2: Vectors and Dot Product

Two points $P = (a, b, c)$ and $Q = (x, y, z)$ in space define a **vector** $\vec{v} = \langle x - a, y - b - z - c\rangle$. It points from P to Q and we write also $\vec{v} = \vec{PQ}$. The real numbers numbers p, q, r in a vector $\vec{v} = \langle p, q, r\rangle$ are called the **components** of $\vec{v}$.

Vectors can be drawn **everywhere** in space but two vectors with the same components are considered **equal**. Vectors can be translated into each other if and only if their components are the same. If a vector starts at the origin $O = (0, 0, 0)$, then the vector $\vec{v} = \langle p, q, r\rangle$ points to the point (p, q, r). One can therefore identify points $P = (a, b, c)$ with vectors $\vec{v} = \langle a, b, c\rangle$ attached to the origin. To make more clear which objects are vectors, we sometimes draw an arrow on top of it and if $\vec{v} = \vec{PQ}$ then P is the "tail" and Q is the "head" of the vector. To distinguish vectors from points, it is custom to different brackets and write $\langle 2, 3, 4\rangle$ for vectors and $(2, 3, 4)$ for points.

The **sum** of two vectors is $\vec{u} + \vec{v} = \langle u_1, u_2\rangle + \langle v_1, v_2\rangle = \langle u_1 + v_1, u_2 + v_2\rangle$. The **scalar multiple** $\lambda\vec{u} = \lambda\langle u_1, u_2\rangle = \langle \lambda u_1, \lambda u_2\rangle$. The difference $\vec{u} - \vec{v}$ can best be seen as the addition of $\vec{u}$ and $(-1) \cdot \vec{v}$.

The vectors $\vec{i} = \langle 1, 0\rangle$, $\vec{j} = \langle 0, 1\rangle$ are called **standard basis vectors** in the plane. In space, one has the basis vectors $\vec{i} = \langle 1, 0, 0\rangle$, $\vec{j} = \langle 0, 1, 0\rangle$, $\vec{k} = \langle 0, 0, 1\rangle$.

Every vector $\vec{v} = \langle p, q\rangle$ in the plane can be written as a combination $\vec{v} = p\vec{i} + q\vec{j}$ of standard basis vectors and every vector $\vec{v} = \langle p, q, r\rangle$ in space can be written as $\vec{v} = p\vec{i} + q\vec{j} + r\vec{k}$. Vectors are abundant in applications. They appear in mechanics: if $\vec{r}(t) = \langle f(t), g(t)\rangle$ is a point in the plane which depends on time t, then $\vec{v} = \langle f'(t), g'(t)\rangle$ will be called the **velocity vector** at $\vec{r}(t)$. Here $f'(t), g'(t)$ are the derivatives. In physics, we often want to determine forces acting on objects. Forces are represented as vectors. In particular, electromagnetic or gravitational fields or velocity fields in fluids are described by vectors. Vectors appear also in computer science: the scalable vector graphics is a standard for the web for describing two-dimensional graphics. In quantum computation, rather than working with bits, one deals with **qbits**, which are vectors. Finally, **color** can be written as a vector $\vec{v} = \langle r, g, b\rangle$, where r is **red**, g is **green** and b is **blue** component of the color vector. An other coordinate system for color is $\vec{v} = \langle c, m, y\rangle = \langle 1 - r, 1 - g, 1 - b\rangle$, where c is **cyan**, m is **magenta** and y is **yellow**. Vectors appear in probability theory and statistics. On a finite probability space, a **random variable** is a vector.

The addition and scalar multiplication of vectors satisfy the laws you know from **arithmetic**. **commutativity** $\vec{u} + \vec{v} = \vec{v} + \vec{u}$, **associativity** $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ and $r * (s * \vec{v}) = (r * s) * \vec{v}$ as well as **distributivity** $(r + s)\vec{v} = \vec{v}(r + s)$ and $r(\vec{v} + \vec{w}) = r\vec{v} + r\vec{w}$, where $*$ denotes multiplication with a scalar.

The **length** $|\vec{v}|$ of a vector $\vec{v} = \vec{PQ}$ is defined as the distance $d(P, Q)$ from P to Q. A vector of length 1 is called a **unit vector**. If $\vec{v} \neq \vec{0}$, then $\vec{v}/|\vec{v}|$ is a unit vector.

