**“How does canonical correlation analysis (CCA) compare to principal component analysis (PCA)?**

Lets say you are trying to understand an event and you do this by taking observations about that particular event. You try to summarize the event in 10 variables and make 100 such observations. PCA tells you weather you really need 10 variables to completely describe the event or weather a lesser number of variables can completely describe it. We find these lesser variable by taking linear combinations of the existing 10 variables which are orthogonal to each other. You can visualize it by first taking a scatter plot in your 10 dimensional space and then finding a vector which captures the most amount of variance in the data. Now find another vector orthogonal to the first vector which again finds the maximum variance in the 9 dimensional space orthogonal to the first vector. The idea is to find k such vectors which completely explains the variance in the data.

Now suppose you are observing an event and you take a certain number of observations for the event. If we have 20 dimensions that were recorded for the given event, and we would like to see how well correlated the dimensions are we can divide the data into 2 sets of 10 dimension each.. Now we wish to understand how these two groups are related. We first take a vector in the 10 dimensional space and project all of the observations of group 1 on to it and take another vector and project all of the observations of group 2, now our objective is to find vector 1 and vector 2 such that the correlation between them is maximized. Ones these vectors are obtained we have to find another vector orthogonal to vector 1, similarly an orthogonal vector for vector 2 and the correlation between these vectors has to be maximized. In essence we are finding the new transformed space which are a linear combination of the existing variables for both events such that the correlation is maximized for individual dimensions.”