Introduction to Vector Space Models

Vector span, Subspaces, and Basis Vectors

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

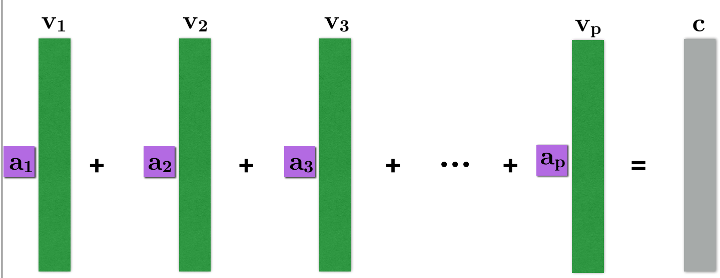
(Algebraically)

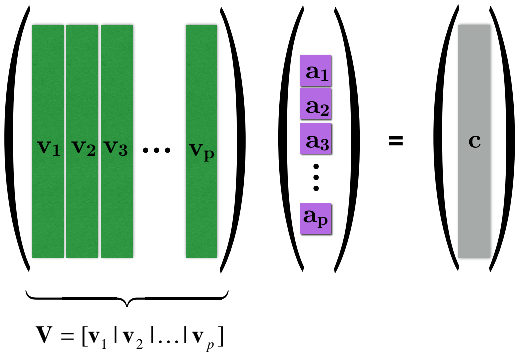
A **linear combination** is constructed from a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ by multiplying each vector by a constant and adding the result:

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

alternatively, we could write

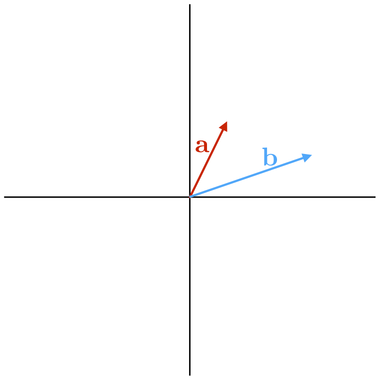
$$\mathbf{c} = \mathbf{V}\mathbf{a} \quad \text{where} \quad \mathbf{V} = [\mathbf{v}_1 | \mathbf{v}_2 | \ldots | \mathbf{v}_p] \quad \text{and} \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$





Linear Combinations

(Geometrically)



Linear Dependence

(Algebraically)

A group of vectors, $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ are **linearly dependent** if there exists corresponding scalars, $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ *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

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

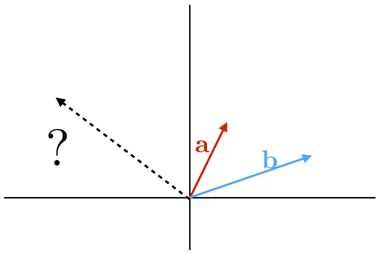
Linear Dependence

(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

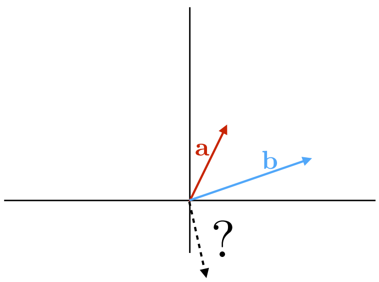
(Geometrically)



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 $\mathbf{v}$:

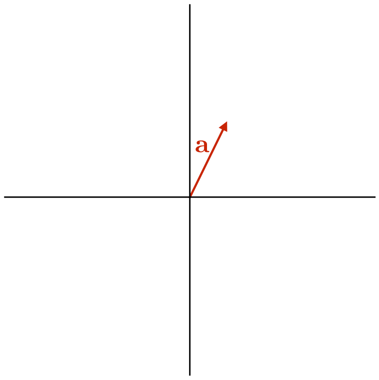
$$span(\mathbf{v}) = \{\alpha\mathbf{v} \quad \text{for all constants } \alpha\}$$

- 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:

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

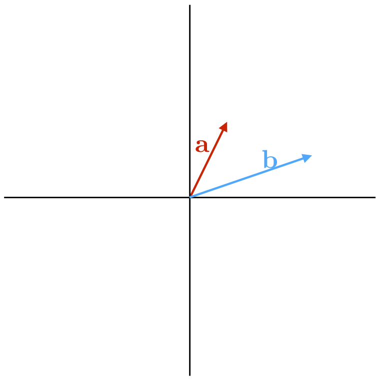
Vector Span

(Example 1)



