Sums and direct sums of subspaces

Let V be a vector space and W_1, W_2 be two subspaces of V. We know that the intersection $W_1 \cap W_2$ is a subspace of V. The union $W_1 \cup W_2$, on the other hand, is not.

There is a way to "combine" two subspaces W_1, W_2 of V which yields a subspace that "contains" both W_1 and W_2 in some sense. It consists in "summing" W_1 and W_2 as follows.

Defn. Given subspaces W_1, W_2 of V, the sum denoted $W_1 + W_2$ of W_1 and W_2 is the set of all vectors v in V that can be expressed as a sum of an element from W_1 and an element of W_2. In other words

$$W_1 + W_2 = \{ v \in V \mid v = w_1 + w_2, \quad w_1 \in W_1, w_2 \in W_2 \}.$$

Ex. $V = \mathbb{R}^3$, $W_1 = \text{span}\{(1,0,0)\}$ "x-axis"

$W_2 = \text{span}\{(0,1,0)\}$ "y-axis"

Then $W_1 + W_2$ is $\text{span}\{(1,0,0),(0,1,0)\}$ "xy-plane".

Lemma Let V be a vector space and W_1, W_2 be two subspaces of V. Then

(a) $W_1 + W_2$ is a subspace of V.

(b) $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$.

proof (a) Let $v, z \in W_1 + W_2$. Then

$$v = v_1 + v_2, \quad v_1 \in W_1, v_2 \in W_2$$
$$z = z_1 + z_2, \quad z_1 \in W_1, z_2 \in W_2.$$

$$\Rightarrow \quad v + z = (v_1 + z_1) + (v_2 + z_2).$$

Now
$v_1, z_1 \in W_1 \Rightarrow v_1 + z_1 \in W_1$ since W_1 is a subspace
$v_2, z_2 \in W_2 \Rightarrow v_2 + z_2 \in W_2$ " W_2 " " "

Hence $v+z \in W_1 + W_2$. Now let α be a scalar and let $v \in W_1 + W_2 \Rightarrow v = v_1 + v_2$, $v_1 \in W_1$, $v_2 \in W_2$.

$$\alpha v = \alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$$

Now

$v_1 \in W_1 \Rightarrow \alpha v_1 \in W_1$ since W_1 is a subspace
$v_2 \in W_2 \Rightarrow \alpha v_2 \in W_2$ " W_2 " "" "

$$\Rightarrow \alpha v = \alpha v_1 + \alpha v_2 \in W_1 + W_2. \checkmark$$

(b) $W_1 \cap W_2$ is a subspace of V, hence has a basis

$S = \{v_1, \dots, v_r\}$.

Since $W_1 \cap W_2$ is a subspace of W_1, we can <u>extend</u> S into a basis $S_1 = \{v_1, \dots, v_r, v_{r+1}, \dots, v_{r+m}\}$ of W_1 by suitably adding elements $v_{r+1}, \dots, v_{r+m}$ into S.

Similarly, $W_1 \cap W_2$ is a subspace of W_2. So we can extend S into a basis $S_2 = \{v_1, \dots, v_r, z_{r+1}, \dots, z_{r+p}\}$ of W_2 by adding elements $z_{r+1}, \dots, z_{r+p}$ into S.
It is clear from this that

$$\dim(W_1 \cap W_2) = r; \quad \dim W_1 = r + m; \quad \dim W_2 = r + p.$$

We will show that the set

$$\tilde{S} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_{r+m}, z_{r+1}, \dots, z_{r+p}\}$$

is a basis for $W_1 + W_2$. The relation in (b) will then follow.
Let $w = w_1 + w_2 \in W_1 + W_2$.

$$w_1 \in W_1 \Rightarrow w_1 = c_1 v_1 + \dots + c_r v_r + c_{r+1} v_{r+1} + \dots + c_{r+m} v_{r+m}$$

$$w_2 \in W_2 \Rightarrow w_2 = d_1 v_1 + \dots + d_r v_r + d_{r+1} z_{r+1} + \dots + d_{r+p} z_{r+p}.$$

$$\Rightarrow w = (c_1 + d_1) v_1 + \dots + (c_r + d_r) v_r + c_{r+1} v_{r+1} + \dots + c_{r+m} v_{r+m}$$
$$+ d_{r+1} z_{r+1} + \dots + d_{r+p} z_{r+p}.$$

This shows readily that $W_1 + W_2 = \text{span}(\tilde{S})$. Hence, it remains to show that $\tilde{S}$ is linearly independent. So consider the equation

$$\underbrace{k_1 v_1 + \cdots + k_r v_r + k_{r+1} v_{r+1} + \cdots + k_{r+m} v_{r+m}}_{v_1} + \underbrace{\xi_{r+1} z_{r+1} + \cdots + \xi_{r+p} z_{r+p}}_{v_2} = 0$$

Note that $v_1 \in W_1$ and $v_2 \in W_2$. We will show that

$$k_1 = \cdots = k_r = k_{r+1} = \cdots = k_{r+m} = \xi_{r+1} = \cdots = \xi_{r+p} = 0.$$

Note that $v_1 = -v_2$. (*)

Since W_1 and W_2 are subspaces, it follows from (*) that $v \in W_1 \cap W_2$.

By the uniqueness of coordinate representation with respect to a basis, in this case $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_{r+m}\}$, we must have

$k_{r+1} = \cdots = k_{r+m} = 0$. It then follows that

$$k_1 v_1 + \cdots + k_r v_r + \xi_{r+1} z_{r+1} + \cdots + \xi_{r+p} z_{r+p} = 0.$$

Finally, since $\{v_1, \ldots, v_r, z_{r+1}, \ldots, z_{r+p}\}$ is a basis (for W_2) it follows that

$$k_1 = \cdots = k_r = \xi_{r+1} = \cdots = \xi_{r+p} = 0. \quad \blacksquare$$

Remark In the special case when $W_1 \cap W_2 = \{0\}$,

we use $W_1 \oplus W_2$ to denote the sum, now called "direct sum" of W_1 and W_2. Also, we have

$$\dim W_1 \oplus W_2 = \dim W_1 + \dim W_2.$$

Lemma The sum $W_1 + W_2$ is direct (i.e. $W_1 \cap W_2 = \{0\}$) if and only if every element w of $W_1 + W_2$ has the unique representation $w = w_1 + w_2$, $w_1 \in W_1$, $w_2 \in W_2$.

proof Suppose $W_1 \cap W_2 = \{0\}$ and let $w_1 + w_2 = \tilde{w}_1 + \tilde{w}_2$

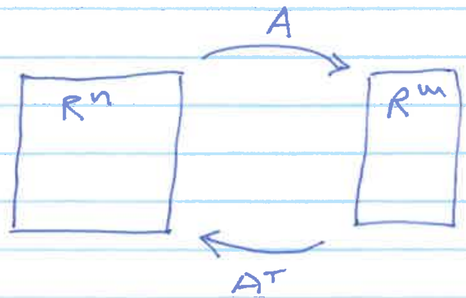
$w_1, \tilde{w}_1 \in W_1$ and $w_2, \tilde{w}_2 \in W_2$. We have $w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2$.
Since W_1, W_2 are subspaces $w_1 - \tilde{w}_1 \in W_1$ and $\tilde{w}_2 - w_2 \in W_2$.
But since $w_1 - \tilde{w}_1 = \tilde{w}_2 - w_2$, $w_1 - \tilde{w}_1$ and $\tilde{w}_2 - w_2$ both belong to $W_1 \cap W_2 = \{0\}$, implying $\tilde{w}_1 = w_1$ and $\tilde{w}_2 = w_2$.

We prove the converse, suppose the representation $w = w_1 + w_2$ is unique but $W_1 \cap W_2 \neq \{0\}$. Let $0 \neq \phi \in W_1 \cap W_2$. Then for $w \in W_1 + W_2$ we have $w = w_1 + w_2 = (w_1 + \phi) + (w_2 - \phi)$.

$w_1 + \phi \in W_1$
$w_2 - \phi \in W_2$

Since $w_1 + \phi \neq w_1$, we have a contradiction to the unique representation assumption. Thus $W_1 \cap W_2 = \{0\}$. ∎

We shall next establish one of the most important facts in Linear Algebra.
It is very much related to a fact seen earlier namely, for an $m \times n$ matrix A



$$\boxed{\begin{aligned} n &= \text{nullity}(A) + \text{rank}(A) \\ m &= \text{nullity}(A^T) + \text{rank}(A) \end{aligned}}$$

Recall the following definitions and facts already established:

nullity(A) is dimension of $\text{Ker}(A)$, which is a subspace of R^n
nullity(A^T) is dimension of $\text{Ker}(A^T)$, which is a subspace of R^m

$\boxed{\text{rank}(A) = \text{row rank}(A) = \text{column rank}(A)}$ (we proved this)

rank(A) is number of nonzero rows in rref(A)

row rank(A) is dimension of row space of A

row space of A = span of rows of A, is a subspace of R^n
= span of nonzero rows of rref(A)

column rank(A) is dimension of column space of A

column space of A = span of columns of A, is a subspace of R^m
= span of "pivot" columns of A

Note: The column space of A is also known as the range of A

$$\text{Range}(A) = \{ y \in R^m \mid y = Ax, \text{ for some } x \in R^n \}.$$

The fact that all 3 "ranks" are equal says that the row space of A and the column space of A are different as subspaces, they nevertheless have the same dimension.

Also, note that

$$\text{row space of } A^T = \text{column space of } A$$
$$\text{column space of } A^T = \text{row space of } A.$$

Theorem. Let A be an $m \times n$ matrix. Then

(a) $R^n = \text{Ker}(A) \oplus \text{Range}(A^T)$

(b) $R^m = \text{Ker}(A^T) \oplus \text{Range}(A)$.

proof (a) We already know that

$\text{Ker}(A)$ and $\text{Range}(A^T)$ are subspaces of R^n.

