Inner Product Spaces (Part II)

Conceptual knowledge

- Understand various contexts of orthogonal and orthonormal.
- See applications in other areas of mathematics of orthogonal sets.
- Understand how least squares is applied.

Procedural knowledge

- Determining whether a set of vectors forms an orthonormal basis or not.
- Find coordinates relative to orthonormal bases.
- Perform Gram-Schmidt orthonormalisation.
- Solve least squares problems.
- Find projections of vectors onto subspaces.

1 Orthonormal Bases

1.1 Fundamentals

Definition

Vectors in an inner product space are ***orthogonal*** if their inner product is 0.

If the vectors are also unit vectors then we call them ***orthonormal***.

A set of vectors in an inner product space is ***orthogonal*** if all vectors are ***mutually orthogonal***. In other words all pairs of vectors are orthogonal pairs.

An orthogonal set of vectors consisting of unit vectors is called ***orthonormal***.

If the set of orthonormal vectors is a basis we call it an ***orthonormal basis***.

An ***orthogonal matrix*** is a matrix whose rows are made from orthonormal vectors.

Example 1.1
Show that the matrix whose rows are the standard basis vectors in $\mathbb{R}^3$, with the Eulcidean inner product, is an orthogonal matrix.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Longrightarrow \begin{array}{l} (1,0,0)\cdot(0,1,0) = 0 \\ (1,0,0)\cdot(0,0,1) = 0 \\ (0,1,0)\cdot(0,0,1) = 0 \end{array}$$

Orthogonal matrix

Mutually orthogonal

Also,
$$\|(1,0,0)\| = 1$$
$$\|(0,1,0)\| = 1$$
$$\|(0,0,1)\| = 1$$
$\therefore$ orthonormal

Note that the set of column vectors is also orthonormal.

Let's look at an example of an orthonormal basis that is non-standard.

Example 1.2
Verify that the following set is an orthonormal basis for $\mathbb{R}^3$ and describe what the basis looks like.

$$S = \{\underset{}{\overset{v_1}{(\cos\theta, \sin\theta, 0)}}, \overset{v_2}{(-\sin\theta, \cos\theta, 0)}, \overset{v_3}{(0,0,1)}\}$$

$$(\cos\theta, \sin\theta, 0)\cdot(-\sin\theta, \cos\theta, 0) = -\cos\theta\sin\theta + \sin\theta\cos\theta + 0$$
$$= 0$$

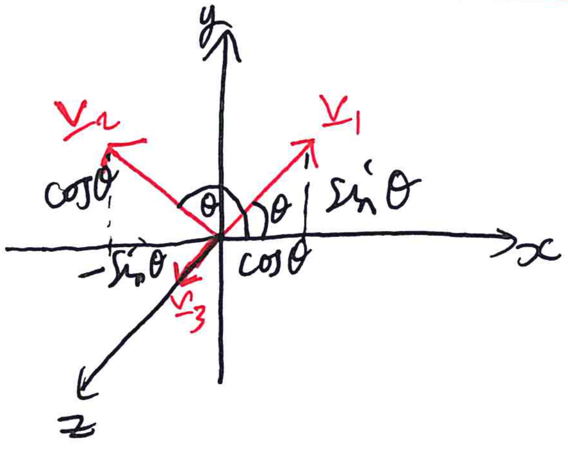
$$(\cos\theta, \sin\theta, 0)\cdot(0,0,1) = 0 = (-\sin\theta, \cos\theta, 0)\cdot(0,0,1)$$

Mutually orthogonal

$$\|(\cos\theta, \sin\theta, 0)\| = \sqrt{\cos^2\theta + \sin^2\theta + 0} = 1 = \|(-\sin\theta, \cos\theta, 0)\|$$

$$\|(0,0,1)\| = 1$$ All unit vectors, $\therefore$ orthonormal

By inspection, the vectors are linearly independent, $\therefore$ they form an orthonormal basis.