Vector Span

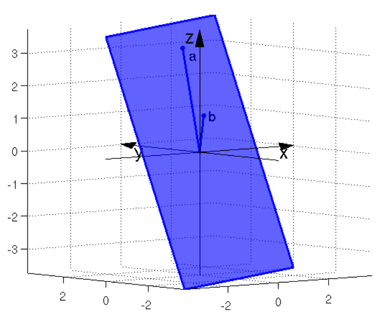
(Example 2)



Vector Span

(Example 3)

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



Subspace

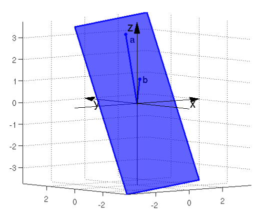
(Definition)

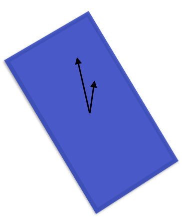
- A **subspace** S of $\mathbb{R}^n$ is thought of as a "flat" (having no curvature) surface within $\mathbb{R}^n$. It is a collection of vectors which satisfies the following conditions:
  - The origin ($\mathbf{0}$ vector) is contained in S.
  - If $\mathbf{x}$ and $\mathbf{y}$ are in S then $\mathbf{x}+\mathbf{y}$ 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

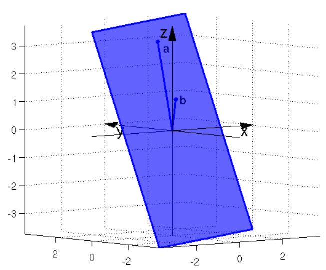
(Definition)

- 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 $\mathbb{R}^3$

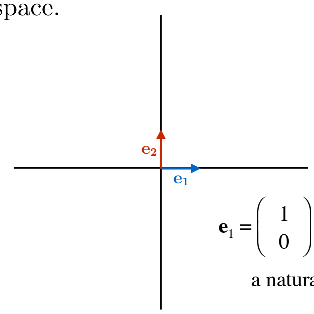
3. Describe the span of two linearly dependent vectors in $\mathbb{R}^3$

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* $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$

Bases and Coordinates

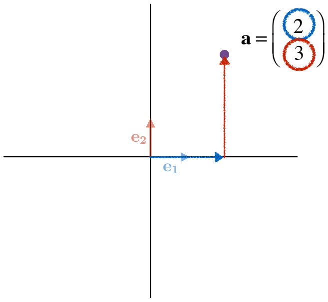
- A collection of vectors makes a **basis** for a space (or a subspace) if they are linearly independent and span the space.



$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are a natural basis for $\mathbb{R}^2$

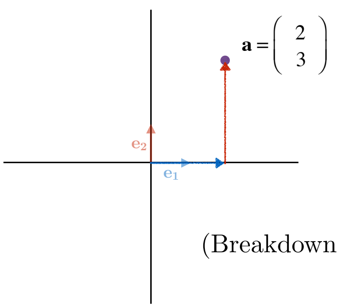
Bases and Coordinates

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



Bases and Coordinates

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



(Breakdown into parts)

