ECE 532: Project Proposal

Blake Mason, Ian Kinesella, Scott Sievert

October 18, 2015

1 LASSO extensions

Exploring LASSO more in depth than what's covered in class (on the syllabus, we're touching on LASSO in week 9, 3 weeks from now). It depends on how much we cover in class, but we can touch on more advanced topics too such as the

- LASSO method [5]
- group lasso [2]
- sparse group lasso [1]
- sparse overlapping set lasso [4]

We plan to investigate these models to fully understand why they work and design a lab around them.

2 Core concepts

In general, some measurements are known for an underdetermined system. LASSO finds a solution by adding a regularization term and finds a solution to the system

$$y = Ax$$

where $y \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$ where $m \ll n$. This system is underdetermined and in general, this equation has infinitely many solutions and can not be solved. However, by assuming certain structure in a unique x can be found.

What structure is assumed about x? Do we assume that x have very little energy? Do we assume that x has very few non-zero elements? Do we assume some mix of the two? These assumptions can be represented by different norms.

The energy in a system can be represented by the ℓ_2 norm and the number of non-zero coefficients can be represented by the ℓ_0 norm. Because ℓ_0 regularization is NP-hard, we replace the ℓ_0 norm with the ℓ_1 norm.

Then we can say one of the following:

$$\widehat{x} = \arg\min_x ||y - Ax||_2^2 + \lambda||x||_1$$
$$\widehat{x} = \arg\min_x ||y - Ax||_2^2 + \lambda||x||_2$$
$$\widehat{x} = \arg\min_x ||y - Ax||_2^2 + \lambda_1||x||_1 + \lambda_2||x||_2$$

ℓ_2 regularization tends to minimize the energy in the solution, leading to a solution that contains many small coefficients.

ℓ_1 regularization tends to give more sparse solutions. This regularization term can be seen as an approximation for a ℓ_0 regularization term, which minimizes the number of non-zero enteries.

It should be noted that including an ℓ_0 regularization term is NP-hard. Qouting [3]

> In general, solving for a given L0 norm is an NP hard problem thus convex relaxation regularization by the L1 norm is often considered

The takeaway is that the ℓ_1 regularization is a tractable problem and that mimics ℓ_0 regularization.

2.1 Goal

We want to implement these algorithms and completely understand both the mathematics behind them and their implementation, instead of relying on third part software (like cvxpy).

References

[1] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. A note on the group lasso and a sparse group lasso. *arXiv preprint arXiv:1001.0736*, 2010.

[2] Lukas Meier, Sara Van De Geer, and Peter Bühlmann. The group lasso for logistic regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):53–71, 2008.

[3] Morten Mrup, Kristoffer Hougaard Madsen, Lars Kai Hansen, and Informatics. Approximate l0 constrained non-negative matrix and tensor factorization.

[4] Nikhil Rao, Christopher Cox, Rob Nowak, and Timothy T Rogers. Sparse overlapping sets lasso for multitask learning and its application to fmri analysis. In *Advances in neural information processing systems*, pages 2202–2210, 2013.

[5] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.