COMP229: Introduction to Data Science
Lecture 25: A change of a linear basis

Olga Anosova, O.Anosova@liverpool.ac.uk
Autumn 2023, Computer Science department
University of Liverpool, United Kingdom

Lecture plan & learning outcomes

On this lecture we should learn

- tips on how to develop critical thinking skills,
- some NLP models basics,
- what is a basis,
- how to translate from one basis to another.

How to develop critical thinking?

OD

is y = kx+c a linear map

The equation $y = kx + c$ represents a linear function, but whether it defines a linear map depends on the context and the specific definitions being used.

In the context of linear algebra and linear maps, a linear map (or linear transformation) has certain properties. Specifically, a function $L : \mathbb{R}^n \to \mathbb{R}^m$ is considered a linear map if it satisfies two properties:

1. **Additivity**: For any vectors $\mathbf{u}$ and $\mathbf{v}$ in the domain of L, the image of their sum is equal to the sum of their images:
   $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$
2. **Homogeneity**: For any vector $\mathbf{v}$ in the domain of L and any scalar c, the image of the scalar multiple of $\mathbf{v}$ is equal to the scalar multiple of the image of $\mathbf{v}$:
   $L(c\mathbf{v}) = c \cdot L(\mathbf{v})$

The equation $y = kx + c$ satisfies both of these properties, making it a linear map in the context of linear transformations. Here, y and x are the variables, and k and c are constants.

However, it's worth noting that in other contexts, the term "linear map" or "linear function" might have a different meaning. For example, in elementary algebra or geometry, the term "linear" might refer to a function that represents a straight line on a graph, and in that case, the equation $y = kx + c$ would indeed represent a linear function but not necessarily a linear map in the specific sense defined in linear algebra.

How to develop critical thinking?

Try asking a large language model and *evaluate its answers*. Slightly non-trivial questions that we've covered in the lectures work particularly well.

For example, *"When is perimeter equal to the area?"*
or *"Is determinant an invariant?"*

Reminder: signed area (volume)

- The determinant of $\begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix}$ is the signed area of the parallelogram spanned by the vectors $\vec{u} = (u_x, u_y)$ and $\vec{v} = (v_x, v_y)$.
- The signed area of any polygon with counter-clockwisely ordered vertices can be found form a shoelace formula.
- The signed volume of a body under a linear map $\vec{v} \mapsto A\vec{v}$ is multiplied by $\det A$.
- The signed volume is an isometric invariant, not an invariant of a linear map (because of scaling).
- The determinant is invariant under many operations of matrices (lecture 23).

A basis of vectors

Definition 25.1. Vectors $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^m$ form a **basis** if those vectors are linearly independent and span $\mathbb{R}^m$.

This happens iff the matrix A with the columns $\vec{v}_1, \ldots, \vec{v}_m$ has $\det A \neq 0$.

In this case by Theorem 23.7 any vector $\vec{v} \in \mathbb{R}^m$ can be written as $\vec{v} = A\vec{c}$ for $\vec{c} = A^{-1}\vec{v}$,

i.e. as a *unique* linear combination $\vec{v} = \sum_{i=1}^{m} c_i \vec{v}_i$, $c_i \in \mathbb{R}$.

The scalars c_i are called **the coordinates** of the vector $\vec{v}$ with respect to this basis.

Images of the basis vectors

The **standard basis** in $\mathbb{R}^m$: each vector $\vec{e}_i$ has coordinate 1 on the i-th place, 0 anywhere else.

We can take as a basis any vectors that satisfy those rules.

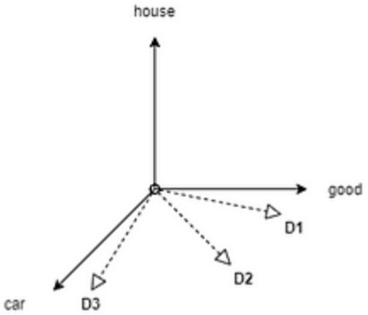
NLP example

Vector space model is an algebraic model for representing text documents in Natural Language Processing (NLP).
The basis is formed by the words in a **corpus** (dataset of words or terms), each corpus item forms one basis vector.
For example, Google News corpus consists of 3mln terms = 3mln axes of basis vectors.

