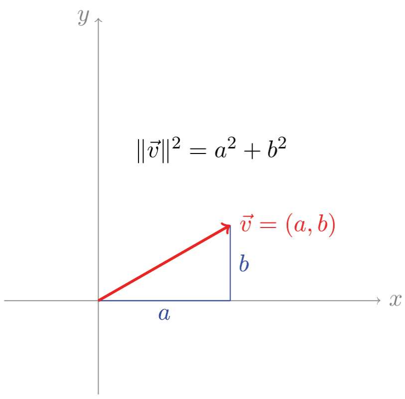
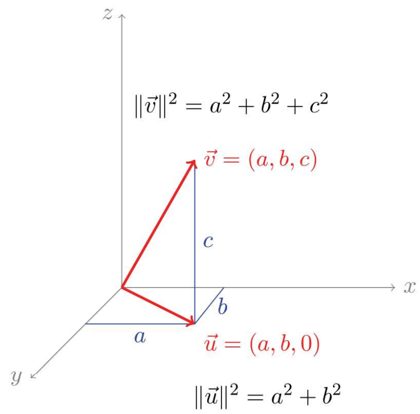
Euclidean geometry

- Length of a vector





- We would like Pythagoras theorem to hold in general, so we define the length of a vector $u = (u_1, u_2, \cdots, u_n) \in \mathbb{R}^n$ as $\|u\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$.
- The distance between two points in $\mathbb{R}^n$ is then defined as

> For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the **distance between u and v**, written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,
>
> $$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

- It is clear that for any scalar $c \in \mathbb{R}^2$

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

- A vector of length 1 is said to be a unit vector. For $v \in \mathbb{R}^n$, it is clear that $v/\|v\|$ is of length 1, and it is also in the direction of v. We call it the normalization of v.

- Given two vector $u = (u_1, u_2, \cdots, u_n)$ and $v = (v_1, v_2, \cdots, v_n)$, we define their inner product（内积）as $\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \cdots u_n v_n$. It is also called dot product（点乘） and denoted by $u \cdot v$. In terms of matrix, it can be written as

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- By definition, $\|u\|^2 = \langle u, u \rangle$.

- The inner product satisfies the properties

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

a. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

b. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

c. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$

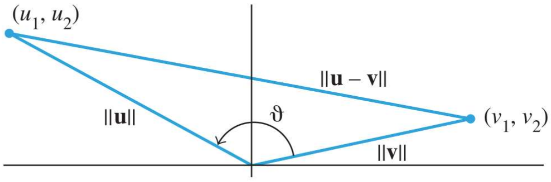
d. $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

- Combining b) and c), we can get

$$(c_1 \mathbf{u}_1 + \cdots + c_p \mathbf{u}_p) \cdot \mathbf{w} = c_1 (\mathbf{u}_1 \cdot \mathbf{w}) + \cdots + c_p (\mathbf{u}_p \cdot \mathbf{w})$$

- The inner product is very useful in describing the angles between two vectors.

- Angle between two vectors



- $\langle u, v\rangle = \|u\|\|v\|\cos\theta$. The formula holds also in R^n as any two vectors lie on a plane.

By the Law of cosine (余弦定理).

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\cdot\|v\|\cdot\cos\theta$$

$$\Rightarrow \quad 2\|u\|\|v\|\cos\theta = \langle u, u\rangle + \langle v, v\rangle - \langle u-v, u-v\rangle$$

$$= \langle u, u\rangle + \langle v, v\rangle - (\langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle)$$

$$= 2\langle u, v\rangle$$

- Reminder on the law of cosines

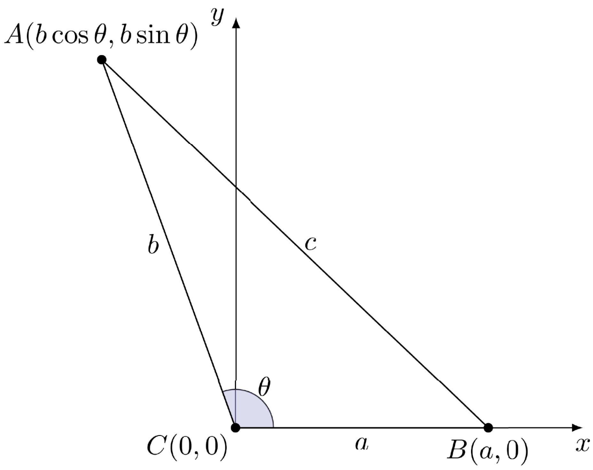
The coordinate of A can be calculated to be $(b\cos\theta, b\sin\theta)$.