Bases and Coordinates

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

Bases and Coordinates

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

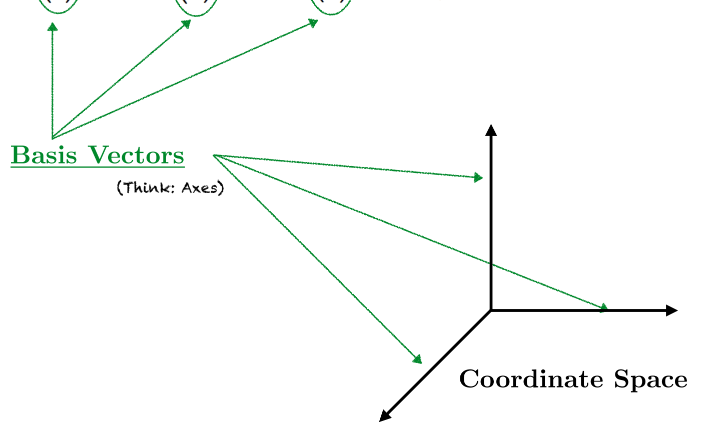
Coordinates

Bases and Coordinates

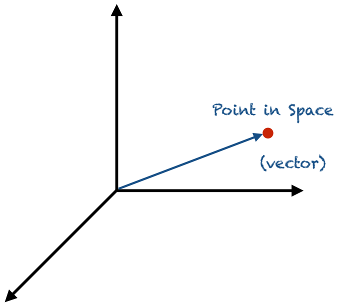
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

Basis Vectors

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



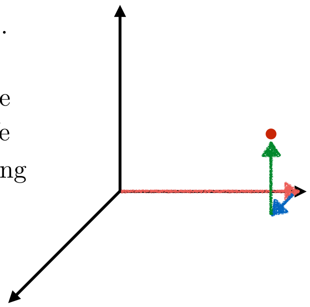
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$



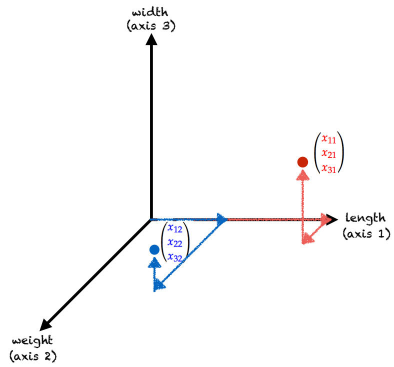
$$x_{11}\begin{pmatrix}1\\0\\0\end{pmatrix} + x_{21}\begin{pmatrix}0\\1\\0\end{pmatrix} + x_{31}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

Coordinates give directions to a point along basis vectors.

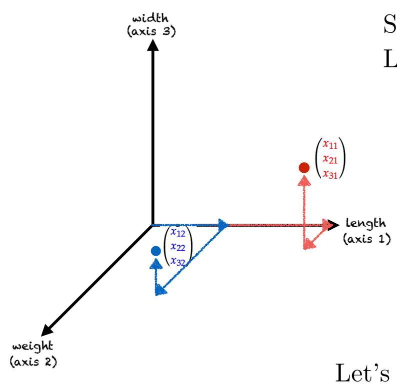
For any set of data points, the basis vectors are the same. We compare the points by comparing their coordinates.



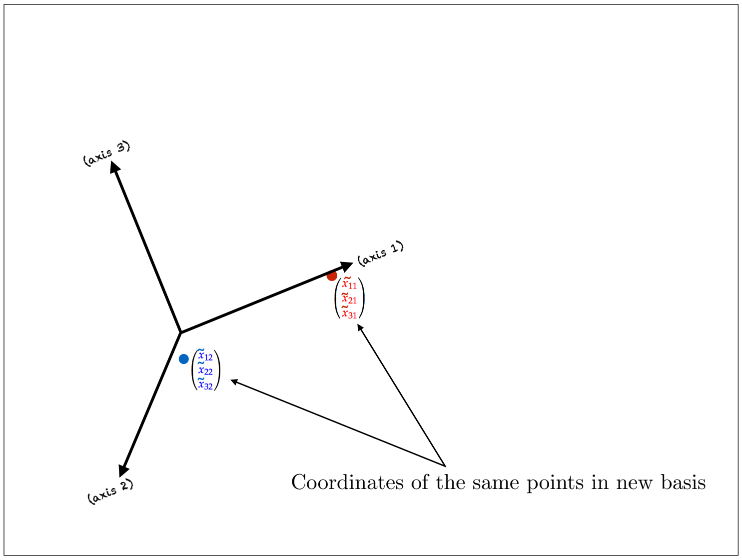
$$x_{12}\underset{\text{length}}{\begin{pmatrix}1\\0\\0\end{pmatrix}} + x_{22}\underset{\text{weight}}{\begin{pmatrix}0\\1\\0\end{pmatrix}} + x_{32}\underset{\text{width}}{\begin{pmatrix}0\\0\\1\end{pmatrix}} =$$



