W07-01

Slide 01: In this unit, we will discuss how to compute the orthogonal projection of a vector onto a subspace.

Slide 02: Let V be a subspace of $\mathbb{R}^n$ that we would like to project onto. Suppose $\boldsymbol{w}$ is a vector in $\mathbb{R}^n$. If we have an orthogonal basis for V, comprising of vectors $\boldsymbol{u_1}$, $\boldsymbol{u_2}$ to $\boldsymbol{u_k}$, then the projection of $\boldsymbol{w}$ onto V

(#)

is given by the following expression. Notice that this expression is a linear combination of the orthogonal basis vectors $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$. This is expected as the projection is a vector in V and thus should be expressible as a linear combination of the basis vectors.

(#)

If we have an orthonormal basis for V instead, comprising of vectors $\boldsymbol{v_1}$, $\boldsymbol{v_2}$ to $\boldsymbol{v_k}$, then the projection of $\boldsymbol{w}$ onto V, written as a linear combination of the orthonormal basis vectors is as shown.

(#)

Take another look at the expression for the projection of $\boldsymbol{w}$ onto V. Do you find this expression familiar? Have you seen the same expression from an earlier unit?

Slide 03: Indeed we have seen this expression before. Recall from an earlier unit, when we introduced the special class of bases known as orthogonal basis, we saw that if $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$ is an orthogonal basis for a vector space V, then to write any vector $\boldsymbol{w}$ in V as a linear combination of the orthogonal basis vectors can be done without having to solve any linear systems to find the coefficients. This expression is precisely how $\boldsymbol{w}$ can be written in terms of $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$.

(#)

Now compare to what we have just seen from the previous slide, which states that if $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$ is an orthogonal basis for the vector space V, then for any $\boldsymbol{w}$ in $\mathbb{R}^n$, the projection of $\boldsymbol{w}$ onto V is given by the same expression as we have seen above.

(#)

Why are the two expressions identical and why does this make sense?

Slide 04: To see why this makes sense, consider the case where we are projecting a vector $\boldsymbol{w}$ onto V.

(#)

The figure here shows $\boldsymbol{w}$ to be outside V,

(#)

so when we project $\boldsymbol{w}$ onto V, we obtain a different vector $\boldsymbol{p}$ whose expression is the linear combination of $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$ as shown.

(#)

Recall that by orthogonal projection, we mean decomposing $\boldsymbol{w}$ and writing it as the sum of two vectors $\boldsymbol{p}$ and $\boldsymbol{n}$ where $\boldsymbol{p}$ belongs to V and $\boldsymbol{n}$ is orthogonal to V.

(#)

We have mentioned earlier what orthogonal projection means if the vector $\boldsymbol{w}$ is already in the subspace V that we would like to project on, as shown in the figure on the left.

(#)

The projection of $\boldsymbol{w}$ onto V in this case is simply $\boldsymbol{w}$ itself. This is where the expression from an earlier unit comes in, where we are writing $\boldsymbol{w}$ as a linear combination of the orthogonal basis vectors $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$.

(#)

If thought of as a projection, we are decomposing $\boldsymbol{w}$ as $\boldsymbol{w} + \boldsymbol{0}$. So indeed there is no conflict in seeing the same expression twice, since the notion of projection can take place whether or not $\boldsymbol{w}$ belongs to V.

Slide 05: We are now ready to present a proof of this orthogonal projection theorem. Remember that this theorem allows us to compute orthogonal projection of any $\boldsymbol{w}$ in $\mathbb{R}^n$ onto the subspace V **provided** we have an orthogonal basis for V.

Slide 06: To prove the theorem, let $\boldsymbol{p}$ be the expression given in the statement of the theorem.

(#)

We will show that $\boldsymbol{p}$ is indeed the projection of $\boldsymbol{w}$ onto V.

(#)

First, observe that $\boldsymbol{p}$ is definitely a vector in V since it is a linear combination of the basis vectors $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$.

(#)

Next, we define the vector $\boldsymbol{n}$ to be the difference between $\boldsymbol{w}$ and $\boldsymbol{p}$. Why do we want to do this?

(#)

By doing this, it would mean that $\boldsymbol{w}$ is in fact decomposed into $\boldsymbol{n}$ plus $\boldsymbol{p}$, where we have already noted that $\boldsymbol{p}$ is a vector in V.

Slide 07: What remains for us to do, in order to conclude that $\boldsymbol{p}$ is the orthogonal projection of $\boldsymbol{w}$ onto V, is to show that the vector $\boldsymbol{n}$, in the way we have defined it, is orthogonal to the space V. If we can do this, then we would have established that $\boldsymbol{p}$ will be the projection of $\boldsymbol{w}$ onto V.