1 $|\langle 3, 4\rangle| = 5$ and $|\langle 3, 4, 12\rangle| = 13$. Examples of unit vectors are $|\vec{i}| = |\vec{j}| = \vec{k}| = 1$ and $\langle 3/5, 4/5\rangle$ and $\langle 3/13, 4/13, 12/13\rangle$. The only vector of length 0 is the zero vector $|\vec{0}| = 0$.

The **dot product** of two vectors $\vec{v} = \langle a, b, c\rangle$ and $\vec{w} = \langle p, q, r\rangle$ is defined as $\vec{v} \cdot \vec{w} = ap + bq + cr$.

Remarks.
a) Different notations for the dot product are used in different mathematical fields. while pure mathematicians write $\vec{v} \cdot \vec{w} = (\vec{v}, \vec{w})$, one can see $\langle \vec{v}|\vec{w}\rangle$ in quantum mechanics or $v_i w^i$ or more generally $g_{ij} v^i w^j$ in general relativity. The dot product is also called **scalar product** or **inner product**.
b) Any product $g(v, w)$ which is linear in v and w and satisfies the symmetry $g(v, w) = g(w, v)$ and $g(v, v) \geq 0$ and $g(v, v) = 0$ if and only if $v = 0$ can be used as a dot product. An example is $g(v, w) = 3v_1 w_1 + 2v_2 w_2 + v_3 w_3$.

The dot product determines distance and distance determines the dot product.

Proof: Lets write $v = \vec{v}$ in this proof. Using the dot product one can express the length of v as $|v| = \sqrt{v \cdot v}$. On the other hand, from $(v + w) \cdot (v + w) = v \cdot v + w \cdot w + 2(v \cdot w)$ can be solved for $v \cdot w$:

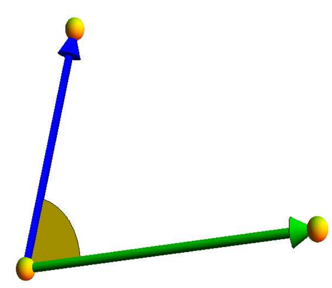
$$v \cdot w = (|v + w|^2 - |v|^2 - |w|^2)/2 .$$

The **Cauchy-Schwarz inequality** tells $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$.

Proof. We can assume $|w| = 1$ after scaling the equation. Now plug in $a = v \cdot w$ into the equation $0 \leq (v - aw) \cdot (v - aw)$ to get $0 \leq (v - (v \cdot w)w) \cdot (v - (v \cdot w)w) = |v|^2 + (v \cdot w)^2 - 2(v \cdot w)^2 = |v|^2 - (v \cdot w)^2$ which means $(v \cdot w)^2 \leq |v|^2$.

Having established this, we have a clean definition of what an **angle** is:

The **angle** between two nonzero vectors is defined as the unique $\alpha \in [0, \pi]$ which satisfies $\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)$.



Al Kashi's theorem: If a, b, c are the side lengths of a triangle ABC and α is the angle opposite to c, then $a^2 + b^2 = c^2 - 2ab\cos(\alpha)$.

Proof. Define $\vec{v} = \vec{AB}, \vec{w} = \vec{AC}$. Because $c^2 = |\vec{v} - \vec{w}|^2 = (\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = |\vec{v}|^2 + |\vec{w}|^2 - 2\vec{v} \cdot \vec{w}$, We know $\vec{v} \cdot \vec{w} = |\vec{v}| \cdot |\vec{w}| \cos(\alpha)$ so that $c^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}| \cdot |\vec{w}| \cos(\alpha) = a^2 + b^2 - 2ab\cos(\alpha)$.

The angle definition works in any space with a dot product. In statistics you have to work with vectors of n components. They are called data or random variables and $\cos(\alpha)$ is called the **correlation** between two random variables $\vec{v}, \vec{w}$ of zero **expectation** $\mathrm{E}[\vec{v}] = (v_1 + \cdots + v_n)/n$. The dot product $v_1 w_1 + \ldots + v_n w_n$ is then the **covariance**, the length $|v|$ is the **standard deviation** and denoted by $\sigma(v)$. The formula $\mathrm{Corr}[v,w] = \mathrm{Cov}[v,w]/(\sigma(v)\sigma(w))$ for the correlation is the familiar angle formula we have seen. It is geometry in n dimensions. We mention this only to convince you that the geometry we do here can be applied to much more. All the computations we have done go through verbatim.

