)

$$\mathbf{B} = \begin{array}{l} \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \overset{\begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array}}{\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}}$$

Basis (Elementary Axes) of our 6-dimensional space

$$doc1 = 1\underbrace{\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}}_{\text{cat}} + 2\underbrace{\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}}_{\text{dog}} + 0\underbrace{\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}}_{\text{tired}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}}_{\text{injured}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix}}_{\text{ankle}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}}_{\text{sprained}}$$

More Complete Example

(Factors in Text)

$$\mathbf{B} = \begin{array}{l} \\ \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

We can approximate this matrix using a matrix factorization

$$\mathbf{B} \approx \begin{array}{l} \\ \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \begin{array}{c} \begin{array}{cc} Factor1 & Factor2 \end{array} \\ \begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix} \end{array}$$

More Complete Example

(Factors in Text)

$$\mathbf{B} = \begin{array}{c} \\ \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

We can approximate this matrix using a matrix factorization

$$\mathbf{B} \approx \begin{array}{c} \\ \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1 & 1.7 & 0 & 0 \\ 1.6 & 2.7 & 0 & 0 \\ 0.4 & 0.72 & 0.36 & 0.44 \\ 0 & 0 & 0.72 & 0.88 \\ 0 & 0 & 0.72 & 0.88 \\ 0 & 0 & 0.72 & 0.88 \end{pmatrix} \end{array}$$

More Complete Example

(Factors in Text)

$$\mathbf{B} = \begin{array}{c} \\ \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

How did I get this?

We'll talk about it later!

$$\mathbf{B} \approx \begin{array}{c} \\ \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \begin{array}{c} \begin{array}{cc} Factor1 & Factor2 \end{array} \\ \begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array} \\ \begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix} \end{array}$$

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \overset{\begin{array}{cc} \textit{Factor1} & \textit{Factor2} \end{array}}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{\begin{array}{cccc} \textit{doc1} & \textit{doc2} & \textit{doc3} & \textit{doc4} \end{array}}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

Each column of B (each document) can be written as a linear combination of factors. These linear combinations are the points' coordinate representations in the new basis.

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \overset{\quad Factor1 \quad Factor2}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{doc1 \quad doc2 \quad doc3 \quad doc4}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

Each column of B (each document) can be written as a linear combination of factors. These linear combinations are the points' coordinate representations in the new basis.

$$\mathbf{B}_{\star 2} \approx 1.7 Factor_1 + 0.1 Factor_2$$

(doc 2)

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{"cat"} \\ \text{"dog"} \\ \text{"tired"} \\ \text{"injured"} \\ \text{"ankle"} \\ \text{"sprained"} \end{array} \overset{\begin{array}{cc} \textit{Factor1} & \textit{Factor2} \end{array}}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{\begin{array}{cccc} \textit{doc1} & \textit{doc2} & \textit{doc3} & \textit{doc4} \end{array}}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

Each column of B (each document) can be written as a linear combination of factors. These linear combinations are the points' coordinate representations in the new basis.

$$\mathbf{B}_{\star 2} \approx 1.7 Factor_1 + 0.1 Factor_2$$

(doc 2)

Conclude: document 2 more aligned with factor 1 than factor 2

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \overset{\begin{array}{cc} Factor1 & Factor2 \end{array}}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{\begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array}}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

$$\mathbf{B}_{\star 2} \approx 1.7 Factor_1 + 0.1 Factor_2$$

(doc 2)

How do we interpret factor 1?

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \overset{\begin{array}{cc} \mathit{Factor1} & \mathit{Factor2} \end{array}}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{\begin{array}{cccc} \mathit{doc1} & \mathit{doc2} & \mathit{doc3} & \mathit{doc4} \end{array}}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

How do we interpret factor 1?

$$Factor_1 = 1\underbrace{\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}}_{\text{cat}} + 1.6\underbrace{\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}}_{\text{dog}} + 0.4\underbrace{\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}}_{\text{tired}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}}_{\text{injured}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix}}_{\text{ankle}} + 0\underbrace{\begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}}_{\text{sprained}}$$

'pets'

More Complete Example

(Factors in Text)

$$\mathbf{B} \approx \begin{array}{l} \text{“cat”} \\ \text{“dog”} \\ \text{“tired”} \\ \text{“injured”} \\ \text{“ankle”} \\ \text{“sprained”} \end{array} \overset{\begin{array}{cc} Factor1 & Factor2 \end{array}}{\begin{pmatrix} 1.0 & 0 \\ 1.6 & 0 \\ 0.4 & 0.4 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}} \overset{\begin{array}{cccc} doc1 & doc2 & doc3 & doc4 \end{array}}{\begin{pmatrix} 1.0 & 1.7 & 0 & 0.0 \\ 0 & 0.1 & 0.9 & 1.1 \end{pmatrix}}$$

Scores/Coordinates.

Allow us to describe data observations according to the new factors.

- Document 1: about pets
- Document 2: about pets
- Document 3: about injuries
- Document 4: about injuries

Loadings.

Allow us to interpret factors.

- Factor 1: pets
- Factor 2: injuries

Why a New Basis?

- We want to use a **subset of the new basis vectors** (i.e. new features/variables/axes) to **reduce the dimensionality** of the data and keep patterns

- We ***hope*** that the new features (being combinations of the old ones) will have some **interpretation**

Interpretation of Features

- The **interpretation** of the new basis vectors (new features/variables) **is subjective.**

- We simply look at the loadings to **find the variables with the highest loading values** (in absolute value) and *try* to **interpret their collective meaning.**

Interpretation of Features

- **Original basis vectors** (features/variables) were: *height, weight, head_ circumference, verbal_ score, quant_ score, household_ income, house_ value.*

- Let's see if we can assign some meaning to our new basis vectors (features/variables)

Axis 1

$$\begin{matrix} height \\ weight \\ head\_circumference \\ verbal\_score \\ quant\_score \\ household\_income \\ house\_value \end{matrix} \begin{pmatrix} 0.7 \\ 0.8 \\ 0.5 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Size?

Interpretation of Features

- **Original basis vectors** (features/variables) were: *height, weight, head_ circumference, verbal_ score, quant_ score, household_ income, house_ value.*
- Let's see if we can assign some meaning to our new basis vectors (features/variables)

Axis 2

$$\begin{matrix} height \\ weight \\ head\_circumference \\ verbal\_score \\ quant\_score \\ household\_income \\ house\_value \end{matrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.7 \\ 0.8 \\ 0.2 \\ 0.1 \end{pmatrix}$$

ability?

Interpretation of Features

- **Original basis vectors** (features/variables) were: *height, weight, head_circumference, verbal_score, quant_score, household_income, house_value.*
- Let's see if we can assign some meaning to our new basis vectors (features/variables)

Axis 3

$$\begin{array}{r} height \\ weight \\ head\_circumference \\ verbal\_score \\ quant\_score \\ household\_income \\ house\_value \end{array} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.1 \\ 0.3 \\ 0.9 \\ 0.7 \end{pmatrix}$$

affluence?

Major Ideas from Section

- linear combinations geometrically
- linear (in)dependence geometrically
- vector span
- subspace
- dimension of subspace
- hyperplane
- basis vectors
- coordinates in different bases
- (generic) factor analysis
- loadings
- scores/coordinates