The next example will provide the background to some of our assumptions regarding the derivation of Fourier series.

Example 1.3
Define the inner product for $C[-L, L]$ to be,

$$\langle f, g\rangle = \int_{-L}^{L} f(x)g(x)\,dx,$$

with $L \in \mathbb{R}$. Show that the following set is orthogonal.

$$\left\{\cos\left(\frac{n\pi x}{L}\right)\right\}_{n=0}^{\infty} = \left\{1, \cos\frac{\pi x}{L}, \ldots, \cos\frac{n\pi x}{L}, \ldots\right\}$$

$$\left\langle \cos\left(\frac{m\pi x}{L}\right), \cos\left(\frac{n\pi x}{L}\right)\right\rangle = \int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right)dx$$

$m \neq n \geqslant 0$

$$= \int_{-L}^{L} \frac{1}{2}\left[\cos\left(\frac{\pi x}{L}(m+n)\right) + \cos\left(\frac{\pi x}{L}(m-n)\right)\right]dx$$

$$= \frac{1}{2}\left[\frac{L}{(m+n)\pi}\sin\left(\frac{\pi x}{L}(m+n)\right) + \frac{L}{(m-n)\pi}\sin\left(\frac{\pi x}{L}(m-n)\right)\right]_{-L}^{L}$$

$$= 0 \quad \therefore \text{ orthogonal}$$

Note that the set is orthogonal but not orthonormal.

It is possible to make an orthonormal set by dividing each term by its respective norm.

So given, $\left\langle \cos\left(\frac{n\pi x}{L}\right), \cos\left(\frac{n\pi x}{L}\right)\right\rangle = \left\|\cos\left(\frac{n\pi x}{L}\right)\right\|^2$

$$= \int_{-L}^{L} \cos^2\left(\frac{n\pi x}{L}\right)dx = \frac{1}{2}\int_{-L}^{L} 1 + \cos\left(\frac{2n\pi x}{L}\right)dx$$

$$= \left[\frac{x}{2} + \frac{L}{4n\pi}\sin\left(\frac{2n\pi x}{L}\right)\right]_{-L}^{L} = \frac{L}{2} - \left(\frac{-L}{2}\right)$$

Not valid for $n=0$

$$= L$$

$$\langle 1,1\rangle = \|1\|^2 = \int_{-L}^{L} dx = [x]_{-L}^{L} = 2L$$

we can make the set orthonormal:

$$\left\{ \frac{1}{a} \cos\left(\frac{n\pi x}{L}\right) \right\}_{n=0}^{\infty}, \quad a = \begin{cases} \sqrt{2L}, & n=0 \\ \sqrt{L}, & n>0 \end{cases}$$

1.2 Linear Independence

We will now show that orthogonal sets are always linearly independent.

> **Theorem**
>
> An orthogonal set of non-zero vectors, $S = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$, is linearly independent.

Proof Show $\sum_{i=1}^{n} c_i \underline{v}_i = \underline{0} \Rightarrow c_i = 0 \quad \forall i \in [1,n]$

Orthogonal: $\langle \underline{v}_i, \underline{v}_i \rangle = \|\underline{v}_i\|^2 \neq 0,$

$\langle \underline{v}_i, \underline{v}_j \rangle = 0, \quad i \neq j$

$$\underline{0} = \langle c_i \underline{v}_i, \underline{v}_j \rangle = \langle -\sum_{\substack{k=1,\\ k\neq i}}^{n} c_k \underline{v}_k, \underline{v}_j \rangle$$

orthogonality

$$= \langle -c_j \underline{v}_j, \underline{v}_j \rangle$$

$$= -c_j \underbrace{\langle \underline{v}_j, \underline{v}_j \rangle}_{\neq 0} \Rightarrow c_j = 0 \quad \square$$

This means that any orthogonal set containing n non-zero vectors is a basis for an inner product space of dimension, n.

We can now use this result to test for bases.

