Use notes to help motivate method for assessing curve quality

**PCA Steps: (**[**Source is Statquest**](https://www.youtube.com/watch?v=FgakZw6K1QQ&t=406s)**)**

1. Compute means of features to calculate center of data
2. Shift center (all data points) to origin of graph
3. Initialize a random line through origin (Fit best line to all points)
   1. Project data points onto line
   2. Either (Equivalent solutions)
      1. Minimize the sum of distances from data points to the projected point on the line (distances are like magnitudes of residuals, buit sideways) (Like Least squares regression)
      2. Maximize the sum of squared distances from projected points on the randomly initialized line to the origin
4. After fitting, this line is PC1
   1. Slope of PC represents a ratio which explains how important one feature is over another for describing how data is spread out
      1. Generally, vector of best fit line will be composed of {features} number of ingredients which contribute differently to how data is spread (largest one is most important for that)
   2. The feature which is the major component of the ratio is the one which the data is more spread out on
   3. PCs are linear combinations of variables (Since they are vectors composed from the “importance” of features)
   4. With **Single Value Decomposition (SVD)**, you need to turn the slope vector into a unit vector (algebra), however you should be left with the same slope ratio (slope value) to make PC1 unit vector (AKA **Singular Vector** or **Eigenvector for PC1**)
   5. **Loading Scores:** The names given to each proportion used to make up the slope (num and denom)
   6. **Eigenvalue for PC1:** sum of squared distances of projected points to the origin on the best-fit line
   7. **Singular Value for PC1:** Square root of eigenvalue for PC1
5. For a 2D graph PC2 is the unit vector perpendicular to PC1
6. **Recall** that the fraction made by Loading Scores (LS) can be interpreted as:
   1. Slope = LS1/LS2 -> {Feature for LS1} is {Slope} times important as {Feature for LS2}
7. To make PCA plot, rotate everything so that PC1 is horizontal
   1. I think this can be done with a linear transformation on all points which make the Eigenvector for PC1 = [[1, 0] [0, 0]] or something like that
8. Make PCA scatter plot using the points projected on each PC per sample as coordinates
   1. Projected points come from a single sample, but are made from one of the features representing PC’s 1 and 2. Take both projected points and use them as coordinates (Makes more sense since each PC is an axis now, after rotation)
9. **Variation accounted for by PC1**: (Sum of square distances for PC1) / n-1
   1. n = number of samples
10. **Total variation** = variation of PC1 + variation of PC2
    1. **PC1** accounts for **{(variation of PC1 / Total variation) \* 100%} of variation around the PCs**
11. Generalize PCA by finding more PCs
    1. Number of PCs correspond to the number of features in data
    2. **Last PC can be found easily since it is the one perpendicular to all other PCs**

Notes:

* PCs = eigenvectors with largest eigenvalues
* PCs correspond to dimensions in which variation is the greatest