Thus we can form the sum

$$\text{Ker}(A) + \text{Range}(A^T)$$

which is a subspace of R^n. We also know that

$$\dim(\text{Ker}(A) + \text{Range}(A^T)) = \dim(\text{Ker}(A)) + \dim(\text{Range}(A^T)) - \dim(\text{Ker}(A) \cap \text{Range}(A^T)),$$

Now,

$$\dim(\text{Ker}(A)) = \text{nullity}(A).$$

Also,

$$\dim(\text{Range}(A^T)) = \dim(\text{column space of } A^T) = \dim(\text{row space of } A) = \text{rank}(A).$$

Thus

$$\dim(\text{Ker}(A) + \text{Range}(A^T)) = \text{nullity}(A) + \text{rank}(A) - \dim(\text{Ker}(A) \cap \text{Range}(A^T))$$

we will show next that $Ker(A) \cap Range(A^T) = \{0\}$.
Indeed, let $v \in Ker(A) \cap Range(A^T)$.

$v \in Ker(A) \Rightarrow Av = 0$
$v \in Range(A^T) \Rightarrow v = A^T w$ for some $w \in R^m$.

$v = A^T w \Rightarrow w^T (A^T)^T = v^T$, i.e. $v^T = w^T A$.
Multiply this from the right by v; to get $v^T v = w^T A v$
Now $Av = 0$, hence $v^T v = w^T A v = w^T 0 = 0$.
On the other hand

$$0 = v^T v = \sum_{i=1}^{n} v_i^2 \Rightarrow v_i = 0, \; i = 1, \dots, n, \text{ i.e. } v = 0.$$

This shows that $Ker(A) \cap Range(A^T) = \{0\}$, implying that $\dim(Ker(A) \cap Range(A^T)) = 0$. Thus

$$\dim(Ker(A) + Range(A^T)) = nullity(A) + rank(A).$$

However, we know that $nullity(A) + rank(A) = n = \dim R^n$,
Thus $Ker(A) + Range(A^T) = R^n$ and the sum is direct.
This justifies writing $R^n = Ker(A) \oplus Range(A^T)$. ✓

(b) The proof is the same. Just replace A by A^T. ∎

We saw in Theorem 4.5.5(b) that any set of linearly independent set of vectors $S = \{v_1, \dots, v_r\}$ of V that does not span V can be extended to a basis $B = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ of V by adding appropriately chosen vectors $v_{r+1}, \dots, v_n$ to S.

As a consequence we have the following

<u>Lemma</u> let W be a subspace of V with $W \neq V$. Then there exists a subspace W^c of V such that $V = W \oplus W^c$.

<u>proof.</u> W must have a basis $\{v_1, \dots, v_r\}$. By Thm 4.5.5(b) we can extend it to a basis $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ of V.

Let $W^c = \text{span}\{v_{r+1}, \dots, v_n\}$.

Clearly $V = W + W^c$. Indeed, since $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ is a basis for V, for any $v \in V$ we can write

$$v = \sum_{i=1}^{n} \alpha_i v_i = \underbrace{\sum_{i=1}^{r} \alpha_i v_i}_{v} + \underbrace{\sum_{i=r+1}^{n} \alpha_i v_i}_{v^c} = v + v^c$$

with $v \in W$ and $v^c \in W^c$.

To show that the sum $V = W + W^c$ is direct, it is enough to show that v and v^c are uniquely defined. So suppose $v = v_1 + v_1^c = v_2 + v_2^c$, $v_1, v_2 \in W$, $v_1^c, v_2^c \in W^c$. We will show that $v_1 = v_2$ and $v_1^c = v_2^c$.

Indeed, we have

$$(*) \qquad (v_1 - v_2) + (v_1^c - v_2^c) = 0$$

Now $v_1 - v_2 \in W \Rightarrow v_1 - v_2 = \sum_{i=1}^{r} \beta_i v_i$

and

$$v_1^c - v_2^c \in W^c \Rightarrow v_1^c - v_2^c = \sum_{i=r+1}^{n} \beta_i v_i.$$

From $(*)$ we have $\sum_{i=1}^{r} \beta_i v_i + \sum_{i=r+1}^{n} \beta_i v_i = 0.$

Hence

$$\sum_{i=1}^{n} \beta_i v_i = 0.$$

Since $\{v_1, \dots, v_n\}$ is a basis for V, we must have $\beta_1 = \dots = \beta_r = \beta_{r+1} = \dots = \beta_n = 0$. ∎

<u>Remark</u> In general, given W, its "complement" W^c is not uniquely defined.

<u>Ex.</u> Let $V = \mathbb{R}^2$, $W = \text{span}\{(1,0)\}$.

Then $W_1^c = \text{span}\{(0,1)\}$ and $W_2^c = \text{span}\{(1,1)\}$

can be shown to satisfy $\mathbb{R}^2 = W \oplus W_1^c = W \oplus W_2^c$ with $W_1^c \neq W_2^c$.

However, keep in mind that <u>once</u> a complement W^c is chosen, every vector in V has a unique representation $v = v_1 + v_1^c$, $v_1 \in W$, $v_1^c \in W^c$.

The dot product, The scalar/inner product, orthogonality

The dot product or Euclidean inner product of two vectors u, v in R^n is denoted by $u \cdot v$ and is defined as

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

other commonly used notations for the dot product are

(u, v) and $u^T v$. For the latter, we are viewing u and v as $n \times 1$ matrices and the quantity $u^T v$ as matrix-matrix product.

Regardless of which notation is used, the important properties characterizing the dot product are

(i) positivity $\quad u \cdot u \geq 0$ and $u \cdot u = 0$ only if $u = 0$

(ii) symmetry $\quad u \cdot v = v \cdot u$

(iii) bilinearity
$$\left.\begin{aligned}(u+v)\cdot w &= u \cdot w + v \cdot w \\ (\alpha u)\cdot w &= u \cdot (\alpha w) = \alpha\, u \cdot v\end{aligned}\right\} \text{linearity in first argument}$$

By symmetry, must be linear in 2nd argument as well.

Defn. we say that two vectors $u, v \in R^n$ are orthogonal if $u \cdot v = 0$

Ex. $\quad u = (1, -1, 2)$, $v = (2, 1, -\frac{1}{2})$

$$u \cdot v = (1)(2) + (-1)(1) + (2)(-\tfrac{1}{2}) = 2 - 1 - 1 = 0.$$

Thus u and v are orthogonal to each other. we also say
u is orthogonal to v
v " " " u

Sometimes we write $u \perp v$

Remark The dot product in $\mathbb{R}^n$ serves as a "template" to define scalar or inner products on general vector spaces V.

Indeed, suppose we have a vector space V and a real-valued function $(\cdot,\cdot)$ acting on pairs of vectors from V having the following properties

(a) positivity $(v,v) \geq 0 \quad \forall v \in V$ and $(v,v)=0$ only if $v=0$

(b) symmetry $(v,w)=(w,v) \quad \forall v,w \in V$

(c) bilinearity $(u+v,w)=(u,w)+(v,w) \quad \forall u,v,w \in V$

$(\alpha u,v)=\alpha(u,v), \quad \forall \alpha \in \mathbb{R}, \ \forall u,v \in V.$

We say that $(\cdot,\cdot)$ is a scalar/inner product on V.

Again, by symmetry, $(\cdot,\cdot)$ is linear in 2nd argument as well.

Ex. $V = C[a,b]$ space of all continuous real-valued functions defined on $[a,b]$.

Define $(\cdot,\cdot)$ on V by $\boxed{(f,g)=\int_a^b f(x)\,g(x)\,dx}$

It is an easy exercise to show that this $(\cdot,\cdot)$ satisfies all 3 properties (a), (b) and (c) of a general scalar (or inner) product

Ex. $V=\mathbb{R}^{m\times n}$ space of all $m\times n$ real matrices

For A, B $m\times n$ real matrices, define

$$(A,B)=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\cdot b_{ij}.$$

Again (A,B) can be shown to satisfy properties (a), (b), (c).

Remark In the light of this generalization, $\mathbb{R}^n$ equipped with the dot product becomes a special case of a general vector space equipped with a scalar product.

The power behind this generalization resides in the fact that many properties that can be established for $\mathbb{R}^n$ equipped with the usual dot product can be "exported" to any general vector space V that is equipped with a scalar product.(*)

For example, we can export the concept of orthogonality to any inner product space $V, (\cdot,\cdot)$

Defn. We say that in an inner product space $V, (\cdot,\cdot)$ two vectors v, w are orthogonal if $(v,w)=0$.

Ex. $V = C[-1,1]$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$

let $f=1$, $g=x$. Then

$$(1,x) = \int_{-1}^{1} 1\cdot x\,dx = \frac{x^2}{2}\Big|_{-1}^{1} = 0.$$

Ex. $V = \mathbb{R}^{2\times 2}$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

$$(A,B) = \sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}b_{ij} = 1\cdot 1 + 0.0 + 0.0 + (1)(-1) = 0.$$

Defn. let $V, (\cdot,\cdot)$ be an inner product space and let $S = \{w_1, \dots, w_r\}$ be a set of vectors in V. We say that S is an orthogonal set if

$$(w_i, w_j) = 0 \qquad i \neq j.$$

(*) Continuing on this theme, once we establish a fact in one inner product space using only the general properties of a vector space and properties (a), (b), (c), then this fact must hold in any inner product space.

we begin with an elementary fact which says that in an inner product space the zero vector is "orthogonal" to all vectors in V.

lemma let $V, (\cdot,\cdot)$ be an inner product space. Then

$$(0,v)=(v,0)=0 \quad \forall v \in V,$$

proof.

$$(0,v)=(0+0,v)=(0,v)+(0,v).$$

$(0,v)$ is a real number, adding $-(0,v)$ to both sides

$$0=(0,v)-(0,v)=(0,v)+(0,v)-(0,v)=(0,v). \checkmark \blacksquare$$

Another important result is

Theorem let $V,(\cdot,\cdot)$ be an inner product space and $S=\{w_1, w_2, \ldots, w_r\}$ a set of nonzero mutually orthogonal vectors. Then S is linearly independent.

