The goal of this homework is to teach you about over fitting and cross validation.

You are to implement regularized polynomial curve fitting. Examples \((x_n, t_n), n = 1..N\)
\(w_i\) is the weight for \(x_n^i\), \((0 \leq i \leq 9)\)

\[
\begin{align*}
\mathbf{w}^* &= \arg\min_{\mathbf{w}} \left( \sum_{n} \left( \sum_{i=0}^{9} w_i x_n^i - t_n \right)^2 + \lambda \sum_{i=0}^{9} w_i^2 \right)
\end{align*}
\]

Let’s rewrite the above in matrix notation: \(\mathbf{X}\) is a \(10 \times N\) matrix, where \(N\) is the number of examples and the \(n\)th column of \(\mathbf{X}\) is \((1, x_n, x_n^2, \ldots, x_n^9)^T\). \(\mathbf{t}\) is the vector of outputs \(\mathbf{t} = (t_1, t_2, \ldots, t_n)^T\).

\[
\begin{align*}
\mathbf{w}^* &= \arg\min_{\mathbf{w}} \left( \| \mathbf{X}^\top \mathbf{w} - \mathbf{t} \|^2 + \lambda \| \mathbf{w} \|^2 \right)
\end{align*}
\]

This is solved by differentiation:

\[
\begin{align*}
\mathbf{X}(\mathbf{X}^\top \mathbf{w}^* - \mathbf{t}) + \lambda \mathbf{w}^* &= 0 \\
(\mathbf{X}^\top + \lambda \mathbf{I})\mathbf{w}^* - \mathbf{X} \mathbf{t} &= 0 \\
\mathbf{w}^* &= (\mathbf{X}^\top + \lambda \mathbf{I})^{-1} \mathbf{X} \mathbf{t}
\end{align*}
\]

- We will provide a training set (\textit{train.txt}) and a test set (\textit{test.txt}) Each row contains one value of \(x\) and the corresponding value of \(t\), separated by space. Use 10-fold cross validation to find the best choice of \(\lambda\) and report the loss on the test set.
- Write a roughly 3 page summary of what you did:
  - at least one plot
  - report the best value of \(\lambda\) and the test error.
- Note that you need to find a suitable way to discretize \(\lambda\).

We encourage you to work in groups of up to three. Every group is to do their own work. Same score for all members of a group.

Hint: First implement a simple split of the \textit{train.txt} into train and validation set. After that implement 10-fold cross validation.

Extra credit:

- Modify the above solution so that the bias term/weight is not regularized. Justify your answer.
• Show variance bars for the losses of the different holdouts.
• Implement leave one out cross validation and show that in some sense this is better than 10-fold cross validation.

Remember you to look at the test score only once to report results!
Strong suggestion: Do this homework in Matlab, Octave or R.