

# MAT 167 Project

Hardy Jones  
999397426  
Professor Cheer  
Winter 2014

## Part 1 Least Squares Approximation

When a problem is overspecified, that is there are more equations than unknowns, we cannot compute an exact solution (or fit the data). The Least Squares Approximation attempts to solve this by projecting the problem onto a suitable space and minimizing the error in this projection.

To put it in symbols, we cannot solve this equation:

$$\mathbf{A}\vec{x} = \vec{b}$$

So we approximate it and minimize the error in the process by solving this equation:

$$\mathbf{A}^T \mathbf{A} \vec{x} = \mathbf{A}^T \vec{b}$$

In order to fit the to a linear curve we want to solve this equation:

$$C + D\vec{x} = \vec{b}$$
$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

In order to fit the to a parabolic curve we want to solve this equation:

$$C + D\vec{x} + E\vec{x}^2 = \vec{b}$$
$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

In order to fit the to a cubic curve we want to solve this equation:

$$C + D\vec{x} + E\vec{x}^2 + F\vec{x}^3 = \vec{b}$$

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & x_m^3 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

When we use a linear fit, Figure 1, we get a *very* rough fit of the data. It seems to just barely provide enough information to find a given datum, and when it does, the error is still fairly large.

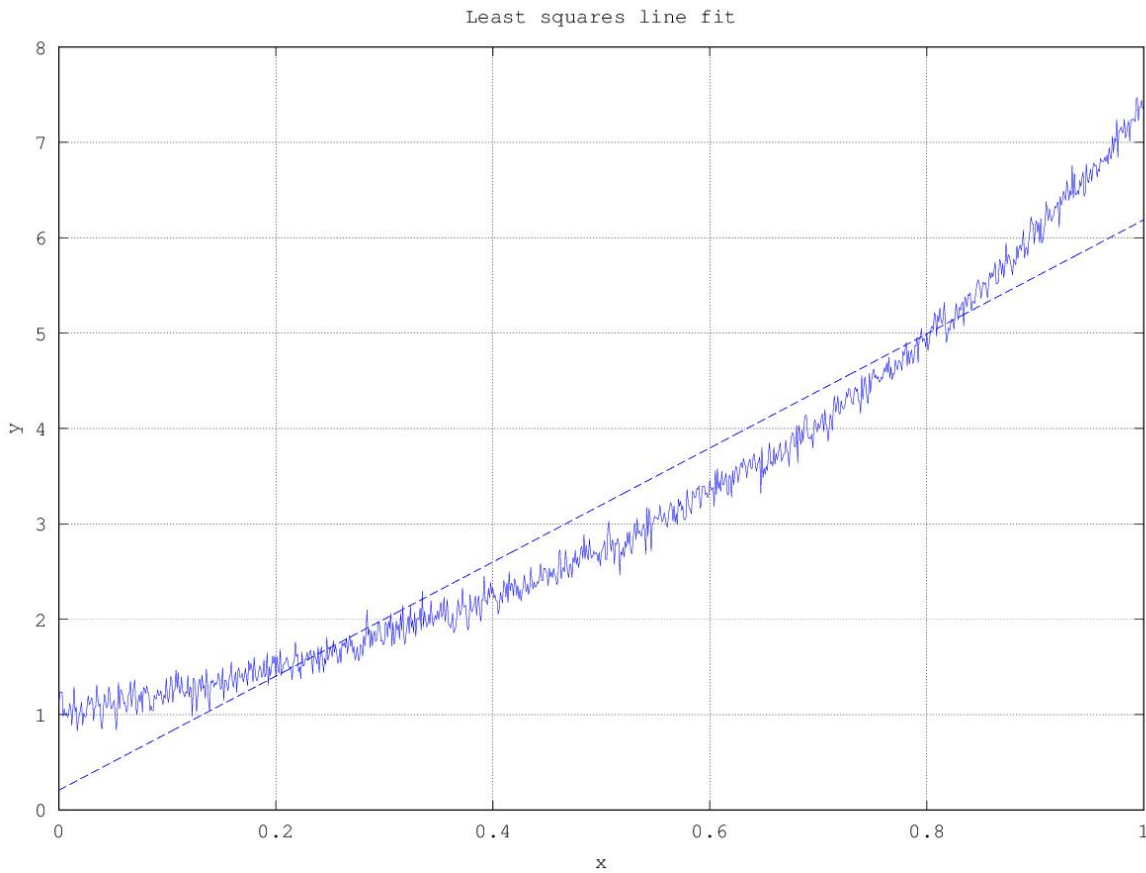


Figure 1: Linear Least Squares

When we use a parabolic fit, Figure 2, we see that our solution much better describes the problem. With the linear fit, it appears as if we're saying the, "The data goes this way". With the parabolic fit, it appears more as though we're saying, "the data is in this region, give or take". There are still some points where the curve is above and below the "meat" of the data. This is a much better approximation because our error has decreased dramatically, but we can still do better.