proof. let

$$k_1 w_1 + k_2 w_2 + \cdots + k_r w_r = 0.$$

we want to show $k_1=k_2=\cdots=k_r=0$. Fix an index j $j=1,\ldots,r$ and take inner product of both sides with w_j

$$(*) \qquad (k_1 w_1 + k_2 w_2 + \cdots + k_r w_r, w_j) = (0, w_j).$$

Now $(0,w_j)=0$ by the preceding lemma. Also, since $(\cdot,\cdot)$ is linear in both arguments

$$(k_1 w_1 + \cdots + k_r w_r, w_j) = k_1 (w_1, w_j) + \cdots + k_r (w_r, w_j).$$

Now the vectors $w_1, \ldots, w_r$ being mutually orthogonal means $(w_i, w_j)=0$ $i \neq j$. Hence $(*)$ reduces to

$$k_j (w_j, w_j) = 0.$$

we also assumed that $w_1, \ldots, w_r$ are nonzero. By the positivity property of $(\cdot,\cdot)$, $(w_j,w_j) \neq 0$. This forces k_j to be zero. Since j was arbitrary, the result follows. $\blacksquare$

The Euclidean norm

Definition The Euclidean norm of a vector $v \in R^n$ is defined by

$$\|v\|_2 = \sqrt{(v,v)} = (v,v)^{1/2}$$

Ex. $\|(1,-2,4)\|_2 = \sqrt{1^2+(-2)^2+4^2} = \sqrt{21}.$

Theorem Properties of the Euclidean norm (dropping the subscript 2)

(a) $\|v\| \geq 0$ and $\|v\| = 0 \Rightarrow v = 0$ "positivity"
(b) $\|\alpha v\| = |\alpha|\,\|v\|$, $\alpha \in R$, $v \in R^n$ "Homogeneity"
(c) $\|v+w\| \leq \|v\| + \|w\|$, $v, w \in R^n$ "Triangle inequality"

proof (a) This is a direct consequence of the positivity of the dot product.

(b) $$\|\alpha v\| = (\alpha v, \alpha v)^{1/2} = \left(\sum_{i=1}^{n} (\alpha v_i)^2\right)^{1/2} = \left(\sum_{i=1}^{n} \alpha^2 v_i^2\right)^{1/2}$$
$$= \left(\alpha^2 \sum_{i=1}^{n} v_i^2\right)^{1/2} = (\alpha^2)^{1/2}\left(\sum_{i=1}^{n} v_i^2\right)^{1/2} = |\alpha|\,\|v\|. \checkmark$$

(c) Let's look at $\|v+w\|^2$ instead.

$$\|v+w\|^2 = (v+w, v+w) = \sum_{i=1}^{n} (v_i+w_i)^2 = \sum_{i=1}^{n} (v_i^2 + 2v_iw_i + w_i^2).$$
$$= \sum_{i=1}^{n} v_i^2 + 2\sum_{i=1}^{n} v_iw_i + \sum_{i=1}^{n} w_i^2$$
$$= \|v\|^2 + 2\, v \cdot w + \|w\|^2.$$

To complete the proof, we will use the

> **Cauchy-Schwarz inequality:** For any two vectors $v, w \in R^n$
>
> $$|v \cdot w| \leq \|v\|\,\|w\|.$$
>
> Furthermore, equality holds if and only if v and w are linearly dependent, i.e. one is a scalar multiple of the other

whose proof follows immediately after this.

Indeed, it follows from the Cauchy-Schwarz inequality that

$$\|v+w\|^2 \le \|v\|^2 + 2|v\cdot w| + \|w\|^2$$

$$\le \|v\|^2 + 2\|v\|\,\|w\| + \|w\|^2$$

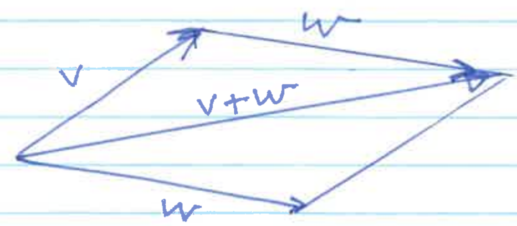
$$= (\|v\| + \|w\|)^2.$$

Hence

$$\|v+w\| \le \|v\| + \|w\|.$$

Remark we call this the triangle inequality since it reflects the well-known geometric fact that
"In any triangle, the length of any one side is less than or equal to the sum of the lengths of the other two sides"

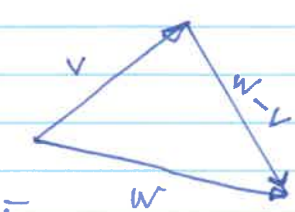
There is a another inequality which is also called the triangle inequality. It is



$$\|v-w\| \ge \big|\,\|v\| - \|w\|\,\big|$$

This reflects the well-known fact that "In any triangle the length of any one side is greater than or equal to the absolute value of the difference of lengths of the other two sides".

proof



write $v = (v-w) + w$.
Then by the 1st triangle inequality

$$\|v\| = \|(v-w)+w\| \le \|v-w\| + \|w\| \Rightarrow \|v\| - \|w\| \le \|v-w\|.$$

Similarly, write $w = (w-v) + v$. Then

$$\|w\| = \|(w-v)+v\| \le \|w-v\| + \|v\| \Rightarrow \|w\| - \|v\| \le \|w-v\| = \|v-w\|$$

So far, we have shown that $\pm(\|w\| - \|v\|) \leq \|v-w\|$.
This is the same as saying that $\big|\|w\| - \|v\|\big| \leq \|v-w\|$. ✓

Proof of the Cauchy-Schwarz inequality

Given the positivity of the dot product: $v \cdot v > 0 \ \forall v \neq 0$, for any $t \in \mathbb{R}$, we have, using also the symmetry and homogeneity properties

$$0 \leq (tv+w)\cdot(tv+w) = t^2\|v\|^2 + 2t\, v\cdot w + \|w\|^2 = q(t)$$

The right-hand-side is a quadratic function of t and is nonnegative for all values of t. In other words $q(t)$ cannot have two distinct real roots. It may have a double real root or two complex conjugate roots. A result of analysis tells us that the discriminant

$$(2\, v\cdot w)^2 - 4\|v\|^2\|w\|^2) \quad \text{cannot be positive, i.e.}$$

$$4(v\cdot w)^2 \leq 4\|v\|^2\|w\|^2, \text{ which proves the inequality.}$$

Finally, if $|v\cdot w| = \|v\|\,\|w\|$, i.e. equality holds, then $(2\, v\cdot w)^2 - 4\|v\|^2\|w\|^2 = 0$ and consequently $q(t)$ has a double real root, i.e. there is value of t for which $(tv+w)\cdot(tv+w) = 0$.
But this means $tv+w=0$ i.e. $w = -tv$. Conversely if

$$w = \alpha v, \text{ then } |v\cdot w| = |v\cdot(\alpha v)| = |\alpha \|v\|^2| = |\alpha|\,\|v\|^2$$
$$= \|v\|\,|\alpha|\,\|v\| = \|v\|\,\|\alpha v\| = \|v\|\,\|w\|$$

or if

$$v = \alpha w, \text{ then } |v\cdot w| = |(\alpha w)\cdot w| = |\alpha \|w\|^2| = |\alpha|\,\|w\|^2$$
$$= |\alpha|\,\|w\|\,\|w\| = \|\alpha w\|\,\|w\| = \|v\|\,\|w\| \ \blacksquare$$

Defn. A vector u is said to be a unit vector if $\|u\| = 1$.

Fact Given a nonzero vector v which is not unit, we can always obtain a unit vector u which has the "same direction" as v, namely

$$\|u\| = \frac{v}{\|v\|}$$

Norms on vector spaces

The concept of the Euclidean norm $\|v\|_2 = \sqrt{v \cdot v}$ in R^n can be generalized to any inner product space V, $(\cdot,\cdot)$ by defining the norm $\|\cdot\|$ by

$$(*) \qquad \|v\| = \sqrt{(v,v)} \,.$$

This is in essence mimicks the relationship between the Euclidean norm and dot product.

For the general norm defined above, it can be shown that the 3 properties

(a) $\|v\| \geq 0, \ \forall v \in V$ and $\|v\| = 0$ only if $v = 0$
(b) $\|\alpha v\| = |\alpha| \, \|v\|, \ \forall v \in V, \ \forall \alpha \in R$
(c) $\|v+w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$

do indeed hold. In addition, the Cauchy-Schwarz inequality

$$|(v,w)| \leq \|v\| \, \|w\| \qquad \forall v, w \in V$$

also holds.

Remark It is important to note that a norm can be defined on a vector space V whenever one has an inner product $(\cdot,\cdot)$ on V.

Nevertheless, one can define norms that are not derived from an inner product. Indeed, define the map $\|\cdot\|_\infty$ on R^n by

$$\|v\|_\infty = \max_{1 \leq i \leq n} |v_i| \,.$$

It can be shown that $\|\cdot\|_\infty$ satisfies properties (a), (b), (c).

The projection into a subspace: The "oblique" case

Let V be a vector space and W a finite dimensional subspace of V. In order to define a projection operator $P: V \to W$ we need a direct sum decomposition of V

$$V = W \oplus W^c.$$

Note that there are more than one choice for W^c, hence the projection into W can be defined only when a choice of W^c is made.

Definition Let $V = W \oplus W^c$. For $v \in V$ we define its projection Pv by

(a) If $v \in W$, then $Pv = v$.
(b) If $v \notin W$, then $Pv = w$ where $v = w + w^c$.

To compute the projection of Pv of v, we need a basis $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ of V which includes the

basis $\{v_1, \dots, v_r\}$ for W and the basis $\{v_{r+1}, \dots, v_n\}$ for W^c

write $$v = \sum_{i=1}^{n} \alpha_i v_i = \underbrace{\sum_{i=1}^{r} \alpha_i v_i}_{w} + \underbrace{\sum_{i=r+1}^{n} \alpha_i v_i}_{w^c} = w + w^c$$

$$\boxed{Pv = w}$$

Ex. Compute the projection of $v = (a, b, c)$ into

$W = \text{span}\{\overset{v_1}{(1,2,1)}, \overset{v_2}{(-1,3,4)}\}$, $W^c = \text{span}\{\overset{v_3}{(1,1,1)}\}$.