Example 1.4
Is the following set a basis for $\mathbb{R}^4$ with the standard Eulcidean inner product?

$$S = \{\underbrace{(2,3,2,-2)}_{\underline{v}_1}, \underbrace{(1,0,0,1)}_{\underline{v}_2}, \underbrace{(-1,0,2,1)}_{\underline{v}_3}, \underbrace{(-1,2,-1,1)}_{\underline{v}_4}\}$$

Our usual method would involve checking linear independence by forming a matrix then using row reduction. Now we can simply check orthogonality.

$\underline{v}_1 \cdot \underline{v}_2 = 2+0+0-2=0$ | $\underline{v}_2 \cdot \underline{v}_3 = -1+0+0+1=0$

$\underline{v}_1 \cdot \underline{v}_3 = -2+0+4-2=0$ | $\underline{v}_2 \cdot \underline{v}_4 = -1+0+0+1=0$

$\underline{v}_1 \cdot \underline{v}_4 = -2+6-2-2=0$ | $\underline{v}_3 \cdot \underline{v}_4 = 1+0-2+1=0$

All vectors mutually orthogonal $\therefore$ Linearly independent $\Longrightarrow$ basis for $\mathbb{R}^4$.

1.3 Coordinates with Orthonormal Bases

One of the advantages of using an orthonormal basis is that we can easily derive a formula for the coordinates of any vector with respect to that basis.

Say we have have an orthonormal basis, $B = \{\mathbf{v_1}, \mathbf{v_2}, , ..., \mathbf{v_n}\}$, for some inner product space, V, and we want to represent any vector, $\mathbf{w} \in V$, using the basis.

In other words, we are looking for the coordinates, c_i, of $\mathbf{w}$ relative to each $\mathbf{v_i}$ ($i \in [1, n]$) such that,

$$\underline{w} = \sum_{j=1}^{n} c_j \underline{v}_j$$

$\underline{w}$ is a linear combination of basis vectors

Using a similar idea to the last proof we looked at we can take an inner product and utilise orthogonality:

$$\langle \underline{w}, \underline{v}_i \rangle = \langle \sum_{j=1}^{n} c_j \underline{v}_j, \underline{v}_i \rangle = \langle c_i \underline{v}_i, \underline{v}_i \rangle \quad \text{orthogonality}$$

$$= c_i \langle \underline{v}_i, \underline{v}_i \rangle$$

Since we also have that the vectors are orthonormal then, norm $= 1$

$$\|\underline{v}_i\|^2 = \langle \underline{v}_i, \underline{v}_i \rangle = 1$$

$$\therefore \quad c_i = \langle \underline{w}, \underline{v}_i \rangle$$

So to get the coordinates we just take the inner product with the respective basis vector.

We call these coordinates ***Fourier coefficients*** of $\mathbf{w}$ relative to B.

Theorem

The ***coordinates*** of a vector, $\mathbf{w}$, in an inner product space, V, with respect to the orthonormal basis, $B = \{\mathbf{v_1}, \mathbf{v_2}, , ..., \mathbf{v_n}\}$, are given by the ***Fourier coefficients***:

$$[\underline{w}]_B = (c_1, c_2, \ldots, c_n)^T$$
$$= (\langle \underline{w}, \underline{v}_1 \rangle, \langle \underline{w}, \underline{v}_2 \rangle, \ldots, \langle \underline{w}, \underline{v}_n \rangle)^T$$

