Journées SMAI MODE 2018

Optimal approximation for regularization and validation path

Eugene Ndiaye
LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France

Tam Le
RIKEN Center for Advanced Intelligence Project

Olivier Fercoq
LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France

Joseph Salmon
LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France

Ichiro Takeuchi
Nagoya Institute of Technology

Résumé. Various machine learning problem can be written as a minimization of a loss function f plus a regularization R calibrated by a positive hyperparameter λ *i.e.,* $\min_\beta f(X\beta) + \lambda R(\beta)$. A major weakness of these methods is the tuning of the regularization parameter λ where the difficulty are both theoretical (for statistical guarantees) as well as algorithmic (for numerical stability). A common approach is to choose the best parameter among the entire solution path on a given range. Unfortunately, these methods have a worst case complexity that is exponential in the number of parameter (Gartner et al 2012). In order to avoid this issue, approximation of the solution path was proposed and an optimal complexity was known to be $(1/\epsilon)$ (Giesen et al 2010). Surprisingly, (Mairal and Yu 2012) show an important improvement to $(1/\sqrt{\epsilon})$ for the special case of the Lasso. Their result was generalized in (Giesen et al 2012) with a lower and upper bound of $(1/\sqrt{\epsilon})$ and derive a generic algorithm by assuming a polynomial lower bound on the objective function. Following those ideas, (Shibagaki et al 2015) have proposed the approximation of the validation path which is more natural for selecting a hyperparameter that are guaranteed to achieve a validation error close to the optimal one.
We revisit the approximation and validation path theory in a unified framework under general regularity assumptions that are commonly verified in machine learning problems. We apply it to examples that encompass classification and regression problem and provide a complexity analysis along with optimality guarantees. We highlight the relationship between the complexity of the approximation path and the regularity of the loss function and show that the previously proposed complexity analysis can be too pessimistic or even insufficient in some situations.

Mots-clefs : Approximation path, parameter programming, cross validation

Références

[1] B. Gartner, M. Jaggi, C. Maria. *An exponential lower bound on the complexity of regularization paths*. Journal of Computational Geometry, 2012.

[2] J. Giesen, M. Jaggi, S. Laue. *Approximating Parameterized Convex Optimization Problems*. Annual European Symposium on Algorithms (ESA), 2010.

[3] J. Mairal, B. Yu. *Complexity analysis of the lasso regularization path*. ICML, 2012.

[4] J. Giesen, J. K. Muller, S. Laue, S. Swiercy. *Approximating concavely parameterized optimization problems*. NIPS, 2012.

[5] A. Shibagaki, Y. Suzuki, M. Karasuyama, I. Takeuchi. *Regularization path of cross- validation error lower bounds*. NIPS, 2015.