COMP229: Introduction to Data Science
Lecture 19: Matrices of linear maps

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

Lecture plan

- From a linear basis to a linear map
- (Re)inventing well-known operations:
  matrix-vector & matrix-matrix multiplication
- Scaling of vectors
- From linear maps to affine maps
- Geometry & algebra $\rightarrow$ Linear algebra

Reminder: Lloyd's k-means clustering

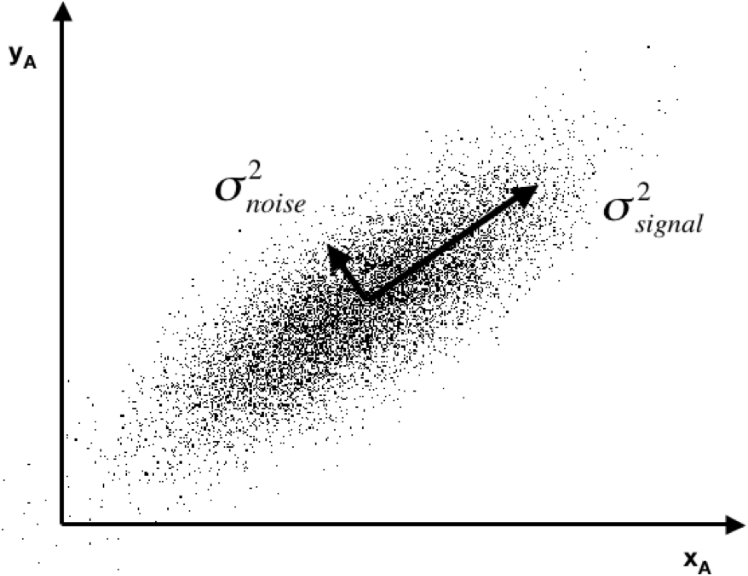
Here are the steps of Lloyd's heuristic algorithm.

- Initialise k centres of clusters in a cloud C.
- For each centre, assign all points of C that are closer to this centre than to all others.
- Re-compute the centre of every cluster that was updated above and re-assign all points.
- Stop when centres of clusters don't change or a maximum number of iterations is reached.

The result is a partitioning of the data space into Voronoi cells.

Most of the improvements that speed up calculations rely on triangle inequality.

Data patterns need transformations



The standard basis $\vec{e}_1, \ldots, \vec{e}_m$

Two thick vectors above form a better coordinate system for the noisy data than the original axes.

Recall that the symbol $\mathbb{R}^m$ denotes a Euclidean space with a fixed basis of orthonormal vectors $\vec{e}_1, \ldots, \vec{e}_m$, where $\vec{e}_i$ has the i-th coordinate 1, all other coordinates 0, e.g. $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in $\mathbb{R}^2$. Then any $\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

What is a linear map?

Definition 19.1. A map $f : \mathbb{R}^m \to \mathbb{R}^k$ is **linear** if $f(s\vec{v}) = sf(\vec{v})$, $s \in \mathbb{R}$, and $f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$.

Claim 19.2. Any linear map f is determined by the images $f(\vec{e}_1), \dots, f(\vec{e}_m)$ of a linear basis $\vec{e}_1, \dots, \vec{e}_m$.

Problem 19.3. For dimensions $m = k = 2$, let $f(\vec{e}_1) = \begin{pmatrix} a \\ c \end{pmatrix}$, $f(\vec{e}_2) = \begin{pmatrix} b \\ d \end{pmatrix}$. Express the image of the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ in terms of a, b, c, d, x, y.

The 2-dimensional case

Solution 19.3. Any vector $\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ has the image

$$f \begin{pmatrix} x \\ y \end{pmatrix} = x \cdot f \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \cdot f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix} =$$

$$\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A\vec{v}, \text{ so}$$

$A = (f(\vec{e}_1), f(\vec{e}_2))$.

This is why matrix-vector multiplication is defined this way: to easily represent a linear map.

The matrix of a linear map

Definition 19.4. For a fixed basis $\vec{e}_1, \dots, \vec{e}_m$ in $\mathbb{R}^m$, the $k \times m$ **matrix of a linear map** $f : \mathbb{R}^m \to \mathbb{R}^k$ consists of m columns that are the images $f(\vec{e}_1), \dots, f(\vec{e}_m)$, i.e. vectors of k coordinates in $\mathbb{R}^k$.

For $m = k = 2$, if $f(\vec{e}_1) = \begin{pmatrix} a \\ c \end{pmatrix}$, $f(\vec{e}_2) = \begin{pmatrix} b \\ d \end{pmatrix}$, then the 2×2 matrix of f is $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Matrix multiplied by a vector

Claim 19.5. Let $A = (a_{ij})$ be the matrix of $f : \mathbb{R}^m \to \mathbb{R}^k$. Any $\vec{v} \in \mathbb{R}^m$ with coordinates $c_1, \ldots, c_m$ has the image $f(\vec{v}) = A\vec{v}$ obtained by multiplying the $k \times m$ matrix A by the vector (column) $\vec{v}$ as follows:
the i-th coordinate of $A\vec{v}$ is the *scalar product* of the i-th row of A and $\vec{v}$.

For example, in $\mathbb{R}^2$: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$.

Examples of linear maps

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is called the **identity** matrix, because the linear map $\vec{v} \mapsto A\vec{v} = \vec{v}$ preserves any vector.

