W07-03

Slide 01: In this unit, we will introduce the notion of approximations to linear systems.

Slide 02: While you may think that so far in our discussions, we have always found exact answers to numerical problems, you should be aware that in many real life computations, exact answers to problems are sometimes not possible, nor necessary. We often are happy with approximate solutions, as long as they are good approximations.

(#)

The concept of orthogonality that we have discussed in earlier units plays a very central role in the study of approximations.

(#)

In subsequent discussions, although the setting will be the usual Euclidean space that you are familiar with, approximations can be extended to other general and abstract vector spaces too.

Slide 03: Some of you may have encountered this simple problem before. Given a point, we would like to find a point on the straight line that is closest to the point.

(#)

In the context of what we have seen, we can think of the straight line as a subspace V.

(#)

The point can be thought of as the end point of a vector $\boldsymbol{u}$, where the initial point of the vector is the origin, which is a point on the subspace V.

(#)

The point on the line that is closest to the given point can be found by projecting the vector $\boldsymbol{u}$ onto V. It is clear that we are referring to orthogonal projection in this case

(#)

and the point we are looking for is the end point of the projection of $\boldsymbol{u}$ onto V.

Slide 04: What if the question is to find a point on the plane that is closest to the given point?

(#)

Once again, we can think of the plane as a subspace V,

(#)

and the point as the end point of a vector $\boldsymbol{u}$.

(#)

The point on the plane that is closest to the given point can again be found by projecting the vector $\boldsymbol{u}$ onto V.

Slide 05: Let us consider a numerical example for this. Here, the plane V is the linear span of $(1, 0, 1)$ and $(1, 1, 1)$. We would like to find the shortest distance from the vector $\boldsymbol{u} = (1, 2, 3)$ to the plane V.

(#)

We know that we need to compute the orthogonal projection of $\boldsymbol{u}$ onto V and in order to do this, we require an orthogonal basis for V.

Slide 06: Although the two vectors $(1, 0, 1)$ and $(1, 1, 1)$ forms a basis for V, they are not an orthogonal basis.

(#)

We need to apply the Gram-Schmidt Process introduced in a previous unit to obtain an orthogonal basis for V, which upon doing so, we have the orthogonal basis $(1, 0, 1)$ and $(0, 1, 0)$ for V.

(#)

By the orthogonal projection theorem, we can now compute the projection $\boldsymbol{p}$ of $\boldsymbol{u}$ onto V as follows.

(#)

The vector $\boldsymbol{p}$ is found to be $(2, 2, 2)$

(#)

and thus the shortest distance from $\boldsymbol{u}$ to V is the distance between $\boldsymbol{u}$ and $\boldsymbol{p}$. This is computed by measuring the length of the vector $\boldsymbol{u} - \boldsymbol{p}$,

(#)

which is found to be $\sqrt{2}$.

Slide 07: This theorem states formally what we have just seen in the previous example. Suppose V is a subspace of $\mathbb{R}^n$.

(#)

Suppose $\boldsymbol{u}$ is a vector in $\mathbb{R}^n$ and $\boldsymbol{p}$ is the orthogonal projection of $\boldsymbol{u}$ onto V. Then

(#)

the distance between $\boldsymbol{u}$ and $\boldsymbol{p}$ is no larger than the distance between $\boldsymbol{u}$ and any vector $\boldsymbol{v}$ in the subspace V. From the diagram, you see that the distance between $\boldsymbol{u}$ and $\boldsymbol{p}$ is the orthogonal distance from the vector $\boldsymbol{u}$ to the space V. If we look at the distance between $\boldsymbol{u}$ and any other vector $\boldsymbol{v}$ in the subspace, the distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ will always be larger than or equal to the orthogonal distance.

(#)

In other words, we say that $\boldsymbol{p}$ is the best approximation of $\boldsymbol{u}$ in V.

Slide 08: Let us look at an example to illustrate how best approximation comes into play. Suppose we believe that three physical quantities r, s and t are related according to the equation $t = cr^2 + ds + e$ where c, d and e are constants to be determined.

