COMP229: Introduction to Data Science
Lecture 13: Operations on vectors

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

Lecture plan

- Multiplication by a scalar
- Vector addition
- Scalar product
- The Cauchy indequality
- Orthogonal vectors

Reminder: free vectors

Bound (or fixed) vectors (with endpoints) $\overrightarrow{AB} = \overrightarrow{CD}$ iff $B_i - A_i = D_i - C_i$ for each i-th coordinate.

A free vector (without fixed endpoints) is a class of equal fixed vectors and can be defined by n coordinates $\vec{v} = (v_1, \dots, v_n)$ and can be represented by using the origin $B = (0, \dots, 0)$ as the tail point and $A = (v_1, \dots, v_n)$ as the head point.

The *length* $|\vec{v}| = \sqrt{\sum_{i=1}^{n} v_i^2}$, the *angle* $\angle(\vec{u}, \vec{v})$ is measured counter-clockwise from $\vec{u}$ to $\vec{v}$.

Equivalences on vectors

On the last lecture we've covered the following equivalences on vectors:

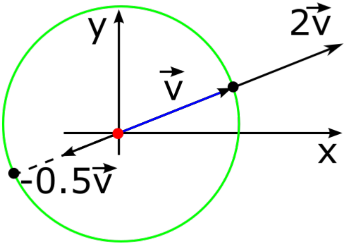
- vector length $\vec{u} \sim \vec{v}$ if $|\vec{u}| = |\vec{v}|$
- angles: all pairs (α, β) such that $\alpha - \beta$ is divisible by 2π
- collinearity (i.e. angle of 0 or π)

Each of them is useful in its own sence.

Multiplication by a scalar

Multiplication by a constant (scalar) $s \in \mathbb{R}$ is defined as $s\vec{v} = (sv_1, \ldots, sv_n)$.

The result is a vector of the length $|s| \cdot |\vec{v}|$ in the same (for $s > 0$) or opposite (for $s < 0$) direction.



For more information, see Khan academy.

Equivalence via multiplication

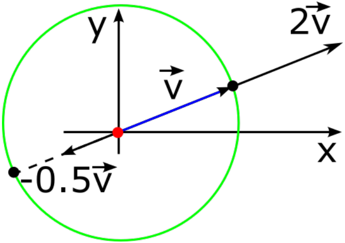
Problem 13.1. Let vectors $\vec{u} \sim \vec{v}$ be related if $\vec{u} = s\vec{v}$ for $s \in \mathbb{R}$. Is it an equivalence relation? What if $s > 0$? What are equivalence classes?

Solution 13.1. Not an equivalence if we allow $s = 0$, e.g. $(0,0) = 0 \cdot (0,1)$, but $(0,1) \neq s(0,0)$ for any s. If we require $s \neq 0$ or $s > 0$, let's check the axioms:
Symmetry: if $\vec{u} = s\vec{v}$ then $\vec{v} = \dfrac{1}{s}\vec{u}$.
Transitivity: if $\vec{u} = s\vec{v}$, $\vec{v} = t\vec{w}$, then $\vec{u} = st\vec{w}$.

So it is an equivalence for non-zero s.

The set of equivalence classes

For the equivalence that allows any $s > 0$, any vector $\vec{v} \neq \vec{0}$ has a canonical representative $\dfrac{\vec{v}}{|\vec{v}|}$ of length 1. The set of equivalence classes is the unit circle S^1 and the class of the zero vector $\{\vec{0}\}$.



If we allow all $s \neq 0$, diametrically opposite points of S^1 should be identified : or identify the endpoints of the angle interval $[0, \pi]$.

Why cannot we divide by 0?

If we set $\dfrac{s}{0} = t$, then multiplying both sides by 0 gives $s = 0 \cdot t = 0$, so s can be only 0.

Assume that $\dfrac{0}{0} = t$ makes sense,
then divide both sides of $1 \cdot 0 = 2 \cdot 0$ by 0,
get $1 \cdot \dfrac{0}{0} = 2 \cdot \dfrac{0}{0}$, hence $t = 2t$, so $0/0 = t = 0$.

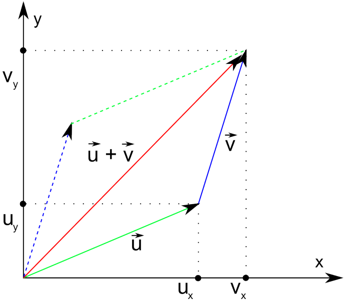
Then $\dfrac{1}{0} = \dfrac{1}{0/0} = \dfrac{1}{0} \cdot 0 = 0$.
Multiplying both sides above by 0 gives $1 = 0 \cdot 0$, a contradiction.

In other words: however we define division by 0, it will lead to collapsing all numbers into one equivalence class.
Hence for any real $s \in \mathbb{R}$, the expression $\dfrac{s}{0}$ is undefined.

