PCA Made Easy

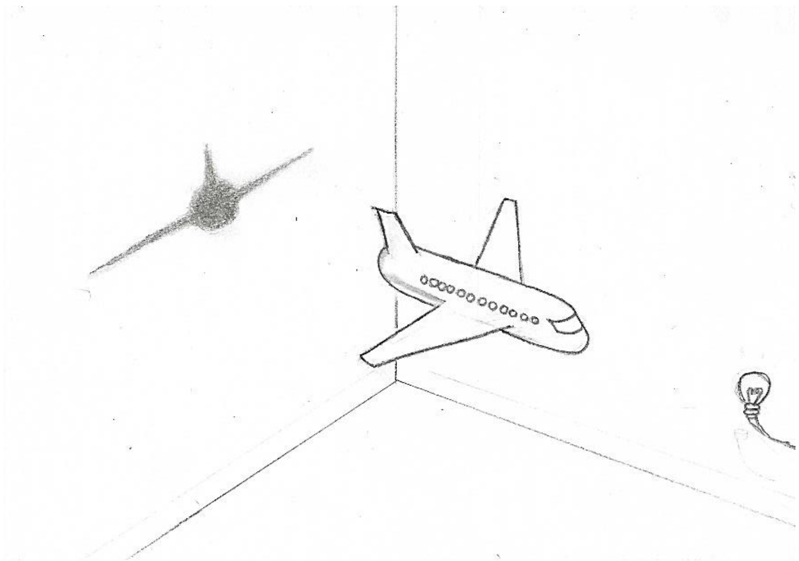
Raunak Banerjee

Most real-life situations are such that numerous factors, big or small, affect the final outcome. For example, when biologists study an ecosystem, there are many factors to consider from something as major as predator to prey ratio, to something seemingly as insignificant as types of microbes in the soil. It is extremely difficult for us to comprehend how all these variables simultaneously affect the final outcome, but data analysts have a trick up their sleeve. They use a dimension reduction technique called Principal Component Analysis (PCA) to identify a handful of the most critical factors that affect the outcome significantly. If you read about PCA, you will come across a lot of mathematical jargon like eigenvalue, eigenvector, standardization, variance, and more. In this article, we will extend our understanding on a more physical interpretation of how PCA reduces dimensionality, and we will define all vocabulary related to it.

Let say we need to describe the shape of an airplane. It has three dimensions, so using some values of x, y, and z variables you can describe the shape completely. But here is the twist. Assume your perception is that of a two-dimensional being that can only observe two-dimensional projections of an object, or in other words, shadows. You cannot comprehend three-dimensional objects, just like it is difficult for us to perceive four or more dimensions physically.

Here we have our "complex" three-dimensional object, an airplane. Let us say you are given a screen attached to a source of light where you can place the 3D object between them, flash the light, and observe the two-dimensional shadow. We have kept the distance between the light and the screen fixed and kept the object in the middle to prevent disproportional shadows from different orientations. In mathematical terms this is called standardization and it prevents misinterpretation due to a deformed shape of the shadow. Now shine the light from the front of the plane to project a shadow on the screen (Figure 1).

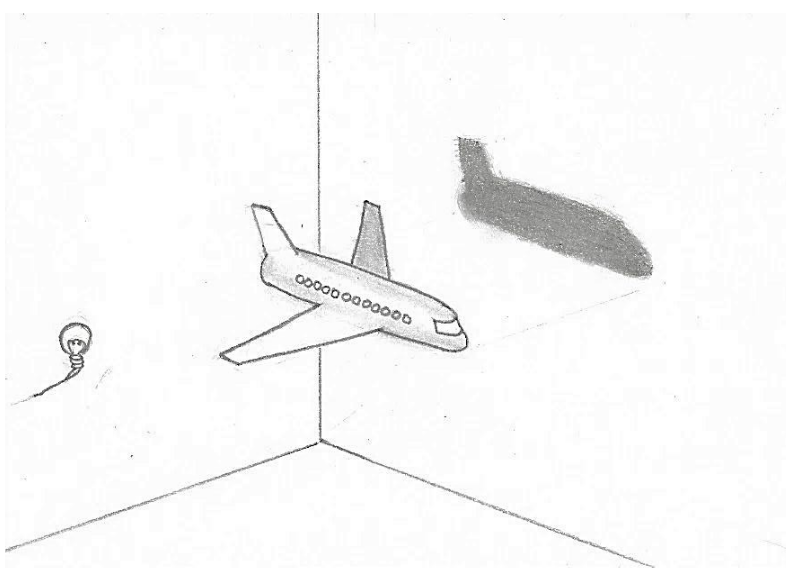
FIGURE 1: LIGHT FROM FRONT



From the shape of the shadow, would you be able to tell that this is an airplane? Probably not. As of now, it just looks like a ball with uneven protrusions.

