MATH 1B03/1ZC3 Winter 2019

Lecture 22: More about bases

Instructor: Dr Rushworth March 26th

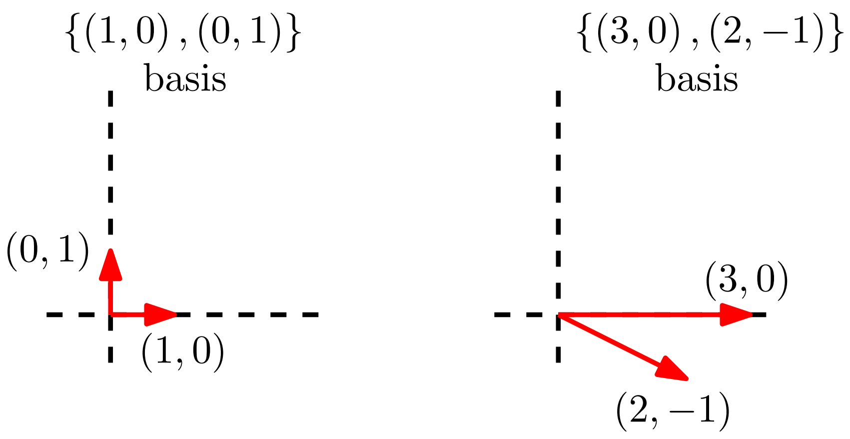
The Gram-Schmidt Orthonormalization process

(from Chapter 6.3 of Anton-Rorres)

We have seen that a vector space can be described using many different bases, in addition to the 'standard' basis we have used in the past.

The standard basis, or standard choice of co-ordinate axes, has two special properties that other bases need not have. First, the basis vectors in the standard basis are orthogonal to one another, and second, they are all of length 1.

For example, we previously considered the two bases of $\mathbb{R}^2$:



Notice that the two basis vectors on the right, $(3, 0)$ and $(2, -1)$, are not orthogonal to one another, and do not have length 1:

$$(3,\,0) \bullet (2,\,-1) = 6 \neq 0$$
$$||\,(3,\,0)\,|| = 3 \neq 1$$
$$||\,(2,\,-1)\,|| = \sqrt{5} \neq 1$$

Despite this, the set $\{(3, 0), (2, -1)\}$ is still a basis for $\mathbb{R}^2$.

However, it is often useful to have a basis for a vector space which does possess the nice properties we saw in the standard basis. We shall now describe a process by which any basis can be converted into new bases possessing these properties.

Important: when reading the corresponding chapter of the textbook, make the following subsitutions

- “Euclidean inner product" = dot product
- $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \bullet \mathbf{v}$

Definition 22.1: Orthogonal set

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$ (not necessarily a basis). We say that S is orthogonal if

$$\mathbf{v}_i \bullet \mathbf{v}_j = 0$$

for all i and j. That is, if every vector in S is orthogonal to all of the others.

Definition 22.2: Orthonormal set

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$ (not necessarily a basis). We say that S is orthonormal if it is orthogonal and

$$||\mathbf{v}_i|| = 1$$

for all i. That is, if every vector in S is orthogonal to all of the others and has length 1.

Fact 22.3: Orthogonal (and non-zero) implies linearly independent

Let V be a vector space and S a set of vectors in $\mathbb{R}^n$, which **does not** contain the zero vector. If S is orthogonal then it is linearly independent.

Notice that the converse is **not true**. That is

$$\text{orthogonal} \implies \text{linearly independent}$$
$$\text{orthogonal} \not\Leftarrow \text{linearly independent}$$

For example, $\{(3, 0), (2, -1)\}$ is linearly independent but it is not orthogonal. It is also important to note that if S contains the zero vector it may be orthogonal but not linearly independent.

Recall that a subset $S \subset \mathbb{R}^n$ is a basis if

1. S is linearly independent
2. span$(S) = \mathbb{R}^n$

The set S is not required to be orthogonal or orthonormal. Given a basis S, we can always use it to produce an orthonormal basis of $\mathbb{R}^n$.

This requires work, however. The following fact provides some evidence why it is worth our while to do this extra work to produce orthonormal bases.