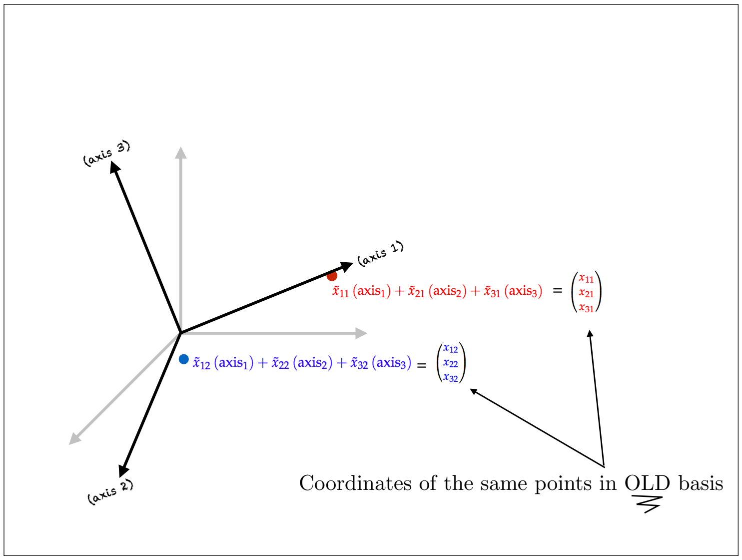
Compared to red point, blue point has:
Smaller length
Smaller width
Larger weight



Let's change the basis...

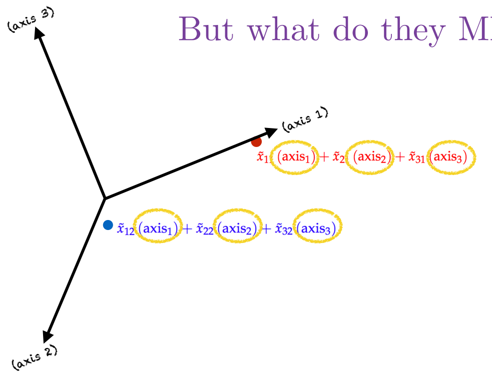


(axis 3)
(axis 1)
(axis 2)
Coordinates of the same points in new basis



(axis 3)
(axis 1)
(axis 2)
Coordinates of the same points in OLD basis

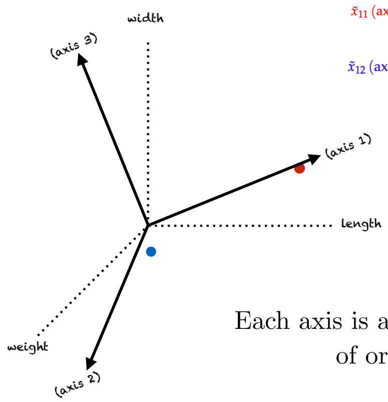
These are the new basis vectors... But what do they MEAN?



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

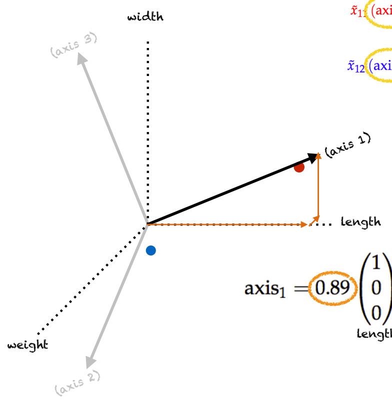
$$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$$

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



Each axis is a linear combination of original axes

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

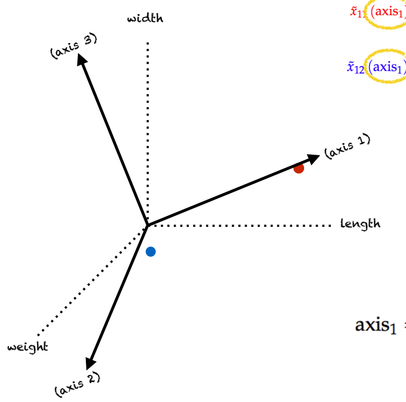


$$\tilde{x}_{11}(\text{axis}_1)+\tilde{x}_{21}(\text{axis}_2)+\tilde{x}_{31}(\text{axis}_3) = \begin{pmatrix}x_{11}\\x_{21}\\x_{31}\end{pmatrix}$$