We first verify that $B = \{\overset{v_1}{(1,2,1)}, \overset{v_2}{(-1,3,4)}, \overset{v_3}{(1,1,1)}\}$ is a basis for $\mathbb{R}^3$. For this, we arrange them as columns:

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{bmatrix}.$$ To show that B is a basis for $\mathbb{R}^3$, it suffices to show that A is invertible.

$$\det(A) = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{vmatrix} \longrightarrow \begin{matrix} 1 & -1 & 1 & 1 & -1 \\ 2 & 3 & 1 & 2 & 3 \\ 1 & 4 & 1 & 1 & 4 \end{matrix}$$

$$\det(A) = (1)(3)(1) + (-1)(1)(1) + (1)(2)(4) - \left[(1)(3)(1) + (1)(1)(4) + (-1)(2)(1)\right]$$
$$= 3 - 1 + 8 - [3 + 4 - 2]$$
$$= 5 \neq 0$$

Next, we need to find coordinates $\alpha_1, \alpha_2, \alpha_3$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = (a, b, c).$$

i.e.

$$\begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

using Matlab we find $A^{-1} = \frac{1}{5}\begin{bmatrix} -1 & 5 & -4 \\ -1 & 0 & 1 \\ 5 & -5 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -1 & 5 & -4 \\ -1 & 0 & 1 \\ 5 & -5 & 5 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$

Hence

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -a + 5b - 4c \\ -a + c \\ 5a - 5b + 5c \end{bmatrix}$$

Thus $P\mathbf{v} = \alpha_1 v_1 + \alpha_2 v_2 = \frac{1}{5}(-a+5b-4c)\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + \frac{1}{5}(-a+c)\begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix}$

$$= \frac{1}{5}\begin{pmatrix} 5b - 5c \\ -5a + 10b - 5c \\ -5a + 5b \end{pmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

We have solved the problem of finding the projection of a vector in a global sense, i.e. for any vector (a,b,c) we have found its projection given by

$P\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ where $P = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}$ is called the projection matrix

One characteristic property of all projection matrices is

$$\boxed{P^2 = P}.$$

procedure for constructing the projection matrix

1) Assume we have a direct sum decomposition $V = W \oplus W^c$

$\{w_1, w_2, \dots, w_r, w_{r+1}, \dots, w_n\}$ basis for V

$\{w_1, w_2, \dots, w_r\}$ basis for W.

2) construct $n \times n$ matrix B using $w_1, \dots, w_n$ as columns

$$B = \begin{bmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_n \\ | & | & & | \end{bmatrix}$$

3) Calculate B^{-1} and partition

$$B^{-1} = \begin{array}{c} n \\ \begin{bmatrix} B_1^{-1} \\ \hline B_2^{-1} \end{bmatrix} \end{array} \begin{array}{l} r \\ n-r \end{array}$$

4)
$$P = \begin{array}{c} r \\ \begin{bmatrix} | & & | \\ w_1 & \cdots & w_r \\ | & & | \end{bmatrix} \end{array} \begin{array}{c} n \\ \begin{bmatrix} & B_1^{-1} & \end{bmatrix} \end{array} r$$

In the previous example, we had

$$P = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 1 & 4 \end{bmatrix} \frac{1}{5} \begin{bmatrix} -1 & 5 & -4 \\ -1 & 0 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 0 & 5 & -5 \\ -5 & 10 & -5 \\ -5 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix}.$$

Let $v = \begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix}$. The projection of v into $\text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix} \right\}$ is given by

$$Pv = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix} = \begin{bmatrix} -11 \\ -20 \\ -9 \end{bmatrix} = -\frac{53}{5} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \frac{2}{5} \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$$

Note that
$$\begin{bmatrix} -53/5 \\ 2/5 \end{bmatrix} = \underbrace{\frac{1}{5} \begin{bmatrix} -1 & 5 & -4 \\ -1 & 0 & 1 \end{bmatrix}}_{B_1^{-1}} \begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix} \in V$$

Summary

1) The projection P (operator/matrix) into a subspace W of a vector space V can be defined only if we have a direct sum decomposition $V = W \oplus W^c$

2) P depends on both choices of W and W^c even though it is the projection into W.

3) It turns out that $I-P$ is the projection operator/matrix into W^c. This can be seen by reversing the roles of W and W^c. Indeed, $(W^c)^c = W$. So it comes as a free bonus!

For instance we saw that the projection matrix P into the subspace $W = \text{span}\left\{\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}-1\\3\\4\end{pmatrix}\right\}$ with $W^c = \text{span}\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$ is given by

$$P = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \\ -1 & 1 & 0 \end{bmatrix} \Rightarrow I-P = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

So, the projection of $\begin{bmatrix}2\\-7\\4\end{bmatrix}$ into W^c is given by

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix}2\\-7\\4\end{bmatrix} = \begin{bmatrix}13\\13\\13\end{bmatrix} = 13\begin{bmatrix}1\\1\\1\end{bmatrix} = 13 \times \text{basis vector of } W^c.$$

Also, by adding $P\begin{bmatrix}2\\-7\\4\end{bmatrix} = \begin{bmatrix}-11\\-20\\-9\end{bmatrix}$ to $13\begin{bmatrix}1\\1\\1\end{bmatrix}$

we get $\begin{bmatrix}2\\-7\\4\end{bmatrix}$, i.e. we recover v!!.

4) Other interesting results are

$$\text{rank}(P) = r, \quad \text{rank}(I-P) = n-r$$

The projection into a subspace: The orthogonal case

In the preceding section we showed how to construct the projection operator/matrix into a space W of a vector space V. Two comments are in order

(1) The process is rather costly since we have to invert the $n \times n$ matrix

$$B = \left[w_1 \,\middle|\, w_2 \,\middle|\, \cdots \,\middle|\, w_n \right].$$

(2) The process as shown works only for $V = \mathbb{R}^n$. We can generalize this to more general vector spaces by working with the coordinates of v and Pv which are vectors in $\mathbb{R}^n$.

Nevertheless, we will now consider the case of a vector space V which is equipped with a scalar product $(\cdot,\cdot)$. We will show that in the presence of orthogonality, many processes can become easier to handle.

Defn. Let S be a subset of an inner product space $V, (\cdot,\cdot)$. The orthogonal complement $S^\perp$ of S in V is defined by

$$S^\perp = \{ w \in V \mid (v, w) = 0 \quad \forall v \in S \}.$$

Ex. $V = \mathbb{R}^n$, equipped with dot product

Let $S = \{ (a, 0, 0),\ a \in \mathbb{R} \}$ x-axis

then $S^\perp = \{ (0, b, c),\ b, c \in \mathbb{R} \}$ yz-plane

We next show that the orthogonal complement $S^\perp$ is a subspace of V, even if S is not a subspace.

Theorem Let S be a subset of the inner product space $V, (\cdot,\cdot)$. Then

(a) If $S = V$, then $S^\perp = \{0\}$.

(b) If $S \neq V$, then $S^\perp$ is a subspace of V.

(c) $\{0\}^\perp = V$

Proof (a) Let $w \in S^\perp$. Then $(w, v) = 0 \quad \forall v \in S$.
but since $S = V$, $w \in S \Rightarrow (w, w) = 0 \Rightarrow w = 0$. ✓

(b) Let $w_1, w_2 \in S^\perp$.

$$\left.\begin{array}{l} w_1 \in S^\perp \Rightarrow (w_1, v) = 0 \quad \forall v \in S \\ w_2 \in S^\perp \Rightarrow (w_2, v) = 0 \quad \forall v \in S \end{array}\right\} \xRightarrow{\text{adding}} (w_1 + w_2, v) = 0$$
$$\Rightarrow w_1 + w_2 \in S^\perp$$

Let $w \in S^\perp$. Then $(w, v) = 0 \quad \forall v \in S$;
for any $\alpha \in \mathbb{R}$, $(\alpha w, v) = \alpha(w, v) = \alpha 0 = 0$
$\Rightarrow \alpha w \in S^\perp$

Hence $S^\perp$ is a subspace of V. ■

Theorem Let W be a subspace of an inner product space $V, (\cdot,\cdot)$. Then

(a) $(W^\perp)^\perp = W$.

(b) $V = W \oplus W^\perp$. □

Remark (a) says that the orthogonal complement of the orthogonal complement of a subspace W is W itself.

(b) This states that the vector space V can be decomposed into the direct sum of W and $W^\perp$.

Some well-known facts are

Theorem Let A be an $m \times n$ matrix. Then

(a) $\text{range}(A^T) = \text{Ker}(A)^\perp$ in $\mathbb{R}^n$

(b) $\text{range}(A) = \text{Ker}(A^T)^\perp$ in $\mathbb{R}^m$. $\square$

Theorem Let $V, (\cdot,\cdot)$ be an inner product space and suppose we have the direct sum decomposition

$$V = W \oplus W^\perp,$$

with $\{w_1, \ldots, w_r\}$ an orthogonal basis for W. Then the projection Pv onto W of any vector v in V can be expressed as

$$Pv = \frac{(v, w_1)}{\|w_1\|^2} w_1 + \frac{(v, w_2)}{\|w_2\|^2} w_2 + \cdots + \frac{(v, w_r)}{\|w_r\|^2} w_r.$$

proof. Since $\{w_1, \ldots, w_r\}$ is a basis for W, we can certainly write

$$Pv = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_r w_r. \quad (*)$$

All we have to do is to show that

$$\alpha_i = \frac{(v, w_i)}{\|w_i\|^2}, \quad i = 1, \ldots, r.$$

Since $V = W \oplus W^\perp$, we can also write $v = w + w^\perp$, $w \in W$, $w^\perp \in W^\perp$ where also $(w, w^\perp) = 0$.

By definition of the projection operator P, $Pv = w$.

For any $i = 1, \ldots, r$, taking inner products with w_i,

$$(v, w_i) = (w + w^\perp, w_i) = (Pv + w^\perp, w_i) = (Pv, w_i) \quad i = 1, \ldots, r$$

since $w_i \in W$ and $w^\perp \in W^\perp$.

$$\Rightarrow (v, w_i) = (\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_r w_r, w_i)$$

$$= \alpha_i (w_i, w_i) = \alpha_i \|w_i\|^2$$

since the vectors $w_1, \ldots, w_r$ are mutually orthogonal. Hence

$$\alpha_i = \frac{(v, w_i)}{\|w_i\|^2}, \quad i = 1, \ldots, r. \quad \blacksquare$$

The Gram-Schmidt orthogonalization Process

Definition Given a set of vectors $S=\{v_1,\dots,v_r\}$ in a vector space equipped with a scalar product $(\cdot,\cdot)$ we say that S is an orthogonal set if

$$(v_i, v_j) = 0 \quad \text{whenever } i \neq j.$$

Ex. $V = \mathbb{R}^2$ $(\cdot,\cdot)$ defined by/as dot product.