The **triangle inequality** tells $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$

Proof: $|\vec{u}+\vec{v}|^2 = (\vec{u}+\vec{v}) \cdot (\vec{u}+\vec{v}) = \vec{u}^2 + \vec{v}^2 + 2\vec{u} \cdot \vec{v} \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u} \cdot \vec{v}| \leq \vec{u}^2 + \vec{v}^2 + 2|\vec{u}| \cdot |\vec{v}| = (|\vec{u}| + |\vec{v}|)^2$.

Two vectors are called **orthogonal** or **perpendicular** if $\vec{v} \cdot \vec{w} = 0$. The zero vector $\vec{0}$ is orthogonal to any vector. For example, $\vec{v} = \langle 2, 3 \rangle$ is orthogonal to $\vec{w} = \langle -3, 2 \rangle$.

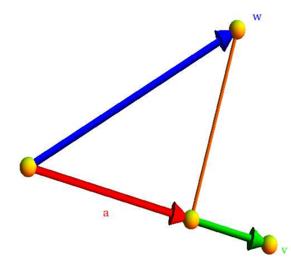
Having given precise definitions of all objects we can now prove **Pythagoras theorem**:

Pythagoras theorem: if $\vec{v}$ and $\vec{w}$ are orthogonal, then $|v - w|^2 = |v|^2 + |w|^2$.

Proof: $(\vec{v} - \vec{w}) \cdot (\vec{v} - \vec{w}) = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w} + 2\vec{v} \cdot \vec{w} = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w}$.

Remarks:
1) You have just seen something very powerful: results like Pythagoras (570-495BC) and Al Khashi (1380-1429) theorems were **derived from scratch** on a space V equipped with a dot product. The dot product appeared much later in mathematics (Hamilton 1843, Grassman 1844, Sylvester 1851, Cayley 1858). While we have used geometry as an intuition, the structure was built algebraically without any unjustified assumptions. This is mathematics: if we have a space V in which addition $\vec{v} + \vec{w}$ and scalar multiplication $\lambda\vec{v}$ is given and in which a dot product is defined, then all the just derived results apply. We have **not used** results of Al Khashi or Pythagoras but we have **derived** them and additionally obtained a **clear definition** what an angle is.
2) The derivation you have seen works in any dimension. Why do we care about higher dimensions? As already mentioned a compelling motivation is **statistics**. Given 12 data points like the average monthly temperatures in a year, we deal with a 12 dimensional space. Geometry is useful to describe data. Pythagoras theorem is the property that the variance of two uncorrelated random variables adds up with the formula $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.
3) A far reaching generalization of the geometry you have just seen is obtained if the dot product $g(v, w)$ is allowed to depend on the place, where the two vectors are attached. This produces **Riemannian geometry** and allows to work with spaces which are intrinsically **curved**. This mathematics is important in **general relativity** which describes gravity in a geometric way and which is one of the pillars of modern physics. But it appears in daily life too. On a hot summer day, if you look close at an object a hot asphalt street, the object can appear distorted or flickers. The dot product depends on the temperature of the air. Light rays no more move on straight lines but gets bent. In extreme cases, when the curvature of light rays is larger than the curvature of the earth, it leads to **Fata morgana** effects: one can see objects which are located beyond the horizon.
4) Why do we not introduce vectors not just as algebraic objects $\langle 1, 2, 3 \rangle$? The reason is that in many applications like physics and even geometry, one wants to work with **affine vectors**, vectors which are attached at points. Forces for example act on points of a body, we will also look at vector fields, where at each point a vector is attached. Considering vectors with the same components as equal gives then the vector space in which we do the algebra. One could define a vector space axiomatically and then build from this affine vectors but it is a bit too abstract and not much is actually gained for the goals we have in mind. An even more modern point of view replaces affine vectors with members of a tangent bundle. But this is only necessary if one deals with spaces which are not flat. Even more general is to allow the space attached at each point to be a more general space like a "group" called fibres. So called "fibre bundles" are the framework of mathematical concepts which describe elementary particles or even space itself. Attaching a circle for example at each point leads to electromagnetism attaching classes of two dimensional matrices leads to the weak force and attaching certain three dimensional matrices leads to the strong force. Allowing this to happen in a curved framework incorporates gravity. One of the main challenges is to include quantum mechanics into that picture. Fundamental physics has become primarily the quest to answer the question "what is space"?