Documents are represented as vectors. If a term occurs in the document, its value in the vector is non-zero.
Often values are given by the TF-idf (term frequency–inverse document frequency) measure.

Document similarity

	good	house	car
D1	0.91	0	0.0011
D2	0.21	0	0.1
D3	0.15	0	0.921



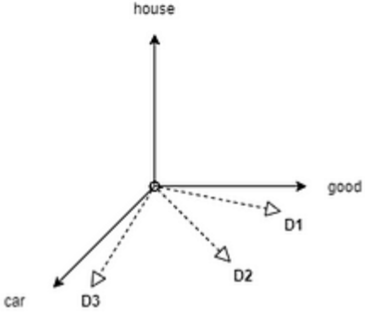
Document similarity is measured from the angles between the vectors (via the dot product).

Advantages: simple, partial match, continuous similarity.

Drawbacks:

Document similarity

	good	house	car
D1	0.91	0	0.0011
D2	0.21	0	0.1
D3	0.15	0	0.921



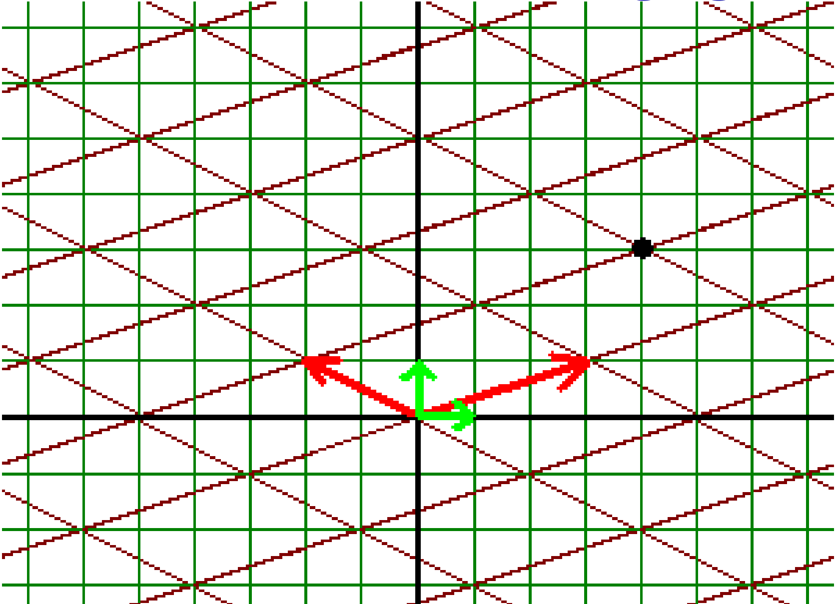
Document similarity is measured from the angles between the vectors (via the dot product).

Advantages: simple, partial match, continuous similarity.

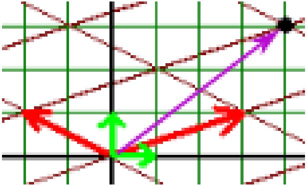
Drawbacks: loss of order, total independence of terms.

What if we would like to take the relation between the terms into account? Or translate the documents?

Same world in different languages



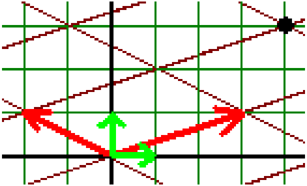
Spanning the plane



What are the coordinates of the purple vector in the standard green basis (GB)? $\vec{v}_e = (4,3)_e$.

What are its coordinates in the new red basis (RB)?
$\vec{v}_e = (2,1)_v$.
How to translate?

The same vector in the two bases



RB is $\vec{v}_1 = 3\vec{e}_1 + \vec{e}_2$, $\vec{v}_2 = -2\vec{e}_1 + \vec{e}_2$ in terms of the GB.

Then to translate vector $\vec{v} = (2,1)_v$ into GB coordinates we should take $2\vec{v}_1 + \vec{v}_2 = (3\vec{e}_1 + \vec{e}_2) + (-2\vec{e}_1 + \vec{e}_2) = 4\vec{e}_1 + 3\vec{e}_2$, hence $\vec{v}$ has the coordinates $(4,3)_e$.

Does it look familiar?

