l_1 regularization path for functional features

Manuel Loth, Philippe Preux*

March 2, 2009

We consider the LASSO problem, and the LARS algorithm to solve it. In this setting, we consider that data are represented by a (possibly large) amount of features. These features may be generated automatically using a kernel function; without loss of generality, for the sake of exemplification, we consider Gaussian kernels here. Typically, N data will then be represented by N features, among which the most important are retained in the expansion corresponding to a certain value of the regularization parameter.

Classically, to that end, the parameters θ of the kernel have to be set *a priori* (θ are the center, and the covariance matrix for a Gaussian kernel). To take into account several such parametrizations, each of them gives rise to a different set of features. So each data is really represented by N times the number of parametrizations; this can obviously be very large. Moreover, setting such parameters is generally left unmentioned in most papers, though an important in any practical application.

Here, we propose to consider functional features, in which a feature is a function of θ, which value is not set *a priori*, thus remains free, within certain bounds. Then, we propose ECON, an algorithm which is an adaptation of the LARS. ECON deals with such functional features; that is, while riding the l_1 regularization path, ECON searches for the best (feature, parametrization) couple to add to the current expansion. The feature remains available during further iterations, to join the expansion with an other parametrization. W.r.t. LARS, a shortcoming of ECON is that while the original LARS algorithm finds an exact solution of the minimization problem, we have to resort to a local optimum.

However, ECON has several interests:

- the parameters are automatically tuned, which is a nice property to get algorithms involving as little human expertise as possible;
- the set of parametrizations is not restricted *a priori*;
- for each feature, all parametrizations within a given domain are considered, and the best(s) enter(s) the expansion;
- experiments show very good performance of ECON on regression problems: state of the art performance are reached without any hand-tuning;
- care have been taken to avoid overfitting, and we clearly observe that the number of features being involved in an expansion generally saturates and that extra iterations do not involve more features;
- the situation varies from a problem to an other, but in many cases, very sparse solutions produce lower test error than previously published results, with less sparse expansions;
- CPU time requirements are far from extremely heavy; basically, ECON runs an order of magnitude slower than the LARS, but this has to be balanced with its advantages.

Current and future works concerning ECON concern a better understanding of the way ECON works, and an analysis of the expansions it yields. We will also evaluate it on supervised classification tasks, and we wish to use it as a function approximator for optimal control (both approximate dynamic programming and reinforcement learning). We also wish to exploit the fact that the selected features provide some insight into the structure of the problem. Finally, the implementation may be improved to get better CPU performances.

A more thorough description of ECON is available as INRIA Research Report RR-6794, available at the url: `http://www.grappa.univ-lille3.fr/~ppreux/papiers/RR-6794.pdf`.

*SequeL team-project, INRIA Lille Nord-Europe, & Computer Science Lab. of the University of Lille, Villeneuve d'Ascq, France; `firstname.name@inria.fr`