The vector $\mathrm{P}(\vec{v}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{w}|^2}\vec{w}$ is called the **projection** of $\vec{v}$ onto $\vec{w}$. The **scalar projection** $\frac{\vec{v} \cdot \vec{w}}{|\vec{w}|}$ is a signed length of the vector projection. Its absolute value is the length of the projection of $\vec{v}$ onto $\vec{w}$. The vector $\vec{b} = \vec{v} - P(\vec{v})$ is a vector orthogonal to the $\vec{w}$-direction.



2 For example, with $\vec{v} = \langle 0, -1, 1 \rangle$, $\vec{w} = \langle 1, -1, 0 \rangle$, $\mathrm{P}(\vec{v}) = \langle 1/2, -1/2, 0 \rangle$. Its length is $1/\sqrt{2}$.

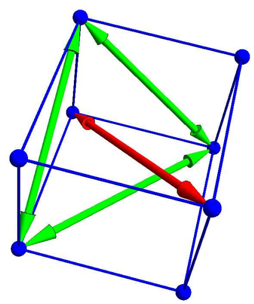
3 Projections are important in physics. For example, if you apply a wind force $\vec{F}$ to a car which drives in the direction $\vec{w}$ and P denotes the projection on $\vec{w}$ then $P(\vec{F})$ is the force which accelerates or slows down the car.

The projection allows to visualize the dot product. The absolute value of the dot product is the length of the projection. The dot product is positive if v points more towards to w, it is negative if v points away from it. In the next lecture we use the projection to compute distances between various objects.

Homework

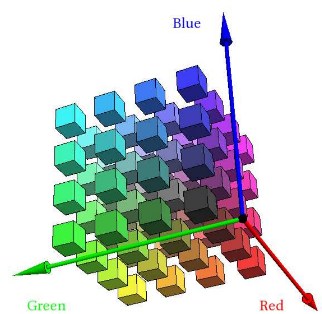
1 Find a unit vector parallel to $\vec{u} - \vec{v}$ if $\vec{u} = \langle 5, 6, 3 \rangle$ and $\vec{v} = \langle 1, 1, 3 \rangle$.

2 An **Euler brick** is a cuboid of dimensions a, b, c such that all face diagonals are integers.
a) Verify that $\vec{v} = \langle a, b, c \rangle = \langle 240, 117, 44 \rangle$ is a vector which leads to an Euler brick.
b) Verify that $\langle a, b, c \rangle = \langle u(4v^2 - w^2), v(4u^2 - w^2), 4uvw \rangle$ leads to an Euler brick if $u^2 + v^2 = w^2$.
If also the space diagonal $\sqrt{a^2 + b^2 + c^2}$ is an integer, an Euler brick is called **perfect**. Nobody has found one, nor proven that it can not exist.



3 **Colors** are encoded by vectors $\vec{v} = \langle$ red , brightgreen , blue $\rangle$. The red, green and blue components of $\vec{v}$ are all real numbers in the interval $[0, 1]$.
a) Determine the angle between the colors yellow and cyan.
b) What is the projection of the mixture $(\vec{v} + \vec{w})/2$ of magenta and orange onto blue?

(0,0,0)	black	(0,0,1)	blue
(1,1,1)	white	(1,1,0)	yellow
$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	gray	(1,0,1)	magenta
(1,0,0)	red	(0,1,1)	cyan
(0,1,0)	green	$(1, \frac{1}{2}, 0)$	orange
$(0, 1, \frac{1}{2})$	spring green	$(1, 1, \frac{1}{2})$	khaki
$(1, \frac{1}{2}, \frac{1}{2})$	pink	$(\frac{1}{2}, \frac{1}{4}, 0)$	brown



4 Find the angle between the diagonal of the unit cube and one of the diagonal of one of its faces. Assume that the two diagonals go through the same edge of the cube. You can leave the answer in the form $\cos(\alpha) = ...$

5 Assume $\vec{v} = \langle -4, 2, 2 \rangle$ and $\vec{w} = \langle 3, 0, 4 \rangle$.
a) Find the vector projection of $\vec{v}$ onto $\vec{w}$.
b) Find the scalar component of $\vec{v}$ on $\vec{w}$.