Hence

$$c^2 = \|AB\|^2 = (b\cos\theta - a)^2 + (b\sin\theta - 0)^2$$

$$= b^2\cos^2\theta - 2ab\cos\theta + a^2 + b^2\sin^2\theta$$

$$= a^2 + b^2 - 2ab\cos\theta.$$



- In particular, two vectors u and v in R^n are orthogonal（垂直）if and only if $\langle u, v\rangle = 0$.
- The formula can be rewritten $\cos\theta = \dfrac{\langle u, v\rangle}{\|u\|\|v\|}$. Notice that both $\dfrac{u}{\|u\|}$ and $\dfrac{v}{\|v\|}$ have length 1.
- Theorem

The Pythagorean Theorem

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Proof

$$\|\vec{u}+\vec{v}\|^2 = \langle \vec{u}+\vec{v}, \vec{u}+\vec{v}\rangle = \langle \vec{u}, \vec{u}\rangle + \langle \vec{u}, \vec{v}\rangle + \langle \vec{v}, \vec{u}\rangle + \langle \vec{v}, \vec{v}\rangle$$

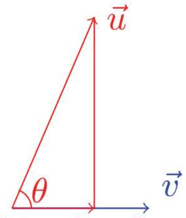
$$= \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\langle \vec{u}, \vec{v}\rangle$$

So $\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|v\|^2$ iff $\langle \vec{u}, \vec{v}\rangle = 0$, i.e. $\vec{u}$ and $\vec{v}$ are orthogonal. $\square$

- The orthogonal projection of u in the direction of v is $\dfrac{\langle u, v\rangle}{\|v\|^2}v$.

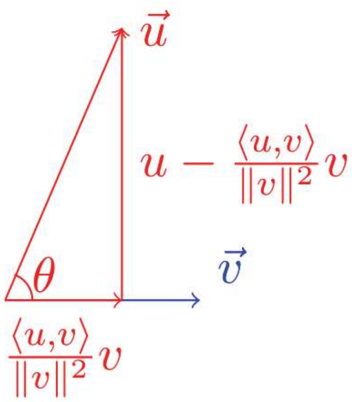
The vector has length $\|u\| \cdot \cos\theta = \|u\| \cdot \dfrac{\langle u, v\rangle}{\|u\| \cdot \|v\|}$, and it is in the direction of v, hence equals

$$\|u\| \cdot \cos\theta \cdot \frac{v}{\|v\|} = \frac{\langle u, v\rangle}{\|v\|^2} \cdot v$$



projection of $\vec{u}$ in the direction of $\vec{v}$

- Consequently, $u - \dfrac{\langle u, v\rangle}{\|v\|^2}v$ is the component of u orthogonal to v.



- The normal vector（法向量） of a hyperplane (超平面)

Observe that

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

$$\Updownarrow$$

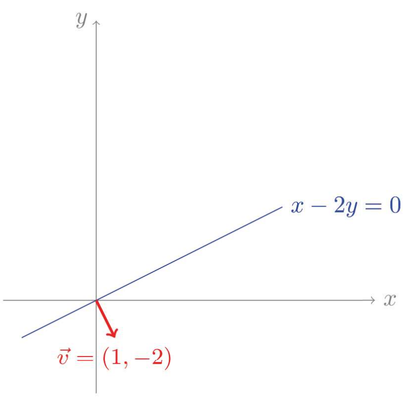
$$\langle \vec{a}, \vec{x} \rangle = 0 \quad \text{with} \quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

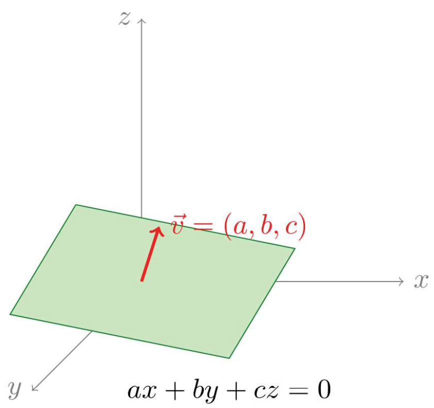
$$\Updownarrow$$

$\vec{x}$ is orthogonal to $\vec{a}$

Let H be the set of solution. it is then a sub vector space of dim $n-1$, consisting of vectors orthogonal to $\vec{a}$. We call $\vec{a}$ the normal vector of H.

