PHY381C HW4 (22111S)

Q1)

$$\mathcal{L}(\theta) = (X\theta - y)^T (X\theta - y) + \lambda \theta^T \theta$$

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left[(X\theta - y)^T (X\theta - y) + \lambda \theta^T \theta\right] = 0$$

$$= X^T(X\theta - y) + X^T(X\theta - y) + \lambda\theta + \lambda\theta = 0$$

Using Eq.(69) Nice. and Eq.(78)

$$= 2X^T(X\theta - y) + 2\lambda\theta = 0$$

$$\Rightarrow X^T(X\theta - y) + \lambda\theta = 0$$

$$X^T X\theta - X^T y + \lambda\theta = 0$$

$$(X^T X + \lambda \mathbb{1})\theta = X^T y$$

$$\boxed{\theta = (X^T X + \lambda \mathbb{1})^{-1} X^T y}$$

Q2)

For learning problems you must keep the testing and training sets separate.

If we mixed/shared data between the two sets we would be ensuring ourselves a good answer, but not necessarily a good model.

Q3)

At higher Reynolds number our model does a worse job of predicting the next time step accurately. For higher Reynolds number our system becomes more chaotic and therefore harder to accurately model.

Q4)

KNeighborsRegressor held up really well, even at higher Re. This might be because our system's local structure is well defined/ordered.

Linear didn't do great, I image because of the nonlinear system we're investigating it didn't have much of a chance.