Example 1.5
Find the coordinates of $\mathbf{w} = (5, -5, 2)$ *relative to the orthonormal basis for* $\mathbb{R}^3$,

$$B = \{(3/5, 4/5, 0), (-4/5, 3/5, 0), (0, 0, 1)\}$$

Assume Euclidean inner product

$$c_1 = \langle \underline{w}, \underline{v}_1 \rangle = \langle (5,-5,2), (\tfrac{3}{5}, \tfrac{4}{5}, 0) \rangle = 3 - 4 + 0 = -1$$

$$c_2 = \langle (5,-5,2), (-\tfrac{4}{5}, \tfrac{3}{5}, 0) \rangle = -4 - 3 + 0 = -7$$

$$c_3 = \langle (5,-5,2), (0,0,1) \rangle = 2$$

$$\therefore [(5,-5,2)]_B = \begin{pmatrix} -1 \\ -7 \\ 2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 5 \\ -5 \\ 2 \end{pmatrix} = -\begin{pmatrix} 3/5 \\ 4/5 \\ 0 \end{pmatrix} - 7\begin{pmatrix} -4/5 \\ 3/5 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

1.4 Gram-Schmidt Orthonormalisation

Since orthonormal bases are useful in terms of their coordinate representation, we now develop a process which turns a general basis of an inner product space into an orthonormal one.

Let's start with a general basis in $\mathbb{R}^2$, $\{\mathbf{v_1}, \mathbf{v_2}\}$. Starting with the first vector, $\mathbf{v_1}$, we can find an orthogonal vector using projections.

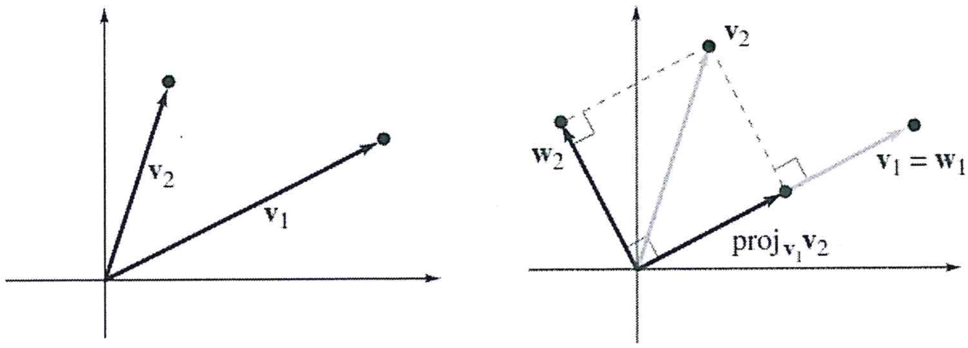


Figure 1: Orthogonal vector from projection in $\mathbb{R}^2$ ©Pearson, 2016.

Let, $\underline{w}_1 = \underline{v}_1$, $\underline{w}_2 = \underline{v}_2 - \text{proj}_{\underline{v}_1} \underline{v}_2 = \underline{v}_2 - \text{proj}_{\underline{w}_1} \underline{v}_2$

$$= \underline{v}_2 - \frac{\underline{v}_2 \cdot \underline{w}_1}{\underline{w}_1 \cdot \underline{w}_1} \underline{w}_1$$

This means that $\{\mathbf{w_1}, \mathbf{w_2}\}$ is an orthogonal basis for $\mathbb{R}^2$.

To make it orthonormal we can just divide each vector by their respective norms.

Let's check that $\mathbf{w_1}$ and $\mathbf{w_2}$ are really orthogonal:

$$\underline{w}_1 \cdot \underline{w}_2 = \underline{w}_1 \cdot \left(\underline{v}_2 - \frac{\underline{v}_2 \cdot \underline{w}_1}{\underline{w}_1 \cdot \underline{w}_1} \underline{w}_1 \right)$$

$$= \underline{w}_1 \cdot \underline{v}_2 - \frac{(\underline{v}_2 \cdot \underline{w}_1)}{\cancel{\underline{w}_1 \cdot \underline{w}_1}} \cancel{\underline{w}_1 \cdot \underline{w}_1}$$

$$= 0 \quad \therefore \text{ orthogonal}$$

Noticing how the inner product can be split into two parts that cancel out with each other suggests a nested type pattern that will all us to make third vector orthogonal to both $\mathbf{w_1}$ and $\mathbf{w_2}$.