$S=\{(1,0),(0,1)\}$ is an orthogonal set. Indeed $(1,0)\cdot(0,1)=0$

$\tilde{S}=\{(1,1),(1,-1)\}$ is an " " . Indeed $(1,1)\cdot(1,-1)=0.$

Ex. $V = \mathbb{R}^3$ $(\cdot,\cdot)$ defined by/as dot product.

$$S=\{\underbrace{(1,2,1)}_{v_1}, \underbrace{(-15, 0, 15)}_{v_2}, \underbrace{(2,-2,2)}_{v_3}\} \text{ is orthogonal.}$$

Indeed

$$(v_1,v_2) = (1,2,1)\cdot(-15,0,15) = -15+0+15=0$$

$$(v_1,v_3) = (1,2,1)\cdot(2,-2,2) = 2-4+2=0$$

$$(v_2,v_3) = (-15,0,15)\cdot(2,-2,2) = -30+0+30=0.$$

A well-known method for constructing a set of orthogonal vectors is the Gram-Schmidt process.

Theorem (Gram-Schmidt) Let $S=\{v_1,v_2,\dots,v_r\}$ be a set of linearly independent vectors in a vector space V equipped with a scalar product $(\cdot,\cdot)$.

Consider the set $\tilde{S}=\{\tilde{v}_1,\dots,\tilde{v}_r\}$ constructed recursively according to the algorithm (Gram-Schmidt process)

$$\tilde{v}_1 = v_1$$

$$\tilde{v}_2 = v_2 - \frac{(v_2,\tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1$$

$$\tilde{v}_3 = v_3 - \frac{(v_3, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \frac{(v_3, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2$$

$$\vdots$$

$$\tilde{v}_r = v_r - \frac{(v_r, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \frac{(v_r, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2 - \cdots - \frac{(v_r, \tilde{v}_{r-1})}{\|\tilde{v}_{r-1}\|^2}\tilde{v}_{r-1}.$$

Then $\tilde{S}$ has the following properties

(a) $\tilde{S}$ is orthogonal, i.e. $(\tilde{v}_i, \tilde{v}_j) = 0$ if $i \neq j$
(b) $\tilde{S}$ is linearly independent.
(c) $\text{span}(\tilde{S}) = \text{span}(S)$.

<u>proof</u> (a) This is equivalent to showing the following:

For $j = 2, \ldots, r$, $(\tilde{v}_j, \tilde{v}_k) = 0$ for $k = 1, \ldots, j-1$. ⊛

We do this by <u>induction</u>.
(i) $(\tilde{v}_2, \tilde{v}_1) = 0$
(ii) suppose ⊛ holds for $2, \ldots, j$. show ⊛ holds for $j+1$.

(i)
$$(\tilde{v}_2, \tilde{v}_1) = \left(v_2 - \frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1, \tilde{v}_1\right)$$
$$= (v_2, \tilde{v}_1) - \frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}(\tilde{v}_1, \tilde{v}_1)$$
$$= (v_2, \tilde{v}_1) - (v_2, \tilde{v}_1) = 0 \Leftarrow (\tilde{v}_1, \tilde{v}_1) = \|\tilde{v}_1\|^2$$

(ii)
$$(\tilde{v}_{j+1}, \tilde{v}_k) = \left(v_{j+1} - \frac{(v_{j+1}, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \cdots - \frac{(v_{j+1}, \tilde{v}_j)}{\|\tilde{v}_j\|^2}\tilde{v}_j, \tilde{v}_k\right)$$

$$= (v_{j+1}, \tilde{v}_k) - \frac{(v_{j+1}, \tilde{v}_1)}{\|\tilde{v}_1\|^2}(\tilde{v}_1, \tilde{v}_k) - \cdots - \frac{(v_{j+1}, \tilde{v}_j)}{\|\tilde{v}_j\|^2}(\tilde{v}_j, \tilde{v}_k)$$

To complete the induction argument, we need to show that $(\tilde{v}_{j+1}, \tilde{v}_k) = 0$ for $k = 1, \ldots, j$. Indeed, for this range of values of k, using the induction hypothesis only one term among $(\tilde{v}_1, \tilde{v}_k), \ldots, (\tilde{v}_j, \tilde{v}_k)$ is nonzero, namely $(\tilde{v}_k, \tilde{v}_k) = \|\tilde{v}_k\|^2$. Hence,

Hence, $(\tilde{v}_{j+1}, \tilde{v}_k) = (v_{j+1}, \tilde{v}_k) - (v_{j+1}, \tilde{v}_k) = 0$.

(b,c) It is clear from the process that we can write

$$v_j = \tilde{v}_j + \frac{(v_j, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 + \frac{(v_j, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2 + \cdots + \frac{(v_j, \tilde{v}_{j-1})}{\|\tilde{v}_{j-1}\|^2}\tilde{v}_{j-1}$$
$j = 1, \ldots, r$.

This shows that each v_j is a linear combination of $\tilde{v}_1, \ldots, \tilde{v}_j$. Thus $\mathrm{span}(\tilde{S}) = \mathrm{span}(S)$ and also that $\tilde{S}$ must be linearly independent. ∎

Defn. A set of vectors $S = \{u_1, \ldots, u_r\}$ is said to be orthonormal if

(i) $(u_i, u_j) = 0$, $i \neq j$
(ii) $\|u_i\| = 1$, $i = 1, \ldots, r$.

Remark Given a linearly independent set of vectors $S = \{v_1, \ldots, v_r\}$, we can transform S into an orthonormal set $\hat{S} = \{u_1, \ldots, u_r\}$ as follows:

(1) First use Gram-Schmidt process to obtain an orthogonal set $\tilde{S} = \{\tilde{v}_1, \ldots, \tilde{v}_r\}$
(2) Transform $\tilde{S}$ into an orthonormal set $\hat{S}$ by dividing each $\tilde{v}_i$ by its norm:

$$u_i = \frac{\tilde{v}_i}{\|\tilde{v}_i\|}, \quad i = 1, \ldots, r.$$

Ex. Consider the set $S = \{\overbrace{(1,2,1)}^{v_1}, \overbrace{(-1,3,4)}^{v_2}, \overbrace{(1,1,1)}^{v_3}\}$

$$\tilde{v}_1 = v_1 = \begin{pmatrix}1\\2\\1\end{pmatrix}$$

$$\tilde{v}_2 = v_2 - \frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 = \begin{pmatrix}-1\\3\\4\end{pmatrix} - \frac{\overbrace{(-1,3,4)\cdot(1,2,1)}^{9}}{\underbrace{\|(1,2,1)\|^2}_{6}}\begin{pmatrix}1\\2\\1\end{pmatrix} = \frac{1}{6}\begin{pmatrix}-15\\0\\15\end{pmatrix}$$

$$\tilde{v}_3 = v_3 - \frac{(v_3, \tilde{v}_1)}{\|\tilde{v}_1\|^2}v_1 - \frac{(v_3, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{(1,1,1)\cdot(1,2,1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \frac{(1,1,1)\cdot(-\frac{15}{6}, 0, \frac{15}{6})}{\|(-\frac{15}{6}, 0, \frac{15}{6})\|^2}\begin{pmatrix}-\frac{15}{6}\\0\\\frac{15}{6}\end{pmatrix}$$

$$= \frac{1}{3}\begin{pmatrix}1\\-1\\1\end{pmatrix}.$$

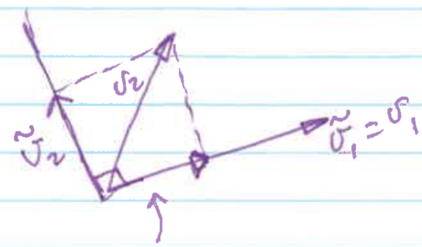
The set $\tilde{v}_1, \tilde{v}_2, \tilde{v}_3$ is orthogonal. To make it orthonormal, we divide each vector by its norm to obtain

$$u_1 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\1\end{pmatrix}; \quad u_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}; \quad u_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\end{pmatrix}.$$

We shall next reveal the "secret" behind the way the Gram-Schmidt process transforms a set of linearly independent vectors into a set of mutually orthogonal vectors with both sets spanning the same subspace.

$$\tilde{v}_1 = v_1$$

$$\tilde{v}_2 = v_2 - \frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1.$$



$$\frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1$$

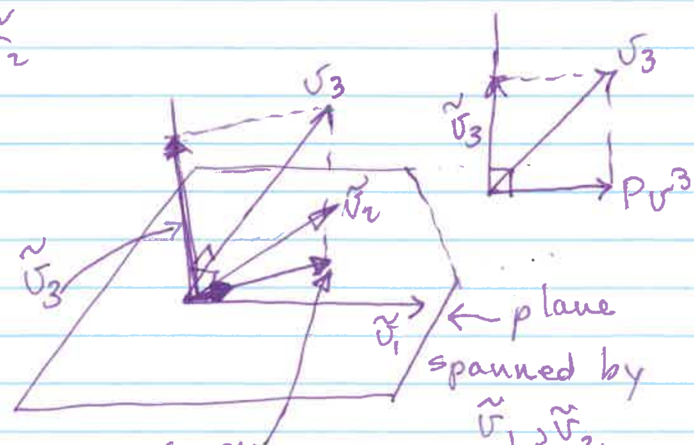
$\frac{(v_2, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1$ is the projection of v_2 into $\text{span}\{\tilde{v}_1\}$. Hence subtracting this projection from v_2 gives a vector $\tilde{v}_2$ which is orthogonal to $\tilde{v}_1$

$$\tilde{v}_3 = v_3 - \frac{(v_3, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \frac{(v_3, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2$$

The vector

$$Pv_3 = \frac{(v_3, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 + \frac{(v_3, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2$$

is the orthogonal projection of v_3 into $\text{span}\{\tilde{v}_1, \tilde{v}_2\}$.



$$Pv_3 = \frac{(v_3, \tilde{v}_1)}{\|\tilde{v}_1\|^2}\tilde{v}_1 + \frac{(v_3, \tilde{v}_2)}{\|\tilde{v}_2\|^2}\tilde{v}_2$$

Subtracting Pv_3 from v_3 gives $\tilde{v}_3$ which is in the orthogonal complement of $\text{span}\{\tilde{v}_1, \tilde{v}_2\}$ i.e. $\tilde{v}_3$ is orthogonal to both $\tilde{v}_1$ and $\tilde{v}_2$.