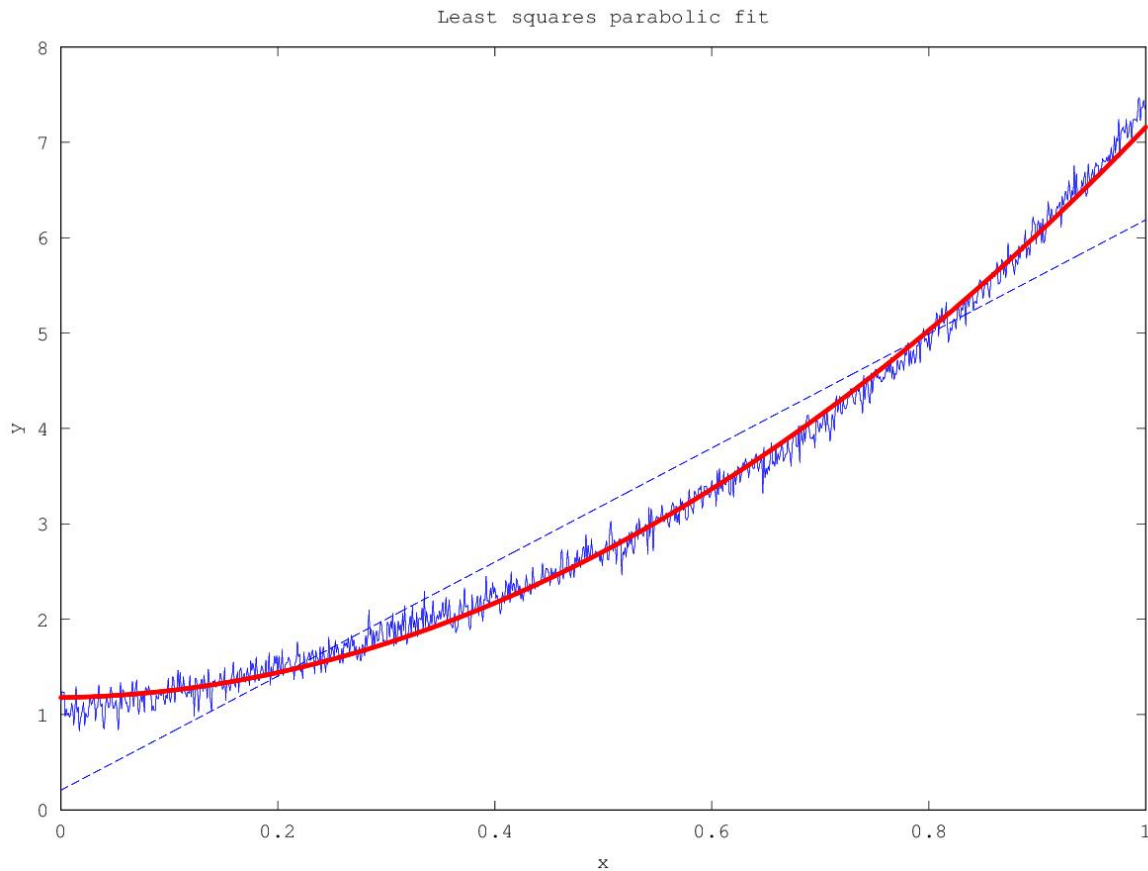


Figure 2: Parabolic Least Squares

When we use a cubic fit, Figure 3, we see a curve that very nearly describes the attitude of the data. It never seems to deviate too far from the “slope” of the data. The error has been reduced even further in comparison with the parabolic fit.