In the matrix from the last identity is

$$\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \text{ or } B\vec{r}_v = \vec{r}_e,$$

where $\vec{r}_e, \vec{r}_v$ represent the same point in the bases, and matrix B consisnts of the coordinates of RB in the GB.

Translating between bases

The translation into GB happens via the transition matrix $B = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$ that represents RB basis in the GB language: $(\vec{v}_1, \vec{v}_2) = B(\vec{e}_1, \vec{e}_2)$

which can be considered as the simple matrix identity:

$$\begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

To summarise: to translate from RB $(\vec{v}_1, \vec{v}_2)$ into GB $(\vec{e}_1, \vec{e}_2)$, apply $\vec{r}_e = B\vec{r}_v$,

where B represents RB vectors in GB terms:

$(\vec{v}_1, \vec{v}_2) = B(\vec{e}_1, \vec{e}_2)$.

Inverse translation

How to translate from GB $\vec{v} = (4,3)_e$ into RB $\vec{v} = (2,1)_v$?

To express GB $(\vec{e}_1, \vec{e}_2)$ in terms of RB $(\vec{v}_1, \vec{v}_2)$, we multiply both sides of $(\vec{v}_1, \vec{v}_2) = B(\vec{e}_1, \vec{e}_2)$ by the inverse matrix B^{-1} on the left: $B^{-1}(\vec{v}_1, \vec{v}_2) = (\vec{e}_1, \vec{e}_2)$.

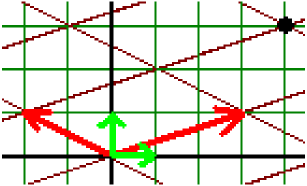
Revision: from Theorem 23.7, the inverse of a 2×2 matrix is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Translating from the GB to the RB

Invert: $B^{-1} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$

$$B^{-1}\vec{v}_e = B^{-1}\begin{pmatrix} 4 \\ 3 \end{pmatrix}_e = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_v.$$



The point with coordinates $(4,3)_e$ in the old basis $(\vec{e}_1, \vec{e}_2)$ has coordinates $(2,1)_v$ in the basis $(\vec{v}_1, \vec{v}_2)$.

To summarise: to translate from GB $(\vec{e}_1, \vec{e}_2)$ into RB $(\vec{v}_1, \vec{v}_2)$, apply $\vec{r}_v = B^{-1}\vec{r}_e$,
where B represents RB vectors in GB terms:
$(\vec{v}_1, \vec{v}_2) = B(\vec{e}_1, \vec{e}_2)$.

Applying linear map

Suppose we have a linear map $\vec{r}_e \mapsto A\vec{r}_e$ that has a known representation in the initial GB basis.
How can we translate this map's action into the new RB?

We need to

- translate from RB coordinates into GB: $B\vec{r}_v$
- apply the map in the GB: $AB\vec{r}_v$
- translate the result back from GB into RB: $(B^{-1}AB)\vec{r}_v$

Applying linear map

Theorem 25.2. The map $\vec{r}_e \mapsto A\vec{r}_e$ becomes $\vec{r}_v \mapsto (B^{-1}AB)\vec{r}_v$ in a new basis $(\vec{v}_1, \vec{v}_2) = B(\vec{e}_1, \vec{e}_2)$.

Proof. When we change to the new basis $(\vec{v}_1, \vec{v}_2)$, the same vectors change their coordinates according to the transition matrix B. If we denote $\vec{s}_e = A\vec{r}_e$, then $\vec{s}_e, \vec{r}_e$ can be expressed as $\vec{r}_e = B\vec{r}_v, \vec{s}_e = B\vec{s}_v$.

The identity $\vec{s}_e = A\vec{r}_e$ with the given matrix A becomes $B\vec{s}_v = A(B\vec{r}_v)$, hence $\vec{s}_v = (B^{-1}AB)\vec{r}_v$.

So the new matrix of the same map is $B^{-1}AB$. □

An example of complete transformation

Let A be a matrix of rotation about 0 through $\pi/2$, i.e. $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_e$, any $\vec{r}_e \in \mathbb{R}^2$ maps to $\vec{s}_e = A\vec{r}_e$.