Fact 22.4: A useful property of an orthonormal basis

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of $\mathbb{R}^n$. Given a vector $\mathbf{u}$

1. if S is orthogonal then

$$\mathbf{u} = \frac{\mathbf{u} \bullet \mathbf{v}_1}{||\mathbf{v}_1||^2}\mathbf{v}_1 + \ldots + \frac{\mathbf{u} \bullet \mathbf{v}_n}{||\mathbf{v}_n||^2}\mathbf{v}_n$$

   and the co-ordinate vector of $\mathbf{u}$ in terms of S is

$$\mathbf{u} = \left(\frac{\mathbf{u} \bullet \mathbf{v}_1}{||\mathbf{v}_1||^2}, \ldots \frac{\mathbf{u} \bullet \mathbf{v}_n}{||\mathbf{v}_n||^2}\right)_S$$

2. if S is orthonormal then

$$\mathbf{u} = (\mathbf{u} \bullet \mathbf{v}_1)\,\mathbf{v}_1 + \ldots + (\mathbf{u} \bullet \mathbf{v}_n)\,\mathbf{v}_n$$

   and the co-ordinate vector of $\mathbf{u}$ in terms of S is

$$\mathbf{u} = (\mathbf{u} \bullet \mathbf{v}_1, \ldots, \mathbf{u} \bullet \mathbf{v}_n)_S$$

We saw in the previous lecture that it can take time to find co-ordinate vectors in bases which are not orthonormal. The fact above shows that it is very easy to find co-ordinate vectors in orthonormal bases.

Example 22.5

Question: Verify that the set S is a basis of $\mathbb{R}^3$ by checking that it is orthonormal, then find the co-ordinate vector of $\mathbf{v} = (-5, 11, 3)$ in terms of S, where

$$S = \left\{(1, 0, 0), \left(0, \frac{5}{13}, -\frac{12}{13}\right), \left(0, \frac{12}{13}, \frac{5}{13}\right)\right\}$$

Answer: First check S is orthogonal:

$$(1, 0, 0) \bullet \left(0, \frac{5}{13}, -\frac{12}{13}\right) = 0$$
$$(1, 0, 0) \bullet \left(0, \frac{12}{13}, \frac{5}{13}\right) 0$$
$$\left(0, \frac{5}{13}, -\frac{12}{13}\right) \bullet \left(0, \frac{12}{13}, \frac{5}{13}\right) = \frac{60}{169} - \frac{60}{169} = 0$$

Next check S is orthonormal

$$\|(1, 0, 0)\| = 1$$
$$\left\|\left(0, \frac{5}{13}, -\frac{12}{13}\right)\right\| = \sqrt{\frac{25}{169} + \frac{144}{169}} = \sqrt{\frac{169}{169}} = 1$$
$$\left\|\left(0, \frac{12}{13}, \frac{5}{13}\right)\right\| = \sqrt{\frac{25}{169} + \frac{144}{169}} = \sqrt{\frac{169}{169}} = 1$$

Therefore S is orthonormal.
To find the co-ordinate vector of $\mathbf{v} = (-5, 11, 3)$, we apply the formula given

in Fact 22.4. First compute the relevant dot products

$$(-5, 11, 3) \bullet (1, 0, 0) = -5$$
$$(-5, 11, 3) \bullet \left(0, \frac{5}{13}, -\frac{12}{13}\right) = \frac{55}{13} - \frac{36}{13} = \frac{19}{13}$$
$$(-5, 11, 3) \bullet \left(0, \frac{12}{13}, \frac{5}{13}\right) = \frac{132}{13} - \frac{15}{13} = \frac{147}{13}$$

then

$$(-5, 11, 3) = \left(-5, \frac{19}{13}, \frac{147}{13}\right)_S$$

As we can see, it is often very useful to have an orthonormal basis of a vector space. The Gram-Schmidt orthonormalization process converts a basis of $\mathbb{R}^n$ into an orthonormal basis.

Recipe 22.6: Gram-Schmidt orthonormalization process

Let $S = \{\mathbf{u}_1 \dots \mathbf{u}_n\}$ be a basis for $\mathbb{R}^n$. Use this recipe to construct an orthonormal basis from S.

Step 1: Compute the vectors

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 \\ \mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \bullet \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \frac{\mathbf{u}_3 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \end{aligned}$$

continue until you have computed

$$\mathbf{v}_n = \mathbf{u}_n - \frac{\mathbf{u}_n \bullet \mathbf{v}_{n-1}}{\|\mathbf{v}_{n-1}\|^2}\mathbf{v}_{n-1} - \frac{\mathbf{u}_n \bullet \mathbf{v}_{n-2}}{\|\mathbf{v}_{n-2}\|^2}\mathbf{v}_{n-2} - \dots - \frac{\mathbf{u}_n \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

Step 2: Compute the vectors

$$\mathbf{e}_1 = \frac{\mathbf{v_1}}{||\mathbf{v_1}||}$$
$$\mathbf{e}_2 = \frac{\mathbf{v_2}}{||\mathbf{v_2}||}$$
$$\vdots$$
$$\mathbf{e}_n = \frac{\mathbf{v_n}}{||\mathbf{v_n}||}$$

The set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthonormal basis of $\mathbb{R}^n$.