Vector addition

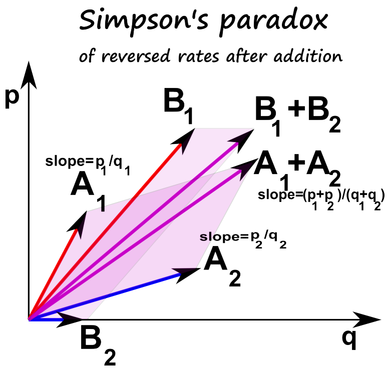
For $\vec{u} = (u_1, \ldots, u_n)$, $\vec{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$,
vector sum $\vec{u} + \vec{v} = (u_1 + v_1, \ldots, u_n + v_n)$.

The vector sum $\vec{u} + \vec{v}$ can be obtained by placing vectors head to tail and drawing the vector from the free tail to the free head (*the parallelogram rule*).



Reminder: vectors and probabilities

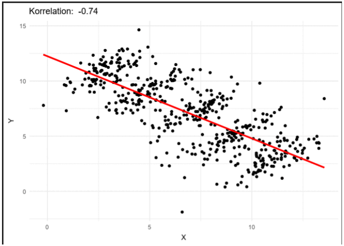
A probability $P_i = \frac{p_i}{q_i}$ is represented by a vector $A_i = (q_i, p_i)$ with the slope $\frac{p_i}{q_i}$. The vector sum is
$A_1 + A_2 = (q_1 + q_2, p_1 + p_2)$ with the slope $\frac{p_1+p_2}{q_1+q_2}$.

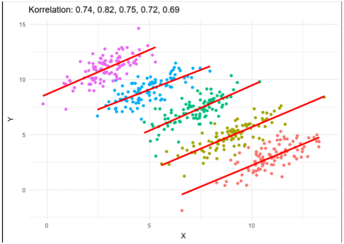


Slope of $\bar{B}_1$ is smaller than the slope of $\bar{A}_1$, and slope of $\bar{B}_2$ is smaller than the slope of $\bar{A}_2$, but the sum $\bar{B}_1 + \bar{B}_2$ has a larger slope than $\bar{A}_1 + \bar{A}_2$.

Simpson's paradox

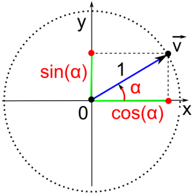
Simpson's (amalgamation) paradox: a trend in different groups can reverse when these groups are combined, see here.





The cosine and sine definitions

Definition 13.2. Fix orthogonal x, y-axes in $\mathbb{R}^2$. For an angle $\alpha \in [0, 2\pi)$, take the unit length vector $\vec{v} \in \mathbb{R}^2$ that has the tail point at the origin $\vec{0} \in \mathbb{R}^2$ and the angle α from the positive x-axis.



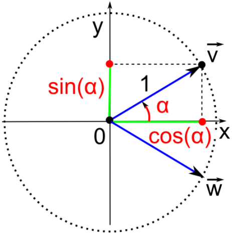
Then $\cos\alpha$ is the x-projection of the head point of $\vec{v}$, $\sin\alpha$ is the y-projection of the head point of $\vec{v}$.

For any $\beta \in \mathbb{R}$, find $\alpha \in [0, 2\pi)$ with $\alpha - \beta$ divisible by 2π. Set $\cos\beta = \cos\alpha$, $\sin\beta = \sin\alpha$ to make them periodic: $\cos(\alpha + 2\pi n) = \cos(\alpha)$ for $n \in \mathbb{Z}$.

The scalar (dot) product of vectors

Definition 13.3. For any vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$, their **scalar product** denoted by $\vec{u} \cdot \vec{v}$ or $\langle \vec{u}, \vec{v} \rangle$ or $(\vec{u}, \vec{v})$ equals $|\vec{u}| \cdot |\vec{v}| \cdot \cos \alpha \in \mathbb{R}$, where α is the angle from $\vec{u}$ to $\vec{v}$ in the plane $P \subset \mathbb{R}^n$ spanned by $\vec{u}, \vec{v}$.

Claim 13.4. The scalar product is symmetric: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$, i.e. α can be measured from $\vec{v}$ to $\vec{u}$.



Proof. $\cos(2\pi - \alpha) = \cos(\alpha)$, because the vector $\vec{w}$ with the angle $2\pi - \alpha$ from the x-axis is symmetric to $\vec{v}$ with respect to the x-axis.

The scalar product in coordinates

The above and other properties are easier to prove by using another definition in terms of coordinates.

Definition 13.5. (algebraic) For $\vec{u} = (u_1, \ldots, u_n)$, $\vec{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$, the **scalar product** is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

A good exercise is to prove that the geometric (13.3) and algebraic (13.5) definitions are equivalent.

Angle between vectors again

In $\mathbb{R}^n$ for $n > 2$, it is convenient to define the angle $\alpha \in [0, \pi]$ between vectors $\vec{u}, \vec{v}$ as the smallest of $\angle(\vec{u}, \vec{v})$ and $\angle(\vec{v}, \vec{u})$, otherwise we need to fix an orientation on the plane spanned by $\vec{u}, \vec{v}$.