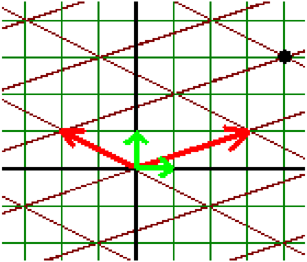
We use the new basis with the same matrix B as before:

$$A \cdot B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 3 & -2 \end{pmatrix}.$$

$$B^{-1}AB = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}.$$

A new rotation matrix

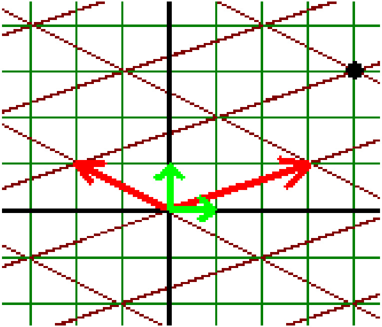
Indeed, we rotate $\vec{v}_1 = (3,1)_e$, $\vec{v}_2 = (-2,1)_e$ as follows: $\vec{v}_1 \mapsto (-1,3)_e = \vec{v}_1 + 2\vec{v}_2 = (1,2)_v$ and $\vec{v}_2 \mapsto (-1,-2)_e = -\vec{v}_1 - \vec{v}_2 = (-1,-1)_v$ as predicted by the matrix $B^{-1}AB$ computed above.



Algebraically (in a computer memory), the matrix will change, because the same vector is expressed differently in different bases.

Geometrically, the map is the same rotation.

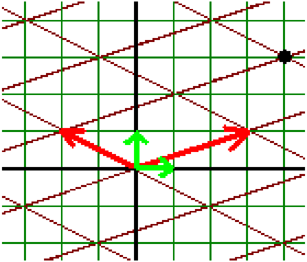
The rotation image in the old basis



Let's see what happened to the point $(4,3)_e = (2,1)_v$ under this rotation.

In the basis $(\vec{e}_1, \vec{e}_2)$, the point $(4,3)$ is rotated to the new point $Ar_e = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix}_e = \begin{pmatrix} -3 \\ 4 \end{pmatrix}_e$.

The rotation image in the new basis



In the new basis $(\vec{v}_1, \vec{v}_2)$, the same point is rotated to the new position:

$$B^{-1}ABr_v = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}_v = \begin{pmatrix} 1 \\ 3 \end{pmatrix}_v,$$ which is indeed

$$\vec{v}_1 + 3\vec{v}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}_e + 3\begin{pmatrix} -2 \\ 1 \end{pmatrix}_e = \begin{pmatrix} -3 \\ 4 \end{pmatrix}_e.$$

Another map in the new basis

Problem 25.3. Find the matrix of the rotation through $-\pi/2$ in the basis $\vec{v}_1 = (3,1)$, $\vec{v}_3 = (1,2)$.

Solution 25.3. $B = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ has columns equal to $\vec{v}_1, \vec{v}_3$ and $B^{-1} = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$. Rotation $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

The matrix of the operator in the new basis is

$$B^{-1}AB = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} =$$

$$\frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}.$$

$\square$

Real life translations

Real life translations

Contemporary NLP models do not really use vector-space models, but those are still widely used for referencing and searches:

Spotify uses the vector-space model in two out of three recommender engines behind its Discover Weekly suggestions.

On the next lecture we continue applying linear algebra to the serach engines, namely to the most popular search engine Google Page Rank, with the introduction of eigenvalues.

Time to revise and ask questions

- When an old basis is replaced by a new basis, the columns of the transition matrix B are the new basis vectors (written in the old basis).
- A linear map with a matrix A in the old basis has the matrix $B^{-1}AB$ in the new basis.
- Any linear map (the same geometric action such as a rotation) can be algebraically represented by many different matrices.

Problem 25.4. Find the image of $\vec{v} = (4,3)_e$ in the basis $(\vec{e}_1, \vec{e}_2)$ under the map with $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ in the new basis $\vec{v}_1 = (3,1)$, $\vec{v}_3 = (1,2)$.

Additional links

- Vector space models in NLP.
- Change of basis video from 3Blue1Brown.
- How to teach multiplication to ChatGPT4.
- Explainable artificial intelligence for assault sentence prediction in New Zealand(2022).