Warning: when using the Gram-Schmidt process be very careful regarding minus signs. One small error will make your entire answer incorrect!

Example 22.7

Question: Apply the Gram-Schmidt process to the basis

$$S = \{(3, 0), (2, -1)\}$$

Question: Set $\mathbf{u}_1 = (3, 0)$ and $\mathbf{u}_2 = (2, -1)$.
Then $\mathbf{v}_1 = \mathbf{u}_1$ with $||\mathbf{v}_1|| = 3$.
Then

$$\begin{aligned}
\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \bullet \mathbf{v}_1}{||\mathbf{v}_1||^2}\mathbf{v}_1 \\
&= (2, -1) - \frac{(2, -1) \bullet (3, 0)}{9}(3, 0) \\
&= (2, -1) - \frac{6}{9}(3, 0) \\
&= (0, -1)
\end{aligned}$$

and $||\mathbf{v}_2|| = 1$.

Finally

$$\mathbf{e}_1 = \frac{(3,\ 0)}{3} = (1,\ 0)$$
$$\mathbf{e}_2 = \frac{(0,\ -1)}{1} = (0,\ -1)$$

and $\{(1,\ 0),\ (0,\ -1)\}$ is an orthonormal basis for $\mathbb{R}^2$.

Question: Apply the Gram-Schmidt process to the basis

$$S = \{(2,\ -1,\ 0),\ (-2,\ 0,\ 1),\ (0,\ -3,\ 0)\}$$

Question: Set $\mathbf{u}_1 = (2,\ -1,\ 0)$, $\mathbf{u}_2 = (-2,\ 0,\ 1)$, and $\mathbf{u}_3 = (0,\ -3,\ 0)$.
Then $\mathbf{v}_1 = \mathbf{u}_1$ and $\|\mathbf{v}_1\| = \sqrt{5}$.
Next

$$\begin{aligned}
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
&= (-2,\ 0,\ 1) - \frac{(-2,\ 0,\ 1) \bullet (2,\ -1,\ 0)}{5}(2,\ -1,\ 0) \\
&= (-2,\ 0,\ 1) - \frac{-4}{5}(2,\ -1,\ 0) \\
&= \left(-\frac{2}{5},\ -\frac{4}{5},\ 1\right)
\end{aligned}$$

with

$$\|\mathbf{v}_2\| = \sqrt{\frac{4}{25} + \frac{16}{25} + 1} = \sqrt{\frac{45}{25}} = \frac{3}{\sqrt{5}}$$

Next

$$
\begin{aligned}
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \bullet \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \frac{\mathbf{u}_3 \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
&= (0, -3, 0) - \frac{(0, -3, 0) \bullet \left(-\frac{2}{5}, -\frac{4}{5}, 1\right)}{\frac{9}{5}} \left(-\frac{2}{5}, -\frac{4}{5}, 1\right) \\
&\qquad - \frac{(0, -3, 0) \bullet (2, -1, 0)}{5} (2, -1, 0) \\
&= (0, -3, 0) - \left(\frac{8}{15}, \frac{16}{15}, -\frac{4}{3}\right) - \left(-\frac{6}{5}, \frac{3}{5}, 0\right) \\
&= \left(-\frac{2}{3}, -\frac{4}{3}, -\frac{4}{3}\right)
\end{aligned}
$$

with

$$
\|\mathbf{v}_3\| = \sqrt{\frac{4}{9} + \frac{16}{9} + \frac{16}{9}} = 2
$$

Finally, we have

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\sqrt{5}}(2, -1, 0) = \left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}, 0\right) \\
\mathbf{e}_2 &= \frac{\sqrt{5}}{3}\left(-\frac{2}{5}, -\frac{4}{5}, 1\right) = \left(-\frac{2}{3\sqrt{5}}, -\frac{4}{3\sqrt{5}}, \frac{\sqrt{5}}{3}\right) \\
\mathbf{e}_3 &= \frac{2}{3}\left(-\frac{2}{3}, -\frac{4}{3}, -\frac{4}{3}\right) = \left(-\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right)
\end{aligned}
$$

Therefore

$$
\left\{\left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}, 0\right), \left(-\frac{2}{3\sqrt{5}}, -\frac{4}{3\sqrt{5}}, \frac{\sqrt{5}}{3}\right), \left(-\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right)\right\}
$$

is an orthonormal basis of $\mathbb{R}^3$.

