The Solution Path of the Generalized LASSO : Application in Statistical Learning

Patrick Tardivel, institut de mathématiques de Bourgognes, Dijon, France

The linear regression model $Y = X\beta + \varepsilon$, where $Y \in \mathbb{R}^n$ is the response of the model, $X \in \mathbb{R}^{n\times p}$ is the design matrix, $\beta \in \mathbb{R}^p$ is the unknown parameter of the regression coefficients, and $\varepsilon \in \mathbb{R}^n$ is the random error, is one of the most popular models in statistics. The parameter β can be estimated via the generalized LASSO estimator, which is defined as a solution to the following convex optimization problem :

$$\min_{b\in\mathbb{R}^p} \left\{ \frac{1}{2}\|Y - Xb\|_2^2 + \lambda\|Db\|_1 \right\}, \tag{1}$$

where $D \in \mathbb{R}^{m\times p}$. The most common choice for the matrix D is the identity, making the penalty term the ℓ_1 norm, and the obtained estimator is the LASSO (acronym for "Least Absolute Shrinkage and Selection Operator") which is known to promote sparsity. Many other matrices D are also relevant for this estimator (the article [3] provides numerous examples). The generalized LASSO estimator depends on the regularization parameter λ ; we denote $\widehat{\beta}(\lambda)$ as a solution of problem (1). A classical approach in statistical learning to choose this regularization parameter is to minimize the residual sum of squares on a validation set $X^{\text{val}}, Y^{\text{val}}$:

$$\lambda > 0 \mapsto \|Y^{\text{val}} - X^{\text{val}}\widehat{\beta}(\lambda)\|_2^2$$

Minimizing this expression requires computing the solution path of the generalized LASSO estimator, i.e., the function $\lambda > 0 \mapsto \widehat{\beta}(\lambda)$. Students opting for this internship can explore the following points :

- Study the theoretical results of the article [2] concerning the solution path of the LASSO.
- Investigate the theoretical results of the articles [1, 3] regarding the solution path the generalized LASSO.
- Implement in R or Python algorithms for computing the solution path of the LASSO or generalized LASSO and apply these algorithms to real datasets.
- Redemonstrate that the mapping $\lambda > 0 \mapsto \widehat{\beta}(\lambda)$ is piecewise linear.
- Explore the solution path in other models (e.g., logistic regression model or graphical model).

Références

[1] Taylor B Arnold and Ryan J Tibshirani. Efficient implementations of the generalized lasso dual path algorithm. *Journal of Computational and Graphical Statistics*, 25(1) :1–27, 2016.

[2] Julien Mairal and Bin Yu. Complexity analysis of the lasso regularization path. In *Proceedings of the 29th International Conference on Machine Learning*, pages 353–360, 2012.

[3] Ryan J. Tibshirani and Jonathan Taylor. The solution path of the generalized lasso. *The Annals of Statistics*, 39(3) :1335 – 1371, 2011.