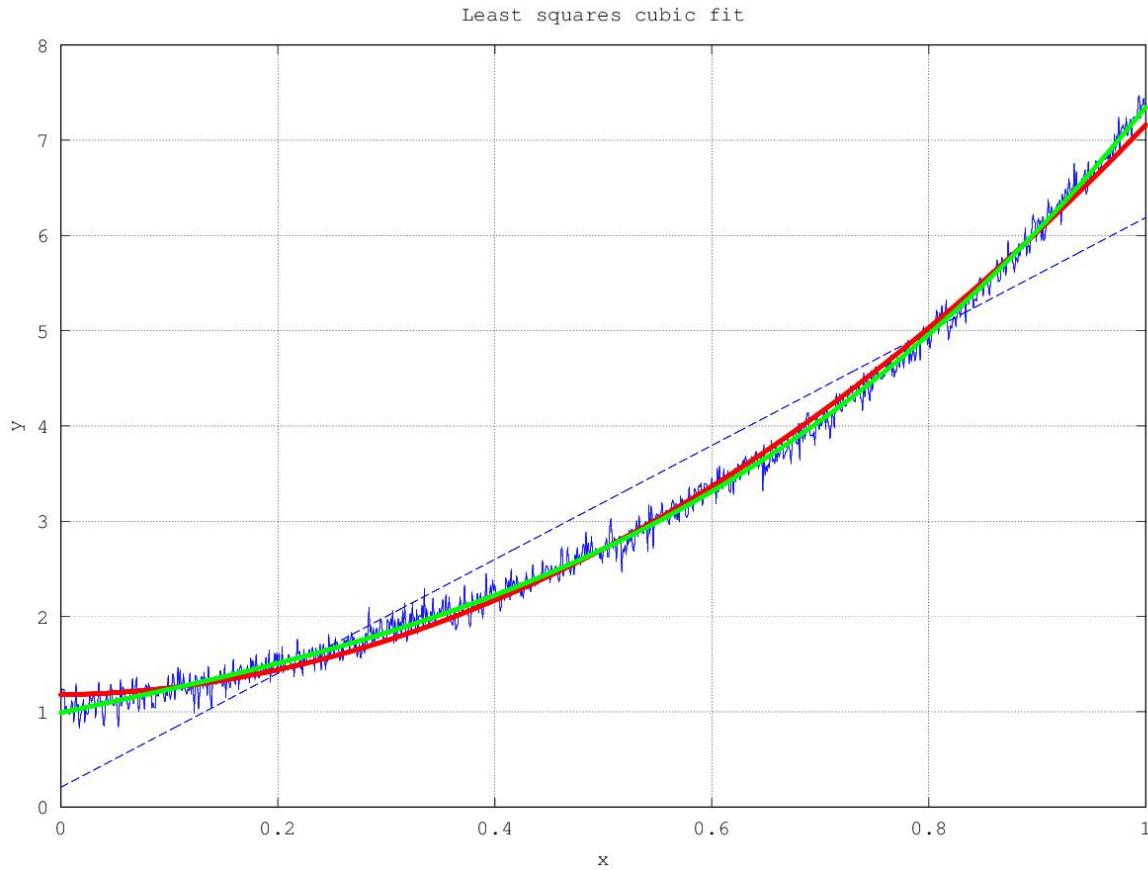


Figure 3: Cubic Least Squares

In each of these examples, the error is the vertical distance between a datum point and the curve.

## Part 2 Singular Value Decomposition

With large data sets a problem arises in storage and transmission of the data. Singular Value Decomposition (SVD) allows us to transform a large data set into a smaller set, and back again. One popular example is signal processing.

SVD says that we can decompose any  $m \times n$  matrix into three matrices:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

.

$\mathbf{\Sigma}$  is a diagonal matrix.

We get the columns of  $\mathbf{U}$  from the eigenvectors of  $\mathbf{A}\mathbf{A}^T$ .

We get the columns of  $\mathbf{V}$  from the eigenvectors of  $\mathbf{A}^T\mathbf{A}$ .

We get the values on the diagonal of  $\mathbf{\Sigma}$  from the non-zero square roots of the eigenvalues of  $\mathbf{A}\mathbf{A}^T$  and  $\mathbf{A}^T\mathbf{A}$ .

