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

Vitaliy Kurlin, vitaliy.kurlin@liverpool.ac.uk

Autumn 2018, Computer Science department

University of Liverpool, United Kingdom

Linear maps (transformations)

For high-dimensional data (not single values, but vectors of features), we need linear maps in $\mathbb{R}^n$.

Definition 19.1. A *linear map* (transformation, operator) on vectors $\vec{v} \in \mathbb{R}^n$ has the form $\vec{v} \mapsto A\vec{v}$, where A is an $n \times n$-matrix. Maps $\vec{v} \mapsto A\vec{v} + \vec{b}$, where $\vec{b} \in \mathbb{R}^n$ is a vector, are called *affine*.

The counterclockwise rotation around 0 in $\mathbb{R}^2$ through an angle β is the linear map $\vec{v} \mapsto A\vec{v}$ with $A = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$.

Transforming data by linear maps

A typical approach to vectorised data is to transform a data sample into a simpler form. The simplest possible transformation is a linear map.

It will be important to know if we lose any data under a linear map. Do the following matrices define invertible maps so that the original sample can be uniquely reconstructed from its image?

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

The determinant of a 3×3 matrix

Claim 19.2. A map $\vec{v} \mapsto A\vec{v}$ is invertible if and only if $\det A \neq 0$ (discussed in Lecture 3 for $n = 2$). For $n = 3$, we expand det along row 1:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} -$$

$$-a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} =$$

$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) +$
$a_{13}(a_{21}a_{32} - a_{22}a_{31})$, the 6 triple products in total.

The determinant of an $n \times n$ matrix

Definition 19.3. In a matrix A of size $n \times n$ for any element a_{ij}, its *minor* is the determinant of the $(n-1) \times (n-1)$ matrix A_{ij} obtained from the $n \times n$ matrix A by removing row i and column j.

The determinant is recursively defined by $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{1j} \det A_{1j}$. This formula is the expansion along row 1 similarly to $n = 3$ above.

This expansion works for any other row or column and gives the same answer (no proof needed).

Adding columns and rows

Claim 19.4. Swapping two columns changes the sign of the determinant. Adding any multiple of one column to another preserves the determinant.

Proof. $\det\begin{pmatrix} b & a \\ d & c \end{pmatrix} = bc - ad = -\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$\det\begin{pmatrix} a & b+ta \\ c & d+tc \end{pmatrix} = a(d+tc) - (b+ta)c =$

$ad - bc = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for any a, b, c, d, t. $\square$

These properties hold for rows and any dimension.

The properties of the determinant

[No proofs are needed] For any upper triangular matrix (having zeros below the diagonal), $\det A$ equals the product $\prod_{i=1}^{n} a_{ii}$ of diagonal elements.

Claim 19.5. $\det(AB) = \det(A)\det(B)$ for any $n \times n$ matrices A, B. Then $\det(AB) = \det(BA)$ even if $AB \neq BA$. If $AB = I$ is the identity matrix, i.e. $B = A^{-1}$ is the inverse, then $\det(A^{-1}) = \dfrac{1}{\det A}$.

So any invertible matrix A should have $\det A \neq 0$.

The transpose of a matrix

Definition 19.6. Any matrix A flipped over its diagonal gives the *transpose* A^T: $(A^T)_{ij} = a_{ji}$.

A matrix A is called *symmetric* if $A^T = A$.

A matrix A is called *orthogonal* if $A^{-1} = A^T$.

In $\mathbb{R}^2$ any rotation matrix $A = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$ is orthogonal. In any $\mathbb{R}^n$ orthogonal maps are high-dimensional analogues of plane rotations.

Properties of the transpose

Claim 19.7. $(A^T)^T = A$, $(A+B)^T = A^T + B^T$, $\det(A^T) = \det(A)$, $(AB)^T = B^T A^T$,$(A^T)^{-1} = (A^{-1})^T$.

Proof. The first two follow from Def 19.5, the 3rd follows from expansion along row 1 (or column 1) in Definition 19.3. The product property: AB has the element $\sum_{k=1}^{n} a_{ik} b_{kj}$ in row i, column j. Then

$$((AB)^T)_{ij} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} (B^T)_{ik} (A^T)_{kj} = (B^T A^T)_{ij}.$$

The last follows: $I = I^T = (AA^{-1})^T = (A^{-1})^T A^T$. $\square$

The trace of a matrix

Definition 19.8. The *trace* of a matrix A is the sum of the diagonal elements $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

Claim 19.9. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$, for any $c \in \mathbb{R}$ $\operatorname{tr}(cA) = c \cdot \operatorname{tr} A$, $\operatorname{tr}(A^T) = \operatorname{tr} A$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof. The first three follow from Def 19.7. Last:

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{ki} \right) =$$

$$= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} b_{ki} a_{ik} \right) = \sum_{k=1}^{n} (BA)_{kk} = \operatorname{tr}(BA). \qquad \square$$

Your questions and the quiz

To benefit from the lecture, now you could

- ask or submit your anonymous questions to the COMP229 folder after the lecture;
- write down your summary in 2–3 phrases, e.g. list key concepts you have learned;
- talk to your classmates to revise the lecture.

Question. Compute $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

Answer to the quiz and summary

Answer. $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 0$, because the 2nd column vanishes after subtracting the average of the 1st and 3rd columns, see Claim 19.4.

- The *determinant*: $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{1j} \det A_{1j}$ satisfies $\det(AB) = \det(A)\det(B)$.
- The *transpose* of a matrix: $(A^T)_{ij} = a_{ji}$.
- The *trace* of a matrix: $\mathrm{tr}\,(A) = \sum_{i=1}^{n} a_{ii}$.