Imagine that we had started with a general basis in $\mathbb{R}^3$, $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$, and we labelled $\mathbf{w_1} = \mathbf{v_1}$ and found an orthogonal vector, $\mathbf{w_2} = \text{proj}_{\mathbf{w_1}}\mathbf{v_2}$ as before.

Following the pattern we can define,

$$\underline{w}_3 = \underline{v}_3 - \text{proj}_{\underline{w}_1} \underline{v}_3 - \text{proj}_{\underline{w}_2} \underline{v}_3$$

I will leave it to you to verify that $\mathbf{w_3}$ is orthogonal to both of the other vectors.

This leads us to the following generalisation.

Definition

Given a general basis, $B = \{\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_n}\}$, for an inner product space, V, we obtain an orthogonal basis, $B' = \{\mathbf{w_1}, \mathbf{w_2}, \ldots \mathbf{w_n}\}$, using the formula,

$$\underline{w}_n = \begin{cases} \underline{v}_n, & n=1 \\ \underline{v}_n - \underbrace{\sum_{i=1}^{n-1} \frac{\langle \underline{v}_n, \underline{w}_i \rangle}{\langle \underline{w}_i, \underline{w}_i \rangle} \underline{w}_i}_{\text{projections}}, & n>1 \end{cases}$$

We create an orthonormal basis, B'', from B' by dividing each vector by its respective norm:

$$\underline{u}_i = \frac{\underline{w}_i}{\|\underline{w}_i\|} \Longrightarrow B'' = \{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n\}$$

The process of going from B to B'' is called ***Gram-Schmidt orthonormalisation***.

Example 1.6

Find an orthonormal basis spanning the same inner product space as the following general basis for $\mathbb{R}^2$ with the standard Euclidean inner product.

$$B = \{\overset{\underline{v}_1}{(1,1,0)}, \overset{\underline{v}_2}{(1,2,0)}, \overset{\underline{v}_3}{(0,1,2)}\}$$

$$\underline{w}_1 = \underline{v}_1 = (1,1,0)$$

$$\underline{w}_2 = \underline{v}_2 - \frac{\underline{v}_2 \cdot \underline{w}_1}{\underline{w}_1 \cdot \underline{w}_1} \underline{w}_1 = (1,2,0) - \frac{3}{2}(1,1,0) = \left(-\tfrac{1}{2}, \tfrac{1}{2}, 0\right)$$

$$\underline{w}_3 = \underline{v}_3 - \frac{\underline{v}_3 \cdot \underline{w}_1}{\underline{w}_1 \cdot \underline{w}_1} \underline{w}_1 - \frac{\underline{v}_3 \cdot \underline{w}_2}{\underline{w}_2 \cdot \underline{w}_2} \underline{w}_2$$

$$= (0,1,2) - \frac{1}{2}(1,1,0) - \frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{2}\right)}\left(-\tfrac{1}{2}, \tfrac{1}{2}, 0\right) = (0,0,2)$$

$\|\underline{w}_1\| = \sqrt{2}$, $\|\underline{w}_2\| = \frac{1}{\sqrt{2}}$, $\|\underline{w}_3\| = 2$

$\therefore B' = \left\{ \frac{1}{\sqrt{2}}(1,1,0), \sqrt{2}(-\frac{1}{2}, \frac{1}{2}, 0), (0,0,1) \right\}$

is an orthonormal basis.

Note: It is also possible to normalise the vectors in the general basis first, before using the formulas. The denominator in the sum formula would not be necessary since it is the norm, which for a unit vector is equal to 1.

2 Orthogonal Subspaces

We will now develop some ideas which can help us to find best possible solutions to inconsistent systems.

> ***Definition***
>
> Two subspaces of $\mathbb{R}^n$, S_1 and S_2, are ***orthogonal*** if every combination of a vector from S_1 with a vector from S_2 results in an orthogonal pair.