It is possible to apply the Gram-Schmidt process to bases of subspaces also.

Example 22.8

Question: Find an orthonormal basis for the plane in $\mathbb{R}^4$ spanned by the vectors

$$(3, 0, 0, -1), (2, 2, 0, 1)$$

Answer: Apply Gram-Schmidt to the set

$$S = \{(3, 0, 0, -1), (2, 2, 0, 1)\}$$

This yields the orthonormal basis

$$S = \left\{ \left(\frac{3}{\sqrt{10}}, 0, 0, -\frac{1}{\sqrt{10}} \right), \left(\frac{1}{\sqrt{26}}, \sqrt{\frac{8}{13}}, 0, \frac{3}{\sqrt{26}} \right) \right\}$$

Orthogonal projection onto a subspace

We have seen how to orthogonally project one vector onto another. We can also orthogonally project a vector onto a subspace.

That is, given a vector $\mathbf{v} \in \mathbb{R}^n$ and a subspace W, we can decompose $\mathbf{v}$ into two parts, one which lies in W and another which is orthogonal to W.

Definition 22.9

Let W be a subspace of $\mathbb{R}^n$. A vector $\mathbf{v} \in \mathbb{R}^n$ is orthogonal to W if

$$\mathbf{v} \bullet \mathbf{w} = 0$$

for all $\mathbf{w} \in W$. That is, if $\mathbf{v}$ is orthogonal to every vector in W.

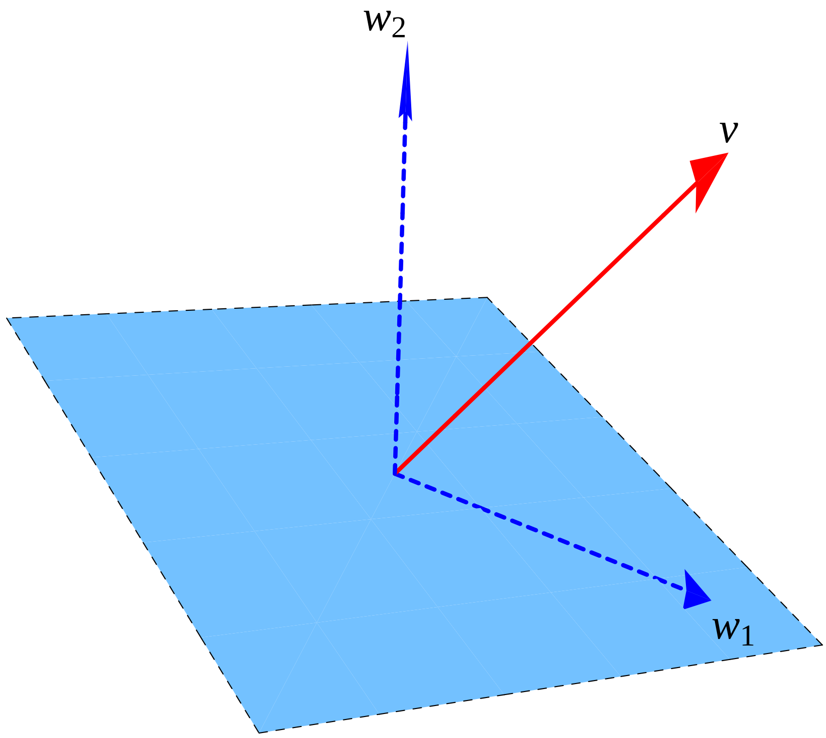
Fact 22.10: Orthogonal decomposition

Let W be a subspace of $\mathbb{R}^n$. Given a vector $\mathbf{v} \in \mathbb{R}^n$, we have

$$\mathbf{v} = \text{proj}_W(\mathbf{v}) + \text{proj}_{W^\perp}(\mathbf{v})$$

where $\text{proj}_W(\mathbf{v}) \in W$ and $\text{proj}_{W^\perp}(\mathbf{v})$ is orthogonal to W.