We have computed the singular values of  $\mathbf{\Sigma}$  in Figure 4

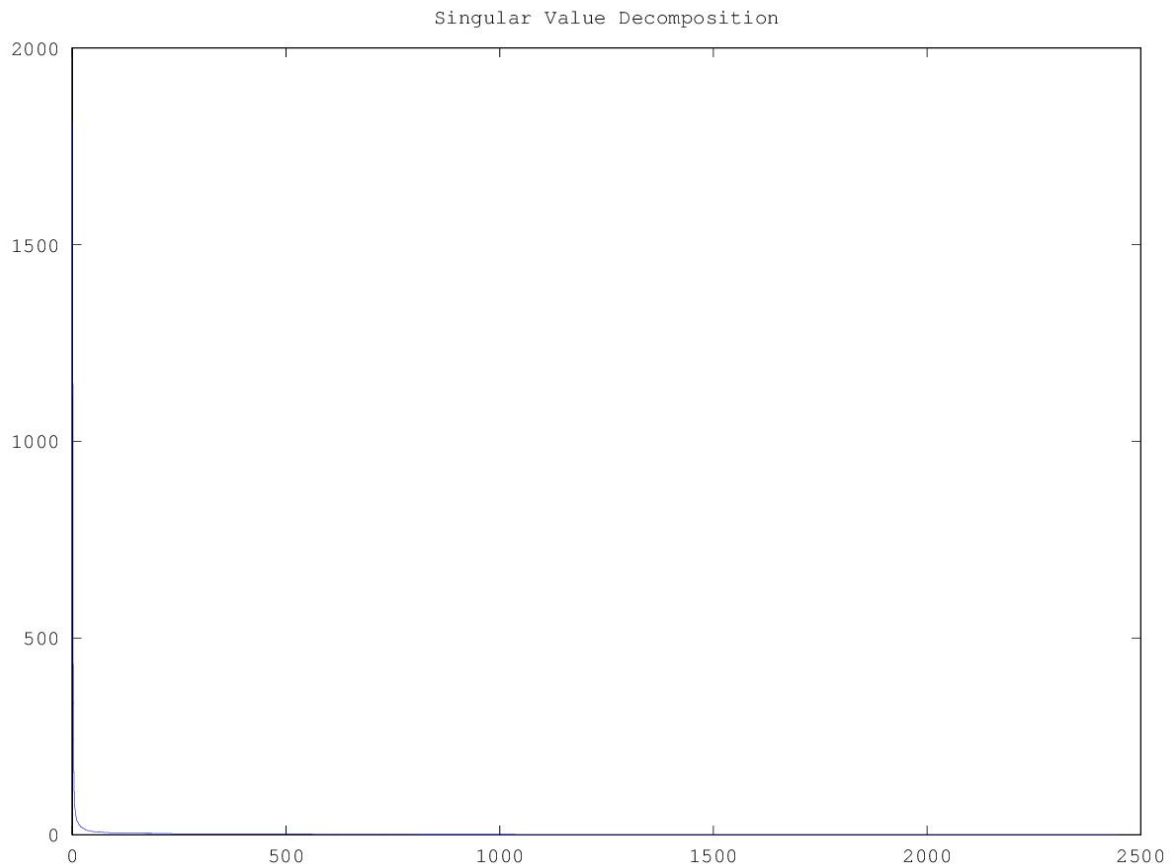


Figure 4: Values of  $\mathbf{\Sigma}$

The Rank-K approximation can be computed using the SVD. This approximation provides a better approximation of data than the Least Squares Approximation for larger values of  $K$ .

We have computed the Rank-K approximation for four different values of  $K$  in Figure 5.

We can see that for  $k = 1$  the approximation is not very indicative of the image.

With  $k = 6$ , it is able to make out that the image is of a person, though lots of the features are still missing.

With  $k = 11$ , this could actually pass as a viable picture. We can make out the facial features: eyes, mouth, beard, etc.

With  $k = 31$ , we have the picture at nearly normal quality. More details are able to be made out, such as the wrinkles in the shirt, and cheekbones.

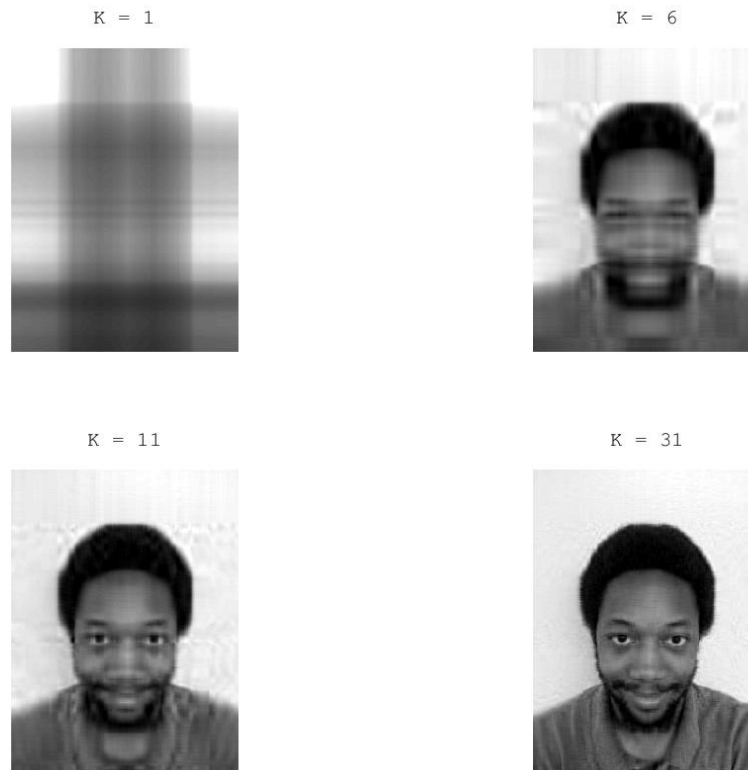


Figure 5: Rank K Approx.