Then the angle α is symmetric with respect to swapping $\vec{u}, \vec{v}$ is computed from the dot product $\vec{u} \cdot \vec{v} = |\vec{u}| \cdot |\vec{v}| \cdot \cos\alpha$ as $\alpha = \arccos \dfrac{\vec{u} \cdot \vec{v}}{|\vec{u}| \cdot |\vec{v}|}$.

For any $t \in [-1, 1]$, $\arccos(t)$ is a unique angle $\alpha \in [0, \pi]$ such that $\cos\alpha = t$, e.g. $\arccos(0) = \dfrac{\pi}{2}$.

The Cauchy inequality

Theorem 13.6. For any vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$, the **Cauchy inequality** states that $|(\vec{u} \cdot \vec{v})| \leqslant |\vec{u}| \cdot |\vec{v}|$.

Geometric proof: By the geometric definition $\vec{u} \cdot \vec{v} = |\vec{u}| \cdot |\vec{v}| \cos \alpha$, where α is the angle between $\vec{u}, \vec{v}$. The Cauchy inequality follows from $|\cos \alpha| \leqslant 1$.

In coordinates $\vec{u} = (u_1, \dots, u_n)$, $\vec{v} = (v_1, \dots, v_n)$, the Cauchy inequality looks more non-trivial:

$$\left| \sum_{i=1}^{n} u_i v_i \right| \leqslant \sqrt{\sum_{i=1}^{n} u_i^2} \cdot \sqrt{\sum_{i=1}^{n} v_i^2}.$$

Algebraic proof of Cauchy's inequality

The quadratic polynomial $f(t) = \sum_{i=1}^{n} (u_i t + v_i)^2 \geqslant 0$ is non-negative for all $t \in \mathbb{R}$ and can be rewritten as

$$f(t) = \left(\sum_{i=1}^{n} u_i^2 \right) t^2 + 2 \left(\sum_{i=1}^{n} u_i v_i \right) t + \left(\sum_{i=1}^{n} v_i^2 \right).$$

A quadratic polynomial $at^2 + 2bt + c \geqslant 0$ is non-negative if and only if $b^2 - ac \leqslant 0$, or $b^2 \leqslant ac$, hence

$$b^2 = \left(\sum_{i=1}^{n} u_i v_i \right)^2 \leqslant \left(\sum_{i=1}^{n} u_i^2 \right) \left(\sum_{i=1}^{n} v_i^2 \right) = ac,$$ which is

equivalent to Cauchy's $\left| \sum_{i=1}^{n} u_i v_i \right| \leqslant \sqrt{\sum_{i=1}^{n} u_i^2} \cdot \sqrt{\sum_{i=1}^{n} v_i^2}$ and

$|(\vec{u} \cdot \vec{v})| \leqslant |\vec{u}| \cdot |\vec{v}|$.

Orthogonal vectors

Claim 13.7. Any non-zero vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ are *orthogonal*, i.e. the angle between them is $\pm\frac{\pi}{2}$ (plus any multiple of 2π) *if and only if* $\vec{u} \cdot \vec{v} = 0$.

Proof. $\vec{u} \cdot \vec{v} = |\vec{u}| \cdot |\vec{v}| \cdot \cos\alpha$ equals 0 if and only if $\cos\alpha = 0$, $\alpha = \pm\frac{\pi}{2} + 2\pi n$ for any integer $n \in \mathbb{Z}$.

Problem 13.8. Write down a vector orthogonal (perpendicular) to a given vector $\vec{v} = (x, y) \in \mathbb{R}^2$.

We'll express the scalar product in coordinates.

Finding an orthogonal vector

Solution 13.8. For any vector $\vec{v} = (x, y) \in \mathbb{R}^2$, the vectors $\vec{u} = (y, -x)$ and $\vec{w} = (-y, x)$ (or any proportional to them) are orthogonal to $\vec{v}$, because
$\vec{v} \cdot \vec{u} = xy + y(-x) = 0$, $\vec{v} \cdot \vec{w} = x(-y) + yx = 0$.

Claim 13.9. The scalar product respects scalar multiplication: $(s\vec{u}) \cdot (\vec{v}) = s(\vec{u} \cdot \vec{v})$ for $s \in \mathbb{R}$, and sums:
$(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$ (distributive rule).

The properties follow from the algebraic definition.

Time to revise and ask questions

- The *scalar* product of vectors $\vec{u} = (u_1, \ldots, u_n)$, $\vec{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ is $\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i = |\vec{u}| \cdot |\vec{v}| \cos \alpha$, here α is the angle between $\vec{u}, \vec{v}$.
- $\vec{u}, \vec{v}$ are orthogonal if and only if $\vec{u} \cdot \vec{v} = 0$.
- Cauchy's inequality $|(\vec{u} \cdot \vec{v})| \leqslant |\vec{u}| \cdot |\vec{v}|$.

Problem 13.10. Compute the pairwise scalar products of $\vec{u} = (2, 1)$, $\vec{v} = (1, 2)$, $\vec{w} = (-1, 2)$.