IE529 Computational Assignment 1: Solutions

November 27, 2017

1. • **Results for de-meaned data**
   Four components for de-meaned data are the following four column vectors:

0.3616	0.6565	-0.5810	0.3173
-0.0823	0.7297	0.5964	-0.3241
0.8566	-0.1758	0.0725	-0.4797
0.3588	-0.0747	0.5491	0.7511

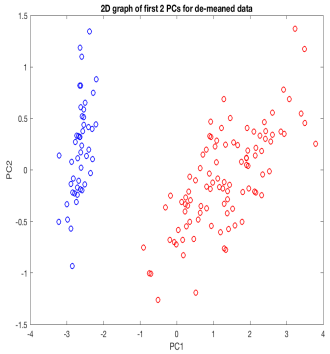
Figure 1: Matrix for four components

Four associated variances:

4.2248	0.2422	0.0785	0.0237

Figure 2: Variances for each components

From the results above, we can only select the first two components to represent the data.
Plot for the first and second components in 2D graph:



There are two virtually apparent clusters, where one is in blue and the other is in red. So we can conclude that the data represent two species.

- **Results for standardized data**
  Four components for standardized data are the following four column vectors:

0.5224	0.3723	0.7210	-0.2620
-0.2634	0.9256	-0.2420	0.1241
0.5813	0.0211	-0.1409	0.8012
0.5656	0.0654	-0.6338	-0.5235

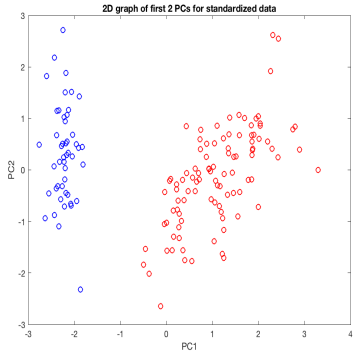
Figure 3: Matrix for four components

Four associated variances:

2.9108	0.9212	0.1474	0.0206

Figure 4: Variances for each components

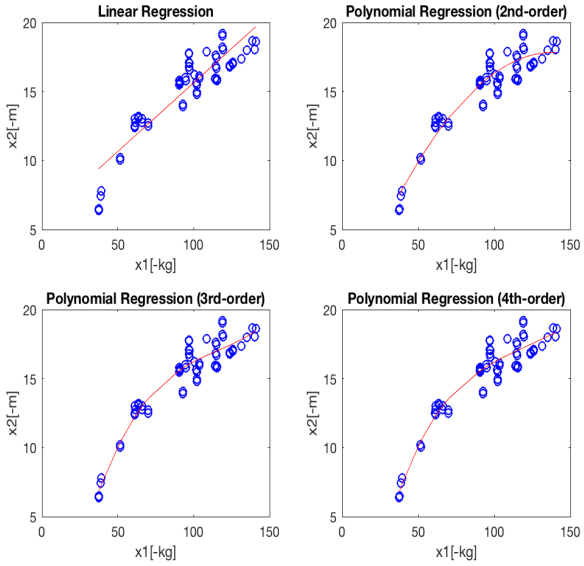
From the results above, we can also only select the first two components to represent the data.
Plot for the first and second components in 2D graph:



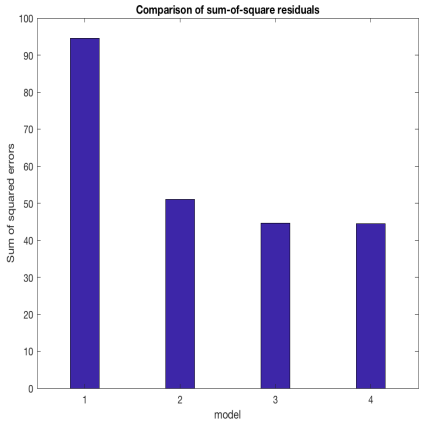
We can also see two virtually apparent clusters in the plot above, where one is in blue and the other is in red. So we can conclude that the data represent two species likewise.

2. (1) From the two figures below, we can see that the second order model fit the data best. Or we can say it's the "elbow". Models with order higher than two do not decrease the sum-of-square residuals too much and may cause overfitting.

   (2) The given data cannot be fitted by logistic regression model. Because neither of vectors contains only integers(labels).



(3) The sum-of-square of residuals for linear model, the 2^{nd} order model, the 3^{rd} order model are 94.5828, 51.0873, 44.5786, respectively.



Note: This problem asks you to write a simple program, so people who directly called some pre-defined functions will have 1 point off. The more approprate way is to compute weights by $(X^T X)^{-1} X^T y$ with your own code.