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

$$\text{axis}_1 = 0.89\underbrace{\begin{pmatrix}1\\0\\0\end{pmatrix}}_{\text{length}} - 0.09\underbrace{\begin{pmatrix}0\\1\\0\end{pmatrix}}_{\text{weight}} + 0.45\underbrace{\begin{pmatrix}0\\0\\1\end{pmatrix}}_{\text{width}}$$

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



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

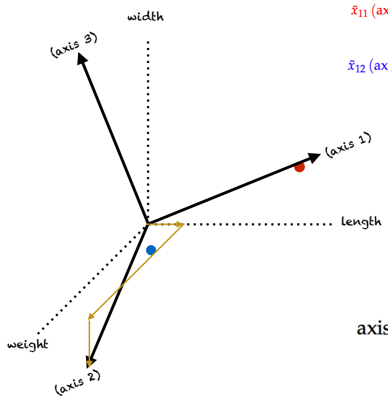
$$\tilde{x}_{12}(\text{axis}_1)+\tilde{x}_{22}(\text{axis}_2)+\tilde{x}_{32}(\text{axis}_3) = \begin{pmatrix}x_{12}\\x_{22}\\x_{32}\end{pmatrix}$$

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

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

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

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



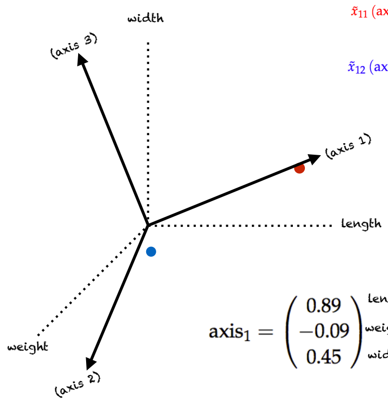
$$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

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

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

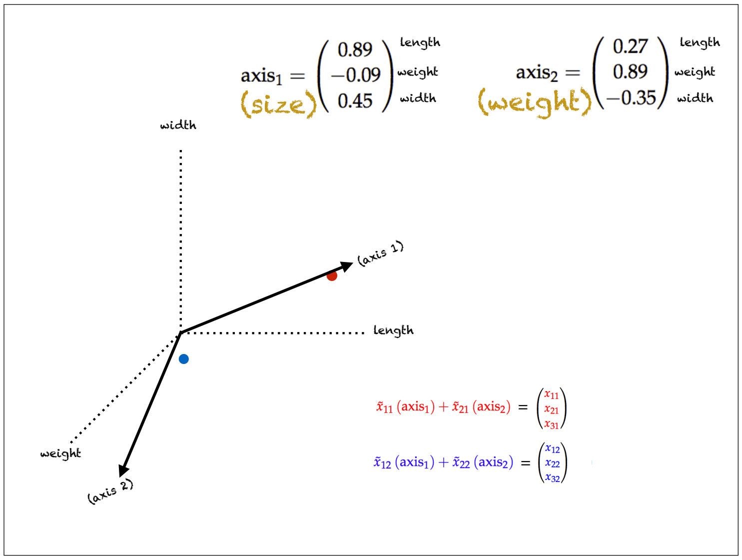
$$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$$

=0

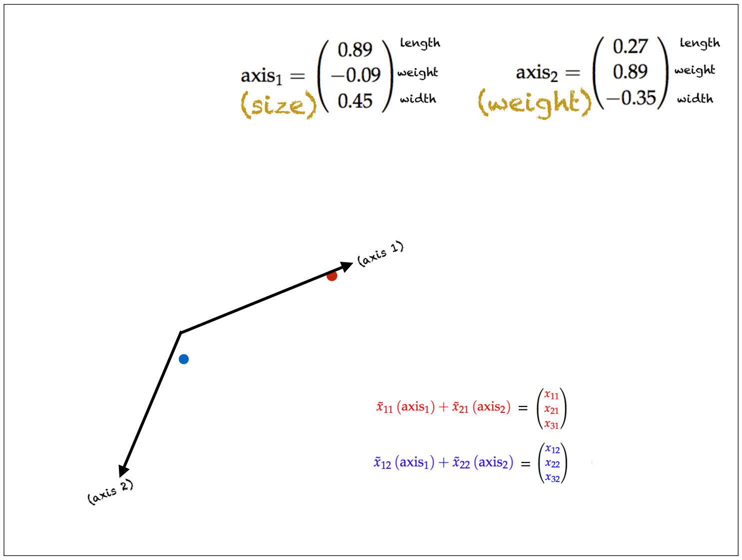


$$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix} \qquad \text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix} \begin{matrix} \text{length} \\ \text{weight} \\ \text{width} \end{matrix}$$

Let's ignore axis 3 for now...



