In this problem you will practice generalizing the analysis of the FIM to nonlinear cases involving heteroscedastic error. The file `NonLinearPoints.txt` has as its first column the samples of the independent variable, the samples of the second column as the samples of the dependent variable, and the uncertainty in the measurements as the third column. Treat the third column as σ_i for each of your data points. Assume that your model is:

$$e^{-\theta_1 x}\sin(\theta_2 x)$$

and that your cost function is one half of the square of the l_2 norm (i.e. $\frac{1}{2}\sum_i (y_i - \hat{y}_i)^2$, where y_i is your data point and $\hat{y}_i$ is your prediction).

(a) Generate a contour plot depicting the cost surface of your model if you do not consider the heteroscedastic errors.

(b) Plot an ellipse defined using the eigenvalues and eigenvectors of the FIM evaluated at the minimum over the cost surface. Scale the ellipse to show that the local curvature of the cost surface around the minimum matches that defined by the FIM. You may use the following procedure:

- Compute the eigen-decomposition of the FIM:
- Assume the semi-axis lengths of the ellipse are given by

  $$a_i = \sqrt{\frac{c}{\lambda_i}},$$

  where λ_i are the eigenvalues of the FIM.
- Generate points on the unit circle
- Transform these circle points into ellipse coordinates using the following formula:

  $$\boldsymbol{\theta}(t) = \hat{\boldsymbol{\theta}} + \mathbf{V}\begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}\mathbf{u}(t).$$

  Where $\hat{\boldsymbol{\theta}}$ is the minimum of your cost surface, $\mathbf{V}$ is the matrix whose columns are the eigenvectors of the FIM, and $\mathbf{u}(t)$ are the points on the unit circle.
- Plot $\boldsymbol{\theta}(t)$.

(c) Generate a second contour plot depicting the cost surface of your model if you **do** consider the heteroscedastic errors. In this case, your cost will be:

$$\frac{1}{2}\sum_i \frac{1}{\sigma_i^2}(y_i - \hat{y}_i)^2$$

(d) Repeat part b using the weighted FIM: $J'SJ$, where S is the matrix whose diagonal elements are $\frac{1}{\sigma_i^2}$.

(e) In your own words, how does the FIM's ability to describe cost basins generalize to the nonlinear case?