- Examples





- In general, it is known that the set of solutions H_b of the equation $a_1x_1 + \cdots + a_nx_n = b$ is a translation of the sub vector space H. In particular, H_b is parallel to H, and they have the same dimension $n-1$. For this reason, we call H_b a hyperplane in $\mathbb{R}^n$ with normal vector $\mathbf{a}$.
- This can be generalized to the set of solutions of a linear system $A\mathbf{x} = \mathbf{b}$.
- Definition: Let W be a sub vector space of $\mathbb{R}^n$, a vector $\mathbf{v}$ is said to be orthogonal to W if $\mathbf{v}$ is orthogonal to all the vectors in W. The set of all the vectors $\mathbf{v}$ which are orthogonal to W is called the orthogonal complement (正交补) of W, denoted $W^{\perp}$.

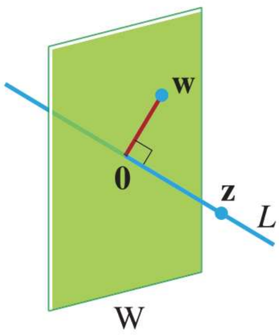


FIGURE 7

A plane and line through $\mathbf{0}$ as orthogonal complements.

- Proposition

1. A vector $\mathbf{x}$ is in $W^{\perp}$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

Proof 1) The necessity is clear. For the sufficiency, suppose that $W = \text{Span}\{\vec{w}_1, \cdots, \vec{w}_r\}$ and $\langle \vec{x}, \vec{w}_i \rangle = 0$ for $i = 1, \cdots, r$, then any vector in W is of the form

$$\vec{w} = \lambda_1 \vec{w}_1 + \cdots + \lambda_r \vec{w}_r \quad \text{for some } \lambda_1, \cdots, \lambda_r \in \mathbb{R}$$

and $\langle \vec{x}, \vec{w} \rangle = \lambda_1 \langle \vec{x}, \vec{w}_1 \rangle + \cdots + \lambda_r \langle \vec{x}, \vec{w}_r \rangle = 0$ by the linearity of inner product.

2) Need to show that $W^{\perp}$ is closed under vector addition and scalar multiplication.

For $\vec{v}_1, \vec{v}_2 \in W^{\perp}$, and any $\vec{w} \in W$, we have

$$\langle \vec{v}_1 + \vec{v}_2, \vec{w} \rangle = \langle \vec{v}_1, \vec{w} \rangle + \langle \vec{v}_2, \vec{w} \rangle = 0 + 0 = 0$$

So $\vec{v}_1 + \vec{v}_2 \in W^{\perp}$.

For $\lambda \in \mathbb{R}$, $\vec{v} \in W^{\perp}$, we have

$$\langle \lambda \vec{v}, \vec{w} \rangle = \lambda \langle \vec{v}, \vec{w} \rangle = \lambda \cdot 0 = 0$$

$\Rightarrow \lambda \vec{v} \in W^{\perp}$. $\square$

- Theorem

> Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:
>
> $$(\text{Row } A)^{\perp} = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

$\mathbb{R}^n$ $\mathbb{R}^m$

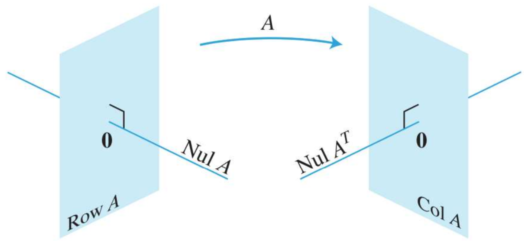


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

Proof The first assertion implies the second because

$$\text{Col}(A) = \text{Row}(A^t).$$

For the first assertion, let $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{bmatrix}$,

then $\text{Row}(A) = \text{Span}\{\vec{a}_1, \cdots, \vec{a}_m\}$.