Planes and affine varieties in R^n

It is most probable that the reader has already encountered the concepts of lines and planes, say in Calculus. But what do we exactly mean by them? We shall use the concepts of linear Algebra developed up to now to give precise definitions to these important objects.

We have defined R^n as the set of all vectors, or n-tuples of the form $(v_1, v_2, \dots, v_n)$, equipped with componentwise addition and componentwise scalar multiplication operations. By thinking of a vector $(v_1, \dots, v_n)$ as an arrow going from the origin identified with $O = (0, \dots, 0)$, to the point $(v_1, \dots, v_n)$, we can identify R^n with the n-dimensional rectangular Cartesian coordinate system E^n familiar to students of Calculus.

Similarly, given two points $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_n)$ in a Cartesian system, we can think of the vector $v = (y_1 - x_1, \dots, y_n - x_n)$ as the arrow going from the point X to the point Y.

With this kind of identification between R^n and the n-dimensional rectangular Cartesian system, we can achieve the following conceptual advances

(1) We can give a geometric interpretation of subsets of R^n by identifying their counterparts in the rectangular Cartesian system.

(2) We can solve problems of geometric nature by using the machinery and tools of Linear Algebra.

Ex. An r-dimensional plane in E^n that contains or "passes through" the origin is just an r-dimensional subspace of R^n.

Ex. what about planes in E^n that do not contain the origin? well these are just planes that are parallel to planes that contain the origin. Their counterparts in R^n are called affine varieties and are sets of the form

$$b + W = \{v \in R^n,\ v = b + w,\ w \in W\}$$

where W is an r-dimensional subspace of R^n and b is some vector.

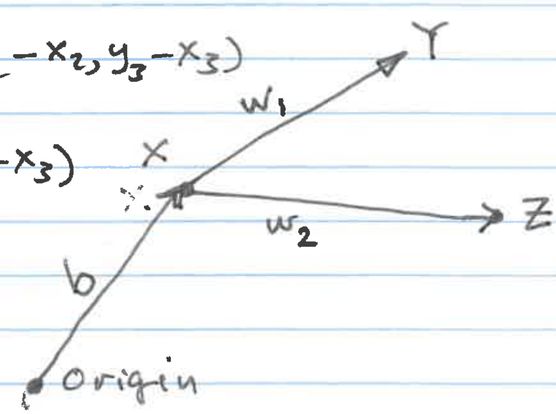
Ex. Find the parametric representation of the 2-dimensional plane in R^3 that contains the three points

$$X = (x_1, x_2, x_3),\ Y = (y_1, y_2, y_3),\ Z = (z_1, z_2, z_3)$$

we let

$$w_1 \equiv \overrightarrow{XY} = (y_1 - x_1, y_2 - x_2, y_3 - x_3)$$

$$w_2 \equiv \overrightarrow{XZ} = (z_1 - x_1, z_2 - x_2, z_3 - x_3)$$



$\text{span}\{w_1, w_2\}$ represents the plane containing the 3 points X, Y, Z. To obtain an equation for the plane in question, we let

$$b \equiv \overrightarrow{OX} = (x_1, x_2, x_3).$$

Hence, the equation of the plane is given by parametrically by

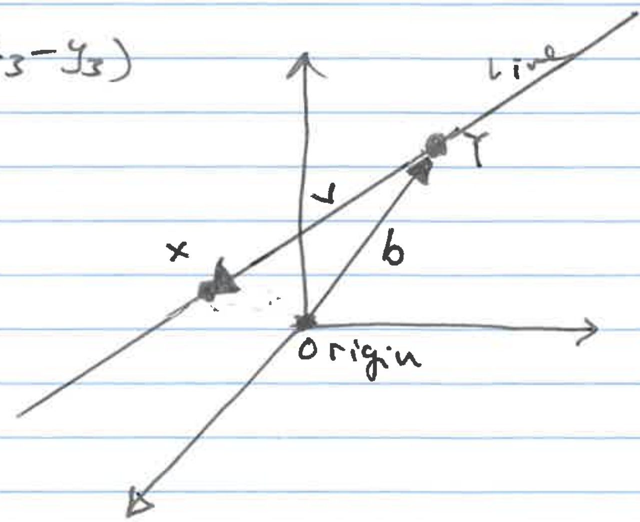
$$\boxed{\begin{array}{l} b + \text{span}\{w_1, w_2\} = b + t w_1 + s w_2,\ \forall t, s \in R \\ \text{or} \\ (x_1, x_2, x_3) + t(y_1 - x_1, y_2 - x_2, y_3 - x_3) + s(z_1 - x_1, z_2 - x_2, z_3 - x_3) \end{array}}$$

Ex. Find the parametric representation of the 1-dimensional line in R^3 containing the 2 points

$$X = (x_1, x_2, x_3), \quad Y = (y_1, y_2, y_3)$$

$$v \equiv \overrightarrow{YX} = (x_1 - y_1, x_2 - y_2, x_3 - y_3)$$

$$b = \overrightarrow{OY} = (y_1, y_2, y_3)$$



The parametric representation of the line is given by

$$b + \text{Span}\{v\}$$

$$= b + tv, \quad \forall t \in R$$

i.e.

$$(y_1, y_2, y_3) + t(x_1 - y_1, x_2 - Y_2, x_3 - y_3) \quad \forall t \in R.$$

Ex. Find the parametric representation of the plane in E^4 that passes through the 3 points

$$X = (1, -3, 4, 5), \ Y = (7, 0, -1, 2), \ Z = (5, 1, 3, 4).$$

we need to choose one of the 3 points to define the translation vector b. It really doesnot matter which of the 3 we choose. Say we use Y

$$b = (7, 0, -1, 2), \quad w_1 = \overrightarrow{YX} = (-6, -3, 5, 3)$$

$$w_2 = \overrightarrow{YZ} = (-2, 1, 4, 2)$$

Thus the parametric representation is

$$b + tw_1 + sw_2 = (7,0,-1,2) + t(-6,-3,5,3) + s(-2,1,4,2) \quad \forall t, s \in \mathbb{R}.$$

<u>Remark</u> We have to be careful when speaking of the dimension of a plane.

When a plane is given as an affine variety $b+W$, where W is a subspace of $\mathbb{R}^n$, then there is no ambiguity: the dimension of the plane is that of W. In the example right above, the dimension is 2 since w_1 and w_2 are <u>linearly independent</u>.

However, when we construct a plane given a number <u>m</u> of points it contains, it may be tempting to conclude that the plane will have dimension $m-1$. It may happen though that the given points may not produce $m-1$ linearly independent vectors. So ultimately, the dimension of the plane is the number of linearly independent vectors generated.

<u>Ex.</u> Construct the parametric representation of the plane in E^3 passing through the 3 points

$$x=(1,1,1), y=(2,2,2) \text{ and } z=(3,3,3)$$

We choose $b = \overrightarrow{OZ} = (3,3,3)$.

$$w_1 = \overrightarrow{ZX} = (-2,-2,-2), w_2 = \overrightarrow{ZY} = (-1,-1,-1).$$

We see that w_1 and w_2 are <u>linearly</u> <u>dependent</u>

Hence

$$\text{span}\{w_1, w_2\} = \text{span}\{(-1,-1,-1)\}$$

That is, the so-called "plane" is really the line
$(1,1,1) + t(-1,-1,-1) \quad t \in \mathbb{R}.$

$\boxed{\text{Exercises}}$

1). Apply the Gram-Schmidt process to the set

$v_1=(1,0,1),\ v_2=(0,1,2),\ v_3=(2,1,0)$

To obtain an <u>orthonormal set</u> u_1, u_2, u_3

<u>Solution</u>

<u>First</u> we get an orthogonal set by Gram-Schmidt:

$$\tilde{v}_1 = v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$$

$$\tilde{v}_2 = v_2 - \frac{v_2\cdot\tilde{v}_1}{\|\tilde{v}_1\|^2}\tilde{v}_1 = \begin{pmatrix}0\\1\\2\end{pmatrix} - \frac{\overset{2}{(0,1,2)\cdot(1,0,1)}}{\underset{2}{\|(1,0,1)\|^2}}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$\tilde{v}_3 = v_3 - \frac{v_3\cdot\tilde{v}_1}{\|\tilde{v}_1\|^2}\tilde{v}_1 - \frac{v_3\cdot\tilde{v}_2}{\|\tilde{v}_2\|^2}\tilde{v}_2$$

$$= \begin{pmatrix}2\\1\\0\end{pmatrix} - \frac{\overbrace{(2,1,0)\cdot(1,0,1)}^{2}}{\underbrace{\|(1,0,1)\|^2}_{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{\overbrace{(2,1,0)\cdot(-1,1,1)}^{-1}}{\underbrace{\|(-1,1,1)\|^2}_{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$= \begin{pmatrix}4/3\\2/3\\-4/3\end{pmatrix} = \frac{2}{3}\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

we then <u>normalize</u> by dividing each vector by its norm

$$u_1 = \frac{\tilde{v}_1}{\|\tilde{v}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} \quad \Leftarrow \quad \|\tilde{v}_1\| = \sqrt{1^2+0^2+1} = \sqrt{2}$$

$$u_2 = \frac{\tilde{v}_2}{\|\tilde{v}_2\|} = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix} \quad \Leftarrow \quad \|\tilde{v}_2\| = \sqrt{1+1+1} = \sqrt{3}$$

$$u_3 = \frac{\tilde{v}_3}{\|\tilde{v}_3\|} = \frac{\cancel{\frac{2}{3}}\begin{pmatrix}1\\2\\-1\end{pmatrix}}{\cancel{\frac{2}{3}}\sqrt{6}} = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\-1\end{pmatrix} \quad \Leftarrow \quad \|\tilde{v}_3\| = \frac{2}{3}\sqrt{1^2+2^2+(-1)^2} = \frac{2}{3}\sqrt{6}$$

2) show that $R^3 = W_1 \oplus W_2$ where

$W_1 = \text{span}\{(1,0,1),(0,1,2)\}$ and $W_2 = \text{span}\{(2,1,0)\}$.

solution

we need to show 2 things

① $S = \{(1,0,1),(0,1,2),(2,1,0)\}$ are linearly independent.
This implies that S is a basis for R^3 and that
$R^3 = W_1 + W_2$

② $W_1 \cap W_2 = \{0\} \Rightarrow R^3 = W_1 \oplus W_2$: the sum is direct.