(#)

How do we go about finding these constants? We can perform a series of experiments for measure the quantity t given different values of r and s.

(#)

Suppose we have six different sets of values for r and s and for each set, the value of t is measured. Thus, we have a total of six data points.

Slide 09: The table here shows the 6 data points. For example, when r and s are both zero, we found experimentally that the value of t was 0.5.

(#)

Now the question is can we find the constants c, d and e such that **all** the 6 data points fits nicely into the equation $t = cr^2 + ds + e$?

(#)

It is extremely unlikely that such c, d and e can be found, for the 6 data points that are obtained experimentally, to all fit into.

Slide 10: What we are saying here is not too different from something that many of you may have observed before. For example, if the 7 points shown here are obtained experimentally, even if the two quatntities x and y are indeed related in a linear manner,

(#)

it is impossible to find a straight line that will pass through all the 7 points. This is usually due to experimental errors done in measurement.

(#)

Back in High School, your teacher may have asked you, even though there isn't a straight line that passes through all the 7 points, how do you then draw the **best** line instead?

Slide 11: Let's return to our experimental example. As mentioned, if there were no experimental errors, the three unknowns c, d and e would satisfy all the 6 equations here, one equation for each data point. For example, we would have, for the first data point, $t_1 = cr_1^2 + ds_1 + e$.

(#)

These six equations can then be rewritten as a linear system as shown, something that you should be familiar with.

(#)

We can represent this linear system as $\boldsymbol{Ax} = \boldsymbol{b}$, where $\boldsymbol{A}$ and $\boldsymbol{b}$ contains the experimental values and $\boldsymbol{x}$ is the variable matrix which we are trying to solve.

Slide 12: As we have agreed earlier, there are likely no solutions for c, d and e. In other words $\boldsymbol{Ax} = \boldsymbol{b}$ is inconsistent.

(#)

So what can we do? Consider the following expression with 6 squared terms. Each term is the square of a t_i,

(#)

which is the observed experimental value,

(#)

minus $cr_i^2 + ds_i + e$, which is the predicted value according to the expression $t = cr^2 + ds + e$.

(#)

We can write down the sum of these 6 terms using the summation sign as shown.

(#)

We call this expression the sum of squares of errors, since each term in the summation is the square of the error, measuring the difference between the observed and the predicted value of t.

(#)

Now here comes the key observation. The 6 components of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ is precisely the 6 terms in the summation and thus the summation is simply the square of the length of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$.

(#)

The choice of c, d and e will then be one where the square of the length of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$, or equivalently, the sum of squares of errors, is minimised.

Slide 13: Let us give a summary of the problem we have just discussed. We started with a linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ that is inconsistent.

(#)

In other words, there is no $\boldsymbol{x}$ such that $\boldsymbol{A}\boldsymbol{x}$ is exactly $\boldsymbol{b}$. We then wish to find the $\boldsymbol{x}$ such that the square of the length of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ is minimised. To minimise the square of the length of the vector is the same as to minimise the length of the vector, so we may consider just minimizing the length of $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$.

(#)

It is important to note that in the event that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent, then to minimise the length of $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$, we simply choose $\boldsymbol{x}$ to be a solution to the system. This would make $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ equal to the zero vector, which obviously has the smallest length.

Slide 14: We are now ready to define what is a least squares solution to a linear system. Consider $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ where $\boldsymbol{A}$ is a $m \times n$ matrix.

(#)

A vector $\boldsymbol{u}$ in $\mathbb{R}^n$ is called a least squares solution to the linear system if the length of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}$ is less than or equals to the length of the vector $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{v}$ for any vector $\boldsymbol{v}$ in the entire $\mathbb{R}^n$.

Slide 15: To summarise this unit.

(#)

We saw that to find the best approximation of a vector $\boldsymbol{u}$ in V, we simply find the projection $\boldsymbol{p}$ of $\boldsymbol{u}$ onto V.

(#)

We defined what is a least squares solution to a linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.