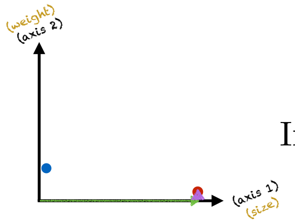
$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix}$ length weight width
(size)
$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$ length weight width
(weight)
width
(axis 1)
length
weight
(axis 2)
$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$
$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$



$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix}$ length weight width
(size)
$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$ length weight width
(weight)
(axis 1)
(axis 2)
$\tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) = \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$
$\tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) = \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix}$

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

(size) (weight)



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

Infer that the red point has:

large size

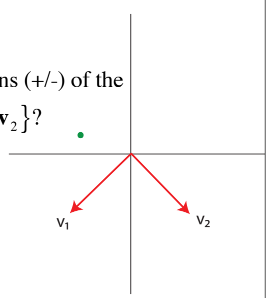
and

small weight

relative to blue point

Practice

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



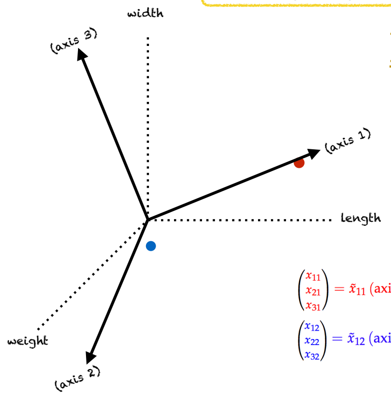
2 Find the coordinates of the vector $\mathbf{x} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ in the basis $\left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

Draw a picture to confirm your answer matches your intuition.