① put the 3 vectors as columns (rows would work also)

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -4 \end{bmatrix}$$

rank $= 3 \Rightarrow$ cols. are lin. indep. ✓

② This can be done using dimensional arguments.

recall

$$\underbrace{\dim(W_1 + W_2)}_{\substack{\shortparallel \\ 3}} = \dim W_1 + \dim W_2 - \dim\{W_1 \cap W_2\}$$

$$= \underbrace{2 + 1}_{3} - \dim\{W_1 \cap W_2\}$$

since $W_1 + W_2 = R^3$

$\Rightarrow \dim(W_1 \cap W_2) = 0 \Rightarrow W_1 \cap W_2 = \{0\}$. ✓

3) Find the Projection matrix $P : R^3 \to \text{span}\{(1,0,1),(0,1,2)\}$ Knowing that (from 2))

$$R^3 = \text{span}\{\overset{w_1}{(1,0,1)},\overset{w_2}{(0,1,2)}\} \oplus \text{span}\{\overset{w_3}{(2,1,0)}\}.$$

<u>Solution</u>

let $v = (a; b; c) \in R^3$

we write

$$\circledast \quad (a, b, c) = \alpha_1 (1,0,1) + \alpha_2 (0,1,2) + \alpha_3 (2,1,0)$$

then

$$Pv = \alpha_1 \overset{w_1}{(1,0,1)} + \alpha_2 \overset{w_2}{(0,1,2)}.$$

$$\circledast \Rightarrow \underset{A}{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}} \underset{\alpha}{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}} \underset{=}{=} \underset{v}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}} \Rightarrow \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \frac{1}{4} \underset{A^{-1}}{\begin{bmatrix} 2 & -4 & 2 \\ -1 & 2 & 1 \\ 1 & 2 & -1 \end{bmatrix}} \underset{v}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}$$

$\boxed{\text{used Matlab for } A^{-1}}$

$$Pv = \overbrace{\frac{1}{4}(2a - 4b + 2c)}^{\alpha_1} \overset{w_1}{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}} + \overbrace{\frac{1}{4}(-a + 2b + c)}^{\alpha_2} \overset{w_2}{\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}}$$

$$= \frac{1}{4} \begin{bmatrix} 2a - 4b + 2c + 0 \\ 0 - a + 2b + c \\ \cancel{2a} - \cancel{4b} + 2c - \cancel{2a} + \cancel{4b} + 2c \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2a - 4b + 2c \\ -a + 2b + c \\ 4c \end{bmatrix}$$

$$= \frac{1}{4} \begin{bmatrix} 2 & -4 & 2 \\ -1 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Hence the projection matrix P into W_1 is given by

$$P = \frac{1}{4} \begin{bmatrix} 2 & -4 & 2 \\ -1 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1 & 1/2 \\ -1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{bmatrix} \quad \leftarrow \text{rank} = 2$$

The projection matrix into W_2 is given by

$$I - P = \begin{bmatrix} 1/2 & 1 & -1/2 \\ 1/4 & 1/2 & -1/4 \\ 0 & 0 & 0 \end{bmatrix} \quad \leftarrow \text{rank} = 1$$

4) Find the projection matrix P into W_1 where

$$R^3 = W_1 \oplus W_2$$

and we have the orthonormal basis (done in problem 1)

$$\underbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}}_{u_1}, \quad \underbrace{\frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}}_{u_2}, \quad \underbrace{\frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\-1\end{pmatrix}}_{u_3}$$

$$W_1 = \operatorname{span}\{u_1, u_2\}, \quad W_2 = \operatorname{span}\{u_3\}$$

In this case, the projection operator $P: R^3 \to W_1$ is given by

$$Pv = \frac{v \cdot u_1}{\underbrace{\|u_1\|^2}_{=1}} u_1 + \frac{v \cdot u_2}{\underbrace{\|u_2\|^2}_{=1}} u_2$$

$$= \left[(a, b, c) \cdot \frac{1}{\sqrt{2}}(1, 0, 1)\right] \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} + \left[(a, b, c)\frac{1}{\sqrt{3}}(-1, 1, 1)\right]\frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$= \frac{1}{2}(a+c)\begin{pmatrix}1\\0\\1\end{pmatrix} + \frac{1}{3}(-a+b+c)\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}a + \frac{1}{2}c + \frac{1}{3}a - \frac{1}{3}b - \frac{1}{3}c \\ 0 - \frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c \\ \frac{a}{2} + \frac{c}{2} - \frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c \end{pmatrix} = \begin{pmatrix} \frac{5}{6}a - \frac{1}{3}b + \frac{1}{6}c \\ -\frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c \\ \frac{1}{6}a + \frac{1}{3}b + \frac{5}{6}c \end{pmatrix}$$

$$= \underbrace{\frac{1}{6}\begin{bmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}}_{P} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \text{rank } P = 2 = \dim W_1.$$

The projection matrix into W_2 is given by $I - P = \frac{1}{6}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix}$ rank $=1$ $= \dim W_2$

5) Find the orthogonal complement $W^\perp$ in R^4 of

$$W = \text{span}\{\overset{w_1}{(1,0,3,4)}, \overset{w_2}{(2,3,-2,5)}\}$$

Solution

Arrange w_1, w_2 as rows of a matrix

$$\Rightarrow \quad A = \begin{bmatrix} 1 & 0 & 3 & 4 \\ 2 & 3 & -2 & 5 \end{bmatrix}$$

clearly row space of $A = \text{span}\{(1,0,3,4), (2,3-2,5)\} = W$
we know, (see notes) that

$$R^n = \ker(A) \oplus \text{range}(A^T) \quad \text{and} \quad \text{range}(A^T) = \ker(A)^\perp.$$

Now range $(A^T) =$ col. space of $A^T =$ row space of $A = W$.
Hence to find $W^\perp$ we need to find $\ker(A)$

$$A = \begin{bmatrix} 1 & 0 & 3 & 4 \\ 2 & 3 & -2 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 3 & -8 & -3 \end{bmatrix}$$

pivot vars. x_1, x_2
free vars. x_3, x_4

basis for $\ker(A)$

$$x_3 = 1, x_4 = 0 \Rightarrow x_2 = \tfrac{8}{3}, \quad x_1 = -3 \Rightarrow z_1 = (-3, \tfrac{8}{3}, 1, 0)$$

$$x_3 = 0, \quad x_4 = 1 \Rightarrow x_2 = 1, \; x_1 = -4 \Rightarrow z_2 = (-4, 1, 0, 1).$$

Hence $W^\perp = \text{span}\{(-3, \tfrac{8}{3}, 1, 0), (-4, 1, 0, 1)\}$.

---

6) Let A be an $m \times n$ matrix, $x \in R^m$, $y \in R^n$.
show that

$$x \cdot Ay = y \cdot A^T x$$

Solution/proof

$$x \cdot Ay = \sum_{i=1}^{m} x_i (Ay)_i = \sum_{i=1}^{m} x_i \sum_{j=1}^{n} a_{ij} y_j$$

interchange sums

$$= \sum_{j=1}^{n} y_j \sum_{i=1}^{m} a_{ij} x_i = \sum_{j=1}^{n} y_j (A^T x)_j$$

$$= y \cdot A^T x. \quad \checkmark$$

7) Find the intersection point(s) (if any) of the

Line: $(1,-1,3)+t(2,3,-5)$

and the plane

$$(2,5,7)+s(1,1,1)+\mu(-3,4,5)$$

$\boxed{\text{At the intersection point(s) equality must hold!}}$

The problem consists in finding values of the parameters t, s, μ such that

$$(1,-1,3)+t(2,3,-5)=(2,5,7)+s(1,1,1)+\mu(-3,4,5)$$

$$\Leftrightarrow \quad -t(2,3,-5)+s(1,1,1)+\mu(-3,4,5)=(1,-1,3)-(2,5,7)$$
$$=(-1,-6,-4)$$

This translates into the system

$$\begin{bmatrix} -2 & 1 & -3 \\ -3 & 1 & 4 \\ 5 & 1 & 5 \end{bmatrix}\begin{bmatrix} t \\ s \\ \mu \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \\ -4 \end{bmatrix} \xRightarrow{\text{Matlab}} \begin{bmatrix} t \\ s \\ \mu \end{bmatrix} = \begin{bmatrix} 1/3 \\ -7/3 \\ -2/3 \end{bmatrix}$$

This shows that there is a unique intersection point P given by

$$(1,-1,3)+\frac{1}{3}(2,3,-5)=(5/3,0,4/3)$$

or

$$(2,5,7)+(-7/3)(1,1,1)+(-2/3)(-3,4,5)=(5/3,0,4/3).$$

as expected! equal

Remark This procedure can be applied to find intersection points between any two affine varieties ("planes") in $\mathbb{R}^n$.

1) Equate the two parametric representations
2) Obtain a linear system.
3) solve linear system.