Example 2.1
Are the following subspaces orthogonal? Assume Euclidean inner product

$$S_1 = Span\{(1,0,1),(1,1,0)\}, \qquad S_2 = Span\{(-1,1,1)\}$$

$(1,0,1)\cdot(-1,1,1) = (1,1,0)\cdot(-1,1,1) = 0$

$\underline{v} \in S_1$, $\underline{v} = a_1(1,0,1) + a_2(1,1,0)$ (a_1, a_2 constant)

$\underline{w} \in S_2$, $\underline{w} = b(-1,1,1)$

$\underline{v} \cdot \underline{w} = [a_1(1,0,1) + a_2(1,1,0)] \cdot b(-1,1,1)$

$= a_1 b (1,0,1)\cdot(-1,1,1) + a_2 b (1,1,0)\cdot(-1,1,1)$

$= 0$ $\therefore S_1$ and S_2 are orthogonal.

The only vector in both subspaces is the zero vector, which is generally true for orthogonal subspaces.

Definition

If S is a subspace of $\mathbb{R}^n$ then the ***orthogonal complement*** is,

$$S^{\perp} = \{ \underline{v} \in \mathbb{R}^n ; \underline{v} \cdot \underline{v} = 0 \quad \forall \underline{v} \in S \}$$

Example 2.2
Find the orthogonal complement of the subspace spanned by the columns of the following matrix.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Col($\underline{\underline{A}}$)

Seek $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ such that $\underline{\underline{A}}^T \underline{x} = \underline{0}$

$$\Rightarrow \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_4 = 0$$

$$x_1 = -2x_2 - x_3$$

$$\therefore \underline{x} = \begin{pmatrix} -2x_2 - x_3 \\ x_2 \\ x_3 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$\therefore$ basis for $S^{\perp}$ is $\{(-2, 1, 0, 0), (-1, 0, 1, 0)\}$

Definition

The ***direct sum*** of 2 subspaces of $\mathbb{R}^n$, S_1 and S_2, is written as,

$$\mathbb{R}^n = S_1 \oplus S_2$$

if every vector, $x \in \mathbb{R}^n$, can be written as a unique sum of a vector in S_1 and a vector in S_2.

We have that,

1. $\dim(S) + \dim(S^{\perp}) = n$.

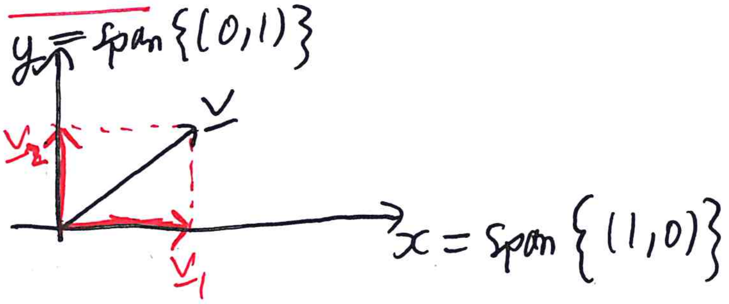
2. $\mathbb{R}^n = S \oplus S^{\perp}$.

3. $(S^{\perp})^{\perp} = S$.

> ***Definition***
>
> Given a subspace of $\mathbb{R}^n$ with its orthogonal complement, a vector, $\mathbf{v} \in \mathbb{R}^n$ can be written as,
>
> $$\underline{v} = \underline{v}_1 + \underline{v}_2 \quad , \quad \underline{v}_1 \in S, \quad \underline{v}_2 \in S^{\perp}$$
>
> where we call $\mathbf{v_1} = \text{proj}_S \mathbf{v}$ the ***projection of* $\mathbf{v}$ *onto the subspace,* S**.

Note that $\mathbf{v} - \text{proj}_S \mathbf{v}$ is orthogonal to S.

Consider this definition with the subspaces that are the Cartesian axes for $\mathbb{R}^2$:



$$\underline{v} = \underline{v}_1 + \underline{v}_2$$

Now say we have an orthonormal basis, $\{\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_p}\}$, for a subspace, S, of $\mathbb{R}^n$.