This can be visualized as follows:



The equation

$$\mathbf{v} = \text{proj}_W(\mathbf{v}) + \text{proj}_{W^\perp}(\mathbf{v})$$

is known as the orthogonal decomposition of $\mathbf{v}$ with respect to W. It is sometimes written alternatively as

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$$

where

$$\mathbf{w}_1 = \text{proj}_W(\mathbf{v})$$

is the component of $\mathbf{v}$ lying in W, and

$$\mathbf{w}_2 = \text{proj}_{W^\perp}(\mathbf{v})$$

is the component of $\mathbf{v}$ orthogonal to W. It is also said that $\mathbf{w}_2$ is perpendicular to W.

Given an orthonormal basis for W we can quickly find the orthogonal decomposition of vectors.

Fact 22.11: Finding the orthogonal decomposition

Let W be a subspace of $\mathbb{R}^n$ with orthogonal basis

$$S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

Given a vector $\mathbf{v} \in \mathbb{R}^n$, we have

$$\text{proj}_W(\mathbf{v}) = \frac{\mathbf{v} \bullet \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{v} \bullet \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{v} \bullet \mathbf{v}_n}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$$

If the basis S is orthonormal then

$$\text{proj}_W(\mathbf{v}) = (\mathbf{v} \bullet \mathbf{v}_1)\,\mathbf{v}_1 + (\mathbf{v} \bullet \mathbf{v}_2)\,\mathbf{v}_2 + \cdots (\mathbf{v} \bullet \mathbf{v}_n)\,\mathbf{v}_n$$

The component of $\mathbf{v}$ orthogonal to W is given by

$$\text{proj}_{W^\perp}(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v})$$

Example 22.12

Question: Let W be the plane in $\mathbb{R}^4$ spanned by the vectors

$$(3, 0, 0, -1),\ (2, 2, 0, 1)$$

Express the vector $\mathbf{v} = (-7, 0, 0, 4)$ as

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$$

where $\mathbf{w}_1$ lies in W and $\mathbf{w}_2$ is orthogonal to W.

Answer: We saw in Example 22.8 that

$$\mathbf{e}_1 = \left(\frac{3}{\sqrt{10}}, 0, 0, -\frac{1}{\sqrt{10}}\right)$$
$$\mathbf{e}_2 = \left(\frac{1}{\sqrt{26}}, \sqrt{\frac{8}{13}}, 0, \frac{3}{\sqrt{26}}\right)$$

are orthonormal basis vectors for the plane W. Then

$$\begin{aligned}
\mathbf{w}_1 = \text{proj}_W(\mathbf{v}) &= (\mathbf{e}_1 \bullet \mathbf{v})\,\mathbf{e}_1 + (\mathbf{e}_2 \bullet \mathbf{v})\,\mathbf{e}_2 \\
&= \left(-\frac{21}{\sqrt{10}} - \frac{4}{\sqrt{10}}\right)\left(\frac{3}{\sqrt{10}}, 0, 0, -\frac{1}{\sqrt{10}}\right) \\
&\quad + \left(-\frac{7}{\sqrt{26}} + \frac{12}{\sqrt{26}}\right)\left(\frac{1}{\sqrt{26}}, \sqrt{\frac{8}{13}}, 0, \frac{3}{\sqrt{26}}\right) \\
&= \left(-\frac{75}{10}, 0, 0, \frac{25}{10}\right) + \left(\frac{5}{26}, \frac{10}{13}, 0, \frac{15}{26}\right) \\
&= \left(-\frac{95}{13}, \frac{10}{13}, 0, \frac{40}{13}\right)
\end{aligned}$$

Next

$$\begin{aligned}
\mathbf{w}_2 = \text{proj}_{W^\perp}(\mathbf{v}) = \mathbf{v} - \text{proj}_W(\mathbf{v}) &= (-7, 0, 0, 4) - \left(-\frac{95}{13}, \frac{10}{13}, 0, \frac{40}{13}\right) \\
&= \left(-\frac{91}{13}, 0, 0, \frac{52}{13}\right) - \left(-\frac{95}{13}, \frac{10}{13}, 0, \frac{40}{13}\right) \\
&= \left(\frac{4}{13}, -\frac{10}{13}, 0, \frac{12}{13}\right)
\end{aligned}$$

Therefore the orthogonal decomposition of $\mathbf{v}$ with respect to W is

$$\mathbf{v} = \left(-\frac{95}{13}, \frac{10}{13}, 0, \frac{40}{13}\right) + \left(\frac{4}{13}, -\frac{10}{13}, 0, \frac{12}{13}\right)$$

Suggested Problems

Practice the material covered in this lecture by attempting the following questions from Chapter 6.3 of Anton-Rorres, starting on page 376

- Questions $7, 9, 23, 29, 31, 35$
- True/False (a), (b), (e)