$\boxed{\text{If no solution exists, then that means the two planes do not intersect.}}$

8). Find the intersection point(s), if any of the two planes

$$(1,-1,3) + t(2,-3,4) + s(5,7,4)$$
and
$$(2,7,5) + \mu(-1,5,4) + \eta(-3,4,9)$$

Note we need 4 parameters. Two for each plane

Equate and obtain system:

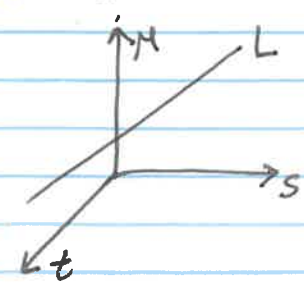
$$\begin{bmatrix} 2 & 5 & 1 & 3 \\ -3 & 7 & -5 & -4 \\ 4 & 4 & -4 & -9 \end{bmatrix} \begin{bmatrix} t \\ s \\ \mu \\ \eta \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{cccc|c} 2 & 5 & 1 & 3 & 1 \\ -3 & 7 & -5 & -4 & 8 \\ 4 & 4 & -4 & -9 & 2 \end{array}\right] \xRightarrow[\text{of Matlab}]{\text{use rref}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -7/9 & 5/9 \\ 0 & 1 & 0 & 37/72 & -19/36 \\ 0 & 0 & 1 & \frac{143}{72} & 19/36 \end{array}\right].$$

We see that we have one free variable, η which can be arbitrary

$$\Rightarrow \boxed{\mu = \frac{19}{36} - \frac{143}{72}\eta}, \quad \boxed{s = \frac{-19}{36} - \frac{37}{72}\eta}, \quad \boxed{t = \frac{5}{9} + \frac{7}{9}\eta}$$

We have thus obtained a one-parameter family of point in R^3, i.e. a line L. To give this line a geometric meaning we choose some ordering of the 3 variables μ, s, t, say (t, s, μ)



$$\Rightarrow L \leftrightarrow \left(\frac{5}{9} + \frac{7}{9}\eta, \frac{-19}{36} - \frac{37}{72}\eta, \frac{19}{36} - \frac{143}{72}\eta\right)$$

$$= \boxed{\left(\frac{5}{9}, \frac{-19}{36}, \frac{19}{36}\right) + \eta\left(\frac{7}{9}, \frac{-37}{72}, \frac{-143}{72}\right)}$$

9) Show that the two lines L_1, L_2 in R^3

$L_1 \Leftrightarrow (1,-3,4) + t(2,5,7)$

$L_2 \Leftrightarrow (-3,5,7) + s(-1,4,5)$

do not intersect.

<u>solution</u>

we equate the two expressions:

$$(1,-3,4)+t(2,5,7) = (-3,5,7) + s(-1,4,5)$$

$$\Rightarrow \begin{bmatrix} 2 & 1 \\ 5 & -4 \\ 7 & -5 \end{bmatrix} \begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \\ 3 \end{bmatrix}$$

$$\Rightarrow \underbrace{\begin{bmatrix} 2 & 1 \\ 5 & -4 \\ 7 & -5 \end{bmatrix}}_{A} \Bigg| \underbrace{\begin{matrix} -4 \\ 8 \\ 3 \end{matrix}}_{b} \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

The system has no solution since $\operatorname{rank}(A) = 2$ but $\operatorname{rank}(A|b) = 3$.

<u>Remark</u> In R^2, two lines that have no intersection point must be parallel. This is not true in R^3 as the above example shows. Indeed, there is no intersection point, yet the two lines are <u>not parallel</u>.

In R^n, two affine varieties

$b_1 + W_1$ and $b_2 + W_2$

are <u>parallel</u> if

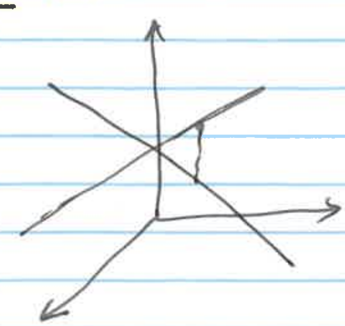
W_1 is a subspace of W_2

or

W_2 is a subspace of W_1

or

$W_1 = W_2$.



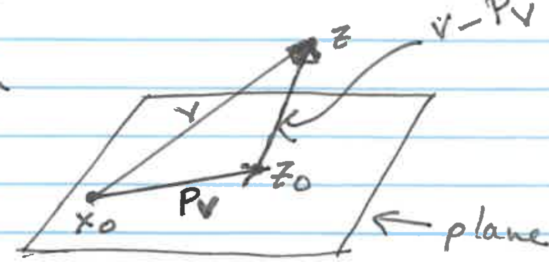
(10) Distance from a point to a plane in R^n

Suppose we are given a plane $b+W$ and a point z not in the plane. It can be shown that there is a unique point z_0 in the plane which is closest to z in the sense that

$$\|z - z_0\| = \min_{x \in b+W} \|z - x\|.$$

We define the distance between z and the plane $b+W$ to be

$$\|z - z_0\|$$



We can find the point z_0 and thus the distance $\|z-z_0\|$ using minimization techniques. However we will outline a method to calculate the distance using Linear Algebra

1) Find a point x_0 in the plane.
2) form the vector $v = z - x_0$
3) Find the orthogonal complement $W^\perp$ of W
4) Calculate the projection Pv of v using $R^n = W \oplus W^\perp$.

Then

distance $= \|z - z_0\| = \|(I-P)v\|$.

Ex. Calculate the distance from the point $z = (7,-1,3)$ to the plane (actually line)

$$L \leftrightarrow (1,-2,3) + t(2,3,-4).$$

1) let $t=0 \Rightarrow x_0 = (1,-2,3)$
2) $v = z - x_0 = (7,-1,3) - (1,-2,3) = (6,1,0)$

3) We already saw how to do this:

$$W = \text{span}\{\overbrace{(2,3,-4)}^{w_1}\}.$$

arrange w_1 as a row of $A = \begin{bmatrix} 2 & 3 & -4 \end{bmatrix}$.

Then $W^{\perp} = \ker(A)$. we next construct a basis for $\ker(A)$:

$x_2 = 1, x_3 = 0 \Rightarrow x_1 = -3/2 \rightarrow w_2 = (-3/2, 1, 0)$

$x_1 = 0, x_3 = 1 \Rightarrow x_1 = 2 \rightarrow w_3 = (2, 0, 1)$.

4) ~~Recall~~ write $v = \alpha_1 w_1 + \alpha_2 w_2 + \alpha_3 w_3$. Then $\boxed{Pv = \alpha_1 w_1}$

To calculate α_1, we need to solve the system

$$\begin{array}{c} \begin{array}{ccc} w_1 & w_2 & w_3 \end{array} \\ \begin{bmatrix} 2 & -3/2 & 2 \\ 3 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \end{array} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}.$$

using Matlab, $\boxed{\alpha_1 = \frac{15}{29}, \ \alpha_2 = -\frac{16}{29}, \ \alpha_3 = \frac{60}{29}}$

Hence $Pv = \frac{15}{29}(2, 3, -4)$

$\Rightarrow \quad v - Pv = (6, 1, 0) - \frac{15}{29}(2, 3, -4) = \frac{1}{29}(144, -16, 60)$

$$\text{distance} = \|v - Pv\| = \frac{1}{29}\left((144)^2 + (-16)^2 + (60)^2\right)^{1/2}$$

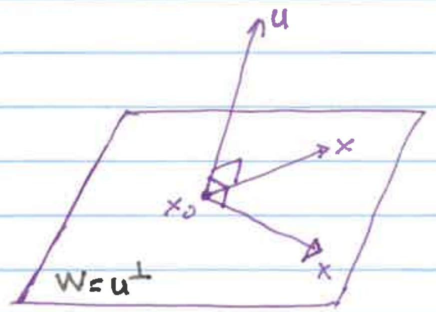
$$= \frac{1}{29}\sqrt{24{,}592} \cong \boxed{5.4075}$$

A special case: (n-1)-dimensional planes in R^n

Let $u \in R^n$ and $d \in R$ be given. Then the plane normal (orthogonal) to u and passing through the point (vector) $x_0 \in R^n$ is given by

$$\{x \in R^n \mid u \cdot (x - x_0) = 0\} = \{x \in R^n \mid u \cdot x = d \equiv u \cdot x_0\}.$$

The plane viewed as a subset of R^n is a subspace of R^n if and only if $d = 0$



Distance from a point to a plane.

Given a point z not in a plane $u \cdot x = d$, it can be shown that there is a unique z_0 in the plane which is closest to z in the sense that

$$\|z - z_0\| = \min_{x \in \text{plane}} \|z - x\| \equiv \text{distance between } z \text{ and the plane } u \cdot x = d.$$

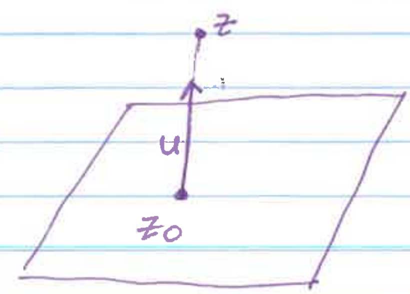
We want to develop a formula for calculating $\|z - z_0\|$ which we define as the distance from z to the plane.

It is clear that

$$z - z_0 = \alpha u$$

for some scalar α.

Now



$$(z - z_0) \cdot u = \alpha u \cdot u = \alpha \|u\|^2$$

$\Rightarrow$

$$z \cdot u - \overbrace{z_0 \cdot u}^{d} = \alpha \|u\|^2 \Rightarrow \alpha \|u\|^2 = z \cdot u - d$$

$$\Rightarrow \alpha = \frac{z \cdot u - d}{\|u\|^2}$$

note that we can compute the value of α since z, u and d are known. Also from $z - z_0 = \alpha u$

$$\|z - z_0\| = \|\alpha u\| = |\alpha| \, \|u\|$$

$$\Rightarrow \boxed{\text{distance} = \frac{|z \cdot u - d|}{\|u\|^2} \|u\| = \frac{|z \cdot u - d|}{\|u\|}}$$

Ex. Calculate the distance between the point $(1, -2, 3) = z$ and the plane $3x - 7y + 12z = 4$.

Here $u = (3, 7, 12)$, $d = 4$.

$$\text{distance} = \frac{|(1,-2,3) \cdot (3,7,12) - 4|}{\|u\| = \|(3,7,12)\|} = \frac{46}{\sqrt{179}}.$$