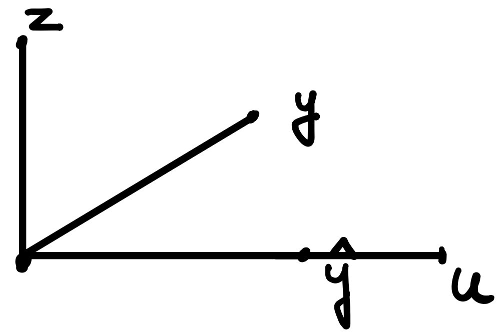
Orthogonal projections

We now introduce a construction that will keep coming up:

Given a nonzero vector u in $\mathbb{R}^n$, we would like to decompose a vector $y \in \mathbb{R}^n$ as $\hat{y} + z$, where $\hat{y}$ is of the form αu for some $\alpha \in \mathbb{R}$, and z is orthogonal to u.



Given any $\alpha \in \mathbb{R}$, let $z = y - \alpha u$.

We want z to be orthogonal to u,

so $z \cdot u = 0$ means that

$$(y - \alpha u) \cdot u = 0$$

i.e., $y \cdot u - \alpha (u \cdot u) = 0$,

so $\alpha = \dfrac{y \cdot u}{u \cdot u}$.

Thus, the required decomposition

of y as $\hat{y} + z$ with $\hat{y} = \alpha u$,
$z \cdot u = 0$

is possible if and only if

$$\alpha = \frac{y \cdot u}{u \cdot u}, \quad \text{and so } \hat{y} = \frac{y \cdot u}{u \cdot u} u .$$

The vector $\hat{y}$ is called the "orthogonal projection of y onto u", also denoted as "$\text{proj}_u y$".

and the vector z is called the "complement of y orthogonal to u".

Observe that what we are doing here is similar to extracting the x-coordinate and y-coordinate from a vector (x, y) in $\mathbb{R}^2$, because $(x, 0)$

is the projection of (x,y) onto the x-axis, or span$\{(1,0)\}$ while $(0,y)$ is the projection of (x,y) onto the y-axis, or span$\{(0,1)\}$, and these two axes are orthogonal to each other.

But we might not be so lucky to be able to use the standard basis for $\mathbb{R}^n$, and so the construction above is a way to perform the decomposition

$$(x,y) = x(1,0) + (0,y)$$

in $\mathbb{R}^n$ with a different orthogonal basis.

In summary:

For two vectors y and u, the "projection of y on u" is denoted by $\text{proj}_u y$ or $\hat{y}$ and can be calculated as

$$\boxed{\hat{y} = \text{proj}_u y = \frac{y \cdot u}{u \cdot u} u.}$$

Then, we can write

$$y = \hat{y} + z.$$

$\hat{y}$: orthogonal projection of y on u → calculated as $\frac{y \cdot u}{u \cdot u} u$

z: complement of y, orthogonal to u → calculated as $y - \hat{y}$, (so compute $\hat{y}$ first)

Example Let $y = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

Find the orthogonal projection of y onto u. Then write y as the sum of two orthogonal vectors, one in Span$\{u\}$ and one orthogonal to u.

Solution $\hat{y} = \dfrac{y \cdot u}{u \cdot u} \cdot u$

$$= \frac{\begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix}}{\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix}} \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$\therefore$ $$= 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$z = y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

$$\therefore\ y = \hat{y} + z$$

$$= \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

and $\hat{y}$, z are orthogonal vectors

(check: $\begin{pmatrix} 8 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \end{pmatrix} = 0$).

Observe that, the scalar α obtained in the course of orthogonal projection is $\dfrac{y \cdot u}{u \cdot u}$, very similar to the scalars obtained earlier when we expressed an arbitrary vector as a linear combination of basis vectors that are orthogonal.

This tells us that when we have $\{u_1, \ldots, u_p\}$, an orthogonal basis for W and $y \in W$ is expressed as

$$y = c_1 u_1 + \cdots + c_p u_p$$

with
$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j},$$

y is being decomposed as a sum of orthogonal projections along $u_1, u_2, \ldots, u_p$.

Orthonormal sets

A set $\{u_1, \dots, u_p\}$ is an orthonormal set if it is an orthogonal set of <u>unit vectors</u>.

If $W = \text{span}\{u_1, \dots, u_p\}$ then $\{u_1, \dots, u_p\}$ automatically forms a basis, because these vectors are lin. dependent.
Moreover u_i's are orthonormal and so they form an <u>orthonormal basis</u> (o.n.b) for W.

<u>Examples</u>

1) The standard basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ is an o.n.b.

2) The vectors $(3,1,1)$, $(-1,2,1)$ and $(-1,-4,7)$ form an orthogonal set in $\mathbb{R}^3$, and hence form a basis. However, these are not unit vectors. Thus, we divide each vector by its norm to get an <u>orthonormal basis</u>:

$$\left\{ \left(\frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right), \left(\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(\frac{-1}{\sqrt{66}}, \frac{-4}{\sqrt{66}}, \frac{7}{\sqrt{66}}\right) \right\}$$

Matrices whose columns form an orthonormal set are important for applications in computer algorithms.

<u>Theorem</u>: An $m \times n$ matrix U has orthonormal columns $\Leftrightarrow$ $U^T U = I$.

Gram - Schmidt process

This is an algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of $\mathbb{R}^n$.

Given a basis $\{x_1, x_2, \ldots, x_p\}$ for a subspace W of $\mathbb{R}^n$, define

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$\vdots$$

$$v_p = x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}.$$

<u>Remark</u>: The above process results in an orthogonal basis. To get an orthonormal basis, we would normalise the vector

v_i by dividing by $\|v_i\|$ so that we get unit vectors.

<u>Example 1</u> Let Π be the plane in $\mathbb{R}^3$ spanned by vectors $x_1 = (1,2,2)$ and $x_2 = (-1,0,2)$.

Find an orthonormal basis for Π.

<u>Solution</u>.

$$v_1 = x_1 = (1,2,2)$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$= (-1,0,2) - \frac{3}{9} \cdot (1,2,2)$$

$$= \left(-4/3, -2/3, 4/3\right)$$

Thus, $\{v_1, v_2\}$ for an orthogonal basis for Π.

Dividing by norms: $w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{3}(1,2,2)$

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{1}{3}(-2,-1,2)$$

$\therefore$ $\{w_1, w_2\}$ form an orthonormal basis for Π.

Exercises based on the above material:

① For the vectors $u = (5,-1,2)$
$v = (2,-1,-3)$

Find $\text{proj}_u v$ and $\text{proj}_u v$.

Ans. $\text{proj}_u v = \frac{1}{6}(5, -1, 2) = \hat{v}$

$\text{proj}_v u = \frac{5}{14}(2,-1,-3) = \hat{u}$

② Using the Gram-Schmidt process, convert the following basis to an orthonormal one:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 1 \end{pmatrix}.$$

Answer: $u_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad u_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$

$$u_3 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}.$$