Geometry of Least Squares

Welcome back.

We saw already the geometrical interpretation of least squares from perspective of linear algebra.

In this video, we will look at an example of this geometrical interpretation on a simple line fitting problem.

Let's assume that we have three observations at time equal to zero, three, and five, respectively.

And we want to fit a line to these three observations, or in other words we want to estimate the initial position at time zero, that is x0 and the velocity v.

We have three observations and two unknowns, and, by now, you all know that the observation equation can be written for this problem as this system of linear equations, y equal Ax.

The system is overdetermined and inconsistent, and the least squares solution for this system can be computed, and it gives us 1.6 and 0.26 estimate of x0 and the velocity v.

Now lets look at the geometry of this problem.

The observation vector is three dimentional, so I can draw it as the red arrow in the plot on the right.

Now let's think about the range space or column space of matrix A.

Recall that the range space is the space that spanned by all the columns of the matrix A.

So let's visualise the columns of A in this plot.

We have two unknown parameters and so two columns.

If I plot them, the green vectors show the first and second column vectors of A.

Now what is the range space of matrix A?

In fact the range space is a plane that goes through these two green vectors.

The plane is visualised as a gray mesh grid in the plot.

We can clearly see that the vector y is outside this plane, which indicates the inconsistency of the system.

From least squares we know that the adjusted observation vector y-hat is a vector in this plane and it has the minimum distance to the observation vector y.

I have already computed the distance between the observation vector y and all the points on this plane.

Let's see the results.

In this plot, the colours show the distance between the observation vector and all the vectors lying in the range space of A, in this case this plane.

The colourbar shows the range of the colour values.

So the minimum distance, is corresponding to a point in the middle of the blue area.

Actually, the vector from zero to this point should be the y-hat vector, which is the orthogonal projection of vector y on the range space of A.

We can also see the blue dashed vector, which demonstrates the vector of residuals.

In order to really see the minimum distance, let's change the colour scale.

With this new colour scale, all the points on the plane with distance larger than five from the observation vector are coloured as dark red.

We can see that indeed, yhat corresponds to the point on the range space of A, with the closest distance to vector y.

So, in summary in this section we learned that the least squares estimation is equivalent to the orthogonal projection of the vector of observation y onto the range space of the design matrix A.

I hope this geometrical interpretation has gave you a better insight and understanding of the concept of least squares.