We will derive a notion of loss of information in the projected space.

Let us say that … be the 1st data point (in D- dimensional space) and be the direction of projection, then the linear combination i.e. gives us a scaler value (i.e. projected value). The scaler value is the projection of first data point … along the direction .

So, we can say that the projection of a d-dimensional data point = = (i.e. a scaler).

Similarly, `mean of the data in the projected space` would be:

Each data point, which is a d-dimensional vector in the original space, gets projected into the projected space.

The idea of information is equal to [variance](https://en.wikipedia.org/wiki/Variance). If, in a dataset, all the columns have constant values, then there is no information gain because every data point would be the same.