The $m \times m$ identity matrix is $A = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$.

Problem 19.6. What geometric transformations are represented by $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$?

Solution 19.6. Projections to the axis x, y in $\mathbb{R}^2$.

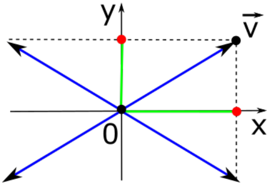
A scaling of vectors

Definition 19.7. The (non-uniform) **scaling** is the multiplication by a diagonal matrix $f(\vec{v}) = A\vec{v}$ for a fixed matrix $A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{pmatrix}$, where only non-diagonal elements don't vanish: $\lambda_i \neq 0$.

Any $\vec{v} = \sum_{j=1}^{m} c_j \vec{e}_j$ has the image $A\vec{v} = \sum_{j=1}^{m} (\lambda_j c_j)\vec{e}_j$, i.e. the j-th coordinate of $\vec{v}$ is multiplied by $\lambda_j \in \mathbb{R}$.

Mirror reflections in axes

Problem 19.8. What maps are represented by $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$?



Solution 19.8. $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}, \begin{pmatrix} x \\ -y \end{pmatrix}, \begin{pmatrix} -x \\ -y \end{pmatrix}$, $x, y \in \mathbb{R}$.

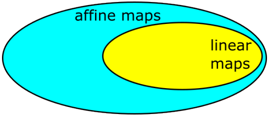
Hence they define the mirror reflections in the axes x, y and their composition, which is the central symmetry with respect to the origin in $\mathbb{R}^2$.

What is an affine map?

Definition 19.9. The **translation** by a fixed vector $\vec{u} = \begin{pmatrix} s \\ t \end{pmatrix}$ is defined for any $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ by $f(\vec{v}) = \vec{v} + \vec{u} = \begin{pmatrix} x+s \\ y+t \end{pmatrix}$, similarly in $\mathbb{R}^m$.

$f : \mathbb{R}^m \to \mathbb{R}^k$ is called **affine** if $f(\vec{v}) = A\vec{v} + \vec{u}$ for a fixed $k \times m$ matrix A and a fixed vector $\vec{u} \in \mathbb{R}^k$.

An affine map is linear if $\vec{u} = \vec{0}$, i.e. the origin is preserved by a linear map.



The matrix multiplication

Claim 19.10. The composition of linear maps $\mathbb{R}^m \to \mathbb{R}^k \to \mathbb{R}^n$ given by matrices $A : \mathbb{R}^m \to \mathbb{R}^k$, $B : \mathbb{R}^k \to \mathbb{R}^n$ is represented by the matrix product BA computed as follows: $(BA)_{ij} = \sum_{s=1}^{k} b_{is} a_{sj}$.

Matrices are added and multiplied by a scalar 'entry-wise' similarly to vectors, but are multiplied together 'rows-by-columns': i-th row by j-th column

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{pmatrix}.$$

The matrix multiplication in 2D

If $n = 2$, any *affine* transformation has the form

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} s \\ t \end{pmatrix} =$$

$$= \begin{pmatrix} ax + by + s \\ cx + dy + t \end{pmatrix} \text{ for any coordinates } x, y \in \mathbb{R}.$$

Historically and only for $n = 1$, we also call the transformation $f(x) = ax + b, x \in \mathbb{R}$, linear, because the fomula coinsides with the equation of a straight line. In general terms of linear algebra, this "linear" transformation is not linear, but affine.

From geometry to algebra

Problem 19.11. Write the algebraic form of this map in $\mathbb{R}^2$: scale the first coordinate by 2 and the second by 3, then translate by the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Solution 19.11. Composition:

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 3y \end{pmatrix} \mapsto \begin{pmatrix} 2x+2 \\ 3y+1 \end{pmatrix},$$

Order of the operations matters: opposite order gives another map,

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+2 \\ y+1 \end{pmatrix} \mapsto \begin{pmatrix} 2(x+2) \\ 3(y+1) \end{pmatrix}.$$

From algebra to geometry

Problem 19.12. Geometrically describe the map given algebraically $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x-3 \\ 2y+4 \end{pmatrix}$.

Solution 19.12. The order is important.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x \\ 2y \end{pmatrix} \mapsto \begin{pmatrix} 3x-3 \\ 2y+4 \end{pmatrix} = \begin{pmatrix} 3(x-1) \\ 2(y+2) \end{pmatrix},$$

In words: we scale by $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, then translate by $\begin{pmatrix} -3 \\ 4 \end{pmatrix}$; or we first translate by $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$, scale by $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$.

Time to revise and ask questions

Let's repeat those cornerstone basics with this video from 3Blue1Brown.

- Any linear map is determined by images of basis vectors and can be represented by the matrix whose columns are these images.
- The matrix multiplication is motivated by the composition of linear maps: $(BA)_{ij} = \sum_{s=1}^{k} b_{is} a_{sj}$.
- Any *affine* map $\mathbb{R}^m \to \mathbb{R}^k$ is $\vec{v} \mapsto A\vec{v} + \vec{u}$ for a fixed $k \times m$ matrix A, a fixed vector $\vec{u} \in \mathbb{R}^k$.

Problem 19.13. Compute the square (composition with itself) of the map given by $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.