Homework 7 - Optimization

NE 697

Due July 25, 2017 at 1:00 pm

Traditional ridge regression uses a single regularization parameter to dampen components in the input space. This regularization parameter, α^2, is determined through a trade-off between model accuracy and our *a priori* belief that the model should be "smooth." The regression parameters, b, are found with a modified pseudo-inverse solution:

$$b = \left(X^T X + \alpha^2 I\right)^{-1} X^T Y.$$

Underlying components in the input space are dampened relative to their associated eigenvalues. Components with eigenvalues much larger than α^2 are passed to the model undampened, while those with eigenvalues close to or smaller than α^2 are dampened. This approach can be problematic when the input space contains significant (read: high eigenvalue) components that do not predict the output. Local ridge regression identifies different α^2 values for each component. The regression parameters are calculated according to

$$b = \left(X^T X + V\alpha_i^2 V^T\right)^{-1} X^T Y$$

where V is the matrix of right-singular vectors from the singular value decomposition. Effectively, $V\alpha_i^2 V^T$ is a diagonal matrix with the appropriate α_i^2 value in each row/column.

Use classical (gradient descent) optimization techniques to determine appropriate regularization parameters using cross-validation error as the objective function to predict the 35^{th} variable in the sim.mat data set. Compare the performance of traditional (single parameter) ridge regression to the local ridge regression method. Compare the results of the local ridge regression optimization (using filter factors) to what you would expect from a standard correlation analysis.

$$filterfactor_i = \frac{s_i^2}{s_i^2 + \alpha_i^2}$$

where s_i is the singular value of the i^{th} component.

Divide data into training, test, and validation data according to:

```
train = Data([1:500 1501:2000 3001:3500 4501:end],:);
test = Data([501:1000 2001:2500 3501:4000],:);
val = Data([1001:1500 2501:3000 4001:4500],:);

x_train = train(:,[1:34 36:end]);
y_train = train(:,35);
x_test = test(:,[1:34 36:end]);
y_test = test(:,35);
x_val = val(:,[1:34 36:end]);
y_val = val(:,35);
```

For this homework, prepare a written report in IEEE format. Include any plots and tables that will support your findings. Make sure you correctly label your figures and tables and refer to them in the text. Include an appropriate citation for the data, both in the text and in the list of references. Your report should include **at a minimum** an abstract, introduction, methodology, results (and discussion!), conclusions, and references. Note that the methodology section of this report (and every report!) should describe the algorithm that you're using – not the implementation in MATLAB. Include all your code in an appendix (single column) at the end of the report. Convert your report to .pdf before submitting it through Canvas.