Now let us reorient the plane in such a way that you get a shadow sideways (Figure 2).

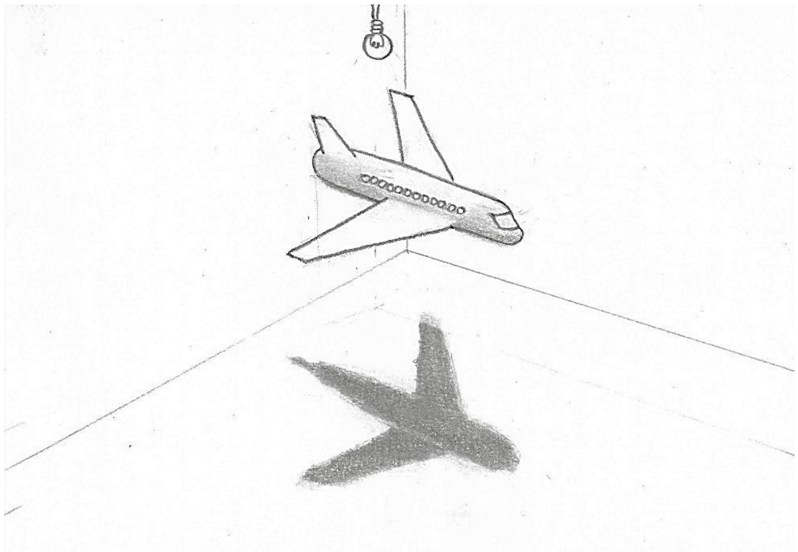
FIGURE 2: LIGHT FROM SIDE



This shadow helps a bit more, but it is hard to tell if it is a side view of a plane or the BFG's shoe, right?

Lastly, let us look at the last orientation where the light flashes from the top of the plane (Figure 3).

FIGURE 3: LIGHT FROM TOP



Now you can see where this is going. You can tell from this shadow that it most likely is an airplane. All these orientations prove to us that there will always be one screen-light plus irregular object setup that best describes the object in a reduced dimension. However, the same cannot be said for regular shapes like spheres. All screen-light orientations will result in a circular shadow. We will not be able to tell if it is a ball, cylinder, or cone if we look at only one shadow.

But what is so special about the last shadow? You as a two-dimensional character could tell that the object was an airplane because the last shadow captures the most variation compared to the others. In other words, there were less features that were being hidden by overlapping parts.

Now we are going to define some vocabulary. The eigenvectors are represented here as the axes of the shadow. The shadow itself is the projection of the data and the area of the projection is the measure of variation in the data. The eigenvalue is a measure of the span along the eigenvectors of the shadow. A higher variance means a higher eigenvalue and for our first principal component, we want an eigenvector with the highest eigenvalue. Since we have already crammed a lot of the information into the first principal component (as it has the highest eigenvalue), we can ignore the remaining principal components and plot our data, thereby achieving dimension reduction. If we wanted a 2D graph, we would take a second principal component along with the first and treat them as the x and y axes of a graph.

PCA aims to reduce dimensionality of a higher dimension space by projecting it to a smaller dimension subspace. The eigenvectors discussed here provide the direction of the axes of these smaller dimensional subspaces, whereas the corresponding eigenvalues decide how much variation (information) that particular projection will capture. The projection involving eigenvectors with the highest eigenvalues, captures the most variation of the original space and therefore is the most informative one (like Figure 3).

After using some matrix operations and other complex mathematical hocus pocus, PCA reduces the number of dimensions in the data so it is a simplified description of the shape without too much attention towards more finer details like, color, windows, etc. Whether it be dimensional reduction from 3 to 2 or from 100 to 2, the underlying mathematics will always be the same.

Unfortunately based on our initial proposition, we as two-dimensional characters cannot conceptualize beyond two dimensions. We are happy in our two-dimensional universe of physical and intellectual existence but should not lose sight of the fact that more than two variables are simultaneously manifesting certain complex phenomenon. For our understanding we need to depend on the best possible description within the dimension we can conceive and ignore finer details to perceive the big picture.