We have that,

$$\underline{v} \in \mathbb{R}^n, \quad \underline{v}_1 \in S, \quad \underline{v}_2 \in S^{\perp} \Rightarrow \underline{v} = \underline{v}_1 + \underline{v}_2$$
$$= \sum_{j=1}^{p} c_j \underline{u}_j + \underline{v}_2$$

Also, $$\underline{v} \cdot \underline{u}_i = \left(\sum_{j=1}^{p} c_j \underline{u}_j + \underline{v}_2\right) \cdot \underline{u}_i$$
$$= c_i \cancel{(\underline{u}_i \cdot \underline{u}_i)}^{1} = c_i$$

which leads us to the following formula for calculating the projection of a vector onto a subspace:

$v_1 \in S$

If $\{\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_p}\}$ is an orthonormal basis for the subspace, S, of $\mathbb{R}^n$, then,

$$(\underline{v}_1) = \text{proj}_S\, \underline{v} = \underline{v} - \underline{v}_2 = \sum_{j=1}^{p} (\underline{v} \cdot \underline{u}_j)\underline{u}_j$$

Example 2.3

Find the projection of $\mathbf{v} = (1, 1, 3)$ *onto the subspace,* S*, of* $\mathbb{R}^3$ *spanned by* $\mathbf{w_1} = (0, 3, 1)$ *and* $\mathbf{w_2} = (2, 0, 0)$.

$\Rightarrow \underline{w}_1 \cdot \underline{w}_2 = 0 \quad \therefore$ orthogonal, $\|\underline{w}_1\| = \sqrt{10}$, $\|\underline{w}_2\| = 2$

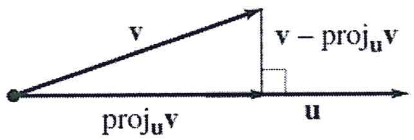
Orthonormal basis: $\left\{\frac{1}{\sqrt{10}}(0,3,1), (1,0,0)\right\}$

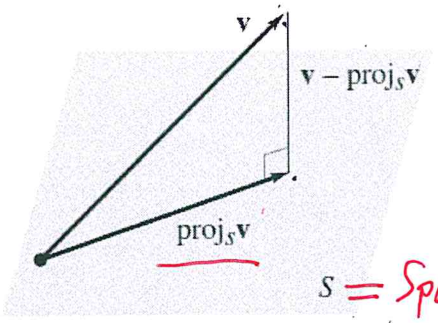
$$\therefore \text{proj}_S\, \underline{v} = (\underline{v} \cdot \underline{u}_1)\underline{u}_1 + (\underline{v} \cdot \underline{u}_2)\underline{u}_2$$

$$= \left[(1,1,3) \cdot \frac{1}{\sqrt{10}}(0,3,1)\right]\frac{1}{\sqrt{10}}(0,3,1) + \left[(1,1,3) \cdot (1,0,0)\right](1,0,0)$$

$$= \left(1, \frac{9}{5}, \frac{3}{5}\right)$$

Remember that orthogonal projections represent shortest distance.





$S = \text{Span}\{\underline{w}_1, \underline{w}_2\}$

Figure 2: Orthogonal projections as shortest distance ©Pearson, 2016.

Before we do the last part let us define the 4 ***fundamental subspaces of a matrix*** as the null spaces and column spaces of the matrix and its transpose ($Nul(\mathbf{A})$, $Nul(\mathbf{A}^T)$, $Col(\mathbf{A})$, $Col(\mathbf{A}^T)$).

We have that for an $m \times n$ matrix, $\mathbf{A}$:

1. $Col(\mathbf{A})$ and $Nul(\mathbf{A}^T)$ are orthogonal subspaces of $\mathbb{R}^m$.
2. $Col(\mathbf{A}^T)$ and $Nul(\mathbf{A})$ are orthogonal subspaces of $\mathbb{R}^n$.
3. $Col(\mathbf{A}) \oplus Nul(\mathbf{A}^T) = \mathbb{R}^m$.
4. $Col(\mathbf{A}^T) \oplus Nul(\mathbf{A}) = \mathbb{R}^n$.