(#)

Recall that to show that a vector is orthogonal to a space, we need to show that it is orthogonal to the vectors that spans the space. Thus, we will check whether $\boldsymbol{n}$ is orthogonal to each of the $\boldsymbol{u_i}$'s.

Slide 08: So for each $i = 1$ to k, consider the dot product between $\boldsymbol{n}$ with $\boldsymbol{u_i}$. Since $\boldsymbol{n}$ is $\boldsymbol{w} - \boldsymbol{p}$, we have the expression on the right side of the equation,

(#)

which can be simplified as follows by applying distributive law.

(#)

We now write down the expression for $\boldsymbol{p}$ as shown.

(#)

Now consider the expression in the red box. Here we have the dot product between $\boldsymbol{u_i}$ and a linear combination of $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$.

(#)

If we apply distributive law on this expression, we should have many terms as a result. Why is it that only one term remain? Namely, the only term that remains seems to be the $\boldsymbol{u_i} \cdot \boldsymbol{u_i}$ term.

(#)

The reason is because that $\boldsymbol{u_1}$ to $\boldsymbol{u_k}$ is an orthogonal set. This means that the vectors in the set are pairwise orthogonal, so the dot product between $\boldsymbol{u_i}$ and $\boldsymbol{u_j}$ will be zero whenever $i \neq j$. This is why only the $\boldsymbol{u_i} \cdot \boldsymbol{u_i}$ term remains.

(#)

Now the dot product of $\boldsymbol{u_i}$ with itself is the square of the length of $\boldsymbol{u_i}$. This will cancel with the denominator of the coefficient,

(#)

which reduces the expression to $\boldsymbol{w} \cdot \boldsymbol{u_i}$ minus $\boldsymbol{w} \cdot \boldsymbol{u_i}$ which is 0.

Slide 09: We have thus established that $\boldsymbol{n}$ is indeed a vector orthogonal to V since it is orthogonal to the vectors that spans V.

(#)

The vector $\boldsymbol{w}$ in $\mathbb{R}^n$, has now been written as the sum of two vectors $\boldsymbol{n}$ and $\boldsymbol{p}$ where

(#)

$\boldsymbol{n}$ is orthogonal to V while

(#)

$\boldsymbol{p}$ belongs to V. So the vector $\boldsymbol{p}$ is indeed the orthogonal projection of $\boldsymbol{w}$ onto V.

(#)

The second part of the theorem actually follows immediately from the first part, since when $\boldsymbol{v_1}$ to $\boldsymbol{v_k}$ is an orthonormal basis for V, they are orthogonal vectors of length 1.

(#)

Thus the denominators of all the coefficients in the expression for $\boldsymbol{p}$ is now 1 and we will obtain the expression for the projection of $\boldsymbol{w}$ onto V accordingly.

Slide 10: Let us go through one example. Let V be a subspace of $\mathbb{R}^3$ spanned by the two vectors $(1, 0, 1)$ and $(1, 0, -1)$. Notice that these two vectors are not multiples of each other, meaning that they are linearly independent. Thus the two vectors form a basis for V.

(#)

In addition, note that the dot product between the two vectors is 0, which means that these two vectors are orthogonal.

(#)

So the two vectorss actually form an orthogonal basis for V. Now that we have an orthogonal basis for V, we are ready to compute projection of vectors onto V.

(#)

What is the projection of the vector $(1, 1, 0)$ onto V?

(#)

Using the orthogonal projection theorem, we have the following linear combination of $(1, 0, 1)$ and $(1, 0, -1)$ as the projection of $\boldsymbol{w}$ onto V.

(#)

This simplifies to the vector $(1, 0, 0)$ which is the projection of $\boldsymbol{w}$ onto V.

Slide 11: What about this subspace V of $\mathbb{R}^3$? It is also spanned by two vectors $(1, 1, 1)$ and $(1, 3, -1)$ which are not multiples of each other. So these two vectors also forms a basis for V.

(#)

However, in this case, the dot product between the two vectors is not zero, which means that this is not an orthogonal basis.

(#)

Remember that to apply the orthogonal projection theorem, we need to have an orthogonal basis for the subspace V that we wish to project on.

(#)

So how can we compute the projection of $\boldsymbol{w}$ onto V in this case?

Slide 12: Clearly, we cannot apply the orthogonal projection theorem in this case.

(#)

So when we are faced with the situation that we do not have an orthogonal basis for the subspace V, what can we do?

(#)

In a subsequent unit, we will introduce a process where we can construct an orthogonal basis for a subspace V. Such a basis, once constructed, will then allow us to use the orthogonal projection theorem to compute projection.

Slide 13: In summary, we have seen in this unit

(#)

how to compute orthogonal projection of a vector onto a vector space V. This can be done provided we have an orthogonal or orthonormal basis for V.