EE 379K Questions

* Does a binary decision stump only look at one feature or can it look at a combination of several?
* If you have more features than instances, is it true that you can always get empirical risk to be zero? i.e. to overfit?
* The gradient points in the direction of greatest increase. Why is that intuitively the solution to the linear regression problem?
* An array of shape (13,1) is 2D. 13 rows and 1 column. An array of shape (13,) is 1-dimensional, with just 13 elements. Which should we use for linear regression?
* On a linear regression graph, what are the variables on the two axes?
* Can beta (in linear regression) be thought of as finding appropriate weights for linear combination of the features?