2.1 Least Squares

The method of least squares is used to get a best fit solution to an inconsistent system.

Say we plot 3 points $(1,0)$, $(2,1)$ and $(3,3)$. Is there a straight line, $y = ax + b$, that best fits the plot?

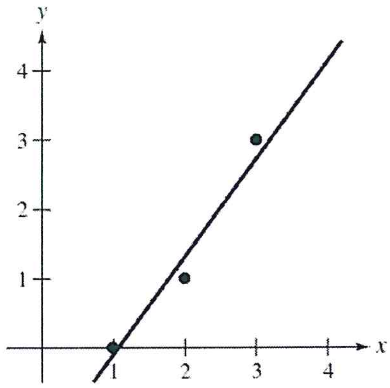


Figure 3: Fitting a line to points.

If the points are ***collinear*** (lying on the same line) then the following system (from the equation) would be consistent:

$$\begin{aligned} a+b&=0 \\ 2a+b&=1 \\ 3a+b&=3 \end{aligned} \implies \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

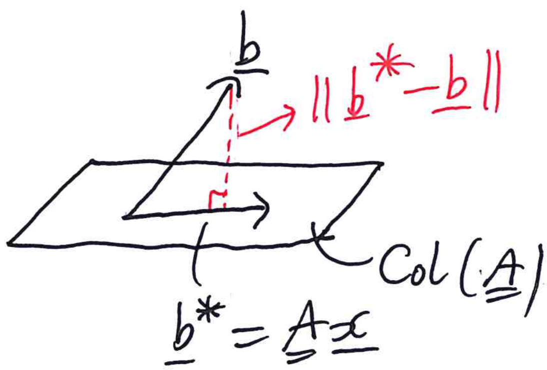
Since they are not collinear the system is inconsistent so we try to find the line such that the norm of the error is minimised.

$$\min \|\underline{A}\underline{x} - \underline{b}\|$$

Recall that all $\mathbf{b}^*$ that satisfy $\mathbf{A}\mathbf{x} = \mathbf{b}^*$ are in the column space of $\mathbf{A}$.
has solution

Our $\mathbf{b}$ is not but we require that,

$$\underline{b}^* = \underline{A}\underline{x} \in Col(\underline{A})$$



To minimise the distance we take the projection of **b** onto $Col(\mathbf{A})$.

$$\underline{b}^* = \underline{A}\underline{x} = \text{proj}_{Col(\underline{A})}\underline{b} \Rightarrow \underline{b}^* - \underline{b} \text{ is orthogonal to } Col(\underline{A})$$

$$\Rightarrow (\underline{b}^* - \underline{b}) \in Nul(\underline{A}^T)$$

$$\therefore \underline{A}^T(\underline{b}^* - \underline{b}) = \underline{0}$$

$$\Rightarrow \underline{A}^T(\underline{A}\underline{x} - \underline{b}) = \underline{0}$$

This results in the system required to solve the least squares problem.

$$\boxed{\underline{A}^T\underline{A}\underline{x} = \underline{A}^T\underline{b} \qquad (*)}$$

Let's now apply this to find our best fit:

$$\underline{A}^T\underline{A} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 14 & 6 \\ 6 & 3 \end{pmatrix}$$

$$\underline{A}^T\underline{b} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 4 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 14 & 6 & 11 \\ 6 & 3 & 4 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \frac{1}{2} & \frac{2}{3} \\ 0 & -1 & \frac{5}{3} \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & \frac{3}{2} \\ 0 & 1 & -\frac{5}{3} \end{array}\right)$$

$$\underline{x} = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \therefore a = \frac{3}{2}, \quad b = -\frac{5}{3}$$

$$\boxed{\therefore y = \frac{3x}{2} - \frac{5}{3}}$$