ECE/CS 532 project proposal

# Scale invariant LASSO by parameter re-weighting and its application to neuroimaging data

Qihong Lu, Zhenyu Zhang, Jiaan Wang

Supervised by Dr. Robert Nowak and Dr. Laurent Lessard

## Brief Overview of Topic and Motivation

The use of machine learning algorithms to explore hidden information from fMRI (functional magnetic resonance imaging) data has been increasingly prevalent. For example, Lewis-Peacock & Postle (2008) established a classifier that reliably classify what kind of images (faces, places or object) that a given participant was viewing, based on the participant's pattern of neural activities on the cortex. By studying the "important" brain regions (in the unit of voxel) that helps the classifier to differentiate the image categories, neuroscientist gain substantial insights about the localization of neural representations of visual information. In particular, if one brain region is very informative of whether a given image is a face or not, it is reasonable to think that this brain region contributes to face processing.

We often determine which brain region is important by comparing their parameters assigned by the classifier. Typically, features that are more influential would have parameters with higher magnitude. However, data normalization or other scaling can change the magnitude of the parameters assigned to different brain regions, which makes the results harder to interpret. In this project, we hope to find a re-weighting scheme that stabilize the parameter estimation under normalization. That is, we hope to make the parameter estimation scale invariant.

**Have your initial investigations led to any interesting questions or challenges? What sort of machine learning or data analysis tasks do you think will arise in your project?**

We view this as a classification problem. Therefore, the performance of an algorithm will be evaluated by cross-validated classification error. Secondarily, the performance will also be accessed based on the “stability” of the parameter estimation. Namely, we want the difference of parameter estimation between raw data and normalized data to be small. However, it is unclear how to optimize error and “stability” simultaneously.

Re-weighting might to make the results harder to interpret. For example, one re-weighting scheme is to let the weights to be inversely proportional to the magnitude of the signals (Candès, Wakin & Boyd, 2008), which would change the magnitude of the parameters. With this reweighting scheme, a feature with a parameter of higher magnitude is not necessarily the “more influential”, compared to a feature with smaller parameter value. Therefore, it no longer make sense to think feature with bigger parameter value is more important. Evaluate the relative contribution of different feature seem to be an issue.

Finally, the main question and challenge is to figure out what kind of re-weighting scheme is best suited for understanding fMRI data. Although there are many proposed ways of re-weight the parameter, it is unclear which one is more appropriate to use in the context of neuroimaging data.

**Core Concepts:**

For our project, related topic covered in our course include: classification with the least square technique, regularization, singular value decomposition and cross validation.

Neuroimaging data are typically highly underdetermined, as there are more features than training examples. When acquiring fMRI data, the device measures the Blood-oxygen-level dependent signal after dividing the brain into tens of thousands of small voxels. The features of the data set are the voxels on the cortical regions. Typically, there are thousands of them. On the other hand, we typically have hundreds of experimental trials, which transfers the number of equations. Because there are many more features than number of equations, we prefer LASSO instead of other methods, such as ridge regression, for interpretability purpose.

Here, we provide a more preliminary envision of this project. To reduce the number of features, ideally, we would like to optimize the following:

$$\min_{\beta}\left||y - X\beta|\right| + \lambda\left||\beta|\right|_0 \qquad where\ \left||\beta|\right|_0 = number\ of\ nonzero\ weights$$

Here, X is the m by n matrix, where each column represents a feature (voxel) and each row represents an experimental trial. The true labels were encoded into m by one vector, y. We want to minimize the difference between X$\beta$ and y, with the fewest non-zero $\beta$. This turns out to be very difficult, but one can optimize the proxy formulated below instead (Candès, Wakin & Boyd, 2008), which is the LASSO procedure:

$$\min_{\beta}\left||y - X\beta|\right| + \lambda\left||\beta|\right|_1$$

We plan to use LASSO to perform the feature selection, which is L1-norm regularized least square problem. Moreover, as the project aims to find a re-weighting scheme allows for stable parameter estimation under normalization, we need to study different re-weighting procedures. For example, a reweighted LASSO can be formulated as follows (Candès, Wakin & Boyd, 2008):

$$\min_{\beta} \Sigma_{i=1}^{n} w_i \left|\left|y - X\beta_i\right|\right| + \lambda \left|\left|\beta_i\right|\right|_1$$

Here, w represents the weights for different parameters $\beta$. There are many ways of constructing the weights. For example, Candès, Wakin & Boyd (2008) proposed a re-weighting scheme to iteratively estimate the parameters and weights, and they let the weights to be inversely proportional to the magnitude of the signal. However, it is currently unclear to us if this is appropriate to application to neuroimaging data.

## Related Papers, Datasets, or Resources:

Our data came from the fMRI study conducted by Lewis-Peacock & Postle (2008). In this study, fMRI signals were recorded while the participants were viewing pictures of famous faces, familiar places or common objects. Previous results show that one can reliably decoding the category of picture that the participants were viewing based on their neural responses, characterized by Blood-oxygen-level dependent signal (Lewis-Peacock & Postle, 2008).

References

Candès, E. J., Wakin, M. B., & Boyd, S. P. (2008). Enhancing Sparsity by Reweighted ℓ 1 Minimization. *Journal of Fourier Analysis and Applications*, *14*(5-6), 877–905. http://doi.org/10.1007/s00041-008-9045-x

Lewis-Peacock, J. A., & Postle, B. R. (2008). Temporary Activation of Long-Term Memory Supports Working Memory. *Journal of Neuroscience*, *28*(35), 8765–8771. http://doi.org/10.1523/JNEUROSCI.1953-08.2008

Figueiredo, M. A. T., Bioucas-Dias, J. M., & Nowak, R. D. (2007). Majorization & Minimization Algorithms for Wavelet-Based Image Restoration. *IEEE Transactions on Image Processing*, *16*(12), 2980–2991. http://doi.org/10.1109/TIP.2007.909318