Some Terminology

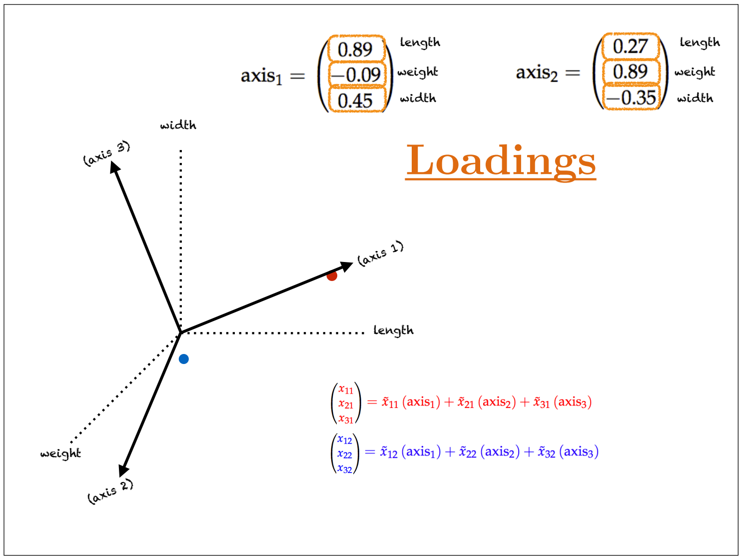
Factors

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

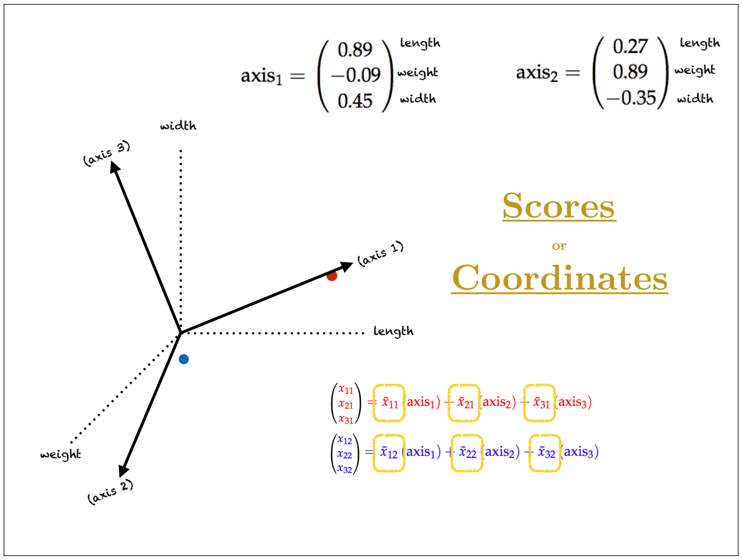


$$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$$

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



$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix}$ length weight width
$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$ length weight width
Loadings
width
(axis 3)
(axis 1)
length
weight
(axis 2)
$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$
$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3)$



$\text{axis}_1 = \begin{pmatrix} 0.89 \\ -0.09 \\ 0.45 \end{pmatrix}$ length weight width
$\text{axis}_2 = \begin{pmatrix} 0.27 \\ 0.89 \\ -0.35 \end{pmatrix}$ length weight width
Scores
or
Coordinates
width
(axis 3)
(axis 1)
length
weight
(axis 2)
$\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \tilde{x}_{11}\,(\text{axis}_1) + \tilde{x}_{21}\,(\text{axis}_2) + \tilde{x}_{31}\,(\text{axis}_3)$
$\begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \tilde{x}_{12}\,(\text{axis}_1) + \tilde{x}_{22}\,(\text{axis}_2) + \tilde{x}_{32}\,(\text{axis}_3)$

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)

$\mathbf{B} =$

	doc1	*doc2*	*doc3*	*doc4*
"cat"	1	2	0	0
"dog"	2	3	0	0
"tired"	0	1	0	1
"injured"	0	0	1	1
"ankle"	0	0	1	1
"sprained"	0	0	1	1

The word "dog" appears 3 times in document 2.

More Complete Example

(Factors in Text)

$\mathbf{B} =$

	doc1	*doc2*	*doc3*	*doc4*
"cat"	1	2	0	0
"dog"	2	3	0	0
"tired"	0	1	0	1
"injured"	0	0	1	1
"ankle"	0	0	1	1
"sprained"	0	0	1	1

Basis (Elementary Axes) of our 6-dimensional space

$$doc1 = 1\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix} + 2\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\0\\0\\1\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}$$

cat dog tired injured ankle sprained

More Complete Example

(Factors in Text)

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

We can approximate this matrix using a matrix factorization

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

More Complete Example

(Factors in Text)

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

We can approximate this matrix using a matrix factorization

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

More Complete Example

(Factors in Text)

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

How did I get this? We'll talk about it later!

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

More Complete Example

(Factors in Text)

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

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)

$\mathbf{B} \approx$

	Factor1	*Factor2*
"cat"	1.0	0
"dog"	1.6	0
"tired"	0.4	0.4
"injured"	0	0.8
"ankle"	0	0.8
"sprained"	0	0.8

doc1	*doc2*	*doc3*	*doc4*
1.0	1.7	0	0.0
0	0.1	0.9	1.1

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)

$\mathbf{B} \approx$

	Factor1	*Factor2*
"cat"	1.0	0
"dog"	1.6	0
"tired"	0.4	0.4
"injured"	0	0.8
"ankle"	0	0.8
"sprained"	0	0.8

doc1	*doc2*	*doc3*	*doc4*
1.0	1.7	0	0.0
0	0.1	0.9	1.1

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)

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

How do we interpret factor 1?

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

'pets'

More Complete Example

(Factors in Text)

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

Loadings.

Allow us to interpret factors.

- Factor 1: pets
- Factor 2: injuries

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

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
height	0.7
weight	0.8
head_circumference	0.5
verbal_score	0
quant_score	0
household_income	0
house_value	0

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
height	0
weight	0
head_circumference	0
verbal_score	0.7
quant_score	0.8
household_income	0.2
house_value	0.1

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

height	0
weight	0
head_circumference	0
verbal_score	0.1
quant_score	0.3
household_income	0.9
house_value	0.7

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