W06-08

Slide 01: In this unit, we will introduce a very important concept that provides the foundation to linear approximation, namely, orthogonal projection.

Slide 02: We first define what is meant by a vector being orthogonal to a subspace. Let V be a subspace of $\mathbb{R}^n$. A vector $\boldsymbol{u}$ is said to be orthogonal to V if $\boldsymbol{u}$ is orthogonal to all the vectors in the subspace V. You can visualise this by putting the vector $\boldsymbol{u}$ with its initial point at the origin, which is certainly contained in V since V is a subspace.

(#)

Any vector in V

(#)

with their initial point likewise placed at the origin

(#)

will be orthogonal or perpendicular to the vector $\boldsymbol{u}$. With this definition, you will now see that the term orthogonal can be used to describe the relationship between two vectors, to describe a set, and now to describe the relationship between a vector and a set.

Slide 03: For example, the set V that contains all the vectors (x, y, z) satisfying $2x + 3y - z = 0$ is a subspace of $\mathbb{R}^3$. You should recognise this immediately as a plane in $\mathbb{R}^3$ that contains the origin.

(#)

A vector in $\mathbb{R}^3$ will be in V if and only if $2x + 3y - z = 0$.

(#)

Or in other words, the dot product between (x, y, z) and $(2, 3, -1)$ is zero.

(#)

So the vector $(2, 3, -1)$ is orthogonal to every vector in the subspace V

(#)

and thus we say that $(2, 3, -1)$ is orthogonal to V.

Slide 04: We can generalise the above example. For any plane in $\mathbb{R}^3$ that contains the origin, we know that the plane is a subspace of $\mathbb{R}^3$ and can be described by the linear equation $ax + by + cz = 0$. So a vector (x, y, z) belongs to the subspace if and only if $(x, y, z) \cdot (a, b, c)$ is zero. (#)

Thus the vector (a, b, c) is orthogonal to the subspace.

Slide 05: Continuing with the subspace V that represents a plane in $\mathbb{R}^3$ containing the origin.

(#)

Let $\boldsymbol{n}$ denote the vector (a, b, c) which we have seen to be a vector orthogonal to V.

(#)

We can now rewrite V as the set of all vectors $\boldsymbol{u}$ in $\mathbb{R}^3$ such that the dot product between $\boldsymbol{n}$ and $\boldsymbol{u}$ is zero.

(#)

The vector $\boldsymbol{n}$ is orthogonal to V and we can also say that $\boldsymbol{n}$ is a normal vector of V.

(#)

Now if $\boldsymbol{n}$ is orthogonal to V, will any non zero multiple of $\boldsymbol{n}$ be also orthogonal to V?

(#)

The answer is definitely yes, since $\boldsymbol{n} \cdot \boldsymbol{u} = 0$ is equivalent to $c\boldsymbol{n} \cdot \boldsymbol{u} = 0$ for any non zero scalar c. You can think of this as $c\boldsymbol{u}$ being just a longer or shorter version of $\boldsymbol{u}$, having the same direction as $\boldsymbol{u}$. So if $\boldsymbol{u}$ is perpendicular to V, so will $c\boldsymbol{u}$.

Slide 06: Let V be a subspace of $\mathbb{R}^4$ spanned by the two vectors $(1, 1, -1, 0)$ and $(0, 1, 1, -1)$. We would like to find all vectors orthogonal to V.

(#)

Remember that a vector is orthogonal to V if and only if it is orthogonal to every vector in V.

(#)

It may seem very challenging to find even one vector that is orthogonal to every single vector in V, let alone finding **all** vectors with such a property which is what the question wants. How can we go about doing this?

Slide 07: Fortunately, if we know that V is spanned by a set of vectors $\boldsymbol{u_1}, \boldsymbol{u_2}$ to $\boldsymbol{u_k}$, then a vector $\boldsymbol{v}$ will be orthogonal to the subspace V if and only if $\boldsymbol{v}$ is orthogonal to each of the vector that spans V.

(#)

In other words, we just need to make sure that the dot product between $\boldsymbol{v}$ and all the $\boldsymbol{u_i}$'s is zero, which is much easier to check.

Slide 08: Let us return to the problem of finding all vectors orthogonal to V.

(#)

We now know that a vector $\boldsymbol{u}$ is orthogonal to V if and only if $\boldsymbol{u}$ is orthogonal to the vectors that spans V.

(#)

Let $\boldsymbol{u} = (w, x, y, z)$. So the dot product between $\boldsymbol{u}$ and $(1, 1, -1, 0)$, as well as the dot product between $\boldsymbol{u}$ and $(0, 1, 1, -1)$ will be zero.

(#)

Thus we now have a homogeneous linear system involving 4 variables and 2 equations.