Recall that $\text{Nul}(A)$ is set of solution of the equation $A\vec{x} = 0$,

i.e.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \iff \begin{matrix} \langle \vec{a}_1, \vec{x} \rangle = 0 \\ \langle \vec{a}_2, \vec{x} \rangle = 0 \\ \vdots \\ \langle \vec{a}_m, \vec{x} \rangle = 0 \end{matrix} \quad (*)$$

By the previous proposition, (*) is equivalent to

$$\vec{x} \in \text{Row}(A)^{\perp}.$$

In other words,

$$\vec{x} \in \text{Nul}(A) \iff x \in \text{Row}(A)^{\perp}$$

So $\text{Nul}(A) = \text{Row}(A)^{\perp}$. $\boxed{\text{S}}$

Orthogonal sets and orthonormal basis

- Definition

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

- Theorem

If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

Proof The second assertion follows from the first one by definition. Suppose that $\lambda_1 \vec{u}_1 + \cdots + \lambda_p \vec{u}_p = 0$ for some $\lambda_1, \cdots, \lambda_p \in \mathbb{R}$, then for $i = 1, \cdots, p$,

$$0 = \langle \vec{u}_i, \lambda_1 \vec{u}_1 + \cdots + \lambda_p \vec{u}_p \rangle$$

$$= \lambda_1 \underbrace{\langle \vec{u}_i, \vec{u}_1 \rangle}_{= 0} + \cdots + \lambda_i \underbrace{\langle \vec{u}_i, \vec{u}_i \rangle}_{\neq 0} + \cdots + \lambda_p \underbrace{\langle \vec{u}_i, \vec{u}_p \rangle}_{= 0}$$

$$= \lambda_i \cdot \underbrace{\| u_i \|^2}_{> 0}$$

$\Rightarrow \lambda_i = 0 \quad \Rightarrow \vec{u}_1, \cdots, \vec{u}_p$ are linearly independent. $\square$

- Definition

An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

- Theorem

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \qquad (j = 1, \ldots, p)$$

Proof: Consider the inner product $\langle \vec{y}, \vec{u}_i \rangle$, we have

$$\langle y, \vec{u}_i \rangle = \langle c_1 \vec{u}_1 + \cdots + c_n \vec{u}_n , \vec{u}_i \rangle$$

$$= c_1 \underbrace{\langle \vec{u}_1 , \vec{u}_i \rangle}_{=0} + \cdots + c_i \underbrace{\langle \vec{u}_i , \vec{u}_i \rangle}_{\neq 0} + \cdots + c_n \underbrace{\langle \vec{u}_n , \vec{u}_i \rangle}_{=0}$$

$$= c_i \langle \vec{u}_i , \vec{u}_i \rangle$$

$$\Rightarrow \quad c_i = \frac{\langle y , \vec{u}_i \rangle}{\| u_i \|^2} .$$

- Definition

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

- Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Proof Let $U = [\vec{u}_1 \cdots \vec{u}_n]$ with $\vec{u}_i \in \mathbb{R}^m$, then $U^t = \begin{bmatrix} \vec{u}_1^{\,t} \\ \vdots \\ \vec{u}_n^{\,t} \end{bmatrix}$

and $$U^t \cdot U = \begin{bmatrix} \vec{u}_1^{\,t} \\ \vdots \\ \vec{u}_n^{\,t} \end{bmatrix} [\vec{u}_1 \cdots \vec{u}_n] \overset{(*)}{=} \begin{pmatrix} \langle \vec{u}_1, \vec{u}_1 \rangle & \langle \vec{u}_1, \vec{u}_2 \rangle & \cdots & \langle \vec{u}_1, \vec{u}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \vec{u}_n, \vec{u}_1 \rangle & \langle \vec{u}_n, \vec{u}_2 \rangle & \cdots & \langle \vec{u}_n, \vec{u}_n \rangle \end{pmatrix}$$

So U has orthonormal columns iff

$$\langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}$$

$$\Leftrightarrow \quad U^t \cdot U = I_n$$

- Remark: (*) is due to the equality

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- Theorem

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$

b. $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$

c. $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$

- Geometrically, this means that U defines a linear transformation which preserves the length of vectors, we call it an isometric transformation (等距变换).
- If moreover U is a square matrix, we call it an orthogonal matrix (正交矩阵). It defines an isometric (等距) isomorphism of $\mathbb{R}^n$. The condition that U has orthonormal columns is then equivalent to the condition $U^t \cdot U = I_n$.

Proof By the previous theorem, we have

$$U^t \cdot U = I_n$$

multiplication of matrix

So $\langle U\vec{x}, U\vec{y} \rangle = (U\vec{x})^t \cdot (U\vec{y})$

$$= (\vec{x}^t \cdot U^t) \cdot (U\vec{y})$$

$$= \vec{x}^t \cdot \underbrace{(U^t U)}_{= I_n} \cdot \vec{y}$$

$$= \vec{x}^t \cdot \vec{y} = \langle \vec{x}, \vec{y} \rangle$$

This implies all the three assertions of the theorem. $\boxed{\text{}}$