Slide 09: Solving the homogeneous linear system by Gauss Jordan elimination,

(#)

we obtain the following general solution involving two arbitrary parameters s and t.

Slide 10: So $\boldsymbol{u}$ is orthogonal to V if and only if $\boldsymbol{u}$ can be written as a linear combination of the two vectors $(2, -1, 1, 0)$ and $(-1, 1, 0, 1)$.

Slide 11: In other words, $\boldsymbol{u}$ is orthogonal to V if and only if $\boldsymbol{u}$ belongs to the linear span of $(2, -1, 1, 0)$ and $(-1, 1, 0, 1)$. We have thus found all such vectors $\boldsymbol{u}$ that are orthogonal to V.

Slide 12: We are now ready to define orthogonal projection. Let V be a subspace of $\mathbb{R}^n$. Every vector $\boldsymbol{u}$ in $\mathbb{R}^n$ can now be written uniquely as the sum of two vectors $\boldsymbol{n}$ and $\boldsymbol{p}$ such that

(#)

$\boldsymbol{n}$ is a vector orthogonal to V

(#)

and $\boldsymbol{p}$ is a vector belonging to V. From the diagram shown, you can see that the red vector $\boldsymbol{u}$ is essentially decomposed into two vectors, one of which is the blue vector $\boldsymbol{n}$ that is orthogonal to the space V, while the other is the green vector $\boldsymbol{p}$ which is in the space V.

(#)

We call $\boldsymbol{p}$ the orthogonal projection of $\boldsymbol{u}$ onto the space V.

(#)

It is also interesting to note that such a decomposition is unique, meaning that every vector $\boldsymbol{u}$ can have only one orthogonal projection onto V. This fact can be proven but we will not do so in this unit.

Slide 13: While the previous slide has shown the red vector $\boldsymbol{u}$ to be sticking out from the space V, making it easy to understand how its orthogonal projection onto V would be, the definition of orthogonal projection does not require $\boldsymbol{u}$ not to be in V before it can be projected.

(#)

So what would orthogonal projection mean if the vector $\boldsymbol{u}$ that we wish to project onto V is already in V itself?

(#)

In this case, we simply decompose $\boldsymbol{u}$ into $\mathbf{0} + \boldsymbol{u}$. Note that this decomposition still satisfies the condition for orthogonal projection since the zero vector is trivially orthogonal to every vector in V, while $\boldsymbol{u}$ is already a vector in V. In other words, the orthogonal projection $\boldsymbol{p}$ of the vector $\boldsymbol{u}$ onto V is $\boldsymbol{u}$ itself.

Slide 14: Let us consider some simple cases of orthogonal projection in $\mathbb{R}^2$ and $\mathbb{R}^3$. For example, suppose V is the y-axis in $\mathbb{R}^2$, that is, V is the set containing all vectors of the form $(0, y)$ where y is a real number.

(#)

Then for any vector $\boldsymbol{u} = (x, y)$ in $\mathbb{R}^2$, we can write $\boldsymbol{u}$ as

(#)

$\boldsymbol{n} + \boldsymbol{p}$ where $\boldsymbol{n}$ is the vector $(x, 0)$

(#)

while $\boldsymbol{p}$ is the vector $(0, y)$. It is easy to check that indeed the vector $(x, 0)$ is orthogonal to every vector in V while the vector $(0, y)$ is certainly a vector in V.

(#)

Thus the projection of (x, y) onto V is $(0, y)$.

Slide 15: Suppose V is the xy-plane in $\mathbb{R}^3$. In other words, V is the subspace of $\mathbb{R}^3$ with the equation $z = 0$, so the vectors in V are of the form $(x, y, 0)$ where x and y are real numbers.

(#)

Then for any vector $\boldsymbol{u} = (x, y, z)$ in $\mathbb{R}^3$, we can write $\boldsymbol{u}$ as

(#)

$\boldsymbol{n} + \boldsymbol{p}$ where $\boldsymbol{n}$ is the vector $(0, 0, z)$

(#)

while $\boldsymbol{p}$ is the vector $(x, y, 0)$. It is easy to check that indeed the vector $(0, 0, z)$ is orthogonal to every vector in V while the vector $(x, y, 0)$ is a vector in V.

(#)

Thus the projection of (x, y, z) onto V is $(x, y, 0)$.

Slide 16: Before proceeding further, note that the two examples on orthogonal projection that we have seen so far are very straightforward, as the subspace V that we are projecting on are simple subspaces that are easy to work with.

(#)

What happens if the subspaces that we are projecting on are not so trivial? How can we then compute the orthogonal projection of a vector onto V? What do we need before the projection can be computed? These questions will be answered in the subsequent units.

Slide 17: Summarising the main points in this unit,

(#)

we defined what is meant by a vector being orthogonal to a vector space.

(#)

We then gave the definition of orthogonal projection onto a vector space.